NONSTATIONARY NORMAL FORMS FOR ANOSOV DIFFEOMORPHISMS AND HYPERBOLIC SKEW PRODUCTS

Thesis by

David A. DeLatte

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1991

Typeset by $\mathcal{AMS}$-TEX

ABSTRACT

The central theorems in my thesis are generalizations of theorems due to G.Birkhoff, S.Sternberg, and J.Moser on local normal forms for invertible mappings. We will consider smooth, area preserving, Anosov diffeomorphisms of the two dimensional torus, $\mathbb{T}^2$. These are among the most fundamental examples of dynamical systems which exhibit extremely complicated (chaotic) behavior. Some geometric consequences and applications to rigidity phenomenon are also explored.

Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a smooth, area preserving mapping for which the origin is a hyperbolic fixed point. Birkhoff considered the formal power series of f at the origin and showed that there is a formal change of coordinates, h, which satisfies $f \circ h = h \circ g$ where g is of the form $g(x, y) = (\lambda x \Phi(xy), \lambda^{-1} y \Phi^{-1}(xy))$ and where λ is the (real) eigenvalue of the linear part ($|\lambda| > 1$) and $\Phi(xy) = 1 + \phi_1 xy + \phi_2(xy)^2 + ...$ is a formal series with $\phi_i \in \mathbb{R}$. The map g is called the local normal form for f. S.Sternberg showed that if the function f is C^∞ then h and hence g can be chosen to be C^∞ also. J. Moser was able to show that if f is analytic then Birkhoff's formal series for h and g converge in a small neighborhood of the origin. Note that the hyperbolae $xy = constant$ are invariant curves for g. One may introduce hyperbolic coordinates (c, θ) where θ, the hyperbolic angle, describes the position on the hyperbola $xy = c$. These coordinates give a clear understanding of the local behavior of f; specifically f shifts points along these

local hyperbolae.

These theorems are generalized by eliminating the necessity of working at a fixed point for the map. Consider a smooth, area preserving, Anosov (i.e., hyperbolic at each point) diffeomorphism ,f, of the two dimensional torus, $\mathbb{T}^2$. There exists a family of local coordinate changes h_p, $p \in \mathbb{T}^2$, which transform f_p (the local representation for f) into the normal form g_p (i.e.$f_p \circ h_p = h_{f(p)} \circ g_p$). Furthermore h_p and g_p are continuous in p.

The first step of the proofs in both the C^∞ and analytic cases is to establish a nonstationary version of the formal theorem above. From the formal solution one can construct a C^∞ representation for h_p which is area preserving and satisfies the conjugacy equation above in a neighborhood of p. In the analytic case a majorization scheme is employed to demonstrate the convergence of h_p and g_p. One should also be able to use a rapidly converging iteration method instead of majorization. Our proofs do not fully exploit the fact that the manifold is the torus (compactness and two dimensionality are used). The theorem above holds if we replace the torus with a fibre bundle which has a compact base and two dimensional fibres and the mapping with a hyperbolic skew product transformation.

Since the hyperbolae $xy = c$ are "preserved" by the nonstationary normal form, hyperbolic coordinates are available in this case also. The map f_p takes the simple form $g_p(c,\theta)_p = (c, \theta + log\lambda_p\Phi_p(c))_{f(p)}$. One can interpret g_p as a nonstationary hyperbolic twist. The higher order terms of the second coordinate of g_p form 1-cocycles (in the sense of group cohomology for $\mathbb{Z}$ actions). Let $\phi_{1,p}$ denote the first nonlinear term of the normal form for Φ_p. By integrating $\phi_{1,p}$

over $\mathbb{T}^2$ one obtains a global invariant of the dynamical system $(f, \mathbb{T}^2)$.

CHAPTER I

HISTORICAL DEVELOPMENT OF THE THEORY OF NORMAL FORMS

The theory of local normal forms has undergone extensive development over the last century. Consider a smooth dynamical system described by a system of differential equations or a smooth mapping. Since the singularities (also known as fixed points, stationary points) have significant impact on the trajectories of the system, one would like to understand the local behavior of trajectories near these points. The goal of local normal form theory is to simplify the local expressions for a dynamical system at a singular point by making a smooth change of coordinates. One hopes that this reduction in complexity of the local representation of the system leads to a greater understanding of the structure of orbits near the fixed point and possibly of the global structure as well.

The exact nature of the normal form varies considerably. A formal solution, where one considers power series without regard for convergence, or a smooth (up to C^∞) solution depends upon resonances (algebraic relations) between the eigenvalues of the linearized system. Analytic solutions require additional assumptions for the eigenvalues such as diophantine conditions, or restrictions on the magnitudes. Needless to say the best possible normal form theorem is one in which the normal form is a linear map. Unfortunately many interesting systems cannot be reduced to linear form. For example any system which preserves volume has resonances, so linearization is impossible.

As with so many ideas in the modern theory of dynamical systems, one finds

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

the origins of the theory of normal forms in the work of Henri Poincaré. In his paper, [P1], and also in his thesis, [P2], Poincaré studied systems of differential equations related to problems in celestial mechanics. He was interested in finding integrals of motion for these systems, and this lead him to consider questions about normal forms.

Since translations do not affect the shape of trajectories there is no loss of generality in formulating his theorem with the assumption that the stationary point is the origin. Let $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and let $F_1(X), \ldots, F_n(X)$ be power series without constant terms which converge in some open neighborhood of the origin. The autonomous system

$$\dot{x}_i = F_i(x) \quad i = 1, \ldots, n \tag{1.1}$$

has a stationary point at $X = 0$. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of the matrix corresponding to the linear part of F, i.e., $\left(\frac{\partial F_i(X)}{\partial x_j}\Big|_0\right)$. Poincaré proved the following theorem about these systems.

Theorem. *Using the notations from above, if the following conditions hold*

(1) *The eigenvalues $\lambda_1, \ldots, \lambda_n$ are distinct.*

(2) *There exists a line through the origin in the complex plane such that all eigenvalues, λ_i, lie on one side on this line.*

(3) *For any n-tuple of nonnegative integers $(k_1, \ldots, k_n)$ with $\sum_{j=1}^{n} k_j > 1$ we have that $\lambda_i \neq \sum_{j=1}^{n} k_j \lambda_j$.*

then there exists an invertible change of coordinates which transforms the system (1.1) into

$$\dot{y}_i = \lambda_i y_i \quad i = 1, \ldots, n \tag{1.2}$$

There are several things to notice about this theorem. First, a linear map is the simplest possible normal form. The objective of understanding the trajectories near the fixed point is completely attained in the normal coordinates since the differential equation is linear and can be solved explicitly. Conditions (1) and (3) are the nonresonance conditions mentioned above. Condition (2) is necessary for analyticity of the change of coordinates. In the case of real eigenvalues, condition (2) requires uniform contraction or expansion for the linear part of the solution of the differential equation.

The proof of Poincare's theorem has two distinct parts. First one considers all functions in the theorem as formal power series. The change of coordinates must satisfy a (formal) differential equation, which places algebraic conditions on the terms of this series. One sees, by an inductive argument, that each of these terms is uniquely defined. (This is where the nonresonance conditions are crucial.)

Next one shows that the formal series obtained in the first step converge in a small neighborhood of the fixed point. One may use the method of majorants to achieve this result, i.e., one finds a positive series which is known to converge, whose terms are larger in magnitude than the terms of the coordinate change.

In the case described in Poincare's theorem there is no question that the normal form is the simplest possible. One may ask how simple we can expect the normal form to be in general. It has been known for some time that smooth linearizaton is not possible when resonances (additive relations as in condition (3)) occur. Resonances are an algebraic obstruction at the formal level to linearization at a fixed point.

Questions related to normal forms appear naturally in the study of celestial mechanics and other areas of interaction with physics. Lyapunov [L], Horn [H], and many other mathematicians considered normal form problems in various guises. Dulac [D] refined the formulation of the normal form problem and proved a nonlinear version of Poincaré's theorem when conditons (1) and (3) are dropped. He found that if one includes the resonance terms in the normal form, the normalizing transformation is well defined as a formal series. Thus a canidate for the nonlinear normal form was discovered. Majorization was employed once again to show convergence of these formal series.

There has been parallel development in the theory of normal forms for maps. Many similarities between these two cases exist; however, there are some important differences as well. While resonances still determine the normal form for maps, the conditions on the eigenvalues are multiplicative instead of additive. This is explained by the exponential map which relates the vector field, F, with the solution of the system of differential equations. It is the solution of the differential equation which is in direct corespondence with the map. The general methodology of proofs is quite similar in both cases. Typically one constructs a formal solution then applies some convergence argument.

Lattés, [La], proved several theorems for two dimensional maps whose linear parts had eigenvalues which were less than one in magnitude, i.e. contraction mappings. In the 1920's Birkhoff did considerable research on the normal form problem for maps. In one of many theorems in [Bi2], Birkhoff considered normal forms for an area preserving tranformation of a surface at a hyperbolic fixed point. He established the following theorem.

Theorem. *Let* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *be an invertible area preserving mapping which* *leaves the origin fixed. Suppose that $Df|_0$ has eigenvalues λ, λ^{-1} with $\lambda > 1$ (hyperbolicity), then there exists a formal area preserving coordinate change h which conjugates f with its normal form, $g(x,y) = (\lambda x\Phi(xy), \lambda^{-1}x\Psi(xy))$, i.e., f, h, and g satisfy $h^{-1} \circ f \circ h = g$.*

Remarks

(1) The power series Φ and Ψ only contain terms involving the product xy.

(2) Resonances for diffeomorphisms occur when $\lambda^{i_1}\lambda^{i_2}\cdots\lambda^{i_n} = \lambda^j$ for some $(i_1,\ldots,i_n) \in (\mathbb{Z}^+)^n$, $j \in \{1,\ldots,n\}$. So the relations $\lambda^{i+1}(\lambda^{-1})^i = \lambda$ and $\lambda^i(\lambda^{-1})^{i+1} = \lambda^{-1}$ for $i = 1, 2, \ldots$ are the resonances for the map above. Note that the terms which are present in the normal form exactly correspond to these resonances. Thus in the first coordinate the terms $x^{i+1}y^i$ are included in the normal form and in the second coordinate terms of the form x^iy^{i+1} are present.

(3) Preservation of area forces the eigenvalues of f to be reciprocals. The same holds for Φ and Ψ.

Proof. Consider the equation $|Jac(g)| = 1$. This becomes

$$\begin{aligned} 1 &= (\lambda\Phi + \lambda xy\Phi'(xy))\left(\lambda^{-1}\Psi + \lambda^{-1}\Psi'(xy)\right) \\ &\quad - \left(\lambda x^2\Phi'\right)\left(\lambda^{-1}y^2\Psi'(xy)\right) \\ &= \Phi\Psi + xy\Phi'\Psi + xy\Phi\Psi' \\ &= \frac{d}{dz}(z\Phi\Psi) \qquad \textit{where } z = xy \end{aligned}$$

The only smooth solution for this differential equation with the correct value at the origin requires $z\Phi\Psi = z$, thus $\Phi\Psi = 1$.

(4) Observe that the curves $xy = constant$ are invariant under the action of the mapping in normal form. These are formal analogues of the first integrals sought by Poincaré.

(5) We will see that the resonance terms of the conjugating map, h, are not uniquely determined by the conditions in the theorem. Similarly the higher order terms of the normal form are nonunique.

Sketch of proof of the theorem:.

Writing formal power series for f, h, and g and assuming that the linear part of f is in normal form, the first coordinate of $f \circ h = h \circ g$ provides the equation

$$\lambda \sum_{ij} h^1_{ij} x^i y^j + \sum_{ij} f^1_{ij} \left(x + \sum_{kl} h^1_{kl} x^k y^l \right)^i \left(y + \sum_{kl} h^2_{kl} x^k y^l \right)^j$$

$$= \lambda \sum_i \Phi_{ii} x^{i+1} y^i + \sum_{ij} h^1_{ij} \left(\lambda \Phi(xy) \right)^{i-j} x^i y^j$$

Any nonresonance term $x^k y^l$, $k - l \neq 1$, requires

$$\lambda h^1_{kl} + f^1_{kl} + l.o.t. = \lambda^{k-l} h^1_{kl} + l.o.t.$$

where l.o.t. stands for lower order terms. Thus we see that

$$h^1_{kl} = \left(\lambda^{k-l} - \lambda \right)^{-1} \left(f^1_{kl} + l.o.t. \right)$$

The coefficient on the right hand side is nonzero since we are not in resonance. The expression above can be used to define the conjugacy inductively (on the order, $k + l$, of the terms). In the resonance case the terms, $h^1_{k+1\,k}$, on both sides of the equation have the same coefficient, λ.

$$\lambda h^1_{k+1\,k} + f^1_{k+1\,k} + l.o.t. = \lambda \Phi_{kk} + \lambda h^1_{k+1\,k} + l.o.t.$$

The equation defines Φ inductively. Note that thus far no condition has been placed on $h^1_{k+1\,k}$.

Similar considerations define the coefficients for the second coordinate, h^2_{kl} and Ψ_{ii} . Once again the resonance coefficients, $h^1_{k\,k+1}$, are free from conditions. We will see later that preservation of area can be satisfied by choosing the resonance terms appropriately but these terms still will not be uniquely determined.

Birkhoff's formal solution waited for Moser, [M1] to show that it was the series of an analytic funtion. Moser's proof uses the formal invariant curves $xy = c$ for the normal form in an essential manner. He defines a cannonical (area preserving) system of formal differential equations whose solution, $s(x, y, t)$, interpolates the map f, i.e., $s(x, y, n) = f^n(x, y)$ for all integers, n, such that $f^n(x, y)$ is sufficiently close to the origin. Then, Moser proves the convergence of the normal form for this formal system of differential equations, from which the result for maps follows. Soon after Moser published his paper, Siegel provided a majorization proof which did not require an excursion into differential equations. Our main theorem in chapter 5 uses a method modeled on Siegel's proof.

Siegel made other major contributions to the theory of normal forms including the first proof of the famous center problem. His theorem is a linearization result for mappings in the one dimensional elliptic ($|\lambda| = 1$) case. The proof required delicate estimates on certain "small divisors" (expressions like $\left(\lambda^{k-l} - \lambda\right)^{-1}$) found in coefficients of the formal solution for the linearizing coordinate change.

Kolmogorov suggested that such a theorem could be proven by a "generalized Newton's method" which lead to the development of KAM (Kolmogorov-Arnold-Moser) theory. Pliss, and others proved normal form theorems using KAM techniques. In his article [Br1], Brjuno (also spelled Bryuno, Bruno by different english translators) masterfully surveyed the state of normal form theory up to 1972 and extended and refined many of the techniques and theorems discussed above. In addition this paper provides a general framework in which one can understand and classify normal forms at a fixed point.

In the 1950's a separate line of normal form results for maps was discovered by Sternberg, [S1]-[S4]. He studied normal from questions for C^∞ and C^k maps, and developed the machinery necessary to find smooth realizations of the formal series of Birkhoff and others. The first step in this process was established in a theorem from E. Borel's thesis, [Bo]. Borel showed that there is a group homomorphism from the local group of C^∞ functions which fix the origin, onto the group of formal power series without constant terms. Thus any formal series could be realized by a C^∞ function. Sternberg used this to show that linearization is possible in the C^∞ category in the absence of resonances, and that Moser's theorem for the case of a hyperbolic fixed point holds in this class of maps as well. We will employ suitable versions of Sternberg's methods to prove the main theorem in chapter 6.

Sternberg is also responsible for approaching the normal form problem from the prespective of infinite dimensional Lie algebras. In [S4] and [S5] he indicates his preference for this point of view. Later Chen, in [C1] and [C2] gave a new proof of Sternberg's result using Lie algebraic techniques. The work of Sternberg and Chen encourages the philosophy that one can always realize formal local normal form results in the C^∞ category. This paradym will be seen to hold in the nonstationary case as well.

The philosophy expoused above for C^∞ maps is in contrast to the state of affairs in the analytic case. While there is no difference in the construction of the formal series, convergence seems to depend in a subtle way on the higher order terms as well as the resonances in the linear part. For example, consider the case of a two dimensional map whose eigenvalues are λ and λ^{-1}. As mentioned previously, if we also assume that the map is area preserving, then there are at least two known methods to prove convergence. However, at the present time no convergence proof exists for the question above without assuming the preservation of area.

There are still many other interesting open problems in the theory of local normal forms at a fixed point. Bruno's survey [Br1] as well as his book [Br2] are good resources for some of the current ideas in this theory. These works also provide a long list of references. It is also noteworthy that this theory has many applications to problems in Hamiltonian systems, celestial mechanics, stability theory, bifurcation theory, fluid dynamics and a variety of other pure and applied areas. A. Katok has used some of these methods in rigidity theory for "large" (e.g. $\mathbb{Z}^n$ or $\mathbb{R}^n$) group actions.

In the chapters which follow we will develop a theory for nonstationary normal forms generalizing the theorems of Birkhoff, Moser, and Sternberg. The term nonstationary indicates that we no longer require the point at which the normal form is found to be stationary (fixed). Compactness of the manifold in question (the two dimensional torus) will be instrumental in dealing with this difficulty. To generalize the hyperbolicity condition to the nonstationary case we will consider Anosov diffeomorphisms, which are hyperbolic at each point. Preservation of area is still required, and as in Moser's case it is important in the convergence proof.

In the final chapter we will place the methods developed in this work in their natural setting. Specifically we will study skew product dynamical systems and interpret the normal forms as cocycles defining a system of smooth invariants.

CHAPTER II

PRELIMINARIES

In this chapter we will attempt to place our results in context by recalling some of the requisite information from the theory of dynamical systems. We begin with a discussion of some definitions and theorems which will clarify the developments of subsequent chapters. For a more complete survey of this material the reader is refered to [An], [K], [M], or [Sh].

Let M be a compact manifold and f, a diffeomorphism. One says that the pair (f, M) is a dynamical system. Two dynamical systems, (f, M) and (g, N), are called equivalent or conjugate if there exists an invertible mapping, h, which satisfies $f \circ h = h \circ g$. A point $x \in M$ such that $f(x) = x$ is called a fixed or stationary point for f. A point which is not fixed will be referred to as a nonstationary point. The classical normal form theorems are local conjugacy results for a neighborhood of a fixed point.

Definition. *A dynamical system, (f, M) is called Anosov if there is a splitting of the tangent bundle TM into invariant subspaces*

$$TM = E^s \oplus E^u$$

$$df(E^s) = E^s$$

$$df(E^u) = E^u$$

and for which there exist constants $c > 0$ and $\lambda > 1$ such that for any $v \in E^s$,

$u \in E^u$,

$$\|df^n(v)\| < c\lambda^{-n}\|v\|$$

$$\|df^{-n}(u)\| < c\lambda^{-n}\|u\|$$

Remarks

(1) The estimates on the invariant subspaces are hyperbolicity conditions.

(2) If M supports a measure, μ, which is invariant under f, then typical (μ *a.e.*) orbits are dense. It is a long standing conjecture that density of orbits holds in general, without the assumption of the existence of an invariant measure.

(3) Anosov systems are stable in the sense that they form an open set in the C^1 topology.

(4) One can show that the above decomposition of the tangent bundle is continuous (and in fact Hölder) for any Anosov dynamical system.

There are a pair of important submanifolds of M associated with a dynamical system, the stable manifold, W^s , and the unstable manifold, W^u . The stable (unstable) manifold of a point p in M is the collection of points q in M which satisfy $d(f^n(p), f^n(q)) \to 0$ as n tends to positive (negative) infinity. The tangent space to $W^s(p)$ at p is $E^s(p)$ and the tangent space to $W^u(p)$ at p is $E^u(p)$. While the structure of these manifolds is quite complicated in general, there is a clear local picture for an Anosov diffeomorphism. Let $W^s_\epsilon(p)$ ($W^u_\epsilon(p)$) denote the intersection of $W^s(p)$ ($W^u(p)$) with a ball of radius ϵ about the point p. These submanifolds are as smooth as the diffeomorphism, f, and for an Anosov system one can show that they intersect transversely at each point p in M. The following theorem is due to Anosov [An] (or see [K], [M]).

Theorem. *Let $f : M \to M$ be an Anosov diffeomorphism of class C^r, $p \in M$, then*

(1) *there exists $\epsilon_0 > 0$ such that for every ϵ, $0 < \epsilon < \epsilon_0$, $W^s_\epsilon(p)$ and $W^u_\epsilon(p)$ are diffeomorphic to disks.*

(2) *for every ϵ, $0 < \epsilon < \epsilon_0$, there is a $\delta = \delta(\epsilon)$ such that for all q with $d(p,q) < \delta$, $W^s_\epsilon(p) \cap W^u_\epsilon(q)$ and $W^u_\epsilon(p) \cap W^s_\epsilon(q)$ each contain a single point.*

The invariance of $W^s_\epsilon(p)$ and $W^u_\epsilon(p)$ under the diffeomorphism, f, the local structure implied by the theorem, and the remarks above make $W^s_\epsilon(p)$ and $W^u_\epsilon(p)$ ideal choices for the axes of the family of coordinate systems which we will need in the nonstationary normal form theorems. By choosing these local coordinates we cause the linear terms of the local representation of f at p, f_p, to be diagonal. In addition this family of local coordinate systems is Hölder continuous as one changes the point of origin of the coordiantes.

Since f is a diffeomorphism and iterates of f or its inverse acts on M, one may think of f as the generator of a $\mathbb{Z}$-action on the manifold M. An integer, n, acts on M by applying the n^{th} power of f to M. It will be convenient to use some notions from group cohomology for $\mathbb{Z}$-actions.

Definition. *Let M be a smooth manifold and G a group. A function $\alpha : M \times \mathbb{Z} \to G$ is called a cocycle over a dynamical system, (f, M), if for every $k, l \in \mathbb{Z}$ and for every $x \in M$*

$$\alpha(x, k+l) = \alpha(f^l x, k) + \alpha(x, l)$$

There are many examples of cocycles for dynamical systems. For example one can form a cocycle from any real valued function on M by considering the sum of the function along an orbit. The Jacobian is a well known example of a cocycle.

Definition. *A cocycle γ is called a coboundary if there is a function $b : M \to G$ such that*

$$\gamma(x,n) = b(f^n(x)) - b(x)$$

for all $n \in \mathbb{Z}$, $x \in M$.

We will see that the terms in the nonstationary normal form are cocycles in the sense of the definition above. S. Hurder and A. Katok, [HK], have used one such cohomology invariant, the Anosov cocycle, to establish rigidity results about the smoothness of the invariant foliations defined by the system of stable and unstable foliations. This cocycle is the first nonlinear term in our nonstationary normal form.

The theorem implied by our title does not indicate the most general setting in which our methods may be applied. In chapter 7 we will formulate a theorem for skew products of dynamical systems.

Definition. *Let M and N be smooth manifolds with $p \in M$, $x \in N$. Also let $T : M \to M$ and $R_p : N \to N$ be a family of diffeomorphisms which vary continuously in p. A transformation $S : M \times N \to M \times N$ is called a skew product if S is of the form*

$$S(p,x) = (f(p), S_p(x)).$$

This generalizes the notion of direct product since the map in the fibres changes as one changes the base point. Any skew product generates a cocycle with values in the group of diffeomorphisms of its second factor (N above). More specifically, if we let Π_2 denote the projection onto the second factor,

$$\begin{aligned} C(p,n) &= \Pi_2 S^n(p,\cdot) \\ &= S_{f^{n-1}(p)} \circ \cdots \circ S_{f^1(p)} \circ S_p \end{aligned}$$

is a cocycle. We will show that one can place the fibre map of certain skew products in normal form.

CHAPTER III

NOTATION AND STATEMENT OF RESULTS

In this chapter we will provide a complete list of the notation employed in this paper as well as concise statememts of the main theorems. Consideration of the nonstationary case has an unfortunate affect on the notation required for theorems similar to those of Chapter 1. In addition to the already abundant subscripts and superscripts, one must include a parameter which designates the location on the torus.

Notation. *We will use the following notations:*

- ◇ $\mathbb{T}^2$ - *the two dimensional torus*
- ◇ f - *an area preserving Anosov diffeomorphism*
- ◇ $W^s(p)$ - *the stable manifold passing through* p
- ◇ $W^u(p)$ - *the unstable manifold passing through* p
- ◇ $W^s_{loc}(p)$ - *the local stable manifold passing through* p
- ◇ $W^u_{loc}(p)$ - *the local unstable manifold passing through* p
- ◇ f_p - *the local power series representation for* f *at* $p \in \mathbb{T}^2$
- ◇ $\lambda(p)$ - *the (real) eigenvalue for the expanding direction of the linear part of* f.
- ◇ $f_p(x,y) = (f_1(p), f_2(p))$
- ◇ g_p - *the local normal form for* f *at* $p \in \mathbb{T}^2$
- ◇ $\Phi(p, xy)$ - *a power series involving terms of the form* $(xy)^i$ *with constant term 1. This is part of the first coordinate of the normal form .*

◇ $\Psi(p, xy)$ - a power series involving terms of the form $(xy)^i$ with constant term 1. This is part of the second coordinate of the normal form .

◇ $g_p(x, y) = (\lambda(p)x\Phi(p, xy), \lambda^{-1}(p)y\Psi(p, xy))$

◇ $\Upsilon(p, xy)$ - the series in the exponent of the exponential form of g_p

◇ $g_p(x, y) = (\lambda(p)x\, e^{\Upsilon}(p, xy), \lambda^{-1}(p)y\, e^{-\Upsilon(p,xy)}$

◇ h_p - the local change of coordinates which conjugates f_p with g_p. A judicious initial choice of coordinates permits us to think of h_p as close to the identity.

Let $A(x, y) = \sum_{i,j} a_{ij}x^iy^j$ and $B(x, y) = \sum_{i,j} b_{ij}x^iy^j$.

◇ $\hat{A}(x, y) = \sum_{i+j\geq 2} a_{ij}x^iy^j$

◇ $[A(x, y)]_n = \sum_{i-j=n} a_{ij}x^iy^j$

◇ $\bar{A}(x, y) = \sum_{i,j} |a_{ij}|x^iy^j$

◇ $A \prec B$, if $|a_{ij}| < b_{ij}$ for all nonnegative integers i and j. B is said to majorize A.

◇ $\mathfrak{h}$, $\mathfrak{H}$, $\mathcal{H}$, k_p^s, k_p^u - various conjugacy maps

◇ $\mathcal{A}_1 = \int\limits_M \Phi_{11}(p)$

◇ $\Gamma(j, n, p)$ - the cocycle defined by the j^{th} order terms of $\Upsilon(p, xy)$.

For the convenience of the reader, we now list our main results. The only difference between the first three theorems is the smoothness category of the mappings involved. While the statements look very much alike the proofs are quite different.

Theorem. *Let f be a smooth, area preserving, Anosov diffeomorphism of the two dimensional torus and f_p its local representation at p as a formal power* *series. Then there exists a family of formal series, h_p, continuous in p, which conjugates f_p with its normal form, i.e.,*

$$f_p \circ h_p = h_{f(p)} \circ g_p$$

where g_p is of the form

$$g(p, x, y) = (\lambda(p)x\Phi(p, xy), \lambda(p)^{-1}y\Psi(p, xy))$$

Theorem. *Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be an analytic, area preserving, Anosov diffeomorphism and let f_p denote the local power series representation for f at $p \in \mathbb{T}^2$. Then there exists a family of analytic coordinate changes, h_q, continuous in q, such that*

$$f_p \circ h_p = h_{f(p)} \circ g_p$$

where g_p is the local normal form for f at p.

$$g(p, x, y) = (\lambda(p)x\Phi(p, xy), \lambda^{-1}(p)x\Psi(p, xy))$$

Theorem. *Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be an C^∞, area preserving, Anosov diffeomorphism and let f_p denote the local power series representation for f at $p \in \mathbb{T}^2$. Then there exists a family of analytic coordinate changes, h_q, continuous in q, such that*

$$f_p \circ h_p = h_{f(p)} \circ g_p$$

where g_p is the local normal form for f at p.

$$g(p, x, y) = (\lambda(p)x\Phi(p, xy), \lambda^{-1}(p)x\Psi(p, xy))$$

Theorem. *Let $F : B \to B$ be a skew product dynamical system. If the eigenvalues $\lambda_1(p), \ldots, \lambda_n(p)$ lie inside the unit circle and are free from resonances at each point $p \in M$, then there is a smooth transformation $H : B \to B$ which satisfies*

$$F \circ H = H \circ \Lambda$$

where $H(p, X) = (p, H_p(X))$ and $\Lambda(p, X) = (f(p), \Lambda_p(x))$. If there are resonances, then the theorem holds with Λ_p replaced by a polynomial containing only resonance terms.

Theorem. *Let $F : M \times \mathbb{R}^2 \to M \times \mathbb{R}^2$ be a skew product transformation of the form*

$$F(p, X) = (f(p), F_p(X)).$$

where the fibre map, F_p is smooth and area preserving, fixes the zero section, and is continuous in the parameter, p. Then there exists $H : B \to B$, $H(p, x) = (p, H_p(x))$ such that

$$F \circ H = H \circ G$$

where $G(p, x) = (f(p), G_p(x))$ and

$$G_p(x, y) = \left(\lambda(p)x\Phi(xy), \lambda^{-1}(p)x\Phi^{-1}(xy)\right)$$

Corollary. *The normal form can be expressed as*

$$G_p(x, y) = \left(\lambda(p)x\mathrm{e}^{\Upsilon(p,xy)}, \lambda^{-1}(p)y\mathrm{e}^{-\Upsilon(p,xy)}\right)$$

The coefficients of $(xy)^j$ in the exponent of the exponential are cocycles for the dynamical system (F, B).

Corollary 7.7. *If the base map $f(p)$ preserves an invariant measure on M then $\mathcal{A}_1 = \int_M \Gamma(1,1,p)$ is an invariant of the skew product (F,B).*

CHAPTER IV

FORMAL CASE

In this section we will establish the formal version of the nonstationary normal form theorem. This result will serve as the first step in the proofs of the C^∞ and analytic cases.

Theorem 4.1. *Let f be a smooth, area preserving, Anosov diffeomorphism of the two dimensional torus and f_p its local representation at p as a formal power series. Then there exists a family of formal series, h_p, continuous in p, which conjugates f_p with its normal form, i.e.,*

$$f_p \circ h_p = h_{f(p)} \circ g_p \tag{4.1}$$

where g_p is of the form

$$g(p,x,y) = (\lambda(p)x\Phi(p,xy), \lambda(p)^{-1}y\Psi(p,xy)) \tag{4.2}$$

Remarks:

(1) Note that $\Phi(p,xy) = 1 + \Phi_{11}(p)xy + \dots$ is a function of the product xy (and the parameter p). Recall that $\lambda(p)$ is the eigenvalue of the expanding direction for f. Since f is area preserving the contracting eigenvalue is the reciprocal of $\lambda(p)$.

(2) In our case (as in the classical case) one can write down formal expansions, compose power series and equate coefficients of like terms. The main difference in the proof of the nonstationary case is that the conjugating map has parameter value p on the left side of the equation and

$f(p)$ on the right. This new difficulty is overcome using hyperbolicity and repeated substitution of equations along the orbit of p.

(3) Also like the classical case the resonance and nonresonance terms require separate analysis.

(4) Preservation of area for $g(p, x, y)$ requires $\Phi(p, xy)$ and $\Psi(p, xy)$ to be reciprocals of eachother. The proof is identical to the one demonstrated in chapter 1 following Birkhoff's theorem.

Proof. We can make things easier by making a good initial choice for the local coordinates for f_p. Let the x-axis be tangent to the unstable manifold W^u and the y-axis be tangent to the stable manifold W^s. This diagonalizes the linear part of f_p. (Any initial choice could be adjusted to fit these criteria by conjugating with any appropriate linear change of coordinates.) Having made this choice we can take h to be close to the identity, i.e., $h(p, x, y) = (x + 2^{nd} order, y + 2^{nd} order)$. Thus we can write the expression for the first coordinate of the conjugacy equation as

$$\lambda(p)\left(x + \sum_{i+j \geqslant 2} h^1_{ij}(p) x^i y^j\right) +$$

$$\sum_{i+j \geqslant 2} f^1_{ij}(p)\left(x + \sum_{k+l \geqslant 2} h^1_{kl}(p) x^k y^l\right)^i \left(y + \sum_{k+l \geqslant 2} h^2_{kl}(p) x^k y^l\right)^j$$

$$= \lambda(p) x \Phi(p, xy) + \sum_{i+j \geqslant 2} h^1_{ij}\left(f(p)\right)\left(\lambda(p)\Phi(p, xy)\right)^{i-j} x^i y^j$$

We have used the fact that $\Phi(p, xy) = \Psi(p, xy)^{-1}$ to simplify the second expression on the right hand side. We will show by induction on the order, $m = i + j$, (not just at p but simultaniously for all $p \in \mathbb{T}^2$) that each coefficient $h^l_{ij}(p)$,

$l = 1, 2$ is well defined and varies continuously with respect to the parameter p. To anchor the induction note that the linear part of h_p is the identity map; so these terms are both well defined and continuous for all p. Suppose all terms of order less than m have been defined and are continuous functions in the parameter, p. By equating coefficients of the term $x^i y^j$ (for $i \neq j+1$, i.e., the nonresonance case) we get the expression

$$\lambda(p) h^1_{ij}(p) + f^1_{ij}(p) + l.o.t. = \lambda(p)^{i-j} h^1_{ij}(f(p)) + l.o.t. \tag{4.3}$$

Collecting all of the lower order terms and the given term $f^1_{ij}(p)$ into one expresion $Q^1_{ij}(p)$ we have

$$h^1_{ij}(p) = \lambda(p)^{i-j-1} h^1_{ij}(f(p)) + Q^1_{ij}(p) \tag{4.4}$$

An equation of the same form as above holds at $f(p)$ namely

$$h^1_{ij}(f(p)) = \lambda(p)^{i-j-1} h^1_{ij}(f^2(p)) + Q^1_{ij}(f(p)) \tag{4.5}$$

and for every point on the torus. If $n = i - j - 1 < 0$ we can write a well defined solution for $h^1_{ij}(p)$ by iteratively substituting along the forward f-orbit of p (i.e., by repeatedly substituting for the h^1_{ij} term on the right in the last equation). This gives us

(4.6)
$$h^1_{ij}(p) = \sum_{k=0}^{\infty} \left(\prod_{l=1}^{i} \lambda^n(f^l(p)) \right) Q^1_{ij}(f^k(p)) + \lim_{N\to\infty} \left(\prod_{l=1}^{N} \lambda^n(f^l(p)) \right) h^1_{ij}(f^N(p))$$

Since we assumed that $\lambda(p) > 1$ and $n < 0$, $\prod_{j=1}^{N} \lambda^n(f^j(p))$ shrinks rapidly to zero as $N \to \infty$. We define

$$h^1_{ij}(p) = \sum_{k=0}^{\infty} \left(\prod_{l=1}^{k} \lambda(f^l(p))^n \right) Q^1_{ij}(f^k(p)) \tag{4.7}$$

Since $Q^1_{ij}(p)$ is a continuous function on the compact space, $\mathbb{T}^2$, it is bounded; hence $h^1_{ij}(p)$ is well defined by the converging summation. As $h^1_{ij}(p)$ is the uniform limit of the continuous functions defined by the partial sums, we see that $h^1_{ij}(p)$ is continuous in p.

If $n = i - j - 1 > 0$ we may consider the equation similar to (4.3) where p is replaced by $f^{-1}(p)$ and $f(p)$is replaced by p.

$$\lambda(f^{-1}(p))h^1_{ij}(f^{-1}(p)) + f^1_{ij}(f^{-1}(p)) + l.o.t. = \lambda(f^{-1}(p))^{i-j}h^1_{ij}(p) + l.o.t. \tag{4.8}$$

We can once again solve for $h^1_{ij}(p)$ to nget an expression similar to (4.4)

$$h^1_{ij}(p) = \lambda(f^{-1}(p))^{i-j-1}h^1_{ij}(f^{-1}(p)) + R^1_{ij}(p) \tag{4.9}$$

Substitution leads to an equation like (4.6) where the limit goes to zero. Finally we have that

$$h^1_{ij}(p) = \sum_{k=0}^{\infty}\left(\prod_{l=1}^{i}\lambda(f^{-l}(p))^n\right)Q^1_{ij}(f^{-k}(p)) \tag{4.10}$$

Thus $h^1_{ij}(p)$ is well defined and continuous in p when $n > 0$. A similar argument treats the nonresonance terms in the second coordinate of h_p.

The above establishes the inductive step for the nonresonance terms. All that remains is to handle the resonance terms (i.e. in the first coordinate, terms of the form $x^{i+1}y^i$). As in the fixed point case there is substantial nonuniqueness in the definition of the resonance terms. The conjugacy equation becomes

$$h^1_{i+1i}(p) = h^1_{i+1i}(f(p)) + \Phi_{ii}(p) + Q^1_{i+1i}(p) \tag{4.11}$$

For simplicity we can choose $h^1_{i+1i}(p) = h^1_{i+1i}(q)$ for all $p, q \in \mathbb{T}^2$, then $\Phi_{ii}(p)$ is determined by $Q^1_{i+1i}(p)$, and hence it is continuous in p. The $h^1_{i+1i}(p)$ terms are also continuous since they are constant with respect to p. All of the above arguments apply for the second coordinate of h_p (in that case the resonance terms have the form $x^i y^{i+1}$) and we may choose $h^2_{ii+1}(p) = h^2_{ii+1}(q)$ for all $p, q \in \mathbb{T}^2$, which determines $\Psi_{ii}(p)$. As in the case of $\Phi_{ii}(p)$ above, $\Psi_{ii}(p)$ is continuous in p since it is also a sum of lower order terms. By induction the formal series h_p and g_p are well defined and continuous in p. At this point one may wonder if the previous choices still allow h_p to be area preserving. This is indeed possible as we will see from the following lemma.

Lemma 4.2 (Nonstationary Sternberg Lemma).

Supose $Jac(h_p)$ has no terms of the form $(xy)^i$, then h_p is (formally) area preserving.

Remark The condition from the lemma only restricts the resonance terms of h_p by requiring that $\forall p \in \mathbb{T}^2$

$$h^1_{i+1i}(p) + h^2_{ii+1}(p) = l.o.t. \tag{4.12}$$

Since $h^1_{i+1i}(p)$ and $h^2_{ii+1}(p)$ are independent and to this point unrestricted, this equation can always be satisfied. Note that it is possible to normalize these terms further. It will be convenient do so when the convergence argument is made in chapter 5.

Proof. First let us establish some notation. Let

$$A(p, \cdot) = \sum a_{ij}(p) x^i y^j = Jac(h(p, \cdot)) - 1$$

$$B(p, \cdot) = \sum b_{ij}(p) x^i y^j = Jac(g(p, \cdot)) - 1$$

We must show that $A(p,\cdot)=0$ and $B(p,\cdot)=0$. Recall that $\forall p \in \mathbb{T}^2 \; Jac(f_p)=1$. Taking the Jacobian of the conjugacy equation we have

$$1 + A(p,\cdot) = (1 + A(f(p),\cdot)(1 + B(p,\cdot))$$

or

$$\sum_{ij} a_{ij}(p)x^i y^j = \sum_{ij} a_{ij}(f(p))x^i y^j + \sum_{ij} b_{ij}(p)x^i y^j + \left(\sum_{ij} a_{ij}(f(p))x^i y^j\right)\left(\sum_{ij} b_{ij}(p)x^i y^j\right)$$

We will use induction to show that $a_{ij}(p)=0$ and $b_{ij}(p)=0$ for all i,j. The coefficients of $x^i y^j$ from above give rise to the equation

$$a_{ij}(p) = \lambda(p)^{i-j} a_{ij}(f(p)) + l.o.t. \tag{4.13}$$

By the induction hypothesis the $l.o.t. = 0$. For $i-j<0$ we substitute along the positive orbit of f to yield

$$a_{ij}(p) = \left(\prod_{k=0}^{N} \lambda(f^N(p))^{i-j}\right) a_{ij}(f^{N+1}(p)) \tag{4.14}$$

Since $a_{ij}(p)$ is a polynomial in the h_{kl}, it is continuous in p. It follows from the compactness of the torus that for each $(i,j) \in \mathbb{Z}^+ \times \mathbb{Z}^+$, there exists a constant, A_{ij}, such that

$$|a_{ij}(p)| < A_{ij}$$

for all $p \in \mathbb{T}^2$. Then we have

$$|a_{ij}(p)| < A_{ij} l^{(i-j)N}$$

for all $N \in \mathbb{Z}^+$. Hence $i - j < 0$ implies that $a_{ij}(p) = 0$. For $i - j > 0$ one simply iterates the analogue of (4.13) (at $f^{-1}(p)$ instead of p) along the negative p-orbit to achieve an equation similar to (4.14). Compactness and continuity then lead to the same conclusion, i.e., $a_{ij}(p) = 0$. If $i - j = 0$ (resonance) we get

$$a_{ii}(p) = a_{ii}(p) + b_{ii}(p) + l.o.t.$$

Since we assumed that $a_{ii}(p) = 0$ and $l.o.t. = 0$, then $b_{ii}(p) = 0$. By induction $a_{ij}(p) = 0$ and $b_{ij}(p) = 0$ for all i, j and for all $p \in \mathbb{T}^2$. $\square$

Even if we assume the hypothesis of the lemma, the resonance terms are not uniquely determined. This nonuniqueness reflects the invariance of the normal form under area preserving transformations which preserve the product xy, i.e., ones of the form

$$(x, y) \mapsto (xU(xy), yU^{-1}(xy))$$

Exponential Form

A slight modification of the normal form reveals a family of cocycles associated to the $\mathbb{Z}$-action generated by f. We will show that one can construct a formal series $\Upsilon(p, xy) = \sum_{i=1}^{\infty} u_i(p)(xy)^i$ which satisfies the equation

$$\Phi(p, xy) = \mathrm{e}^{\Upsilon(p,xy)}$$

The normal form becomes

$$g(p, x, y) = \left(\lambda(p)x\,\mathrm{e}^{\Upsilon(p,xy)}, \lambda^{-1}(p)y\,\mathrm{e}^{-\Upsilon(p,xy)}\right)$$

Proposition 4.3. *Given any Φ constructed as above the formal equation*

$$\Phi(p, xy) = \mathrm{e}^{\Upsilon(p,xy)}$$

has a well defined formal solution, Υ.

Proof. Let $z = xy$ and recall that Φ is of the form $1 + \Phi_{11}z^i + \ldots$. For the moment let us suppress the parameter p. We have that

$$\mathrm{e}^{\Upsilon(z)} = 1 + \sum_{n=1} \frac{1}{n!}\left(\sum_i u_i z^i\right)^n$$

so the constant terms agree. Equating coefficients one obtains

$$\Phi_1 = u_1$$

$$\Phi_2 = u_2 + {u_1}^2$$

$$\vdots$$

$$\Phi_k = u_k + l.o.t$$

By induction we see that $\Upsilon(p, xy)$ is a well defined formal series.

Corollary 4.4. *The coefficients of $\Upsilon(p, xy)$ are formal cocycles for the dynamical system (f, M).*

Proof. Consider the n^{th} iterate of the normal form, ${g_p}^n = g_{f^{n-1}(p)} \circ \cdots \circ g_p$. In exponential coordinates one computes that the first coordinate of ${g_p}^n$ is

$$\left(\prod_{i=0}^{n-1} \lambda(f^i(p))\right) x\,\mathrm{e}^{\sum_{i=0}^{n-1} \Upsilon\left(f^i(p), xy\right)}$$

Let $\Gamma(s,n,p)$ denote the coefficient of $(xy)^s$ in the exponent of the above expression. Carefull inspection tells us that

$$\Gamma(s,n,p) = \sum_{i=0}^{n-1} u_s\left(f^i(p)\right)$$

If $k+l=n$ we observe that

$$\Gamma(s,n,p) = \Gamma\left(s,k,f^l(p)\right) + \Gamma(s,l,p)$$

Thus $\Gamma(s,n,p)$ is an additive cocycle for the dynamical system (f,M). Further discussion of these cocycles will appear in chapter 7.

The curves $xy = constant$ are preserved by the normal form g_p. The exponential form provides a geometric characterization of the mapping properties of the normal form. If we think of parameterizing the hyperbolæ $xy=c$ by a hyperbolic angle, θ, the action of g_p^n is simply to add $log\lambda(p) + \Upsilon(p,c)$ to θ. In our special coordinates the Anosov diffeomorphism, f has a formal (later analytic) reduction to a map which shifts points along these hyperbolæ.

CHAPTER V

ANALYTIC NORMAL FORMS

In this chapter we will establish our main result, the nonstationary normal form theorem for area preserving, Anosov diffeomorphisms of the two dimensional torus. The proof of the analytic version of this theorem will rely on the groundwork laid in the previous chapter for the formal case. Analyticity will be demonstrated using a majorization scheme similar to the one developed by C. L. Siegel in his proof of Moser's theorem for the analytic fixed point case. Compactness will be instrumental in handling the complexities introduced by the nonstationary parameter.

Theorem. *Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be an analytic, area preserving, Anosov diffeomorphism and let f_p denote the local power series representation for f at $p \in \mathbb{T}^2$. Then there exists a family of analytic coordinate changes, h_q, continuous in q, such that*

$$f_p \circ h_p = h_{f(p)} \circ g_p \tag{5.1}$$

where g_p is the local normal form for f at p.

$$g(p, x, y) = (\lambda(p)x\Phi(p, xy), \lambda^{-1}(p)x\Psi(p, xy)) \tag{5.2}$$

Before we begin the proof let us recall some of the notation which we will use. Let $A(x, y) = \sum a_{ij}x^iy^j$ and $B(x, y) = \sum b_{ij}x^iy^j$.

$$A \prec B \quad if \ \ |a_{ij}| \leq b_{ij} \ \ \forall i, j$$

$$\bar{A} = \sum |a_{ij}| x^i y^j \quad \hat{A} = \sum_{i+j \geq 2} a_{ij} x^i y^j$$

$$[A]_n = \sum_{i-j=n} a_{ij} x^i y^j$$

Since $\Phi = 1 + \ldots$ is a series in the product xy and $\Psi = \Phi^{-1}$, we have

$$A(\lambda x \Phi, \lambda^{-1} y \Psi) = \sum_{ij} a_{ij} \left(\lambda \Phi\right)^{i-j} x^i y^j$$

which leads to the following important identity

$$[A(g)]_n = \left(\lambda \Phi\right)^n [A]_n \tag{5.3}$$

It will be more convenient to write expressions such as $(c-x)^{-1}$ instead of $\frac{1}{c} \sum_{n=0}^{\infty} \left(\frac{x}{c}\right)^n$. Notice that if $0 < c' < c$, then $(c-x)^{-1} < (c'-x)^{-1}$. The following lemmas will make the manipulations of these expressions more transparent. They will be used repeatedly in the majorization argument.

Lemma 5.1. *Let $a, b \in \mathbb{R}$ with $0 < a < b$ and let A be a series with nonnegative coefficients, then there exist c such that*

$$(a-A)^{-1}(b-A)^{-1} \prec (c-A)^{-1}$$

Proof. Expressing $(a-A)^{-1}$ and $(b-A)^{-1}$ as power series we have

$$\begin{aligned}(a-A)^{-1}(b-A)^{-1} &= \frac{1}{ab} \sum_{i=0}^{\infty} \left(\frac{A}{a}\right)^i \sum_{i=0}^{\infty} \left(\frac{A}{b}\right)^i \\ &= \frac{1}{ab} \sum_{i=0}^{\infty} \left(\sum_{k+l=i} \frac{1}{a^k b^l}\right) x^i\end{aligned}$$

Notice that the coefficient of x^i satisfies

$$\begin{aligned}\sum_{k+l=i} \frac{1}{a^k b^l} &< \frac{i+1}{a^i} \\ &< \frac{1}{d^i}\end{aligned}$$

for d sufficiently smaller than a. By choosing $c < \min(d, ab)$ we have satisfied the conditions of the lemma. $\square$

Lemma 5.2. *Let $c \in \mathbb{R}$ and A and B be nonnegative series. Then*

$$(c-A)^{-1}(c-B)^{-1} \prec c^{-1}(c-(A+B))^{-1}$$

Proof. Writing down the series we see that

$$\begin{aligned}(c-A)^{-1}(c-B)^{-1} &= c^{-2}\sum_{i=0}^{\infty}\left(\frac{A}{c}\right)^i\sum_{i=0}^{\infty}\left(\frac{B}{c}\right)^i \\ &= c^{-2}\sum_{ij}\frac{i+j}{c^{i+j}}A^iB^j \\ &\prec c^{-2}\sum_{i=0}^{\infty}\left(\frac{A+B}{c}\right)^i \\ &= c^{-1}(c-(A+B))^{-1} \quad \square\end{aligned}$$

Lemma 5.3. *Let A and $\hat{A}$ be as above and let $r > 0$ denote the radius of convergence of A, then there exists a positive real number, c, $0 < c < r$, such that*

$$\hat{A} \prec \frac{(x+y)^2}{c-(x+y)}$$

Proof. Let $A(x,y) = \sum_{ij} a_{ij}x^iy^j$. We wish to show that

$$\sum_{i+j\geq 2} a_{ij}x^iy^j \prec \sum_{i=2}^{\infty}\left(\frac{x+y}{c}\right)^i$$

For a_{ij} such that $i+j=k$, we have the Cauchy estimates

$$|a_{ij}| < \frac{K}{r^k}$$

where K is a positive real number. Coefficients of the k^{th} term on right hand side are of the form

$$\binom{k}{i}\frac{1}{c^k}$$

So for $k = i + j \geq 2$, we must show that

$$|a_{ij}| < \binom{k}{i}\frac{1}{c^k}$$

From the Cauchy estimates for a_{ij}, it suffices to show

$$\frac{K}{r^k} < \frac{1}{c^k}$$

If we choose $c < \frac{r}{\max(1,\sqrt{K})}$, then this inequality is satisfied for all integers, $k \geq 2$, and the lemma is verified. □

Lemma 5.4. *Let A be a power series composed of terms with nonnegative coefficients, and let c be a positive real constant. For any positive real number, d, we have*

$$\frac{d}{(c-A)} \prec \frac{1}{(c'-A)}$$

for all $c' < \min(c, \frac{c}{d})$.

Proof. If $d < 1$ take $c = c'$; otherwise

$$\begin{aligned} \frac{d}{(c-A)} &= \frac{d}{c}\sum_{i=0}^{\infty}\left(\frac{A}{c}\right)^i \\ &\prec \frac{1}{\frac{c}{d}}\sum_{i=0}^{\infty}\left(\frac{A}{\frac{c}{d}}\right)^i \\ &\prec \frac{1}{(c'-A)} \end{aligned}$$ □

Lemma 5.5. *Let $P(x)$ be a polynomial in x and c a positive real number, then there exists a positive real, $c' < c$ such that*

$$(c-x)^{-1} + P(x) \prec (c'-x)^{-1}$$

Proof. Let $P(x) = \sum_{i=0}^{n} p_i x^i$.

$$(c-x)^{-1} + P(x) = \sum_{i=0}^{n} \left(p_i + \left(\frac{1}{c}\right)^i\right) x^i + \sum_{i=n+1}^{\infty} \left(\frac{x}{c}\right)^i$$

Choose c_i' such that

$$\left(\frac{1}{c_i'}\right)^i > \left(\frac{1}{c}\right)^i + |p_i|.$$

Equivalently such that

$$\left(\frac{c}{c_i'}\right)^i > 1 + |p_i|c^i.$$

This is satisfied for c_i' sufficiently small. Now take $c' = \min_{0 \leq i \leq n}(c_i')$. □

The following two lemmas will help to simplify the end of the majorization process.

Lemma 5.6. *Let $A(x,y)$ be a nonnegative formal series. $A(x,y)$ converges if and only if $A(x,x)$ converges.*

Proof. The convergence of $A(x,x)$ given that of $A(x,y)$ is apparent. Suppose $A(x,x)$ converges.

$$A(x,x) = \sum_{i=0}^{\infty} \left(\sum_{k+l=i} a_{kl}\right) x^i$$

So by Cauchy's estimates

$$\sum_{k+l=i} a_{kl} < \frac{K}{r^i} \quad \textit{for some} \quad r, K$$

But the a_{kl}'s are nonnegative so each one satisfies the same estimates. Hence, $A(x,y)$ converges. □

Lemma 5.7(Cauchy). *Let A and B be nonnegative series without constant terms. Suppose A satisfies*

$$A \prec \frac{2x + A}{c - A}$$

and B satisfies

$$B = \frac{2x + B}{c - B}$$

then B majorizes A, i.e., $A \prec B$.

Proof. The argument proceeds by induction. First note that A and B satisfy

$$\sum_{i=1}^{\infty} b_i x^i = \left(2x + \sum_{i=1}^{\infty} b_i x^i\right)\left(\frac{1}{c}\sum_{i=1}^{\infty}\left(\frac{B}{c}\right)^i\right)$$

and

$$\sum_{i=1}^{\infty} a_i x^i = \left(2x + \sum_{i=1}^{\infty} a_i x^i\right)\left(\frac{1}{c}\sum_{i=1}^{\infty}\left(\frac{A}{c}\right)^i\right)$$

Consider the expressions derived from equating coefficients of the powers of x. It is important to observe that the left hand side of both expressions contains the only term in which the index matches the power of x. Note that $b_1 = \frac{2}{c}$ while $a_1 < \frac{2}{c}$ so induction may begin. Suppose $a_k < b_k$ for all $k = 1, 2, \ldots, n$, then $b_{n+1} = P_{n+1}(b_1, b_2, \ldots, b_n)$ where P_{n+1} is a polynomial with positive coefficients. Thus $a_k < b_k$ implies that $P_{n+1}(a_1, a_2, \ldots, a_n) < P_{n+1}(b_1, b_2, \ldots, b_n)$, hence $A \prec B$. □

We will need some global (on $\mathbb{T}^2$) constants to use in estimates in the majorization proof. We list them here for reference. Recall that f_p is the local power series representation for f at p, and $\lambda(p)$ is the eigenvalue for the expanding direction, i.e. $\lambda(p) > 1$. Let r_p be the radius of convergence for f_p.

$$r = \min_{p \in \mathbb{T}^2} r_p$$
$$l = \min_{p \in \mathbb{T}^2} \lambda(p) \qquad (l > 1)$$
$$L = \max_{p \in \mathbb{T}^2} \lambda(p)$$

Based on the previous chapter we may define the following formal series.

$$H_1(x,y) = \sum_{i,j} \max_{p \in \mathbb{T}^2} |h^1_{ij}(p)| x^i y^j$$
$$H_2(x,y) = \sum_{i,j} \max_{p \in \mathbb{T}^2} |h^2_{ij}(p)| x^i y^j$$
$$\Phi(xy) = \sum_{i,j} \max_{p \in \mathbb{T}^2} |\Phi_{ii}(p)| x^i y^i$$
$$F_1(x,y) = \sum_{i,j} \max_{p \in \mathbb{T}^2} |f^1_{ij}(p)| x^i y^j$$
$$F_2(x,y) = \sum_{i,j} \max_{p \in \mathbb{T}^2} |f^2_{ij}(p)| x^i y^j$$

All of the formal series above are well defined since $\mathbb{T}^2$ is compact. Further F_1 and F_2 converge in a disk of radius r. By Lemma 5.3 there is a small positive real number, c_1, for which $\hat{F}_1$ and $\hat{F}_2$ satisfy

$$\hat{F}_1 \prec \frac{(x+y)^2}{c_1 - (x+y)}$$
$$\hat{F}_2 \prec \frac{(x+y)^2}{c_1 - (x+y)}$$

We will show that the series H_1, H_2, and Φ converge, from which the convergence of h_p and g_p is immediate. The theorem will follow from a collection of estimates on the formal series developed so far. We will manipulate these estimates until we have a series which is known to be convergent as an upper bound for our (previously) formal series.

Proof. We begin with a slightly different formulation of the formal solutions developed in the previous chapter. Taking $[\ ]_n$ of the conjugacy equation (5.1) in each coordinate we obtain

(5.4a)

$$\lambda(p)[\hat{h}_1(p,\cdot)]_n + [\hat{f}_1(p,h(p,\cdot)]_n = \lambda(p)[x\hat{\Phi}(p,\cdot)]_n + (\lambda(p)\Phi(p,\cdot))^n\,[\hat{h}_1(f(p),\cdot)]_n$$

and

(5.4b)

$$\lambda^{-1}(p)[\hat{h}_2(p,\cdot)]_n + [\hat{f}_2(p,h(p,\cdot)]_n = \lambda^{-1}(p)[y\hat{\Psi}(p,\cdot)]_n + (\lambda(p)\Phi(p,\cdot))^n\,[\hat{h}_1(f(p),\cdot)]_n$$

Note that the property (5.3) was used to simplify the second term on the right hand side of both equations. Since Φ only has terms involving the product xy, $[x\Phi]_n = 0$ unless $n = 1$, and $[x\Phi]_1 = x\Phi$. Similarly $[y\Psi]_n = 0$ unless $n = -1$, and $[y\Psi]_{-1} = y\Psi$. The terms in the series $x\Phi$ and $y\Psi$ are called resonance terms analogous to the classical fixed point case. As in chapter 4 we can iteratively substitute for $[h]_n$ these equations along the orbit of f to get a well defined expression in the nonresonance case. (We have a slightly more complicated remainder term to work with this time which will require more careful estimates.)

$$[h_1(p,\cdot)]_n = -\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i-1}\lambda^n(f^j(p))\prod_{j=0}^{i-1}\Phi^n(f^j(p))}{\prod_{j=0}^{i}\lambda(f^j(p))}[\hat{f}_1(f^i(p),h(f^i(p),\cdot)]_n +$$

$$\lim_{N\to\infty}\prod_{j=1}^{N}\lambda^{n-1}(f^N(p)\prod_{j=1}^{N}\Phi^N(f^N(p))[\hat{h}_1(f^N(p),\cdot)]_n \tag{5.5a}$$

$$[h_2(p,\cdot)]_n = -\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i-1} \lambda^n(f^j(p)) \prod_{j=0}^{i-1} \Phi^n(f^j(p))}{\prod_{j=0}^{i} \lambda(f^j(p))} [\hat{f}_2(p, h(f^i(p),\cdot)]_n +$$

$$\lim_{N\to\infty} \prod_{j=1}^{N} \lambda^{n-1}(f^N(p) \prod_{j=1}^{N} \Phi^n(f^N(p))[\hat{h}_2(f^N(p),\cdot)]_n \tag{5.5b}$$

By showing that the above limit approaches zero we will have a convenient formula for $[h_1(p)]_n$ when $n \leq 0$. The uniform hyperbolicity of f plays a significant role in obtaining the value of the limit and in the future utility of this expression in the majorization process. If we "forget" all of the dependence on the nonstationary parameter, p, equation (5.5a) becomes

$$[h_1]_n = ((\lambda\Phi) - \lambda)^{-1} [f_1]_n$$

as in the fixed point case.

Since $\lambda(p) > 1$ for all $p \in \mathbb{T}^2$, $l > 1$. Therefore

$$\prod_{j=1}^{N} \lambda^n(f^j(p)) < (l^n)^N \quad \forall p \in \mathbb{T}^2 \quad \forall n \leq 0$$

The exponential shrinkage of this term counteracts the growing number of terms of a given order generated by $\prod_{j=1}^{N} \Phi^n(f^j(p))$ as N increases. It suffices to work with the global majorizing series $\mathit{\Phi}$ and the smallest value for $\lambda(p)$, l. We may consider

$$\lim_{N\to\infty} l^{(n-1)N} \mathit{\Phi}^{nN}$$

since the expression $[h_1(p)]_n$ does not involve N.

Let $z = xy$ and recall that $\mathit{\Phi}$ has leading term 1. Let $\mathit{\Phi}^n = \sum_{i=0}^{\infty} a_i z^i$ then the k^{th} term of $\left(\sum_{i=0}^{\infty} a_i z^i\right)^N$ is

$$\binom{N}{k} a_1^k + \binom{N}{k-2 \;\; 1} a_1^{k-2} a_2 + \cdots + \binom{N}{1} a_k$$

where

$$\binom{N}{i_1 \dots i_l} = \frac{N!}{i_1! \dots i_l!(N-(i_1+\cdots+i_l)!}$$

The number of terms is bounded by the number of ways one can write k as a sum of positive integers. Denote this function of k the partition function by $P(k)$. All of these combinatorial coefficients satisfy the estimate

$$\binom{N}{i_1 \dots i_l} < N(N-1)\dots(N-(k-1)) < N^k$$

So if $\Phi^N = \sum_{i=0}^{\infty} b_i z^i$ we have the estimate

$$|b_k| < N^k P(k) M_k$$

where

$$M_k = max \prod a_{i_l}$$

such that $i_1, \dots, i_l$ are nondecreasing, nonnegative integers with $\sum_l i_l = k$.

Since we work with formal series we may fix $k \in \mathbb{Z}^+$ and take the limit. We see that

$$|(l^n)^N| < l^{nN} N^k P(k) M_k$$

So

$$\lim_{N\to\infty} \prod_{j=1}^{N} l^{n-1} \Phi^N b_k = 0$$

As previously mentioned the above considerations only hold if we assume $n \leq 0$. The case $n = 1$ (resonance) will be dealt with separately. If $n > 1$ we return our attention to (5.4a). By replacing p with $f^{-1}(p)$ and $f(p)$ with p in (5.4a) and iteratively substituting along the negative orbit of p, we derive a similar expression involving a limit which converges to zero and a well defined sum. All of the above applies equally to the second coordinate, h_2, with the caviat that resonance occurs when $n = -1$.

Thus we have the following expressions for h_p.

(5.6a)

$$[h_1(p,\cdot)]_n = -\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i-1} \lambda^n(f^j(p)) \prod_{j=0}^{i-1} \Phi^n(f^j(p))}{\prod_{j=0}^{i} \lambda(f^j(p))} [\hat{f}_1(f^i(p), h(f^i(p),\cdot))]_n$$

if $n \leq 0$.

(5.7a)

$$[h_1(p,\cdot)]_n = \sum_{i=1}^{\infty} \frac{\prod_{j=1}^{i-1} \lambda(f^{-j}(p))}{\prod_{j=1}^{i} \lambda^n(f^{-j}(p)) \prod_{j=1}^{i} \Phi^n(f^{-j}(p))} [\hat{f}_1(f^{-i}(p), h(f^{-i}(p),\cdot))]_n$$

if $n > 1$.

(5.6b)

$$[h_2(p,\cdot)]_{-n} = \sum_{i=1}^{\infty} \frac{\prod_{j=1}^{i} \lambda^n(f^{-j}(p)) \prod_{j=1}^{i} \Phi^n(f^{-j}(p))}{\prod_{j=1}^{i-1} \lambda(f^{-j}(p))} [\hat{f}_2(f^i(p), h(f^{-i}(p),\cdot))]_{-n}$$

if $n \leq 0$.

(5.7b)

$$[h_2(p,\cdot)]_{-n} = -\sum_{i=0}^{\infty} \frac{\prod_{j=0}^{i} \lambda(f^j(p))}{\prod_{j=0}^{i-1} \Phi^n(f^j(p)) \prod_{j=0}^{i-1} \lambda(f^j(p))} [\hat{f}_1(f^i(p), h(f^i(p),\cdot))]_{-n}$$

if $n > 1$.

It is also useful to have these expressions in the form

(5.8a)

$$[h_1(p,\cdot)]_n = -\sum_{i=0}^{\infty} \frac{1}{\lambda(f^i(p))} \prod_{j=0}^{i-1} \lambda^{n-1}(f^j(p)) \Psi^{-n}(f^j(p)) [\hat{f}_1(f^i(p), h(f^i(p),\cdot))]_n$$

if $n \leq 0$.

(5.9a)

$$[h_1(p,\cdot)]_n = \sum_{i=1}^{\infty} \frac{1}{\lambda(f^{-i}(p))} \prod_{j=1}^{i} \lambda^{1-n}(f^{-j}(p)) \Psi^n(f^{-j}(p)) [\hat{f}_1(f^{-i}(p), h(f^{-i}(p),\cdot))]_n$$

if $n > 1$.

(5.8b)

$$[h_2(p,\cdot)]_{-n} = \sum_{i=1}^{\infty} \lambda(f^{-i}(p)) \prod_{j=1}^{i} \lambda^{n-1}(f^{-j}(p))\Psi^{-n}(f^{-j}(p))[\hat{f}_2(f^{-i}(p), h(f^{-i}(p),\cdot))]_{-n}$$

if $n \leq 0$.

(5.9b)

$$[h_2(p,\cdot)]_{-n} = \sum_{i=0}^{\infty} \lambda(f^{i}(p)) \prod_{j=0}^{i-1} \lambda^{1-n}(f^{j}(p))\Psi^{n}(f^{j}(p))[\hat{f}_2(f^{i}(p), h(f^{i}(p),\cdot))]_{-n}$$

if $n > 1$.

Normalizations.

We have some freedom in the choice of the resonance terms, $[h_1(p)]_1$ and $[h_2(p)]_{-1}$. However we must ensure that the restrictions placed by the conjugacy equation and the preservation of area are maintained. We have the equations

(5.10a) $$\lambda(p)[\hat{h}_1(p)]_1 + [\hat{f}_1(h(p))]_1 = \lambda(p)x(\Phi(p)-1) + \lambda(p)\Phi(p)[\hat{h}_1(f(p))]_1$$

(5.10b)

$$\lambda^{-1}(p)[\hat{h}_2(p)]_{-1} + [\hat{f}_2(p)]_{-1} = \lambda^{-1}(p)(\Psi(p)-1) + \lambda^{-1}(p)\Psi(p)[\hat{h}_2(f(p))]_{-1}$$

As in the formal case we may choose $h^1_{i+1i}(p) = h^1_{i+1i}(q)$ and $h^2_{ii+1}(p) = h^2_{ii+1}(q)$ for all $p, q \in \mathbb{T}^2$. This choice determines $\Phi_{ii}(p)$ uniquely in terms of lower order terms. By also choosing $h^1_{i+1i}(p) + h^2_{ii+1}(p) = l.o.t.$ we have (by the Nonstationary Sternberg Lemma) that h, and hence g, are area preserving. We may normalize further by choosing

(5.11) $$[h_2(q)]_{-1} = y \quad \forall q \in \mathbb{T}^2$$

This defines each term of $[h_1(p)]_1$ uniquely in terms of lower order terms. The normalization from (above) yields the equation

$$[f_2(p,\cdot)]_{-1} = \lambda^{-1}(p)(\Psi(p,\cdot) - 1) \tag{5.12}$$

which will be helpful later.

We now derive the final expression necessary to begin our majorization process. Since h preserves area the following condition on the Jacobian must be satisfied.

$$\begin{aligned} 1 &= \frac{\partial h_1}{\partial x}\frac{\partial h_2}{\partial y} - \frac{\partial h_2}{\partial x}\frac{\partial h_1}{\partial y} \\ &= \sum_{i,j,k,l} (ik - jl) h^1_{ij} h^2_{kl} x^{i+k-1} y^{j+l-1} \quad \textit{taking } [\ \ ]_0 \\ 1 &= \sum_{j,l,n} n(l+j) h^1_{j+nj} h^2_{l-nl} (xy)^{l+j-1} \end{aligned}$$

Integrating with respect to the product xy we see that

$$\begin{aligned} xy &= \sum_{j,l,n} n h^1_{j+nj}(p) h^2_{l-nl}(p)(xy)^{l+j} \\ &= \sum_{n=-\infty}^{\infty} [h_1(p)]_n [h_2(p)]_{-n} \end{aligned}$$

so

$$xy - [h_1(p)]_1 [h_2(p)]_{-1} = \sum_{n \neq 1} [h_1(p)]_n [h_2(p)]_{-n}$$

Finally by employing our normalization (5.11) we have

$$xy - [h_1(p)]_1 y = \sum_{n \neq 1} [h_1(p)]_n [h_2(p)]_{-n} \tag{5.13}$$

Estimates.

With these formulæ in hand we can begin to make the estimates necessary for the majorization scheme. Throughout the remainder of this chapter we will employ a sequence of decreasing positive constants $c_1, c_2, \dots$. Recall the definitions of the constants c_1, l, and L. Let $\mathcal{H} = H_1 + H_2$ and $\Delta = l^{-1}\Psi$. We have the following estimates for $[h_i(p)]_{\pm n}$.

If $n \leq 0$ (5.8a) and (5.8b) yield

$$\begin{aligned} [h_1(p)]_n &\prec \frac{1}{l}\sum_{i=0}^{\infty}\left(l^{n-1}\Psi^{-n}\right)^i\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_n \\ &= \frac{1}{l}\sum_{i=0}^{\infty}\left(l^{-1}\Delta^{-n}\right)^i\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_n \\ &= \frac{1}{l}\left(1-l^{-1}\Delta^{-n}\right)^{-1}\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_n \end{aligned} \tag{5.14a}$$

$$\begin{aligned} [h_2(p)]_{-n} &\prec L\sum_{i=1}^{\infty}\left(l^{n-1}\Psi^{-n}\right)^i\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_{-n} \\ &\prec L\sum_{i=1}^{\infty}\left(l^{-1}\Delta^{-n}\right)^i\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_{-n} \\ &= \frac{L}{l}\Delta^{-n}\left(1-l^{-1}\Delta^{-n}\right)^{-1}\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_{-n} \end{aligned} \tag{5.14b}$$

If $n > 1$ (5.9a) and (5.9b) imply

$$\begin{aligned} [h_1(p)]_n &\prec \frac{1}{l}\sum_{i=1}^{\infty}\left(l^{1-n}\Psi^{n}\right)^i\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_n \\ &= \frac{1}{l}\sum_{i=1}^{\infty}\left(l\Delta^{n}\right)^i \\ &= \Delta^{n}\left(1-l\Delta^{-n}\right)^{-1}\left[\frac{\mathcal{H}^2}{c_1-\mathcal{H}}\right]_n \end{aligned} \tag{5.15a}$$

$$[h_2(p)]_{-n} \prec L \sum_{i=0}^{\infty} \left(l^{1-n}\Psi^n\right)^i \left[\frac{\mathcal{H}^2}{c_1 - \mathcal{H}}\right]_{-n}$$

$$(5.15b) \qquad = L\left(1 - l\Delta^n\right)^{-1} \left[\frac{\mathcal{H}^2}{c_1 - \mathcal{H}}\right]_{-n}$$

Looking ahead to combining (5.14) and (5.15) with (5.13) we observe that we will need to estimate

$$n\Delta^{-n}\left(1 - l^{-1}\Delta^{-n}\right)^{-2}$$

for $n \leq 0$, and

$$n\Delta^n\left(1 - l\Delta^n\right)^{-2}$$

for $n > 1$. Let $\Omega = \Delta - l^{-1}$.

If $n \leq 0$,

$$\begin{aligned}
(1 - l^{-1}\Delta^{-n})^{-1} &= \sum_{i=0}^{\infty}\left(l^{-1}\Delta^{-n}\right)^i \\
&\prec \sum_{i=0}^{\infty}\left(\Delta^{-n}\right)^i \\
&\prec \sum_{i=0}^{\infty}(\Delta)^i \\
&= (1-\Delta)^{-1} \\
&= (c_2 - \Omega)^{-1}
\end{aligned}$$

where $c_2 = 1 - l^{-1}$.

$$\begin{aligned}
n\Delta^{-n} &\prec \sum_{i=0}^{\infty}(i+1)\Delta^i \\
&= (1-\Delta)^{-2} \\
&\prec (c_3 - \Omega)^{-1}
\end{aligned}$$

where $0 < c_3 < c_2$ is chosen as in Lemma 5.1.

So

$$n\Delta^{-n}(1 - l\Delta^n)^{-2} \prec (c_3 - \Omega)^{-1}(c_2 - \Omega)^{-2}$$

$$\prec (c_4 - \Omega)^{-1} \tag{5.17}$$

for some c_4 defined by Lemma 5.1.

If $n > 1$ we can make the following estimates.

$$n\Delta^n \prec (c_3 - \Omega)^{-1}$$

exactly as above, and

$$\begin{aligned}(1 - l\Delta^n)^{-1} &\prec \sum_{i=0}^{\infty}\left(l^{\frac{1}{n}}\Delta\right)^{ni} \\ &\prec \sum_{i=0}^{\infty}\left(l^{\frac{1}{2}}\Delta\right)^{ni} \\ &\prec \sum_{i=0}^{\infty}\left(l^{\frac{1}{2}}\Delta\right)^{i} \\ &\prec (1 - l^{\frac{1}{2}}\Delta)^{-1} \\ &\prec (c_5 - \Omega)^{-1}\end{aligned}$$

where $0 < c_5 < 1 - l^{\frac{1}{2}}$.

Combining these results we may estimate the desired quantity.

$$\begin{aligned}n\Delta^n(1 - l\Delta^n)^{-2} &\prec (c_3 - \Omega)^{-1}(c_5 - \Omega)^{-2} \\ &\prec (c_6 - \Omega)^{-1}\end{aligned}$$

for some c_6 constructed by Lemma 5.1 and (for convenience) smaller than c_4.

Substituting the above relations into (5.13) we obtain the following.

$$
\begin{aligned}
xy - [h_1(p)]_1 y &\prec \frac{L}{l^2} \sum_{n\leq 0} n\Delta^n (1 - l\Delta^n)^{-2} [h_1(p)]_n [h_2(p)]_{-n} \\
&\quad + L \sum_{n>1} n\Delta^n (1 - l\Delta^n)^{-2} [h_1(p)]_n [h_2(p)]_{-n} \\
&\prec \left(\frac{\mathcal{H}^2}{c_1 - \mathcal{H}} \right)^2 (c_6 - \Omega)^{-1} \\
&\prec \frac{\mathcal{H}^4}{c_7 - (\mathcal{H} + \Omega)}
\end{aligned} \tag{5.18}
$$

where we have used Lemmas 5.1, 5.2, and 5.4 to choose c_7.

From (5.14a), (5.15a), (5.16), and (5.17) we have

$$
\begin{aligned}
h_1(p) - [h_1(p)]_1 &= \sum_{n\neq 1} [h_1(p)]_1 \\
&\prec \frac{\mathcal{H}^2}{c_1 - (\mathcal{H} + \Omega)}
\end{aligned} \tag{5.19}
$$

Due to (5.12), (5.14b), (5.15b), (5.16), and (5.17) we obtain

$$
h_2(p,\cdot) - y + y\lambda^{-1}(p)(\Psi(p,\cdot) - 1) \prec \frac{\mathcal{H}^2}{c_1 - (\mathcal{H} + \Omega)} \tag{5.20}
$$

The estimates (5.11) - (5.20) hold at each point $p \in \mathbb{T}^2$ so taking the maximum in p term by term and recalling the definitions of the series H_1 and H_2 and the positivity of $[H_1]_1 - x$ we see that

$$y([H_1]_1 - x) \prec \frac{\mathcal{H}^4}{c_7 - (\mathcal{H} + \Omega)} \tag{5.21}$$

$$H_1 - [H_1]_1 \prec \frac{\mathcal{H}^2}{c_7 - (\mathcal{H} + \Omega)} \tag{5.22}$$

$$H_2 - y + y\Omega \prec \frac{\mathcal{H}^2}{c_7 - (\mathcal{H} + \Omega)} \tag{5.23}$$

where c_1 has been replaced by the smaller quantity c_7.

Since all of the series (5.21) – (5.23) are composed entirely of positive terms, it suffices to consider the case when $x = y$ (by Lemma 5.6). We want to combine (5.21) and (5.22) in such a way that the $[H_1]_1$ terms cancel. Eventually we hope to put the resulting expression in the form of Lemma 5.7. The series used in the lemma must majorize $\mathcal{H} + \Omega$ as well as the expression obtained from adding (5.21) – (5.23).

Multipling (5.21) by x^{-2} and (5.22) and (5.23) by x^{-1} we arrive at the expression

$$x^{-1}\mathcal{H} - 2 + \Omega \prec \frac{x^{-2}\mathcal{H}^4}{c_9 - \mathcal{T}} + 2\frac{x^{-1}\mathcal{H}^2}{c_9 - \mathcal{T}} \tag{5.24}$$

Let $\mathcal{T} = \mathcal{H} + x^{-1}\mathcal{H} - 2 + \Omega$. Note that $\mathcal{T}$ achieves the goals stated above and that $\mathcal{T}$ is a positive series without constant terms. Let $\mathcal{A} = x^{-1}\mathcal{H} - 2$, then $\mathcal{H} = 2x + x\mathcal{A}$ and $\mathcal{T} = 2x + x\mathcal{A} + \mathcal{A} + \Omega$. Hence $\mathcal{H} \prec 2x + \mathcal{T}^2$. We also have

$$x^{-1}\mathcal{H}^2 \prec 4x + \mathcal{T}^2 \prec 2\mathcal{T}(1 + \mathcal{T})$$

and so

$$\begin{aligned} x^{-2}\mathcal{H}^4 &\prec 4\mathcal{T}^2(1 + \mathcal{T})^2 \\ &\prec \frac{4\mathcal{T}^2}{1 - 2\mathcal{T}} \end{aligned}$$

Combining these estimates we have

$$\mathcal{T} \prec \frac{4\mathcal{T}^2}{(c_8 - \mathcal{T})(1 - 2\mathcal{T})} + \frac{4x + \mathcal{T}^2}{c_8 - \mathcal{T}} + 2x + \mathcal{T}^2$$

By choosing a smaller constant c_9 we may eliminate the 4 in the numerator of the first term and combine the denominator into a single expression $c_9 - \mathcal{T}$. Replacing c_8 with c_9 in the second term as well and adding we achieve the following relation.

$$\mathcal{T} \prec \frac{2(2x + \mathcal{T}^2)}{c_9 - \mathcal{T}} + 2x + \mathcal{T}^2$$

By replacing c_9 with a smaller constant, c, we can absorb the extra terms (by applying Lemmas 5.4 and 5.5) to attain the final majorization inequality which is required.

$$\mathcal{T} \prec \frac{2x + \mathcal{T}^2}{c - \mathcal{T}}$$

By Cauchy's Lemma (5.7),

$$\mathcal{T} \prec \frac{c}{4} - \sqrt{\left(\frac{c}{4}\right)^2 - x}$$

which converges in a neighborhood of the origin. The convergence of $\mathcal{T}$ implies the convergence of $\mathcal{H}$ and Ω from which it follows that h_p and g_p are analytic for every $p \in \mathbb{T}^2$. $\square$

CHAPTER VI

C^∞ NORMAL FORMS

In this chapter we will prove a C^∞ version of the analytic result from the previous chapter. The similarities between the two proofs end with the referral to the formal result from chapter 4. Our proof is modeled on the work of S. Sternberg who succeeded in showing the stationary case of this theorem.

Theorem. *Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be a C^∞, area preserving, Anosov diffeomorphism and let f_p denote the local Taylor series representation for f at $p \in \mathbb{T}^2$. Then there exists a family of C^∞ coordinate changes, h_q, continuous in q for all $q \in \mathbb{T}^2$, such that*

$$f_p \circ h_p = h_{f(p)} \circ \mathfrak{G}_p \tag{6.1}$$

where $\mathfrak{G}_p$ is the local normal form for f at p. The Taylor series for $\mathfrak{G}_p$ at p is denoted by g_p.

$$g(p, x, y) = (\lambda(p)x\Phi(p, xy), \lambda^{-1}(p)x\Psi(p, xy)) \tag{6.2}$$

As in the formal and analytic cases compactness is the key ingredient in dealing with the nonstationary parameter.

The proof proceeds roughly as follows.

(1) Use the formal result.

(2) By a general principle we can realize the formal conjugacy and normal form by C^∞ functions which have the correct behavior at the origin.

(3) Step (2) can be achieved in an area preserving manner using basic existence and uniqueness for a linear partial differential equation ($Jac(h) = 1$).

(4) Make another preparatory change of coordinates which normalizes our mapping along $W^s_{loc}(p)$ and $W^u_{loc}(p)$. This requires a nonstationary version of Sternberg's result for contractions, [S2].

(5) Construct a (not necessarily area preserving) conjugacy which normalizes our diffeomorphism in a small neighborhood of p and verify that the mapping under construction is smooth.

(6) Show that one can perturb the constructed conjugacy in a way which makes it area preserving.

Proof. We will use $\mathfrak{h}, \mathfrak{H}$ to denote the sequence of preparatory normalizing transformations mentioned in Remarks (1) - (6). We have already constructed a formal solution to the problem at hand. It remains to show that this formal series can be realized by a family of C^∞ functions which are area preserving in the neighborhoods defined by the family of local coordinate systems at each $p \in \mathbb{T}^2$. The first step is Borel's lemma [Bo].

Lemma 6.1. *Let $A(x,y) = \sum_{i,j} a_{ij} x^i y^j$ be any formal power series. Then there exists a C^∞ function whose Taylor series at the origin is $A(x,y)$.*

Proof. We introduce a C^∞ bump function $\alpha : \mathbb{R} \to \mathbb{R}$, such that $\alpha = 1$ if $|x| < 1$ and $\alpha = 0$ if $|x| > 2$. Then let

$$\mathfrak{A}(x,y) = \sum_{ij} a_{ij} x^i y^j \alpha\left((i+j)! M_m (x^2 + y^2)\right) \tag{6.3}$$

where

$$M_m = \sum_{k=0}^{m} \sum_{i+j=k} |a_{ij}|.$$

The function $\mathfrak{A}$ is a C^∞ function with the desired Taylor series at the origin.

With this lemma in mind we begin by representing the formal series h_p as a C^∞ function.

Lemma 6.2. *There exists a C^∞ function, $\mathfrak{h}_p$ which has h_p as its Taylor series at the origin.*

Proof. Using the previously established notation $h_p = (h_1(p), h_2(p))$, we begin by representing $h_1(p)$ by a C^∞ function, $\mathfrak{h}_1(p)$ as is possible by lemma 6.1. The series $h_2(p, x, 0)$ can also be represented by a C^∞ function, $b_1(x)$ defined on the x-axis. Since we want $\mathfrak{h}$ to be area preserving at the origin we seek a solution to the partial differential equation defined by $Jac(\mathfrak{h}) = 1$, i.e.,

$$\frac{\partial \mathfrak{h}_1}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial \mathfrak{h}_1}{\partial y}\frac{\partial b}{\partial x} = 1$$

A straightforward application of the method of characteristics (see [Zw] or any standard text on partial differential equations) indicates that there exists a unique solution $b(p)$ with the stated initial conditions.

Let $\mathfrak{h}(p) = (\mathfrak{h}_1(p), b(p))$ and $\mathfrak{h}_p$ is a C^∞ realization of the formal transformation, h_p.

Remark. *The same result holds for the formal series, g_p.*

In the case of g_p it suffices to represent Φ_p (or Υ_p) by a C^∞ function. Then $\mathfrak{g}_p$ is C^∞, area preserving, and still preserves the curves $xy = constant$.

Let $\mathfrak{E}_p = f_p \circ \mathfrak{h}_p - \mathfrak{h}_{f(p)} \circ \mathfrak{g}_p$ denote the error introduced into the conjugacy equation by the construction in lemma 6.2. Note that $\mathfrak{E}_p$ vanishes to all orders as its argument tends to the origin (in the local coordinates at p). We will call such a function infinitely flat. The precise estimate for $\mathfrak{E}_p$ is that for all $j \in \mathbb{Z}^+$ there exist constants E_j such that

$$\|\mathfrak{E}_p(x,y)\| < E_j \|(x,y)\|^j \tag{6.4}$$

for (x, y) near the origin.

Unfortunately our construction has only assured us that the conjugacy equation is satisfied at p. Our next step is to normalize along W_p^s and W_p^u by an area preserving C^∞ diffeomorphism. Recall that our initial choice of coordinates made $W^u_{loc}(p)$ and $W^s_{loc}(p)$ the x and y axes. To achieve the desired normalization we require a nonstationary version of Sternberg's theorem for contracting maps, [S1].

Theorem 6.3. *Let B be a line bundle over $\mathbb{T}^2$ and $F : B \to B$ a function such that*

(1) $F(p, x) = (f(p), F_p(x))$
(2) $F_p(x) = \mu(p)\,x + \ldots$ *where* $\mu(p) < 1$ *for all* $p \in \mathbb{T}^2$.
(3) F_p *is* C^∞ *for each* $p \in \mathbb{T}^2$ *and continuous as one moves the base point* p.

Then there exists $H : B \to B$ with the properties:

(1) $H(p, x) = (p, H_p(x))$
(2) $F \circ H = H \circ G$ *where* $G(p, x) = (p, \mu(p)x)$
(3) H_p *is a* C^∞ *diffeomorphism which is continuous in* p.

Remarks.

(1) *This theorem provides a linearizing transformation, denoted* k_p^s *for* $\mathfrak{g}_p$ *along* $W_{loc}^s(p)$.

(2) *The same theorem applies to* $\mathfrak{g}_p^{-1}$ *along* $W_{loc}^u(p)$.

Now consider $\mathfrak{g}_p$ restricted to $W_{loc}^s(p)$, i.e., $\mathfrak{g}_p(0,y) = \lambda^{-1}(p)y + \ldots$. By theorem 6.3 we can construct a C^∞ function k_p^s such that

$$\mathfrak{g}_p\big|_{W_{loc}^s(p)} \circ k_p^s = K_{f(p)}^s \circ \Lambda_p^{-1}$$

where $\Lambda_p^{-1} y = \lambda^{-1}(p)y$. The same theorem applies to $\mathfrak{g}_p^{-1}$ along $W_{loc}^u(p)$ which in turn linearizes $\mathfrak{g}_p$ on $W_{loc}^u(p)$. Thus we may construct k_p^u such that

$$\mathfrak{g}_p\big|_{W_{loc}^u(p)} \circ k_p^u = k_{f(p)}^u \circ \Lambda_p \tag{6.5}$$

Now let

$$K_p^u = \left(k_p^u(x), \frac{y}{\left|\frac{\partial k_p^u}{\partial x}\right|} \right) \tag{6.6}$$

and

$$K_p^s = \left(\frac{x}{\left|\frac{\partial k_p^s}{\partial y}\right|}, k_p^s(y) \right)$$

Note that $Jac(K_p^s) = 1 = Jac(K_p^u)$. Conjugation by $K_p^u \circ K_p^s$ places $\mathfrak{g}_p$ in normal form along $W_{loc}^u(p)$ and $W_{loc}^s(p)$. Let us use the same symbol, $\mathfrak{g}_p$ for this new C^∞ partial normal form.

Theorem 6.3 is a special case of a more general theorem about skew products of $\mathbb{R}^n$-bundles over compact manifolds to be discussed in chapter 7.

Summing up the previous paragraphs we have the following lemma.

Lemma 6.4. *By the constructions above we can consider $\mathfrak{g}_p$ to be in local normal form at p and along $W^s_{loc}(p)$ and $W^u_{loc}(p)$ for all $p \in \mathbb{T}^2$.*

We will now show how to construct a C^∞ conjugacy, $\mathfrak{H}_p$ which normalizes $\mathfrak{g}_p$ in a neighborhood of p. The size of this neighborhood will have a uniform lower bound over all points on the torus. Take $\mathfrak{f}_p$ to be in partial normal form. Thus the next family of local changes of coordinates may be chosen to be the identity plus an infinitely flat part at p.

Lemma 6.5. *There exists a family of local diffeomorphisms, $\mathfrak{H}_p$, such that*

(1) *The restriction of $\mathfrak{H}_p$ to $W^s_{loc}(p)$ and $W^u_{loc}(p)$ is the identity.*

(2) *The diffeomorphisms, $\mathfrak{H}_p$ are infinitely flat as they approach $W^s_{loc}(p)$ or $W^u_{loc}(p)$.*

(3) $\mathfrak{H}_{f(p)} \circ \mathfrak{f}_p = \mathfrak{G}_p \circ \mathfrak{H}_p$

Remarks.

(1) *Note the conspicuous absence of preservation of area from lemma 6.5. After constructing $\mathfrak{H}_p$ a separate argument (and construction) will yield an area preserving family of conjugacies.*

(2) *The purpose of the construction of $\mathfrak{H}_p$ is to eliminate the error term, $\mathfrak{E}_p$, from lemma 6.2.*

Before we begin the proof we need to establish some notation. Let $X = (x, y) \in W^u_{loc}(p) \times W^s_{loc}(p) \subset \mathbb{T}^2$. Occasionally we may write X_p or $(x, y)_p$ to emphasize the location of the local coordinates. Let

$$\|X\|^u = x^2 \quad \text{and} \quad \|X\|^s = y^2$$

$$\|X\| = \|X\|^u + \|X\|^s$$

Let $\mathbb{B}_p(r)$ denote the ball of radius r in the local coordinates at p with the metric defined above, i.e.,

$$\mathbb{B}_p(r) = \{X \mid \|X\| < r\}$$

Let

$$\begin{aligned} \mathcal{C}_p &= \{(x,y)_p \mid |x| = |y|\} \\ \mathcal{C}_p^u &= \{(x,y)_p \mid |x| > |y|\} \qquad (6.7) \\ \mathcal{C}_p^s &= \{(x,y)_p \mid |x| < |y|\} \end{aligned}$$

The cone $\mathcal{C}_p^u$ contains the unstable axis while $\mathcal{C}_p^s$ contains the stable axis. Note that

$$\mathfrak{f}^n_{f^{-n}(p)}\left(\mathcal{C}^u_{f^{-n}(p)}\right) \longrightarrow W^u_{loc}(p) \text{ as } n \to \infty \text{ and}$$

$$\mathfrak{f}^n_{f^{-n}(p)}\left(\mathcal{C}^s_{f^{-n}(p)}\right) \longrightarrow W^s_{loc}(p) \text{ as } n \to -\infty$$

To construct the desired conjugacy we choose a fundamental domain, $\mathcal{S}_p^{0,r}$ and define $\mathfrak{H}_p$ there, then use the iterated conjugacy equation to define $\mathfrak{H}_p$ throughout a neighborhood of p. Let $\mathcal{S}_p^{0,r}$ denote the wedge of $\mathbb{B}_p(r)$ bounded by $\mathcal{C}_p$ and $f(\mathcal{C}_{f^{-1}(p)})$. Choose r_0 small enough that

$$\|\hat{\mathfrak{f}}_p(X)\| < \epsilon\|X\| \text{ for all } X \in \mathbb{B}_p(r_0) \qquad (6.8)$$

Recall that $\hat{f}_p(X)$ denotes the higher order terms of f_p. Compactness allows us to take r_0 of uniform size over all local coordinate systems.

In $\mathbb{B}_p(r_0)$ we have the following estimates.

$$\|\mathfrak{f}_p\|^u > (\lambda(p) - \epsilon)\|X\|^u > (l - \epsilon)\|X\|^u$$

$$\|\mathfrak{f}_p\|^s > (\lambda^{-1}(p) - \epsilon)\|X\|^s > (l^{-1} - \epsilon)\|X\|^s$$

where l denotes the minimum of $\lambda(p)$ over $\mathbb{T}^2$. By choosing ϵ small enough we can find a (global) constant $c > 1$ which satisfies

$$\|\mathfrak{f}_p\|^u > c\|X\|^u \quad \text{and}$$

$$\|\mathfrak{f}_p\|^s < c^{-1}\|X\|^s \tag{6.9}$$

Thus in $\mathbb{B}_p(r_0)$, c controls the local expansion in $\|\cdot\|^u$ and the local contraction in $\|\cdot\|^s$ for each $p \in \mathbb{T}^2$. Now let $r < r_0$ and $\mathcal{S}_p^{0,r}$ be the fundamental domain described above. Define

$$\mathcal{S}_p^{n,r} = \mathfrak{f}_{f^{-1}(p)}\left(\mathcal{S}_{f^{-1}(p)}^{n-1,r}\right) \cap \mathbb{B}_p(r) \text{ if } n > 0 \text{ and}$$

$$\mathcal{S}_p^{n,r} = \mathfrak{f}_{f(p)}^{-1}\left(\mathcal{S}_{f(p)}^{n+1,r}\right) \cap \mathbb{B}_p(r) \text{ if } n < 0 \tag{6.10}$$

Lemma 6.6. *The wedges $\mathcal{S}_p^{n,r}$ cover a neighborhood of p (except for the axes).*

Proof. We must verify the previous claim that

$$\mathfrak{f}_{f^{-n}(p)}^n\left(\mathcal{C}_{f^{-n}(p)}\right) \longrightarrow W_{loc}^u(p) \qquad \textit{as } n \to \infty \tag{6.11}$$

The estimate (6.9) indicates that

$$\|\mathfrak{f}_{f^{-n}(p)}^n(X)\|^s \leq \frac{1}{c^n}\|X\|^s$$

where $c > 1$. Hence the assertion holds. Similarly

$$\mathfrak{f}_{f^{-n}(p)}^n\left(\mathcal{C}_{f^{-n}(p)}\right) \longrightarrow W_{loc}^u(p) \qquad \textit{as } n \to -\infty$$

By construction there are no "gaps" between $\mathcal{S}_p^{i,r}$ and $\mathcal{S}_p^{i+1,r}$ thus

$$\cup_{n=-\infty}^{\infty}\mathcal{S}_p^{n,r} \supset \mathbb{B}_p(r) - (W_{loc}^s(p) \cup W_{loc}^u(p))$$

We will also need a lemma which regulates how $\mathfrak{f}(\mathcal{S}_{f^{-1}(p)}^{n-1,r})$ covers $\mathcal{S}_p^{n,r}$.

Lemma 6.7. *There exists a constant $\tau < 1$ and an integer κ such that for $n > \kappa$ and r sufficiently small we have*

$$\mathfrak{f}\left(\mathcal{S}^{n-1,\tau r}_{f^{-1}(p)}\right) \supset \mathcal{S}^{n,r}_{p} \tag{6.12}$$

Proof. Fix $\epsilon > 0$. From the proof of lemma 6.5 we can choose κ_p such that for all $k > \kappa_p$ and $X \in \mathcal{S}^{k,r}_p$

$$\|X\|^u > (1-\epsilon)\|X\|$$

Then

$$\|\mathfrak{f}_p(X)\|^u > (1-\epsilon)\|\mathfrak{f}(X)\| > c(1-\epsilon)\|X\| > 1$$

For ϵ small enough $c(1-\epsilon) > 1$ thus we may choose $\tau = \frac{1}{c(1-\epsilon)}$. Let $\kappa = \sup_{p\in\mathbb{T}^2}\kappa_p < \infty$, and the lemma is verified. □

We will construct $\mathfrak{H}_p$ by defining it on the wedge $\mathcal{S}^{0,r}_q$ for all $q \in \mathbb{T}^2$. The conjugacy equations will determine $\mathfrak{H}$ in the neighborhoods, $\mathbb{B}_q(r)$. Begin by defining $\mathfrak{H}_p\big|_{\mathcal{C}_p} = Id_p$ for all $p \in \mathbb{T}^2$ where Id_p denotes the identity map. Then $\mathfrak{H}$ is defined on $\mathfrak{f}(\mathcal{C}_{f^{-1}(p)}) \subset \mathcal{C}^u_p$ by the equation

$$\mathfrak{H}_p \circ \mathfrak{f}_{f^{-1}(p)} = \mathfrak{g}_{f^{-1}(p)} \circ \mathfrak{H}_{f^{-1}(p)} \tag{6.13}$$

The equation

$$\mathfrak{H}_p \circ \mathfrak{f}^n_{f^{-n}(p)} = \mathfrak{g}^n_{f^{-n}(p)} \circ \mathfrak{H}_{f^{-n}(p)} \tag{6.14}$$

defines $\mathfrak{H}_p$ on a family of curves tending uniformly (as $n \to \infty$) to $W^u_{loc}(p)$ in a small neighborhood of p. To define $\mathfrak{H}_p$ on $\mathcal{S}^{0,r}_p$ we choose a C^∞ function which tends to the identity in an infinitely flat manner as it approaches $\mathcal{C}_p$ and which agrees with the requirements of equation (6.13) in a smooth manner (This is possible as in [S1].). Equation (6.14) defines $\mathfrak{H}_p$ in each wedge, $\mathcal{S}_p^{n,r}$, and at each $p \in \mathbb{T}^2$. Moreover the construction makes $\mathfrak{H}_p$ a C^∞ function over the union of all of the wedges. All that remains is to check that $\mathfrak{H}_p$ is C^∞ at $W^s_{loc}(p)$ and $W^u_{loc}(p)$. (Due to our previous normalizations we require $\mathfrak{H}_p$ to be the identity on $W^s_{loc}(p)$ and $W^u_{loc}(p)$ so that $\mathfrak{f}$ remains normalized there.) We begin by demonstrating the continuity of $\mathfrak{H}_p$.

Lemma 6.8. *The definition above makes $\mathfrak{H}_p$ continuous at $W^s_{loc}(p)$, $W^u_{loc}(p)$, and the origin.*

Proof. Assume that X_n is a sequence of points in $\mathbb{B}_p(r)$ approaching $X \in W^u_{loc}(p)$. Since our definition makes $\mathfrak{H}_p$ the identity on $W^u_{loc}(p)$ we have

$$\begin{aligned} \|\mathfrak{H}_p(X_n) - \mathfrak{H}_p(X)\| &= \|\mathfrak{H}_p(X_n) - X\| \\ &< \|(\mathfrak{H}_p - Id_p)(X_n)\| + \|X - X_n\| \end{aligned} \tag{6.15}$$

So it suffices to show that $(\mathfrak{H}_p - Id_p)(X_n)$ vanishes as $n \to \infty$.

Since the wedges $\mathcal{S}_p^{k,r}$ cover $\mathbb{B}_p(r)$ we can define a function $k(n) : \mathbb{Z}^+ \to \mathbb{Z}^+$ where $X_n \in \mathcal{S}_p^{k(n),r}$. Note that $k(n) \to \infty$ as $n \to \infty$.

Let $\mathfrak{R}_p = \mathfrak{G}_p - \mathfrak{f}_p$ and recall that our previous normalizations make $\mathfrak{R}_p$ infinitely flat at the origin in all of the local coordinate systems. From its construction

$$\mathfrak{H}_p|_{\mathcal{S}_p^{k,r}} = \mathfrak{G}^k_{f^{-k}(p)} \circ \mathfrak{H}_{f^{-k}(p)} \circ \mathfrak{f}_p^{-k}$$

Let

$$\mathfrak{D}_p^k = \mathfrak{H}_p|_{\mathcal{S}_p^{k,r}} - Id_p \tag{6.16}$$

Observe that

$$\mathfrak{D}_p^0 = \mathfrak{H}_p|_{\mathbb{S}_p^{0,r}} - Id_p \qquad \text{is infinitely flat}$$

and that $\mathfrak{D}_p^k$ satisfies

$$\mathfrak{D}_p^k = \mathfrak{G}_{f^{-1}(p)} \circ \left(Id_{f^{-1}(p)} + \mathfrak{D}_{f^{-1}(p)}^{k-1}\right) \circ \mathfrak{f}_p^{-1} - Id_p \tag{6.17}$$

By the mean value theorem there exists ξ_k such that

$$\begin{aligned}\mathfrak{D}_p^k &= \mathfrak{G}_{f^{-1}(p)} \circ \mathfrak{f}_p^{-1} + d\mathfrak{G}_{f^{-1}(p)}(\xi_k)\cdot\mathfrak{D}_{f^{-1}(p)}^{k-1} \circ \mathfrak{f}_p^{-1} - Id_p \\ &= \mathfrak{R}_{f^{-1}(p)} \circ \mathfrak{f}_p^{-1} + d\mathfrak{G}_{f^{-1}(p)}(\xi_k)\cdot\mathfrak{D}_{f^{-1}(p)}^{k-1} \circ \mathfrak{f}_p^{-1}\end{aligned} \tag{6.18}$$

Repeated application of the mean value theorem yields the relation

(6.19)

$$\mathfrak{D}_p^k = \sum_{i=0}^{k-1}\left(\prod_{j=0}^{i-1} d\mathfrak{G}_{f^{-j}(p)}(\xi_{k-j})\right)\mathfrak{R}_{f^{-1}(p)} \circ \mathfrak{f}_p^{-i} + \left(\prod_{i=1}^{k} d\mathfrak{G}_{f^{-i}(p)}(\xi_i)\right)\mathfrak{D}_{f^{-k}(p)}^0 \circ \mathfrak{f}_p^{-k}$$

Since $\mathfrak{G}_p = \mathfrak{f}_p + \mathfrak{R}_p$, we can find $\mathcal{M}$ such that

$$\mathcal{M} = \max_{p\in\mathbb{T}^2}\max_{\xi\in\mathbb{B}_p(r)} \|d\mathfrak{G}(\xi)\| \tag{6.20}$$

As $\mathfrak{R}_p$ and $\mathfrak{D}_p^0$ are infinitely flat, we can find constants D_j for any $j \in \mathbb{Z}^+$ such that

$$\sup_{p\in\mathbb{T}^2}\sup_{X\in\mathbb{B}_p(r)} \mathfrak{D}_p^0(X) \leq D_j\|X\|^j \tag{6.21}$$

$$\sup_{p\in\mathbb{T}^2}\sup_{X\in\mathbb{B}_p(r)} \mathfrak{R}_p(X) \leq D_j\|X\|^j \tag{6.22}$$

The continuity of $\mathfrak{H}_p$ at $W_{loc}^u(p)$ depends upon the amount of control which can be exercised over $\mathfrak{D}_p^k$ as $k \to \infty$ (i.e. as $X_n \to X$). Recall the definition of τ from lemma 6.6 and choose j such that $\sigma = \tau^j \mathcal{M} < 1$. Using lemma 6.6 we have the estimates

$$\sup_{p\in\mathbb{T}^2} \sup_{X\in\mathbb{B}_p(r)} \|\mathfrak{D}_p^0 \circ \mathfrak{f}^{-i}(X)\| \le D_j \|\mathfrak{f}^{-i}(X)\|^j$$

$$\le D_j \left(2r\tau^{i-\kappa}\right)^j \tag{6.23}$$

If $i < k - \kappa$ then

$$\sup_{p\in\mathbb{T}^2} \sup_{X\in\mathbb{B}_p(r)} \|\mathfrak{R}_p \circ \mathfrak{f}^{-i}(X)\| \le D_j \left(r\tau^{i}\right)^j \tag{6.24}$$

and if $i > k - \kappa$, then

$$\sup_{p\in\mathbb{T}^2} \sup_{X\in\mathbb{B}_p(r)} \|\mathfrak{R}_p \circ \mathfrak{f}^{-i}(X)\| \le D_j \left(2r\tau^{k-\kappa}\right)^j \tag{6.25}$$

We can use (6.21) and (6.23) to estimate the last term of (6.19), and (6.20) - (6.25) to estimate the sum.

$$\|\mathfrak{D}_p^k\| \le \sum_{i=0}^{\kappa+l-1} \mathcal{M}^i \|\mathfrak{R}_p \circ \mathfrak{G}^{-i}\| + D_j r^j \frac{\sigma^{\kappa+l}}{1-\sigma} + D_j 2^j \mathcal{M}^\kappa \sigma^{k-\kappa}$$

Fix some $\epsilon > 0$. By choosing l sufficiently large the middle term can be bounded above by ϵ. Having fixed l we see that as $n \to \infty$

$$\mathfrak{G}^{-i}(X_n) \to W_{loc}^u(p)$$

and since $\mathfrak{R}$ is infinitely flat as it approaches $W_{loc}^u(p)$, the first term and the last term ($\sigma < 1$) tends to zero. Thus

$$\|\mathfrak{D}_p^{k(n)}(X_n)\| \to 0 \qquad as\ n \to \infty$$

and $\mathfrak{H}_p$ is continuous at $W_{loc}^u(p)$. It follows from a similar argument that $\mathfrak{H}_p$ is continuous at $W_{loc}^s(p)$ and the origin for all $p \in \mathbb{T}^2$.

Lemma 6.9. *The family of conjugacies $\mathfrak{H}_p$ is C^∞ in $\mathbb{B}_p(r)$.*

Proof. The only issue is the smoothness at $W^s_{loc}(p)$ and $W^u_{loc}(p)$. One may argue by induction, anchored by the continuity lemma. Suppose $\mathfrak{H}_p$ and all of its derivatives up to order $n-1$ are continuous at $W^u_{loc}(p)$ for all $p \in \mathbb{T}^2$. The induction step is verified in a manner analogous to the continuity lemma. The main difference is that the objects are contravarient tensors of various orders instead of vector valued functions and Jacobians. Let ∇ denote the differential operator on the space of symmetric tensors composed of partial derivatives of smooth functions on $\mathbb{T}^2$. That is ∇ takes the tensor containing all partial derivatives of order n of a vector valued function to partial derivatives of order $n+1$. Let $A(X) = \left(A^1(X), \ldots, A^n(X)\right)$.

$$\nabla^n A = \left(A^i_{j_1 \cdots j_n}\right)$$

where

$$A^i_{j_1 \cdots j_n} = \frac{\partial^n A^i}{\partial x_{j_1} \cdots \partial x_{j_n}}$$

Taking ∇ of (6.19) we get

$$\begin{aligned}
\nabla \mathfrak{D}^k_p = & \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i-1} \nabla \mathfrak{G}_{f^{-j}(p)} \right) \nabla \mathfrak{R}_p \prod_{i=0}^{i-1} \nabla \mathfrak{f}^{-1}_{f^{-j}(p)} \\
& + \sum_{i=0}^{k-1} \left(\sum_{l=0}^{i-1} \left(\prod_{j=0, j\neq l}^{i-1} \nabla \mathfrak{G}_{f^{-j}(p)} \right) \nabla^2 \mathfrak{G}_{f^{-l}(p)} \right) \mathfrak{R}_{f^{-i}(p)} \circ \mathfrak{f}_p^{-i} \\
& + \sum_{l=0}^{k} \left(\prod_{j=0, j\neq l}^{k} \nabla \mathfrak{G}_{f^{-j}(p)} \right) \left(\nabla^2 \mathfrak{G}_{f^{-l}(p)}\right) \mathfrak{D}^0_{f^{-j}(p)} \circ \mathfrak{f}_p^{-k} \\
& + \prod_{j=0}^{k} \nabla \mathfrak{G}_{f^{-j}(p)} \nabla \mathfrak{D}^0_{f^{-j}(p)} \prod_{j=0}^{k} \nabla \mathfrak{f}^{-1}_{f^{-j}(p)}
\end{aligned}$$

Split the summations as in the continuity lemma and redefine $\mathcal{M}$ as

$$\mathcal{M} = \sup_{p\in\mathbb{T}^2} \sup_{\xi\in\mathbb{B}_p(r)} \left(\|\nabla\mathfrak{G}_p(\xi)\|, \|\nabla\mathfrak{f}_p^{-1}\|, \|\nabla^2\mathfrak{G}_p(\xi)\|, \|\mathcal{H}_p\|\right).$$

The continuity of $\mathcal{H}_p$ implies $\mathcal{M} < \infty$ for r small. Since $\mathfrak{R}_p$ is infinitely flat, $\nabla\mathfrak{R}_p$ is also. Hence for all $j \in \mathbb{Z}^+$, and all *ptor*, there is a (uniform) constant R_j such that

$$\|\nabla\mathfrak{R}_p(X)\| \leq R_j\|X\|^j$$

$$\|\mathfrak{R}_p(X)\| \leq R_j\|X\|^j$$

$$\|\nabla\mathcal{D}_p^0(X)\| \leq R_j\|X\|^j$$

$$\|\mathcal{D}_p^0(X)\| \leq R_j\|X\|^j$$

The continuity of $\mathcal{D}_p^k$ follows from estimating the tail of the summations on the right hand side. Choosing X_n close to X makes the remaining terms ϵ-small. By applying ∇^{n+1} to (6.19) we get a complicated but similar expression for the derivatives of $\mathcal{H}_p - Id_p$ of order $n+1$ in terms of derivatives of $\mathcal{H}_p - Id_p$ of order less than or equal to n, and derivatives of $\mathfrak{f}$ and $\mathfrak{G}$ of order lessthan or equal to $n+1$. The induction Hypothesis allows us to redefine $\mathcal{M}$ and R_j based on the continuity of the derivatives of order less than or equal to n. The same procedure as above implies that $\nabla\mathcal{H}_p$ vanishes as $X_n \to X$ and by induction $\mathcal{H}_p - Id_p$ vanishes to all orders and is C^∞ at $W^u_{loc}(p)$. The proof for $W^s_{loc}(p)$ is almost identical. □

Now we will indicate that the preservation of area can be regained by perturbing the construction of $\mathfrak{H}_p$. Note that $\mathfrak{G}_p$ maps the curve $xy = c$ at p into $xy = c$ at $f(p)$. Define $f^t(p) = f^{[t]}(p)$ where $[t]$ denotes the greatest integer function. If

$0 < t < 1$ interpret $\lambda^t(p)$ as usual. For $t \geq 1$ let

$$\lambda^t(p) = \left(\prod_{i=0}^{[t]-1} \lambda(f^i(p)) \right) \lambda^{t-[t]}(f^{[t]}(p))$$

Consider the family of mappings $\mathfrak{G}_p^t$, which can be expressed in exponential form.

$$\mathfrak{G}_p^t(x, y) = \left(\lambda^t(p) x \, \mathrm{e}^{t\Upsilon(f^t(p), xy)}, \lambda^{-t}(p) x \, \mathrm{e}^{-t\Upsilon(f^t(p), xy)} \right)$$

By conjugating with $\mathfrak{H}_p$ we embed $\mathfrak{f}_p$ in a similar one parameter family $\mathfrak{f}_p^t$. Although $\mathfrak{f}_p^t$ is not area preserving for each t, it is area preserving for $t \in \mathbb{Z}$. As in [S3] we can construct a time change, $u(t)$, for the family of diffeomorphisms which makes $\mathfrak{f}_p^{u(t)}$ area preserving for each $p \in \mathbb{T}^2$. Now we can construct a new conjugacy $\mathcal{H}_p$ which preserves area as follows. Let $x \in \mathcal{C}_p$ and $X_t = \mathfrak{f}^{u(t)}(X)$ (for $t < 1$ so that $X_t \in \mathbb{B}_p(r)$). Define $\mathcal{H}_p$ on the fundamental domain, $\mathcal{S}_p^{0,r}$ by

$$\mathcal{H}_p(X^t) = \mathfrak{G}_p^t(X).$$

As before the conjugacy extends $\mathcal{H}_p$ to $\mathbb{B}_p(r)$. The considerations of smoothness are no different than before since $\mathfrak{H}_p$ was only modified inside of the wedges (hence the asymptotics of (6.19) and (6.26) are identical). Thus we have constructed the desired C^∞ realization of the conjugacy for theorem 6.1. $\square$

CHAPTER VII

NORMAL FORMS FOR SKEW PRODUCT TRANSFORMATIONS

The nonstationary normal form theorem for area preserving Anosov diffeomorphisms is actually a special case of a more general theorem about skew product dynamical systems. Essentially all of the theorems for local normal forms for diffeomorphisms at a fixed point can be carried out in the skew product setting. We begin by establishing a generalization of theorem 6.3 to higher dimensions.

Let M be a compact manifold and B an $\mathbb{R}^n$-bundle over M. Let $(p, X) \in M \times \mathbb{R}^n$. Recall that a skew product $F : B \to B$ is a mapping of the form

$$F(p, X) = (f(p), F_p(X)) \tag{7.1}$$

We will assume that the fibre map, F_p, changes continuously in the parameter $p \in M$ and that F_p is a diffeomorphism for each p. In addition suppose that F_p has a fixed point at the origin of $\mathbb{R}^n$, i.e., F_p fixes the zero section of our bundle B. Let $\lambda_1(p), \ldots, \lambda_n(p)$ denote the eigenvalues of the linear part of F_p at the origin (for simplicity we assume that the linear part of F_p at the origin is diagonalizable). Let Λ_p denote the diagonal matrix with λ_i's as entries.

Theorem 7.1. *Let $F : B \to B$ be a skew product as above. If all of the eigenvalues of F_p, $\lambda_1(p), \ldots, \lambda_n(p)$ lie inside the unit circle and are free from resonances at each point $p \in M$, then there is a smooth transformation $H : B \to B$ which satisfies*

$$F \circ H = H \circ \Lambda \tag{7.2}$$

where $H(p,X) = (p, H_p(X))$ and $\Lambda(p,X) = (f(p), \Lambda_p(x))$.

Remarks.

(1) *The word smooth can be replaced by analytic, C^k or C^∞. Since we have used a C^∞ version in chapter 6 we will verify this explicitly. The C^k case follows by combining the nonstationary formulation of the C^∞ version found in this chapter with the considerations from [S2], and the analytic case essentially follows from our formulation along with the methods of [La] or [S2].*

(2) *The procedure follows the established pattern. We will cite a formal theorem then argue that there is a smooth realization of the formal series. The case at hand is much more simple than those of the previous chapters since we are dealing with a contracting map.*

(3) *An analogous theorem holds if the eigenvalues lie outside of the unit circle (just consider F^{-1}).*

Proof. Write $F_p = (F_1(p), \ldots, F_n(p))$ and let

$$F_i(p) = \sum_\alpha F^i_\alpha x^\alpha$$

where α is a multiindex $\alpha = (\alpha_1, \ldots, \alpha_n)$ with the usual conventions (e.g.$|\alpha| = \alpha_1 + \cdots + \alpha_m$). Also assume that the eigenvalues have been ordered by their magnitudes i.e., $|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n| < 1$. Composition of series yields

$$\lambda_i(p) H^i_\alpha(p) = \lambda^\alpha H^i_\alpha(f(p)) + l.o.t. \tag{7.3}$$

The nonresonance condition allows a unique solution for $H^i_\alpha(p)$ as an expression of lower order terms. Uniform contraction and compactness make H_p well defined and continuous in p.

By lemma 6.1 we can realize H_p by a C^∞ function so

$$E_p = F_p \circ H_p - H_{f(p)} \circ \Lambda_p \tag{7.4}$$

is infinitely flat.

Let $l = \max_{p \in M} |\lambda_1(p)|$ and $l' = \max_{p \in M} |\lambda_n(p)|$ so that $0 < l < l' < 1$.

Let $\mathcal{O}_p$ be a neighborhood of the origin in $\{p\} \times \mathbb{R}^m$ and $V_{\mathcal{O}}^k$ denote the set of C^k functions on $M \times \mathcal{O}_p$ which vanish up to order k at the origin for each $p \in M$. For $Z = (Z_1(p), \ldots, Z_m(p))$, we have the metric,

$$\|Z\|_{\mathcal{O}}^k = \sup_{p \in M} \sup_{\mathcal{O}_p} \sum_{i=l}^{n} \sum_{|\alpha|=k} |D^\alpha Z_i(p)| \tag{7.5}$$

The mean value theorem yields the following lemma.

Lemma 7.2. *Given any $Z \in V_{\mathcal{O}}^k$ and any real number $\epsilon > 0$, there exists a neighborhood $\mathcal{O}^\epsilon$ such that*

$$\|Z\|_{\mathcal{O}^\epsilon}^1 + \|Z\|_{\mathcal{O}^\epsilon}^2 + \cdots + \|Z\|_{\mathcal{O}^\epsilon}^{k-1} < \epsilon \|Z\|_{\mathcal{O}^\epsilon}^k \tag{7.6}$$

Proof. Applying the mean value theorem $k-1$ times to the first term (on the left) and one less time for each subsequent term in a δ-ball we see that

$$\|Z\|_{\mathcal{O}^\epsilon}^1 + \|Z\|_{\mathcal{O}^\epsilon}^2 + \cdots + \|Z\|_{\mathcal{O}^\epsilon}^{k-1} < (\delta^{k-1} + \cdots + \delta)\|Z\|^k$$

Given $\epsilon > 0$ choose δ Such that $\frac{\delta}{1-\delta} < \epsilon$ and the [illegible] satisfied. □

Define $\mathcal{L}_f : V_{\mathcal{O}}^k \to V_{\mathcal{O}}^k$ by

$$(\mathcal{L}_f Z)' \; [illegible] \; (p, Jac(F_{f(p)})^{-1} Z_{f(p)} \circ F_p) \tag{7.7}$$

We will [illegible] $\mathcal{L}_F$ is a contraction on $V_{\mathcal{O}}^k$.

Lemma 7.3. *Let Z be a function in $V_{\mathcal{O}}^k$ and F a C^∞ function as in theorem 7.1, then there exists a neighborhood $\mathcal{O}' \subset \mathcal{O}$ and a positive constant $\kappa < 1$ such that*

$$\|\mathcal{L}_f(Z)\|_{\mathcal{O}'}^k < \kappa \|Z\|_{\mathcal{O}'}^k \tag{7.8}$$

Proof. Let $A = (a_{ij}(p,x))$ denote the Jacobian of F_p and $|\alpha| = k$.

$$D^\alpha(Z_{f(p)}F_p)(X) = \sum D^\alpha Z_{f(p)}F_p(X) \cdot A^k + Q_\alpha \tag{7.9}$$

where Q_α is a polynomial in the derivatives of Z of order less than k and the derivatives of F_p of order less than or equal to k. Continuity allows us to choose $\mathcal{O}^\delta$ such that lemma 7.2 is satisfied and hence

$$\|Q_\alpha\|_{\mathcal{O}^\delta} < \epsilon \|Z\|_{\mathcal{O}^\delta}^k$$

Thus

$$\|\mathcal{L}_F Z\|_{\mathcal{O}^\delta}^k < l^{-1}\left((m+\delta)^k + \epsilon\right)\|Z\|_{\mathcal{O}^\delta}^k \tag{7.10}$$

By choosing δ and ϵ small enough $\kappa = l^{-1}\left((m+\delta)^k + \epsilon\right) < 1$ and the lemma is verified. □

We will complete the proof of theorem 7.1 by constructing the C^∞ function $\mathfrak{H}$ by successive approximations. Let $\mathfrak{H}^O = \mathfrak{h}$, the solution of (7.4) Then $\mathfrak{H}^O$ satisfies the conjugacy equation up to an infinitely flat error, i.e.,

$$\Lambda_{f(p)}^{-1} \circ \mathcal{H}_p^0 \circ F_p - \mathcal{H}_p^0 \in V_{\mathcal{O}}^\infty$$

Let

$$\mathcal{H}_p^n = \Lambda_{f^n(p)}^{-n} \circ \mathcal{H}_{f^n(p)}^0 \circ F_p^n \tag{7.11}$$

Since $\mathcal{L}_F$ contracts by a factor of κ, $\mathcal{L}_{F^j}$ contracts by a factor of κ^j. Thus the function

$$\mathcal{H}_p^n - \mathcal{H}_p^0 = \sum_{i=0}^{n-1} \mathcal{L}_{F^i}\left(\Lambda_{f(p)}^{-1} \circ \mathcal{H}_p^0 \circ F_p - \mathcal{H}_p^0\right) \tag{7.12}$$

is uniformly convergent, hence $\mathfrak{H}_p^n \longrightarrow \mathfrak{H}_p \in C^\infty$. By construction $\mathcal{H}(p, X) = (p, \mathcal{H}_p(X))$ satisfies

$$F \circ \mathcal{H} = \mathcal{H} \circ \Lambda. \quad \square$$

If the fibre maps F_p are contractions but there are resonances between the eigenvalues, a similar proof establishes the following theorem.

Theorem 7.4. *Let $F : B \to B$ be a skew product whose eigenvalues lie inside the unit circle. Then there exists smooth transformation, H, which satisfies*

$$F \circ H = H \circ G$$

where the normal form G_p is a polynomial containing only those terms which correspond to the resonances among the eigenvalues of F.

Remarks.

(1) *Slight modifications of the proof of theorem 7.1 (along the lines of [S2]) verify this theorem.*

(2) *Since the eigenvalues have magnitude less than one, resonances terms can only occur once for each relation in the normal form (in contrast to the hyperbolic area preserving case where $\lambda\lambda^{-1} = 1$ generates an infinite number of terms). Thus G_p is a polynomial.*

The theorems of Chapters 4, 5, and 6 for Anosov diffeomorphisms can be expressed for certain skew products as well.

Let M be a compact manifold and B a trivial $\mathbb{R}^2$-bundle over M. Let $F : B \to B$ have the form.

$$F(p, x) = (f(p), F_p(x))$$

With F_p continuous in p and smooth and area preserving in each fibre. As usual let $\lambda(p)$ and $\lambda^{-1}(p)$ be the eigenvalues at $(p, 0)$.

Theorem 7.5. *Let $F : B \to B$ be a skew product with the fibre map smooth area preserving fixing the zero section and continuous in the parameter for the base. Then there exists $H : B \to B$, $H(p, x) = (p, H_p(x))$, such that*

$$F \circ H = H \circ G$$

where $G(p, x) = (f(p), G_p(x))$ and

$$G_p(x, y) = \left(\lambda(p)x\Phi(xy), \lambda^{-1}(p)x\Phi^{-1}(xy)\right)$$

(1) *The word smooth can be replaced by formal, analytic, C^∞, to generalize the theorems of chapter 4, 5, 6 respectively.*

(2) *There are no essential differences between the proof of this theorem and those proofs given in the previous chapters.*

(3) *As before we have an exponential form of the normal form which leads us to a corollary about cocycles.*

Corollary 7.6. *The normal form can be expressed as*

$$G_p(x, y) = \left(\lambda(p)x\mathrm{e}^{\Upsilon(p,xy)}, \lambda^{-1}(p)y\mathrm{e}^{-\Upsilon(p,xy)}\right)$$

The coefficients of $(xy)^j$ in the exponent of the exponential are cocycles for the dynamical system (F, B).

Proof. The proof follows exactly as in chapter 4.

Let $\Gamma(s, m, p)$ denote the cocyles defined via the exponential normal form. Let

$$\mathcal{A}_1 = \int_M \Gamma(1, 1, p) d\nu$$

Corollary 7.7. *If the base map $f(p)$ has an invariant measure, ν, on M then $\mathcal{A}_1$ is an invariant of the skew product (F, B).*

Proof. Recall that $\Gamma(1, 1, p) = \Phi_{11}(p)$. From (4.11), $\Phi_{11}(p)$ is defined in terms of $\lambda(p)$ and is unique up to a coboundary $\left(h^1_{21}(f(p)) - h^1_{21}(p)\right)$. Integration over M eliminates the coboundary since f preserves measure and yields a uniquely defined invariant. □

References

[A] Anosov, D.V., *Geodesic Flows on Closed Riemannian Manifolds with Negative Curvature*, Tr. Mat. Inst. Steklov **90** (1967) (Russian); , Proc. Steklov Inst. Math.; **90**; (1976).

[Bi] Birkhoff, G.D., *Dynamical Systems*, vol. 9, Amer. Math Soc., 1927.

[Bi2] Birkhoff, G.D., *Surface Transformations and Their Dynamical Applications*, Acta Math **43 pages 1–119** (1922).

[Bi3] Birkhoff, G.D., *Collected Mathematical Papers*, vol. 2, American Mathematical Society, 1950.

[Br1] Brjuno, A.D., *Analytic Form of Differential Equations*, Tr. Mos. Math. Obsc. **25** (1971).

[Br2] Bruno, A.D., *Local Methods in Nonlinear Differential Equations*, Springer-Verlag, 1980.

[C1] Chen, K.T., *Equivalence and Decomposition of Vector Fields About an Elementary Critical Point*, Amer. J. Math **85** (1963), 693–722.

[C2] Chen, K.T., *Diffeomorphisms-C^∞ Realization of Formal Properties*, Amer. J. Math **87** (1965), 140–157.

[D] Dulac, H., *Solutions Dun Systeme Differentielles Dans Le Voisinage Des Valeurs Singulieres*, Bull. Soc. Math. France **40**, 324–383.

[H1] Hartman, P., *Ordinary Differential Equations*, Wiley, 1964.

[HK] Hurder, S. and Katok, A.B., *Differentiability, Rigidity and Godbillion Vey Classes for Anosov Flows*, Institute des Hautes Études Scientifiques **72** (1990).

[Ka] Katok, A. B., *Dynamical Systems with Hyperbolic Structure*, Three Papers on Dynamical Systems, vol. 116, American Mathematical Society, 1981 pages 125 - 211.

[L1] Lattes, S., *Sur Les Formes Reduites Des Transformations Ponstuelles a Deux Variables*, Comptes Rendus **152** (1911), 1566–1569.

[Ma] Mane, R., *Ergodic Theorey and Differentiable Dynamics*, Springer-Verlag.

[M1] Moser, J., *The Analytic Invariants of an Area Preserving Mapping Near a Hyperbolic Fixed Point*, Comm. Pure. Appl. Math. **9** (1956), 673–692.

[M2] Moser, J., *A Rapidly Converging Iteration Method and Nonlinear Differential Equations I, II, Ann Scuola*, Norm. Sup. Pisa 3 **20** (1966), 265–315, 499–535.

[M3] Moser, J., *Stable and Random Motions in Dynamical Systems, Annals of Mathematical Studies*, Princeton University Press, 1973.

[P1] Poincaré, H., *Surles Proprieties Des Fonctions Definies Par Les Equations Aux Differences Partielles*, Thesis, Paris, 1979; Reprinted in Oeuvres De Henri Poincare. Vol 1, Gauthier-Villars, Paris, 1928.

[Sh] Shub, M., *Global Stability of Dynamical Systems*, springer-Verlag, 1981.

[S1] Siegel, C.L., *Vereinfachter Beweis Eines Satzes von J. Moser*, Comm. Pure Appl. Math. **10**, 305–309.

[SM] Siegel, C.L. and Moser, J., *Lectures on Celestial Mechanics*, Springer-Verlag, 1956, 1971.

[S1] Sternberg, S., *On Local C^∞ Contractions of the Real Line*, Duke Math. J. **24** (1957), 97–102.

[S2] Sternberg, S., *Local Contractions and a Theorem of Poincaré*, Amer. J. Math **79** (1957), 97–102.

[S3] Sternberg, S., *The Structure of Local Homeomorphisms, II*, Amer. J. Math. **80** (1958), 623–632.

[S4] Sternberg, S., *The Structure of Local Homeoeomorphisms, III*, Amer. J. Math **81** (1958), 578-604.

[S5] Sternberg, S., *Infinite Lie Groups and The Formal Aspects of Dynamical Systems*, J. Math. Mech. **10** (1961), 451–474.

[Z] Zimmer, R.J., *Ergodic Theory and Semisimple Lie Groups*, Birkhauser, 1984.

[Zw] Zwillinger, D., *Handbook of Differential Equations*, Academic Press, 1986.