SECTION 6

THE HOPF BIFURCATION THEOREM FOR DIFFEOMORPHISMS

Let $X$ be a vector field and let $\gamma$ be a closed orbit of the flow $\phi_t$ of $X$. Let $P$ be a Poincaré map associated with $\gamma$. (See §2B). Suppose there is a circle $\sigma$ that is invariant under $P$. Then it is clear that $\bigcup_t \phi_t(\sigma)$ is an invariant torus for the flow of $X$ (see Figure 6.1).

![Figure 6.1](image)

If we have a one parameter family of vector fields and closed
orbits $X_\mu$ and $Y_\mu$, it is quite conceivable that for small $\mu$, $Y_\mu$ might be stable, but for large $\mu$ it might become unstable and a stable invariant torus take its place. Recall that $Y_\mu$ is stable (unstable) if the eigenvalues of the derivative of the Poincaré map $P_\mu$ have absolute value $< 1$ ($> 1$). (See §2B). The Hopf Bifurcation Theorem for diffeomorphisms gives conditions under which we may expect bifurcation to stable invariant tori after loss of stability of $Y_\mu$. The theorem we present is due to Ruelle-Takens [1]; we follow the exposition of Lanford [1] for the proof.

In order to apply these ideas, one needs to know how to compute the spectrum of the Poincaré map $P$. Fortunately this can be done because, as we have remarked earlier, the spectrum of the time $\tau$ map of the flow is that of $P \cup \{1\}$. (See §2B).

**Reduction to Two Dimensions**

We thus turn our attention to bifurcations for diffeomorphisms. The first thing to do is to reduce to the two dimensional case.* This is done by means of the center manifold theorem exactly as we did in the previous case; i.e., assume we have a one parameter family of diffeomorphisms $\phi_\mu : Z \to Z$, $\phi_\mu(0) = 0$ and assume a single complex conjugate pair of simple non-real eigenvalues crosses the unit circle as $\mu$ increases past zero. Then the center manifold theorem applied to $\psi : (x,\mu) \to (\phi_\mu(x),\mu)$ yields a locally invariant three manifold $M$; the $\mu$ slices $M_\mu$ then give a family of

---

*As remarked before, for partial differential equations, $P$ may become a diffeomorphism only after the reduction, and be only a smooth map before.
manifolds which we can identify by some fixed coordinate chart. Then on $M_\mu$ we have induced a family of diffeomorphisms containing all the recurrence.

Thus we are reduced to the following case: (modulo "global" stability problems as in the last section).

We have a one parameter family $\phi_\mu : \mathbb{R}^2 \to \mathbb{R}^2$ of diffeomorphisms satisfying:

a) $\phi_\mu(0,0) = (0,0)$

b) $d\phi_\mu(0,0)$ has two non-real eigenvalues $\lambda(\mu)$ and $\overline{\lambda(\mu)}$ such that for $\mu < 0$ $|\lambda(\mu)| < 1$ and for $\mu > 0$ $|\lambda(\mu)| > 1$

c) $\frac{d|\lambda(\mu)|}{d\mu}|_{\mu=0} > 0$.

We can reparametrize so that the eigenvalues of $d\phi_\mu(0,0)$ are $(1+\mu)e^{\pm i\theta(\mu)}$. By making a smooth $\mu$-dependent change of coordinates, we can arrange that:

$$d\phi_\mu(0,0) = (1+\mu) \begin{pmatrix} \cos \theta(\mu) & -\sin \theta(\mu) \\ \sin \theta(\mu) & \cos \theta(\mu) \end{pmatrix}.$$  

The Canonical Form

The next step is to make a further change of coordinates to bring $\phi_\mu$ approximately into appropriate canonical form. To be able to do this, we need a technical assumption:

$$e^{im\theta(0)} \neq 1 \quad m = 1,2,3,4,5. \quad (6.1)$$

(6.1) Lemma. Subject to Assumption (6.1), we can make a smooth $\mu$-dependent change of coordinates bringing $\phi_\mu$

*As D. Ruelle has pointed out, only $m = 1,2,3,4$ is needed for the bifurcation theorems as can be seen from the proof in §6A.
into the form:

\[ \phi_\mu(x) = N_\mu(x) + O(|x|^5) \]

where, in polar coordinates,

\[ N_\mu: (r, \phi) \mapsto ((1+\mu)r - f_1(\mu)r^3, \phi + \theta(\mu) + f_3(\mu)r^2). \]

The proof of this proposition uses standard techniques and may be obtained, for example, from §23 of Siegel and Moser [1]. We give a straightforward and completely elementary proof in Section 6A. As indicated above, we think of \( N_\mu \) as an approximate canonical form for \( \phi_\mu \). Note two special features of \( N_\mu \):

i) The new \( r \) depends only on the old \( r \), not on \( \phi \).

ii) The new \( \phi \) is obtained from the old \( \phi \) by an \( r \)-dependent rotation. We now add a final assumption:

\[ f_1(0) > 0. \]  \hspace{1cm} (6.2)

This assumption implies that for small positive \( \mu \), \( N_\mu \) has an invariant circle of radius \( r_0 \), where \( r_0 \) is obtained by solving

\[ (1+\mu)r_0 - f_1(\mu)r_0^3 = r_0, \]

i.e., \( r_0^2 = \frac{\mu}{f_1(\mu)}. \)

*The canonical form is important in celestial mechanics for proving the existence and stability of closed orbits near a given one; i.e., for finding fixed points or invariant circles for the Poincaré map. In the Hamiltonian case this map is symplectic (see Abraham-Marsden [1]) so Birkhoff's theorem applies, as are the results of Kolmogorov, Arnold and Moser if it is a "twist mapping".*
We shall verify shortly that this circle is attracting for $N_{\mu}$. Since $\phi_{\mu}$ differs only a little from $N_{\mu}$, it is not surprising that $\phi_{\mu}$ has a nearby invariant circle.

**The Main Result**

(6.2) **Theorem.** (Ruelle-Takens, Sacker, Naimark). Assume (6.1) and (6.2). Then for all sufficiently small positive $\mu$, $\phi_{\mu}$ has an attracting invariant circle.

Before giving the proof, let us look at what happens if (6.2) is replaced by the assumption $f_1(0) < 0$. Then, for a small positive $\mu$, $N_{\mu}$ has no invariant sets except $\{0\}$ and $\mathbb{R}^2$. For $\mu < 0$, $N_{\mu}$ does have an invariant circle, but it is repelling rather than attracting. By applying the result of Ruelle and Takens to $\phi_{\mu}^{-1}$, we prove that, in this case, $\phi_{\mu}$ has a nearby invariant circle. Thus, we again find the usual situation that supercritical branches are stable and subcritical branches unstable. If

$$\phi_{\mu}: (y, \phi) \rightarrow ((1-2\mu)y - \mu(3y^2 + y^3) + \mu^20(1),$$

$$\phi + \theta(\mu) + \mu \frac{f_3(\mu)}{f_1(\mu)} (1+y)^2 + \mu^20(1)).$$

Finally, we scale $y$ again by putting

$$y = \sqrt{\mu} z;$$

then

* It would be interesting to explicitly compute $f_1(0)$ in terms of $\phi_{\mu}$ directly as we did in §4 for the bifurcation to invariant circles. However the labor involved in the earlier calculations, and the promise of a harder, if not impossible computation has left the present authors sufficiently exhausted to leave this one to the ambitious reader. The calculation of $f_1(0)$ in terms of $\phi_{\mu}$ rather than $X_{\mu}$ is not so hard and has been done by Wan [preprint] and Iooss [6].
\[ \Phi_\mu: (z, \phi) \mapsto ((1-2\mu)z - \mu^{3/2}(3z^2 + \mu^{1/2}z^3) + \mu^{3/2}0(1), \]
\[ \phi + \theta(\mu) + \mu \frac{f_3(\mu)}{f_1(\mu)} (l + \mu^{1/2}z)^2 + \mu^20(1); \]

we rewrite this last formula as

\[ (z, \phi) \mapsto ((1-2\mu)z + \mu^{3/2}H_\mu(z, \phi), \phi + \theta(\mu) + \mu^{3/2}K_\mu(z, \phi)). \]

The functions \( H_\mu(z, \phi), K_\mu(z, \phi) \) are smooth in \( z, \phi, \mu \) on

\[ -1 \leq z \leq 1, \quad 0 \leq \phi \leq 2\pi, \quad 0 \leq \mu \leq \mu_0 \]

for some sufficiently small \( \mu_0 \); the region \(-1 \leq z \leq 1,\)
\( 0 \leq \phi \leq 2\pi \) corresponds to an annulus of width \( O(\mu) \) about the invariant circle for \( N_\mu \) (which has radius \( O(\mu) \)). We are going to produce an invariant circle inside this annulus.

The qualitative behavior of \( \Phi_\mu \) is now easy to read off: \( \Phi_\mu \) can be written as

\[ (z, \phi) \mapsto ((1+2\mu)z, \phi + \theta_1(\mu)) \]

plus a small perturbation. The approximate \( \phi \) is simply a rotation in the \( z \) direction and a contraction in the \( z \) direction. Note, however, that the strength of the contraction goes to zero with \( \mu \). If this were not the case, we could simply invoke known results about the persistence of attracting invariant circles under small perturbations. As it is, we need to make a slightly more detailed argument, exploiting the fact that the size of the perturbation goes to zero faster than the strength of the contraction.

We are going to look for an invariant manifold of the form

\[ \{ z = u(\phi) \}, \]
where

i) \( u(\phi) \) is periodic in \( \phi \) with period \( 2\pi \)

ii) \( |u(\phi)| \leq 1 \) for all \( \phi \)

iii) \( u(\phi) \) is Lipschitz continuous with Lipschitz constant \( 1 \) (i.e., \( |u(\phi_1) - u(\phi_2)| \leq |\phi_1 - \phi_2| \)).

The space of all functions \( u \) satisfying i), ii) and iii) will be denoted by \( \mathcal{U} \).

We shall give a proof based on the contraction mapping principle. In outline, the argument goes as follows: We start with a manifold

\[ M = \{ z = u(\phi) \}, \]

with \( u \in \mathcal{U} \), and consider the new manifold \( \phi_\mu M \) obtained by acting on \( M \) with \( \phi_\mu \). We show that, for \( \mu \) sufficiently small, \( \phi_\mu M \) again has the form \( \{ z = \hat{u}(\phi) \} \) for some \( \hat{u} \in \mathcal{U} \).

Thus, we construct a non-linear mapping \( \mathcal{F} \) of \( \mathcal{U} \) into itself by

\[ \mathcal{F}u = \hat{u}. \]

We then prove, again for small positive \( \mu \), that \( \mathcal{F} \) is a contraction on \( \mathcal{U} \) (with respect to the supremum norm) and hence as a unique fixed point \( u^* \). The manifold \( \{ z = u^*(\phi) \} \) is the desired invariant circle. As a by-product of the proof of contractivity, we prove this manifold is attracting in the following sense: Pick a starting point \((z,\phi)\) with \( |z| \leq 1 \), and let \((z_n,\phi_n)\) denote \( \phi_\mu^n(z,\phi) \). Then

\[ \lim_{n \to \infty} z_n - u^*(\phi_n) = 0. \]

It is not hard to see that the domain of attraction is much
larger than the annulus \(|z| < 1\). In particular, it contains everything inside the annulus except the fixed point at the center, but we shall not pursue this point.

To carry out the argument outlined above, we must first construct the non-linear mapping \( F \). To find \( F(u(\phi)) \), we should proceed as follows:

A) Show that there is a unique \( \phi \) such that the \( \phi \)-component of \( \phi \mu(u(\phi), \phi) \) is \( \phi \), i.e., such that

\[
\phi \equiv \tilde{\phi} + \theta_1(\mu) + \mu^{3/2} K_\mu(u(\tilde{\phi}), \tilde{\phi})(2\pi). \tag{6.3}
\]

and

B) Put \( F(u(\phi)) \) equal to the \( z \)-component of \( \phi \mu(u(\tilde{\phi}), \tilde{\phi}) \), i.e.,

\[
F(u(\phi)) = (1-2\mu)u(\tilde{\phi}) + \mu^{3/2} H_\mu(u(\tilde{\phi}), \tilde{\phi}). \tag{6.4}
\]

In the estimates we are going to make, it will be convenient to introduce

\[
\lambda = \sup_{0 < \phi < 2\pi} \left\{ |H_\mu| + |K_\mu| \right\} \sup_{-1 < z < 1} \left\{ \left| \frac{\partial H_\mu}{\partial z} \right| + \left| \frac{\partial K_\mu}{\partial z} \right| + \left| \frac{\partial H_\mu}{\partial \phi} \right| + \left| \frac{\partial K_\mu}{\partial \phi} \right| \right\};
\]

so defined, \( \lambda \) depends on \( \mu \) but remains bounded as \( \mu \to 0 \).

We now prove that (6.3) has a unique solution. To do this, it is convenient to denote the right-hand side of (6.3) temporarily by \( x(\phi) \):

\[
x(\phi) = \tilde{\phi} + \theta_1(\mu) + \mu^{3/2} K_\mu(u(\tilde{\phi}), \tilde{\phi}).
\]

We want to show that, as \( \phi \) runs from \( 0 \) to \( 2\pi \), \( x(\phi) \) runs *i.e., \( \phi \) differs from \( \tilde{\phi} + \theta_1(\mu) + \mu^{3/2} K_\mu(u(\tilde{\phi}), \tilde{\phi}) \) by an integral multiple of \( 2\pi \).
exactly once over an interval of length $2\pi$. From the periodicity of $u(\tilde{\phi}), K_\mu(\tilde{z}, \tilde{\phi})$ in $\tilde{\phi}$, it follows that

$$x(2\pi) = x(0) + 2\pi.$$  

We therefore only have to show that $x$ is strictly increasing. Let $\tilde{\phi}_1 < \tilde{\phi}_2$. Then

$$x(\tilde{\phi}_2) - x(\tilde{\phi}_1) = \tilde{\phi}_2 - \tilde{\phi}_1 + \mu^{3/2} [K_\mu(u(\tilde{\phi}_2), \tilde{\phi}_2) - K_\mu(u(\tilde{\phi}_1), \tilde{\phi}_1)].$$

Now

$$|K_\mu(u(\tilde{z}, \tilde{\phi}) - K_\mu(u(\tilde{z}, \tilde{\phi}_1))| \leq \lambda \left[ |u(\tilde{\phi}_2) - u(\tilde{\phi}_1)| + |\tilde{\phi}_2 - \tilde{\phi}_1| \right]$$

$$\leq 2\lambda |\tilde{\phi}_2 - \tilde{\phi}_1| = 2\lambda (\tilde{\phi}_2 - \tilde{\phi}_1).$$

(The second inequality follows from the Lipschitz continuity of $u$.) Thus

$$x(\tilde{\phi}_2) - x(\tilde{\phi}_1) \geq (1 - 2\lambda \mu^{3/2})(\tilde{\phi}_2 - \tilde{\phi}_1), \text{ so, provided}$$

$$1 - 2\lambda \mu^{3/2} > 0,$$

then $x$ is strictly increasing and (6.3) has a unique solution. We thus get $\tilde{\phi}$ as a function of $\phi$, and it follows from our above estimates that $\tilde{\phi}$ is Lipschitz continuous:

$$|\tilde{\phi}(\phi_1) - \tilde{\phi}(\phi_2)| \leq (1 - 2\lambda \mu^{3/2})^{-1} |\phi_1 - \phi_2|.$$  

The definition (6.4) of $\mathcal{F}_u$ therefore makes sense, and we next have to check that $\mathcal{F}_u \in \mathcal{U}$. Condition (i), corresponds to 6.7 is immediate. For (ii), note that

$$|\mathcal{F}_u(\phi)| \leq (1 - 2\mu) |u(\tilde{\phi})| + \mu^{3/2} |H_\mu(u(\tilde{\phi}), \tilde{\phi})|$$

$$\leq 1 - 2\mu + \mu^{3/2}.$$
Thus, $|\mathcal{F}u(\phi)| \leq 1$ for all $\phi$ provided

$$2\mu - \mu^{3/2}\lambda \geq 0. \tag{6.7}$$

Finally,

$$|\mathcal{F}u(\phi_1) - \mathcal{F}u(\phi_2)| \leq (1-2\mu)|u(\tilde{\phi}_1) - u(\tilde{\phi}_2)|$$

$$+ \mu^{3/2}\lambda \left[ |u(\tilde{\phi}_1) - u(\tilde{\phi}_2)| + |\tilde{\phi}_1 - \tilde{\phi}_2| \right]$$

$$\leq (1-2\mu + 2\mu^{3/2}\lambda)|\tilde{\phi}_1 - \tilde{\phi}_2|$$

by the Lipschitz continuity of $u$. Inserting estimate (6.6) for $|\tilde{\phi}_1 - \tilde{\phi}_2|$, we get

$$|\mathcal{F}u(\phi_1) - \mathcal{F}u(\phi_2)| \leq (1-2\mu + 2\mu^{3/2}\lambda)(1-2\mu^{3/2}\lambda)^{-1}|\tilde{\phi}_1 - \tilde{\phi}_2|,$$

so $\mathcal{F}u$ is Lipschitz continuous with Lipschitz constant $1$ provided

$$(1-2\mu + 2\mu^{3/2}\lambda)(1-2\mu^{3/2}\lambda)^{-1} \leq 1. \tag{6.8}$$

Evidently (6.8) holds for all sufficiently small positive $\mu$, so (iii) holds.

The next step is to prove that $\mathcal{F}$ is a contraction. Thus, let $u_1, u_2 \in U$, choose $\phi$, and let $\tilde{\phi}_1, \tilde{\phi}_2$ denote the solutions of

$$\phi = \tilde{\phi}_1 + \theta_1(\mu) + \mu^{3/2}\kappa_\mu(u_1(\tilde{\phi}_1), \tilde{\phi}_1)$$

$$\phi = \tilde{\phi}_2 + \theta_1(\mu) + \mu^{3/2}\kappa_\mu(u_2(\tilde{\phi}_2), \tilde{\phi}_2),$$

respectively. Subtracting these equations, transposing, and taking absolute values yields
\[
\begin{aligned}
|\tilde{\phi}_1 - \tilde{\phi}_2| & \leq \mu^{3/2} |K_\mu (u_1(\tilde{\phi}_1), \tilde{\phi}_1) - K_\mu (u_2(\tilde{\phi}_2), \tilde{\phi}_2)| \\
& \leq \mu^{3/2} \lambda [ |u_1(\tilde{\phi}_1) - u_2(\tilde{\phi}_2)| + |\tilde{\phi}_1 - \tilde{\phi}_2| ].
\end{aligned}
\] (6.9)

Now
\[
|u_1(\tilde{\phi}_1) - u_2(\tilde{\phi}_2)| \leq |u_1(\tilde{\phi}_1) - u_2(\tilde{\phi}_2)| + |u_2(\tilde{\phi}_1) - u_2(\tilde{\phi}_2)|
\]
\[
\leq |u_1 - u_2| + |\tilde{\phi}_1 - \tilde{\phi}_2|.
\]

Inserting this inequality into (6.9), collecting all the terms involving \( |\tilde{\phi}_1 - \tilde{\phi}_2| \) on the left, and dividing yields
\[
|\tilde{\phi}_1 - \tilde{\phi}_2| \leq (1 - 2\mu^{3/2} \lambda)^{-1} \mu^{3/2} \lambda \cdot |u_1 - u_2|. \tag{6.10}
\]

Now we use the definition (6.4) of \( \mathcal{F} u \):
\[
|\mathcal{F} u_1(\phi) - \mathcal{F} u_2(\phi)| \leq (1 - 2\mu) |\bar{u}_1(\tilde{\phi}_1) - \bar{u}_2(\tilde{\phi}_2)|
\]
\[
+ \mu^{3/2} |H_\mu (u_1(\tilde{\phi}_1), \tilde{\phi}_1) - H_\mu (u_2(\tilde{\phi}_2), \tilde{\phi}_2)|
\]
\[
\leq (1 - 2\mu) |u_1 - u_2| + |\tilde{\phi}_1 - \tilde{\phi}_2|
\]
\[
+ \mu^{3/2} \lambda [ |u_1 - u_2| + 2 |\tilde{\phi}_1 - \tilde{\phi}_2|]
\]
\[
\leq |u_1 - u_2| \left( 1 - 2\mu \right) (1 + \mu^{3/2} \lambda (1 - 2\mu^{3/2} \lambda)^{-1})
\]
\[
+ \mu^{3/2} \lambda [ 1 + 2 \mu^{3/2} \lambda (1 - 2\mu^{3/2} \lambda)^{-1} ].
\]

Let \( \alpha \) denote the expression in braces. Then
\[
\alpha = 1 - 2\mu + O(\mu^{3/2}),
\]
so we can make \( \alpha < 1 \) by making \( \mu \) small enough. If this is done, we have
\[
||\mathcal{F} u_1 - \mathcal{F} u_2|| \leq \alpha \cdot |u_1 - u_2| \text{ with } \alpha < 1, \tag{6.11}
\]
i.e., \( \mathcal{F} \) is a contraction on \( U \) and hence has a unique fixed
point $u^*$.

To prove that the invariant manifold $\{z = u^*(\phi)\}$ is attracting, we pick a point $(z, \phi)$ in the annulus $|z| \leq 1$, and we let $(z_1, \phi_1)$ denote $\phi_\mu(z, \phi)$. Note that

$$|z_1| \leq (1 - 2\mu)|z| + \mu^{3/2} \lambda \leq 1 - 2\mu + \mu^{3/2} \lambda \leq 1$$

(by (6.7)), so $(z_1, \phi_1)$ is again in the annulus. Now let $\tilde{\phi}_1$ denote the solution of

$$\tilde{\phi}_1 = \phi_1 + \theta_1(\mu) + \mu^{3/2} K_\mu(u^*(\tilde{\phi}_1), \tilde{\phi}_1).$$

The definition of $\phi_1$, on the other hand, needs

$$\phi_1 = \phi + \theta_1(\mu) + \mu^{3/2} K_\mu(z, \phi).$$

Subtracting these equations and then estimating and re-arranging as in the proof of (6.8), we get

$$|\tilde{\phi}_1 - \phi| \leq \mu^{3/2} \lambda (1 - 2\mu^{3/2} \lambda)^{-1} |z - u^*(\phi)|.$$

Now subtract the equations

$$u^*(\phi_1) = F u^*(\phi_1) = (1 - 2\mu) u^*(\tilde{\phi}_1) + \mu^{3/2} H_\mu(u^*(\tilde{\phi}_1), \tilde{\phi}_1)$$

$$z_1 = (1 - 2\mu) z + \mu^{3/2} H_\mu(z, \phi)$$

and again imitate the proof that $F$ is a contraction to get

$$|z_1 - u^*(\phi_1)| \leq \alpha \cdot |z - u(\phi)|,$$

with the same $\alpha$ as in (6.11). By induction,

$$|z_n - u^*(\phi_n)| \leq \alpha^n |z - u(\phi)| \to 0 \text{ as } n \to \infty.$$

In our proof, we used only the continuity of $H_\mu, K_\mu$.
and their first derivatives, and we obtained a Lipschitz continuous \( u^* \). Closer examination of the argument shows that we needed only Lipschitz continuity of \( H_\mu, K_\mu \). If we have more differentiability of \( H_\mu, K_\mu \), we would expect to obtain more differentiability for \( u^* \). This is indeed the case. Specifically, let \( U_k \) denote the set of periodic functions \( u(\phi) \) of class \( C^k \) satisfying

1) \[ |u^{(j)}(\phi)| < 1, j = 0,1,...,k; \] all \( \phi \).

2) \[ |u^{(k)}(\phi)| \] is Lipschitz continuous with Lipschitz constant one.

If \( H_\mu, K_\mu \) have Lipschitz continuous \( k \)th derivatives, a straightforward generalization of the estimates we have given shows that for \( \mu \) sufficiently small, \( \mathcal{F} \) maps \( U_k \) into itself. It may be shown that \( U_k \) is complete in the supremum norm (as in the proof of the center manifold theorem), so the fixed point of \( \mathcal{F} \) must be in \( U_k \), i.e., \( u^* \) has Lipschitz continuous \( k \)th derivative. If we make the weaker assumption that \( H_\mu, H_\mu \) have continuous \( k \)th derivatives, slightly more complicated arguments show that \( u^* \) also has a continuous \( k \)th derivative; we proceed with the proof by showing that the set of \( u \)'s, whose \( k \)th derivatives have an appropriately chosen modulus of continuity, is mapped into itself by \( \mathcal{F} \).
SECTION 6A
THE CANONICAL FORM

We shall give here, following Lanford [1], an elementary and straightforward derivation of the canonical form for the mapping \( \Phi_\mu : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \). Recall that we have already arranged things so that

\[
\Phi_\mu \begin{pmatrix} x \\ y \end{pmatrix} = (1+\mu) \begin{pmatrix} \cos \theta(\mu) & -\sin \theta(\mu) \\ \sin \theta(\mu) & \cos \theta(\mu) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + O(r^2).
\]

We want to organize the second, third, and fourth degree terms by making further coordinate changes. It will be convenient to identify \( \mathbb{R}^2 \) with the complex plane by writing

\[
z = x + iy.
\]

Then

\[
\Phi_\mu (z) = \lambda(\mu)z + O(|z|^2), \quad \lambda(\mu) = (1+\mu)e^{i\theta(\mu)}.
\]

From now on we shall leave \( \mu \) out of our notation as much as possible.

The higher-order terms in the Taylor series for \( \Phi \) may
be written as polynomials in $z$ and $\bar{z}$, i.e.,

$$\phi(z) = \lambda z + A_2(z) + A_3(z) + \ldots,$$

where, for example,

$$A_2(z) = \sum_{j=0}^{2} a_j z^{2-j}.$$

Let us begin in a pedestrian way with $A_2$. We choose a new coordinate $z' = z + \gamma(z)$, where $\gamma$ is homogeneous of degree 2, i.e., has the same form as $A_2$. We can invert the relation between $z$ and $z'$ as

$$z = z' - \gamma(z') + \text{higher order terms}.$$

Since for the moment we are only concerned with terms of degree 2 or lower, we calculate modulo terms of degree 3 and higher and replace equality signs by congruence signs ($\equiv$). Thus we have

$$z \equiv z' - \gamma(z') = (1 - \gamma)(z').$$

In terms of the new coordinate we have

$$\phi'(z') \equiv (1 + \gamma) \phi(z' - \gamma(z')).$$

$$\equiv (1 + \gamma)[\lambda z' - \lambda \gamma(z') + A_2(z' - \gamma(z'))]$$

$$\equiv (1 + \gamma)[\lambda z' - \lambda \gamma(z') + A_2(z')]$$

$$\equiv \lambda z' - \lambda \gamma(z') + A_2(z') + \gamma(\lambda z' - \lambda \gamma(z') + A_2(z'))$$

$$\equiv \lambda z' + A_2(z') + \gamma(\lambda z') - \lambda \gamma(z').$$

Now

$$\gamma(z') = \gamma_2 z'^2 + \gamma_1 z' \bar{z'} + \gamma_0 \bar{z'}^2$$

$$\gamma(\lambda z') - \lambda \gamma(z') = \gamma_2 (\lambda^2 - \lambda) z'^2 + \gamma_1 (|\lambda|^2 - \lambda) z' \bar{z'} + \gamma_0 (\lambda^2 - \lambda) \bar{z'}^2.$$
\[ A_2(z') = a_2 z'^2 + a_1 z' z + a_0 z'^2, \]
so, if we put
\[
\gamma_2 = \frac{-a_2}{\lambda^2 - \lambda}, \quad \gamma_1 = \frac{-a_1}{|\lambda|^{2} - \lambda}, \quad \gamma_0 = \frac{-a_0}{\lambda^2 - \lambda},
\]
we get
\[
\phi'(z') = \lambda z' + o(|z'|^3).
\]

We must, of course, make sure that the denominators in our expressions for the \( \gamma_i \) do not vanish. Since \( |\lambda| = 1+\mu \), there is no problem for \( \mu \neq 0 \), but we want our \( \mu \)-dependent coordinate change to be well-behaved as \( \mu \to 0 \). This will be the case provided
\[
ed^{2i\theta(0)} \neq e^{i\theta(0)}, \quad 1 \neq e^{i\theta(0)}, \quad e^{-2i\theta(0)} \neq e^{i\theta(0)},
\]
\[ i.e., \text{provided} \]
\[
ed^{i\theta(0)} \neq 1, \quad e^{3i\theta(0)} \neq 1.
\]
Thus, if these conditions hold, we can make a smooth \( \mu \)-dependent coordinate change, bringing \( A_2 \) to zero. We assume that we have made this change and drop the primes:
\[
\phi(z) = \lambda z + A_3(z) + \cdots.
\]
(\( A_3 \) is not the original \( A_3 \)) and see what we can do about \( A_3 \).

This time, we take a new coordinate \( z' = z + \gamma(z) \), \( \gamma \) homogeneous of degree 3, and we calculate modulo terms of degree 4 and higher. Just as before, we have
\[
\phi'(z') \equiv (I + \gamma)\phi(z' - \gamma(z'))
\]
\[ \equiv \lambda z' + A_3(z') + \gamma(\lambda z') - \lambda \gamma(z'). \]

Again, we write out
\[ \gamma(z') = \gamma_3 z^3 + \gamma_2 z^2 z' + \gamma_1 z z' + \gamma_0 \]

\[ \gamma(\lambda z') - \lambda \gamma(z') = \gamma_3 (\lambda^3 - \lambda) z^3 + \gamma_2 (|\lambda|^{-1} - 1) \lambda z^2 z' + \gamma_1 (|\lambda|^{-2} - \lambda) z' z'' + \gamma_0 (\lambda^3 - \lambda) z'' z'' + \gamma_0 z''^2. \]

\[ A_3(z') = a_3 z^3 + a_2 z^2 z' + a_1 z z' + a_0 z'. \]

By an appropriate choice of \( \gamma_3, \gamma_1, \gamma_0 \), we can cancel the \( a_3, a_2, a_1 \) terms provided

\[ e^{2i\theta(0)} \neq 1, \quad e^{4i\theta(0)} \neq 1. \]

The \( a_2 \) term presents a new problem. For \( \mu \neq 0 \), we can, of course, cancel it by putting

\[ \gamma_3 = \frac{-a_2}{\lambda (|\lambda|^{-1} - 1)}. \]

This expression, however, diverges as \( \mu \to 0 \), independent of the value of \( \theta(0) \). For this reason, we shall not try to adjust this term and simply put \( \gamma_3 = 0 \). Then, in the new coordinates (dropping the primes)

\[ \phi(z) = (\lambda + a_2^2 |z|^2) + O(|z|^4). \]

We next set out to cancel the 4th degree terms by a coordinate change \( z' = z + \gamma(z), \gamma \) homogeneous of degree 4. A straightforward calculation of a by now familiar sort shows that such a coordinate change does not affect the terms of degree \( \leq 3 \) and that all the terms of degree 4 can be cancelled provided

\[ e^{5i\theta(0)} \neq 1. \]

Thus we get
\[ \psi(z) = \left( \lambda + a_2^3 |z|^2 \right)z + O(|z|^3). \]

This is still not quite the desired form. To complete the argument, we write

\[ \lambda + a_2^3 |z|^2 = (1+\mu) e^{i \theta(\mu)} \left[ 1 - \frac{f_1(\mu)}{1+\mu} |z|^2 + i f_3(\mu) |z|^2 \right] \]

(where \( f_1, f_3 \) are real)

\[ = \left( 1+\mu - f_1(\mu) |z|^2 \right) e^{i [\theta(\mu) + f_3(\mu) |z|^2]} + O(|z|^4). \]

Thus

\[ \phi(z) = \left( 1+\mu - f_1(\mu) |z|^2 \right) e^{i [\theta(\mu) + f_3(\mu) |z|^2]} z + O(|z|^5); \]

when we translate back into polar coordinates, we get exactly the desired canonical form.