IMM-NYU 333
October 1964

NEW YORK UNIVERSITY

COURANT INSTITUTE OF MATHEMATICAL SCIENCES

ON INVARIANT SURFACES AND BIFURCATION OF PERIODIC SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS

Robert John Sacker

This work was sponsored by the Office of Naval Research, Contract Nonr-285(46), the National Science Foundation and the Sloan Foundation.

CHAPTER II
BIFURCATION - MAPPING METHOD

1. Introduction

In this chapter we treat the bifurcation problem by considering the mapping induced by the vector field near the periodic solution: the method of surfaces of section mentioned earlier.

We consider the differential equation

$$\frac{dx}{dt} = F(x, \mu) \tag{2.1}$$

where x and F are real n-vectors and μ is a real parameter. Suppose that for all sufficiently small μ, $|\mu| < \mu_*$, (2.1) has a periodic solution $x = \psi(t, \mu)$ with period 2π. Assume that F has $\rho \geq 5$ continuous derivatives with respect to x and μ in some neighborhood of ψ and that in this neighborhood

$$|F|_\rho \leq M_0$$

where differentiation is with respect to x and μ. (See Appendix for definition of $|F|_\rho$.)

Suppose that ψ is asymptotically stable for $\mu < 0$ in the sense that n-1 of the Floquet multipliers lie inside the unit circle in the complex plane. As μ increases through zero we assume that a conjugate pair of the multipliers

$$\lambda(\mu) = e^{2\pi\alpha(\mu)},\ \overline{\lambda(\mu)} = e^{2\pi\overline{\alpha(\mu)}}$$

leaves the unit circle with

$$Re\ \alpha(0) = 0,\ Re\ \alpha^{'}(0) > 0$$

while the remaining n-3 stay inside. Hence the periodic solution ψ becomes unstable.

We do not require that ψ is obtained by a bifurcation from an equilibrium as discussed in Chapter I but just consider any periodic solution which loses its stability in the above manner as μ passes through a critical value (assumed to be $\mu = 0$).

2. Reduction to a Mapping Problem

We now obtain the mapping described by the flow near the periodic solution ψ. Let C be the curve described by ψ. Choose a point P on C and let S be the n-1 dimensional hyperplane normal to C at P. Let P be the origin of coordinates $x = (x_{(1)}, \ldots, x_{(n-1)})$ describing S. It is known [1; p.50] that any solution of (2.1) starting on S and close enough to P will return to S for the first time after a time lapse of $2\pi + \epsilon(x)$ where $\epsilon \to 0$ as $x \to 0$. In a small neighborhood of P this defines a mapping T of S into itself having P as a fixed point

$$T : x_1 = Dx + \ldots \tag{2.2}$$

It is known [1; p.163] that the eigenvalues of D are just the $n-1$ previously mentioned Floquet multipliers. Thus for $\mu < 0$, P is an asymptotically stable fixed point which becomes unstable as μ increases through zero.

The existence of a torus will be established by showing that P bifurcates into a closed curve $\bar{C}$, lying in S which is invariant under T, i.e., if q is a point of $\bar{C}$, then $Tq = q_1$ is on $\bar{C}$.

3. Normal Form for the Mapping

By appropriately transforming coordinates we will put the mapping in a certain normal form which is needed in the existence theorem. We assume that a real linear transformation of coordinates has been carried out so that D is in the real form

$$D = \begin{bmatrix} Re\ \lambda(\mu) & -Im\ \lambda(\mu) & 0 \\ Im\ \lambda(\mu) & Re\ \lambda(\mu) & \\ 0 & & S(\mu) \end{bmatrix}$$

where $S(\mu)$ is a matrix whose eigenvalues $\sigma(\mu)$ satisfy

$$|\sigma(0)| < 1 \tag{2.3}$$

and

$$\sup_{|Y|=1} |S(0)Y| = \sigma_0 < 1$$

where $Y = Col(x_{(3)}, \ldots, x_{(n-1)})$ and $|Y|$ is the ordinary Euclidean length of a vector $(Y,Y)^{\frac{1}{2}}$. Let $z = x_{(1)} + ix_{(2)}$. Then in the z, Y coordinates the mapping is

$$T: \begin{array}{rcl} z_1 & = & \lambda(\mu)z + U(z, \bar{z}, Y, \mu) \\ Y_1 & = & S(\mu)Y + V(z, \bar{z}, Y, \mu) \end{array} \tag{2.4}$$

where U is complex and V is real. The functions U, V, λ and S have five continuous derivatives in a neighborhood of $(z, Y, \mu) = 0$ which we assume to be

$$|z|^2 + |Y|^2 \leq 2, \qquad |\mu| < \mu_*.$$

We assume U and V to be expanded into Taylor series up to polynomials of degree 4 plus remainder.

In order to motivate the following theorem consider the following example of a mapping

$$\begin{array}{rcl} z_1 & = & e^{i+\mu+\beta|z|^2} z \\ Y_1 & = & SY + b|z|^2 \end{array}$$

where β is complex, b is a real constant vector and S is a matrix which satisfies (2.3). If $Re\ \beta < 0$, this mapping has the invariant curve

$$|z| = \left(\frac{\mu}{-Re\ \beta}\right)^{\frac{1}{2}}, \quad Y = (S-I)^{-1} b \frac{\mu}{Re\ \beta}$$

contained in a neighborhood $|z| \leq c\sqrt{\mu}$, $|Y| \leq c^2\mu$, c a constant. We now consider the effect of perturbing this example by adding more nonlinear terms. In such an oblate neighborhood Y is small of order μ whereas z is small only to order $\mu^{\frac{1}{2}}$. More precisely consider the neighborhood

$$N: z = a\zeta, \quad Y = a^2\tilde{Y}$$

where $|\zeta|^2 + |\tilde{Y}|^2 \leq 2$ and a is a small parameter, $0 \leq a \leq 1$. In N the monomial $M_\tau = Y_{(i)}^p z^q \bar{z}^r$, $\tau = 2p + q + r$, satisfies $|M_\tau|_k \leq c(k)a^\tau$ where differentiation is with respect to $\tilde{Y}_{(i)}$, ζ, and $\bar{\zeta}$ and $c(k)$ is a constant depending only on k.

Definition. We call τ the weight of M_τ. In the following theorem we will transform the mapping (2.4) into a form similar to the above example in which monomials of weight 2 and 3 will serve to determine the invariant curve in the first approximation while the remaining terms will act as small perturbations. The transformations used are of the type treated by B. Segre [2]. See also A. Kelly [3] for an excellent bibliography.

Theorem 1. If $\lambda^4(0) \neq 1$, $\lambda^3(0) \neq 1$ and (2.3) is satisfied, then there exists a $\mu_0 \leq \mu_*$ such that for $|\mu| < \mu_0$ there exists a transformation

$$\begin{aligned} z &= cw + P(w, \bar{w}, W, c, \mu) \\ Y &= cW + \tilde{P}(w, \bar{w}, W, c, \mu) \end{aligned} \qquad 0 < c \leq 1$$

with P and $\tilde{P}$ polynomials in $w, \bar{w}$, and W, which carries (2.4) into the form

$$\begin{aligned} w_1 &= e^{2\pi\alpha(\mu) + c^2\beta(\mu)|w|^2} w + R_4(w, \bar{w}, W, c, \mu) \\ W_1 &= S(\mu)W + R_3(w, \bar{w}, W, c, \mu) \end{aligned} \tag{2.5}$$

where R_k are function whose Taylor expansions[1] about $(w, \bar{w}, W) = 0$ contain no terms of weight $< k$. P and $\tilde{P}$ contain no constant or linear terms. For proper choice of c the transformation is one-to-one in the neighborhood $|w|^2 + |W|^2 \leq 2$.

Proof: In this proof subscripts on functions will carry the same significance as for R_k defined above. We write the first equation of (2.4) as

$$z_1 = \lambda(\mu)z + u(z, \bar{z}, \mu) + v(z, \bar{z}, Y, \mu) + F_4 \tag{2.6}$$

where u contains quadratic and third order terms in z and $\bar{z}$ and $v = Y'[p(\mu)z + q(\mu)\bar{z}]$ with p, q vectors and prime denoting transpose. We transform (2.6) by

$$z = w + r(\mu)w^k\bar{w}^\ell, \qquad k + \ell = 2, 3 \tag{2.7}$$

to obtain

$$w_1 = \lambda(\mu)w + u(w, \bar{w}, \mu) - g(\mu)w^k\bar{w}^\ell + v(w, \bar{w}, Y, \mu) + \tilde{F}_4$$

where

$$g(\mu) = (\lambda^k\bar{\lambda}^\ell - \lambda)r(\mu), \qquad \lambda = \lambda(\mu)$$

Unless $k = 2$, $l = 1$ we see that $r(\mu)$ may be determined so that (2.7) removes the term $w^k\bar{w}^\ell$ from u. By successive applications of (2.7), we obtain the form

$$w_1 = \lambda(\mu)w + b(\mu)|w|^2 w + v(w, \bar{w}, Y, \mu) + G_4. \tag{2.8}$$

The transformation

$$w = \zeta + Y'r(\mu)\zeta^k\bar{\zeta}^\ell, \quad k + \ell = 1 \tag{2.9}$$

[1] By the smoothness assumption made after (2.4), R_k is well defined provided $k \leq 4$.

with $r(\mu)$ a vector carries (2.8) into

$$\zeta_1 = \lambda(\mu)\zeta + b(\mu)|\zeta|^2\zeta + v(\zeta, \bar{\zeta}, Y, \mu) - Y' g(\mu)\zeta^k\bar{\zeta}^\ell + \tilde{G}_4$$

where

$$g(\mu) = [S'(\mu)\lambda^k\bar{\lambda}^\ell - \lambda I]r(\mu), \quad \lambda = \lambda(\mu)$$

and I is the identity matrix. Since $\lambda(0)$ is on the unit circle and from (2.3) the eigenvalues of $S'(0)$ are inside the unit circle, we see that $r(\mu)$ may be determined for μ sufficiently small such that the term $\zeta^k\bar{\zeta}^\ell$ is removed from v. The mapping now has the form

$$\zeta_1 = \lambda(\mu)\zeta + b(\mu)|\zeta|^2\zeta + H_4. \tag{2.10}$$

The second equation of (2.4) may be written

$$Y_1 = S(\mu)Y + s(\zeta, \bar{\zeta}, \mu) + H_3 \tag{2.11}$$

where s is quadratic in ζ and $\bar{\zeta}$. The transformation

$$Y = X + r(\mu)\zeta^k\bar{\zeta}^\ell, \quad k + \ell = 2$$

reduces (2.11) to

$$X_1 = S(\mu)X + s(\zeta, \bar{\zeta}, \mu) - g(\mu)\zeta^k\bar{\zeta}^\ell + \tilde{H}_3$$

where

$$g(\mu) = [\lambda^k\bar{\lambda}^\ell I - S(\mu)]r(\mu), \quad \lambda = \lambda(\mu).$$

Just as before $r(\mu)$ may be determined for μ sufficiently small. Thus the mapping has the form (2.10) together with

$$X_1 = S(\mu)X + V_3.$$

Using $\lambda(\mu) = e^{2\pi\alpha(\mu)}$, (2.10) may be written

$$\begin{aligned} \zeta_1 &= e^{2\pi\alpha(\mu)}[1 + \beta(\mu)|\zeta|^2]\zeta + H_4 \\ &= e^{2\pi\alpha(\mu) + \beta(\mu)|\zeta|^2}\zeta + \tilde{H}_4 \end{aligned}$$

with $\beta = b/\lambda$. All the previous transformations may be combined into one; $z = \zeta + \ldots, Y = X + \ldots$ where the dots represent polynomials in $\zeta, \bar{\zeta}, X$ without constant and linear terms. It is one-to-one in a neighborhood $|\zeta|^2 + |X|^2 \leq 2c$. The transformation $\zeta = cw$, $X = cW$ then gives the desired results.

4. Existence of an Invariant Curve

Having the mapping in the form (2.5) we may now establish the existence of an invariant curve and hence, according to the discussion of Section 2, an invariant torus of the differential equation (2.1).

Theorem 2. In (2.4) assume

1. $\lambda^4(0) \neq 1, \ \lambda^3(0) \neq 1$ (non-resonance)
2. $\lambda(\mu), \ S(\mu), \ U, \ V \in C^\ell(z, \bar{z}, Y, \mu)$ for ℓ an integer $\geq 5, \ |\mu| \leq \mu_*, \ |z|^2 + |Y|^2 \leq 2$ and assume $S(\mu)$ satisfies (2.3).

In (2.5) assume

3. $A = 2\pi Re\ \alpha'(0) > 0,\ B = Re\ \beta(0) < 0.$

Let $r,\ 1 \leq r \leq \ell$ be a positive integer. Then there exists $\mu_r \leq \mu_*$ such that for $0 \leq \mu < \mu_r$, (2.4) has an asymptotically stable invariant curve

$$\begin{aligned} z &= a_0\sqrt{\mu}e^{i\theta} + \mu f(\theta,\mu)e^{i\theta} \\ Y &= \mu g(\theta,\mu) \end{aligned} \tag{2.12}$$

where $a_0 = \sqrt{-A/B}$. f and g are defined for all θ and $0 < \mu < \mu_r$, have period 2π in θ,

$$f, g \in C^{r-1}(\theta) \cap Lip^{r-1}(\theta) \cap C(\mu)$$

and

$$|f|_{r-1}, |g|_{r-1} \leq \tilde{c}(r)$$

where $\tilde{c}(r)$ is a constant which depends only on r and differentiation is with respect to θ only. In general $\mu_r \to 0$ as r increases. μ_r depends on the coefficients of terms of degree ≤ 3 in(2.4), i.e., on terms of degree ≤ 3 in the expansion of (2.1) about ψ. μ_r also depends on the constant M_0 defined after (2.1).

Proof. In (2.5) we assume c to be fixed such that the transformation is one-to-one for $|w|^2 + |W|^2 \leq 2$. Note that c depends only on the coefficients of terms of degree ≤ 3 in (2.5).
We restrict attention to a neighborhood,

$$w = az,\ W = a^2Y \tag{2.13}$$

where $|z|^2 + |Y|^2 \leq 2$ and $0 \leq a \leq 1$. Then (2.5) becomes

$$\begin{aligned} z_1 &= e^{2\pi\alpha(\mu)+a^2c^2\beta(\mu)|z|^2}z &+ f_1(z,\bar{z},Y,a,\mu) \\ Y_1 &= S(\mu)Y &+ f_2(z,\bar{z},Y,a,\mu) \end{aligned}$$

and

$$|f_1|_r \leq c(r)a^3, \qquad |f_2|_r \leq c(r)a$$

where differentiation is with respect to z, $\bar{z}$ and Y. Note that the subscripts on the functions are used simply for identification and have nothing to do with "weight" as they did in the previous theorem.

Expanding α, β, and S, we obtain

$$\begin{aligned} z_1 &= \exp[2\pi\alpha(0) + \mu 2\pi\alpha'(0) + a^2c^2\beta(0)|z|^2]z + f_3(z,\bar{z},Y,a,\mu) \\ Y_1 &= S(0)Y + f_4(z,\bar{z},Y,a,\mu) \end{aligned}$$

where

$$|f_3|_r \leq c(r)(\mu^2 + a^2\mu + a^3), \quad |f_4|_r \leq c(r)(a+\mu).$$

Letting

$$\begin{aligned} S(0) &= S, & 2\pi\alpha(0) &= i\tau \\ 2\pi\alpha'(0) &= A + i\tilde{A}, & \beta(0) &= B + i\tilde{B} \end{aligned}$$

and choosing

$$a = \sqrt{\frac{-A\mu}{c^2 B}}, \qquad 0 \le \mu < \mu_0 = \min\left(\mu_*, \frac{c^2 B}{-A}\right)$$

we obtain

$$\begin{aligned} z_1 &= \exp[i(\tau + \mu\phi) + \mu A(1 - |z|^2)]z + f_5(z, \bar{z}, Y, \mu) \\ Y_1 &= SY + f_6(z, \bar{z}, Y, \mu) \end{aligned} \tag{2.14}$$

where

$$\phi = \tilde{A} - \frac{A\tilde{B}}{B}|z|^2$$

and

$$|f_5|_r \le c(r)\mu^{\frac{3}{2}}, \quad |f_6|_r \le c(r)\mu^{\frac{1}{2}}.$$

The unperturbed mapping ($f_5, f_6 \equiv 0$) has an invariant curve $|z| = 1,\ Y \equiv 0$. We introduce coordinates in a small toriod (see Fig. 6) surrounding this curve.

$$z = (1+\rho)e^{i\theta},\ |\rho|^2 + |Y|^2 \le \frac{1}{4}. \tag{2.15}$$

The first equation of (2.14) becomes

$$\rho_1 = (1 - 2\mu A)\rho - 3\mu A\rho^2 - \mu A\rho^3 + f_7(\rho, \theta, Y, \mu)$$

$$\theta_1 = \theta + \tau + \mu\left[\tilde{A} - \frac{A\tilde{B}}{B}(1+\rho)^2\right] + f_8(\rho, \theta, Y, \mu)$$

where

$$|f_7|_r, \quad |f_8|_r \le c(r)\mu^{\frac{3}{2}}.$$

Letting $\rho = \mu^{\frac{1}{2}}u,\ Y = \mu^{\frac{1}{2}}v$, we finally obtain (2.16)

$$\begin{aligned} u_1 &= (1 - 2\mu A)u &+\ & \mu[g(\theta) + \tilde{G}(\theta, u, v, \mu)] \\ \theta_1 &= \theta + \tau &+\ & \mu[f_0 + \tilde{F}(\theta, u, v, \mu)] \\ v_1 &= Sv &+\ & h(\theta) + \tilde{H}(\theta, u, v, \mu) \end{aligned} \tag{2.17}$$

where f_0 is a constant, all functions have period 2π in θ and

$$|g|_r,\ |h|_r \le \hat{c}(r);\ |\tilde{G}|_r,\ |\tilde{H}|_r,\ |\tilde{F}|_r \le \tilde{c}(r, K)\mu^{\frac{1}{2}} \tag{2.18}$$

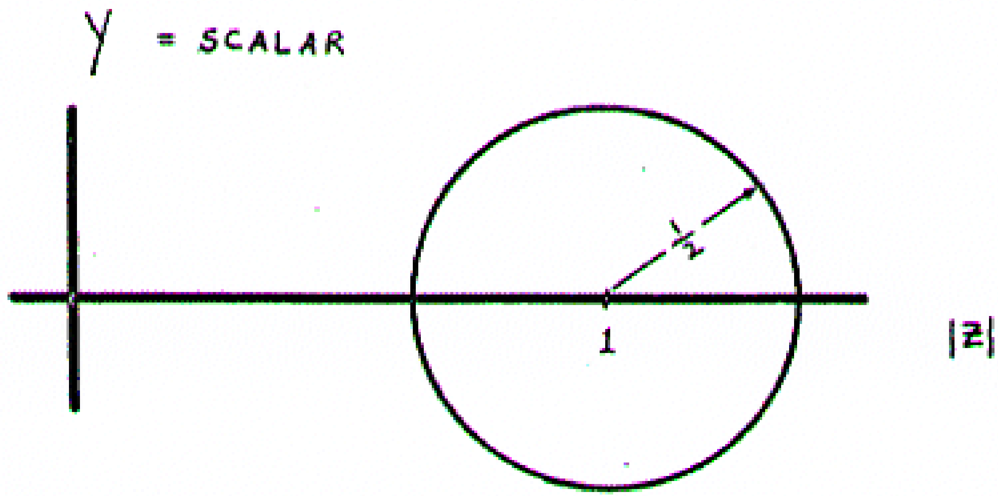
for

$$|u|^2 + |v|^2 \le K^2, \qquad 0 \le \mu \le \frac{1}{4K^2}$$

where differentiation is with respect to all variables except μ. $\hat{c}$ and $\tilde{c}$ are constants which depend only on the arguments shown. See Fig. 7 for a description of the toroidal neighborhood in the $W,\ w$ coordinates. Finding an invariant curve $u = u(\theta, \mu),\ v = v(\theta, \mu)$ of (2.17) is equivalent to solving the functional equation

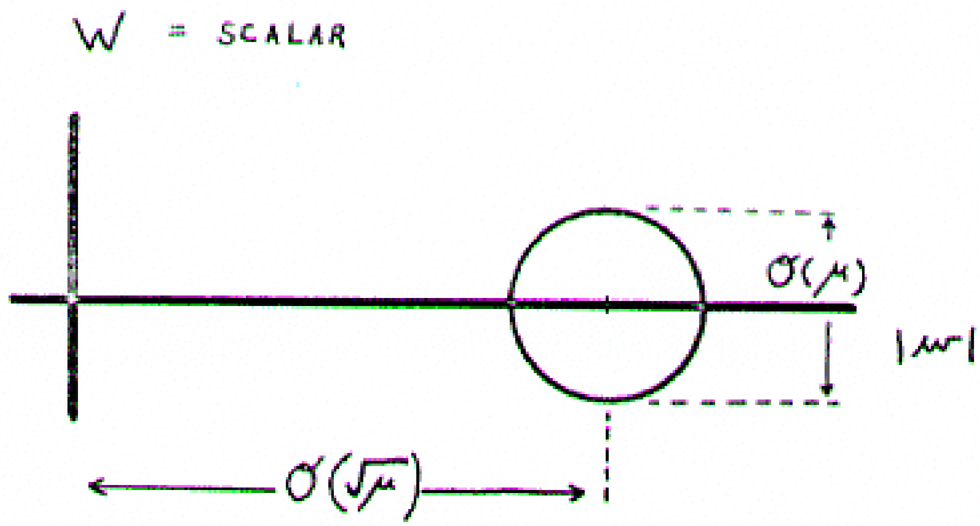
$$\begin{aligned} u(\theta_1) - (1 - 2\mu A)u(\theta) &= \mu G(\theta, u, v, \mu) \\ v(\theta_1) - Sv(\theta) &= H(\theta, u, v, \mu) \\ \theta_1 &= \theta + \tau + \mu F(\theta, u, v, \mu) \end{aligned} \tag{2.19}$$

FIG. 6



In each figure the toroidal neighborhood is obtained by rotating the circle about the vertical axis.

FIG. 7



where $F = f_0 + \tilde{F},\ f_0$ a constant, $G = g(\theta) + \tilde{G}$ and $H = h(\theta) + \tilde{H}$ satisfy (2.18). Theorem 3 which follows guarantees the existence of a unique solution $u = u(\theta,\mu),\ v = v(\theta,\mu)$ of (2.19) satisfying all our requirements. The form (2.12) is obtained by applying the transformations (2.16), (2.15), (2.13) and finally, the transformation of Theorem 1.

We now prove the stability of the invariant curve. If $u(\theta,\mu),\ v(\theta,\mu)$ is the solution of (2.19) for $0 < \mu < \bar{\mu}_r$, let $u = u(\theta,\mu) + \delta u,\ v = v(\theta,\mu) + \delta v$ in (2.17) to obtain the nonlinear variational mapping

$$\begin{aligned} \delta u_1 &= (1 - 2\mu A)\delta u + \mu(G_u - u' F_u)\delta u + \mu(G_v - u' F_v)\delta v \\ \delta v_1 &= S\delta v + (H_u - \mu v' F_u)\delta u + (H_v - \mu v' F_v)\delta v \end{aligned}$$

where prime is $d/d\theta$ and F_u etc. are evaluated at $u(\theta,\mu) + \epsilon\delta u,\ v(\theta,\mu) + \epsilon\delta v,\ 0 < \epsilon < 1$. From (2.18) we see that the derivatives of F, G, and H with respect to u and v are $\leq$ (constant) $\mu^{\frac{1}{2}}$ uniformly in a small tube $R:\ |\delta u|^2 + |\delta v|^2 \leq$ constant. Hence in R for μ sufficiently small, $0 < \mu < \mu_r \leq \bar{\mu}_r$, the above mapping is a contraction. This completes the proof of the theorem.

5. Solution of Functional Equation

The above stability argument is based on the principle of contraction, i.e., the eigenvalues of the linear part of the mapping are less than unity in modulus. This same principle will now be utilized to prove existence of a solution of equation (2.19).

Theorem 3. In (2.19) assume

1. $A > 0,\ |Sv| \leq \sigma_0|v|,\ 0 < \sigma_0 < 1\ \sigma_0$ constant
2. $F, G, H \in C^r(\theta, u, v) \cap C(\mu),\ r \geq 1,$

for $0 \leq \mu < \mu_0$ and $|u|^2 + |v|^2 \leq K^2,\ K$ sufficiently large and assume these functions satisfy (2.18) and have period 2π in θ.

Then there exists $\mu_r \leq \mu_0$ such that for $0 < \mu < \mu_r$ (2.19) has a unique solution $u(\theta,\mu),\ v(\theta,\mu)$ having period 2π in θ.

$$u, v \in C^{r-1}(\theta) \cap Lip^{r-1}(\theta) \cap c(\mu)$$

and $|u|_{r-1},\ |v|_{r-1} \leq K$ uniformly in μ where differentiation is with respect to θ only. μ_r depends on A, σ_0, r and the constants $\hat{c}$ and $\tilde{c}$ in (2.18). In general, $\mu_r \to 0$ as r increases.

Proof. Define $b = \min(2A, 1 - \sigma_0)$. Then $0 < b < 1$. Also define

$$K = b^{-r-1} c_0(r)\ \hat{c}(r) 2^{\delta(r+1)+3}$$

where $c_0(r)$ and δ are defined in Lemma 2 [2] and $\hat{c}(r)$ comes from (2.18). We will construct successive approximations u_n and v_n with $u_0 \equiv 0, v_0 \equiv 0$. For convenience let $w_n = \binom{u_n}{v_n}$. As an induction assumption suppose w_{n-1} satisfies

$$w_{n-1} \in C^r(\theta) \cap C(\mu), \qquad 0 < \mu < \mu_0. \tag{2.20}$$

[2] Lemmas 1, 2 and 3 are given later

$$|w_{n-1}|_r < K \tag{2.21}$$

Inserting this approximation in (2.19) we obtain

$$\begin{aligned} u(\theta_1) &- (1-2\mu A)u(\theta) &= \mu G_n(\theta,\mu) \\ v(\theta_1) &- Sv(\theta) &= H_n(\theta,\mu) \end{aligned} \tag{2.22}$$

$$\theta_1 = \theta + \tau + \mu F_n(\theta,\mu)$$

to be solved for $w_n(\theta,\mu)$, where

$$F_n(\theta,\mu) = F(\theta, w_{n-1}(\theta,\mu),\mu) \tag{2.23}$$

and so on. We now verify the conditions of Lemma 1. We have $F_n, G_n, H_n \in C^r(\theta) \cap C(\mu)$ for $0 < \mu < \mu_0' = \min(\mu_0, 1/4K^2)$. Also if we define $m = \inf_{\theta,\mu} F_{n\theta}$ we have, using (2.18), $b + rm \geq b - cr\mu^{\frac{1}{2}} > b/2$ for μ small, $0 < \mu < \mu' \leq \mu_0'$, where c depends on K and $\tilde{c}(r,K)$. Unless $F \equiv 0$ we see that $\mu' \to 0$ as r increases. By Lemma 1, (2.22) has a solution $w_n(\theta,\mu)$ which satisfies (2.20) for $0 < \mu < \mu'' \leq \mu'$. To verify (2.21) for w_n define $Q_n = \binom{G_n}{H_n}$ and apply Lemma 2:

$$|w_n|_r \leq \frac{c_0(r)|Q_n|_r[1 + |F_{n\theta}|_{r-1}]^{\delta(r)}}{(b+rm)^{r+1}}.$$

From (2.18) we have

$$|F_{n\theta}|_{r-1} \leq c_1\mu^{\frac{1}{2}} < 1$$

$$|Q_n|_r \leq 2\hat{c}(r) + c_1\mu^{\frac{1}{2}} < 4\hat{c}(r)$$

for μ small, $0 < \mu < \tilde{\mu} < \mu''$. Thus

$$|w_n|_r < (2/b)^{r+1} 4\hat{c}(r) c_0(r) 2^{\delta(r)} = K.$$

Repeating this procedure we generate a sequence $\{w_n\}$ which satisfies (2.20) and (2.21) for $0 < \mu < \tilde{\mu}$.

To show convergence define $\delta u_{n+1}(\theta) = u_{n+1}(\theta,\mu) - u_n(\theta,\mu)$, $\delta G_{n+1} = G_{n+1} - G_n$ and so on. Then from (2.22)

$$\begin{aligned} \delta u_{n+1}(\theta_1) - (1-2\mu A)\delta_{n+1}u(\theta) &= \mu[\delta G_{n+1} - u_n'\delta F_{n+1}] \\ \delta v_{n+1}(\theta_1) - S\delta v_{n+1}(\theta) &= \delta H_{n+1} - v_n'\delta F_{n+1}. \end{aligned}$$

$$\theta_1 = \theta + \tau + \mu F_{n+1}(\theta,\mu)$$

Defining Q_n as before, we have from the first estimate of Lemma 2

$$|\delta w_{n+1}|_0 \leq \frac{2}{b}\{|Q_{n+1}|_0 + |w_n|_1\ |\delta F_{n+1}|_0\}.$$

Using (2.21), (2.23) and (2.18) we have

$$|\delta w_{n+1}|_0 \leq \bar{c}\mu^{\frac{1}{2}}|\delta w_n|_0 \leq \frac{1}{2}|\delta w_n|_0$$

for μ sufficiently small, $0 < \mu < \bar{\mu} \leq \tilde{\mu}$, i.e., uniform convergence of the w_n. For the θ derivative $w_n^{(\lambda)}$, $0 \leq \lambda \leq r-1$, we have from Lemma 3 (with $m=1,\ x_1=\theta,\ \ell=r$) and (2.21)

$$|w_{n+p}^{(\lambda)} - w_n^{(\lambda)}|_0 \leq 2K\ c(\lambda, r)|w_{n+p} - w_n|_0^{1-(x/r)}.$$

Hence $\{w_n\}$ converges to a solution w of (2.19) and $w(\theta,\mu) \in C^{r-1}(\theta) \cap C(\mu)$. The Lipschitz condition follows from passing to the limit in

$$|w_n^{(\lambda)}(\theta',\mu) - w_n^{(\lambda)}(\theta,\mu)| \leq |w_n|_{\lambda+1}|\theta'-\theta| \leq K|\theta'-\theta|.$$

Uniqueness follows by applying the convergence argument to the sequence $\{w_1, w_2, w_1, w_2, \ldots\}$ where w_1 and w_2 are two supposed solutions. Finally we note that $|w|_{r-1} \leq K$. Thus the theorem is proved with $\mu_r = \bar{\mu}$.

Lemma 1. (Solution of Linear Functional Equation)
Consider the equation

$$\begin{aligned} u(\theta_1) \quad - \quad (1-2\mu A)u(\theta) \quad &= \quad \mu G(\theta) \\ v(\theta_1) \quad - \quad Sv(\theta) \quad &= \quad H(\theta) \end{aligned} \tag{2.25}$$

$$\theta_1 = \theta + \tau + \mu F(\theta,\mu) \tag{2.26}$$

where τ is constant, $0 < \mu < \mu'$, $A > 0$ and $|Sv| \leq \sigma_0|v|$ for some $\sigma_0,\ 0 < \sigma_0 < 1$. Define $b = \min(2A, 1-\sigma_0)$ and $m = \inf_{\theta,\mu} F_\theta$. Assume

1. $b + rm > 0$ for some integer $r \geq 0$
2. $F, G, H \in C^r(\theta) \cap C(\mu)$ and have period 2π in θ.

Then for

$$0 < \mu < \mu'' = \min\left[\frac{1}{2A}, \frac{1}{|m|}, \mu', 1\right]$$

(2.25-6) has a solution $u(\theta,\mu), v(\theta,\mu) \in C^r(\theta) \cap C(\mu)$ having period 2π in θ.

Remark. Since F is periodic in θ, m will be ≤ 0 and hence if 1 is true for some integer $r = r_0$ then it is true for any smaller integer $r < r_0$.

Proof. Each equation in (24) has the form

$$Y(\theta_1) - D(\mu)Y(\theta) = P(\theta,\mu) \tag{2.27}$$

with

$$|D(\mu)Y| \leq (1-\mu b)|Y| \tag{2.28}$$

where $D(\mu)$ represents either $S(\mu)$ or $1 - 2\mu A$. Let I be any closed subinterval of $(0, \mu'')$. We may solve (2.26) for the negative iterates $\theta_{-1}, \theta_{-2}, \ldots$ where $\theta_{-n+1} = \theta_{-n} + \tau + \mu F(\theta_{-n}, \mu)$ since

$$\left|\frac{d\theta_1}{d\theta}\right| \geq 1 + \mu m > 0 \text{ for } \mu \in I.$$

We allow D to depend on θ also and write

$$Y(\theta) - D(\theta_{-1}, \mu)Y(\theta_{-1}) = P(\theta_{-1}, \mu)$$

$$D(\theta_{-1}, \mu)Y(\theta_{-1}) - D(\theta_{-1}, \mu)D(\theta_{-2}, \mu)Y(\theta_{-2}) = D(\theta_{-1}, \mu)P(\theta_{-2}, \mu)$$
$$\vdots$$

Adding and passing to the limit

$$Y(\theta, \mu) = \sum_{j=1}^{\infty} \prod_{k=1}^{j-1} D(\theta_{-k}, \mu)P(\theta_{-j}, \mu)$$

where $\prod_{k=1}^{0} \equiv 1$. From (2.28) we see that the series converges uniformly for all θ and for $\mu \in I$. If $r = 0$ we are finished. If $r = 1$ we differentiate term by term using

$$\frac{d}{d\theta}\theta_{-j} = \frac{d\theta_{-j}}{d\theta_{-j+1}} \cdots \frac{d\theta_{-1}}{d\theta}$$

and

$$\left|\frac{d}{d\theta}\theta_{-j}\right| \leq \left(\frac{1}{1 + \mu F_\theta}\right)^j.$$

The differentiated series converges uniformly for the same values of θ and μ since for $|y| = 1$,

$$\frac{|D(\mu)y|}{1 + \mu F_\theta} \leq \frac{1 - \mu b}{1 + \mu m} \leq c < 1 \tag{2.29}$$

where c depends only on I, b and m. Hence $Y(\theta, \mu) \in C^1(\theta) \cap C(\mu)$. If $r > 1$ assume for $s \leq r$ that $Y \in C^{s-1}(\theta) \cap C(\mu)$. Then differentiate (2.27) $s - 1$ times with respect to θ

$$Y^{(s-1)}(\theta_1) - \frac{D(\mu)Y^{(s-1)}(\theta)}{(1 + \mu F_\theta)^{s-1}} = \tilde{P}(\theta, \mu) \tag{2.30}$$

where $\tilde{P}$ contains derivatives $Y^{(\lambda)},\ 0 \leq \lambda \leq s - 2$ considered as known functions. Thus $\tilde{P} \in C^1(\theta) \cap C(\mu)$. We now consider this as an equation for $Y^{(s-1)}$ and repeat the above procedure with $D(\mu)[1 + \mu F_\theta]^{1-s}$ in place of $D(\mu)$. Using the fact that for $x > -1, (1 + x)^s \geq 1 + sx$ we see that the corresponding condition (2.29) is

$$\frac{|D(\mu)y|}{[1 + \mu F_\theta]^s} \leq \frac{1 - \mu b}{1 + \mu sm} \leq c_1 < 1$$

hence the solution of (2.30) is in $C^1(\theta) \cap C(\mu)$, i.e.,

$$Y(\theta, \mu) \in C^r(\theta) \cap C(\mu). \qquad \text{Q.E.D.}$$

Lemma 2. "A-priori" Estimates. Consider equation (2.25-6) of Lemma 1 and assume everything in the statement of that lemma. Letting $Q = \binom{G}{H}$ we then have the following estimates for the solution $w = \binom{u}{v}$ of (2.25-6).

$$|w|_0 \leq \frac{2}{b}|Q|_0$$

$$|w|_1 \leq \frac{2|Q|_1}{(b+m)^2}$$

and in general for $1 \leq s \leq r$, s an integer,

$$|w|_s \leq \frac{c_0(s)|Q|_s[1+|F_\theta|_{s-1}]^{\delta(s)}}{(b+sm)^{s+1}}$$

where, as before, $m = \inf_{\theta,\mu} F_\theta$. $c_0(s)$ is a constant which depends only on s, $\delta(1) = 0$ and $\delta(n+1) = \delta(n) + n$.

Proof. We will use the fact that for $x > -1$ and $n \geq 0$ an integer

$$(1+x)^n \geq 1 + nx. \tag{2.30a}$$

From the first equation of (2.25)

$$|u(\theta_1)| \leq (1-\mu b)|u|_0 + \mu|G|_0$$

for all θ_1 and in particular at the point where $|u|_0$ is assumed. Thus $|u|_0 \leq b^{-1}|G|_0$. Similarly for v

$$|v|_0 \leq (1-b)|v|_0 + |H|_0$$

and therefore $|v|_0 \leq b^{-1}|H|_0$. Adding we obtain the first estimate. For the second estimate we differentiate to obtain

$$u^{(1)}(\theta_1)(1+\mu F_\theta) - (1-2\mu A)u^{(1)}(\theta) = \mu G_\theta.$$

Noting the restrictions imposed on μ in Lemma 1,

$$(1+\mu m)|u^{(1)}|_0 \leq (1-\mu b)|u^{(1)}|_0 + \mu|G|_1$$

and similarly for v. Thus

$$|w^{(1)}|_0 \leq 2(b+m)^{-1}|Q|_1.$$

Using $b+m \leq b < 1$ we have $|w|_1 \leq \dfrac{2|Q|_1}{(b+m)^2}$. To prove the general statement for $s \geq 2$ assume it true for $s-1$ and differentiate the first equation in (2.25) s times

$$u^{(s)}(\theta_1)(1+\mu F_\theta)^s - (1-2\mu A)u^{(s)}(\theta) = \mu[G^{(s)} + g_s]$$

where g_s is a linear combination of terms of the form

$$u^{(p)} \prod_{i=1}^{q} F^{(\lambda_i)} \tag{2.31}$$

where $0 \le p,\ q \le s-1$ and $1 \le \lambda_i \le s$.

At the point θ_1 where $|u^{(s)}(\theta_1)| = |u^{(s)}|_0$ we have, using (2.30a), $1 + \mu sm \le (1 + \mu F_\theta)^s$ and therefore

$$|u^{(s)}|_0(1 + \mu sm) \le (1 - \mu b)|u^{(s)}|_0 + \mu[|G|_s + |g_s|_0]$$

or

$$|u^{(s)}|_0 \le \frac{|Q|_s + |g_s|_0}{b + sm}.$$

Using the estimate for $s-1$ and the fact that $b + sm \le \ldots \le b + m \le b \le 1$, we obtain

$$|u^{(s)}|_0 \le \frac{c(s)|Q|_s[1 + |F_\theta|_{s-1}]^{\delta(s)}}{(b + sm)^{s+1}}.$$

In estimating g_s, since the maximum number of terms in the product in (2.31) is $s-1$, we choose $\delta(s) = \delta(s-1) + s - 1$. Repeating this for the second equation in (2.25) and adding we obtain the desired result. Q.E.D.

We now establish an inequality which is useful in proving the convergence of the derivatives $g_n^{(\lambda)}$, $0 \le \lambda < l$, of a sequence whenever the original sequence $\{g_n\}$ converges and the sequence $\{g_n^{(l)}\}$ is uniformly bounded.

Lemma 3. Suppose $f(x) = (f_1, \ldots, f_n) \in C^\ell(x)$ for all $x = (x_1, \ldots, x_m)$. Then for $0 \le \lambda < \ell$

$$|D^\lambda f|_0 \le c(\lambda, \ell)|f|_0^{1-(\lambda/\ell)}|f|_\ell^{\lambda/\ell}$$

for positive integers λ and ℓ, where D_λ is any partial derivative of order λ. c depends on m and n but not f.

Proof. We define the norm

$$[f]_r = \sup_{x,i} |D^r f_i|$$

where the supremum is taken over all r^{th} order derivatives. We introduce the smoothing operator (J. Moser [4])

$$T_N f(x) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K_N(x - x')f(x')dx'_1 \ldots dx'_m \tag{2.33}$$

valid for any $N > 0$ and all x. The kernel is given by

$$K_N(x) = N^m K(Nx_1) \ldots K(Nx_m)$$

where $K(t) \in C^\infty$ for all t and $K(t) \equiv 0$ for $|t| > 1$, and

$$\int_{-\infty}^{\infty} t^k\ K(t)dt = \begin{cases} 1 & \text{for } k = 0 \\ 0, & 0 < k < s \end{cases}$$

with s an integer to be fixed later. K is constructed in [4] from a C^∞ function ϕ. We assume that the mode of construction as well as ϕ are fixed. Therefore K depends only on t and s. For (2.33) the following inequalities can be verified for $\sigma,\ s,\ r \geq 0$:

$$[T_N f]_{r+\sigma} \leq c_1(\sigma, s) N^\sigma [f]_r \tag{2.34}$$

$$[f - T_N f]_r \leq c_2(s)\ N^{-s} [f]_{r+s} \tag{2.35}$$

where c_1 and c_2 depend on m but not on N and f. Letting $N = ([f]_\ell / [f]_0)^{1/\ell}$, we use (2.35) with $r = \lambda,\ s = \ell - \lambda$ and (2.34) with $r = 0,\ \sigma = \lambda$ to obtain

$$[f]_\lambda \leq \tilde{c}(\lambda, \ell) [f]_0^{1-(\lambda/\ell)} [f]_\ell^{\lambda/\ell}.$$

The final result follows using

$$c_3(n) |D^p f|_0 \leq [f]_p \leq |f|_p$$

where c_3 is a constant depending only on n. Q.E.D.

APPENDIX

The following material is taken from Chapter 1. (RJS 2008)

Notation: We now introduce some notation which will be used throughout the rest of our work. For r a positive integer

$$F(x, y, \mu) \in C^r(x, y) \cap C(\mu)$$

means that F has r derivatives with respect to x and y which are continuous in (x, y, μ).

$$F(x, y, \mu) \in C^r(x, y) \cap Lip^r(x, y) \cap C(\mu)$$

means the same as above and the derivatives are Lipschitz continuous in x and y uniformly in μ.

For a vector $f = (f_i)$ define

$$|f| = \left(\sum_i f_i^2 \right)^{\frac{1}{2}} = (f, f)^{\frac{1}{2}}$$

We will consider functions $f(x, \mu)$ defined for x a vector in some domain G and μ a small parameter. For $r \geq 0$ an integer define the norm

$$|f|_r = \max_{0 \leq \rho \leq r} \ \sup_{x \in G} |D^\rho f|$$

where D^ρ is any ρ^{th} order derivative of f with respect to the components x_i of x. Unless otherwise stated, differentiation will never be performed with respect to μ.

For two real vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, the symbol (u, v) will denote the usual inner product

$$(u, v) = \sum_{k=1}^n u_k v_k.$$

REFERENCES

This a truncated set of references containing only those cited in Chapter 2. (RJS 2008)

1. Lefschetz, Solomon, Differential Equations: Geometric Theory - Second Edition. New York: Interscience Publishers, 1963.

2. Segre, Beniamino, Some Properties of Differentiable Varieties and Transformations. Berlin-Gottingen-Heidelberg: Springer-Verlag, 1957, Part 2.

3. Kelly, A., “Using Change of Variables to Find Invariant Manifolds of Systems of Ordinary Differential Equations in a Neighborhood of a Critical Point, Periodic Orbit, or Periodic Surface - History of the Problem,” Technical Report No. 3, Office of Naval Research, Dept. of Math., Univ. of Calif., Berkeley, May 1963.

4. Moser, Jürgen K., “On Invariant Curves of Area-Preserving Mappings of an Annulus,” Akad. Wiss. Göttingen, Math.-Phys. Klasse Nachrichten. No. 1, 1962, pp. 1-20, (In particular pp.10-12).