JOURNAL OF DIFFERENTIAL EQUATIONS 26, 112–159 (1977)

Integral Averaging and Bifurcation*

SHUI-NEE CHOW

Mathematics Department, Michigan State University, East Lansing, Michigan 48824

AND

JOHN MALLET-PARET

Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

Received September 1, 1976

1. INTRODUCTION

One of the simplest topological variations of the phase space of a one-parameter family of differential equations (vector fields, flows) is the creation of periodic orbits from equilibria as the parameter crosses a critical value. The study of such topological variations about an equilibrium was initiated and developed by Poincaré perhaps 90 years ago and belongs today to the classical theory of periodic solutions. It was Hopf [23] who presented the bifurcation theorem in 1942 and it is now commonly known as the Hopf bifurcation theorem. Specifically, consider a one-parameter family of ODE (ordinary differential equations)

$$\dot{x}(t) = f(\alpha, x(t)), \qquad \alpha \in R,\ x \in R^n.$$

Suppose that $f(\alpha, 0) \equiv 0$ and f admits the linearization

$$\dot{y}(t) = A(\alpha)\, y(t).$$

Assume that $A(\alpha)$ has a pair of complex conjugate eigenvalues $\lambda(\alpha)$ and $\bar{\lambda}(\alpha)$ such that

$$\operatorname{Re} \lambda'(0) > 0, \qquad \operatorname{Re} \lambda(0) = 0, \qquad \text{and} \qquad \operatorname{Im} \lambda(0) \neq 0.^1$$

* This research was supported by the National Science Foundation under GP 28931X3 and in part by the United States Army under DA-ARO-D-31-124-73-G-130.

[1] If $\operatorname{Im} \lambda(0) = 0$, then it is not periodic but equilibrium states which bifurcate from the zero solution. Such bifurcation theory is not within the scope of the present paper. In [8, 9], we consider bifurcation problems under precisely the condition $\operatorname{Im} \lambda(0) = 0$.

ISSN 0022-0396

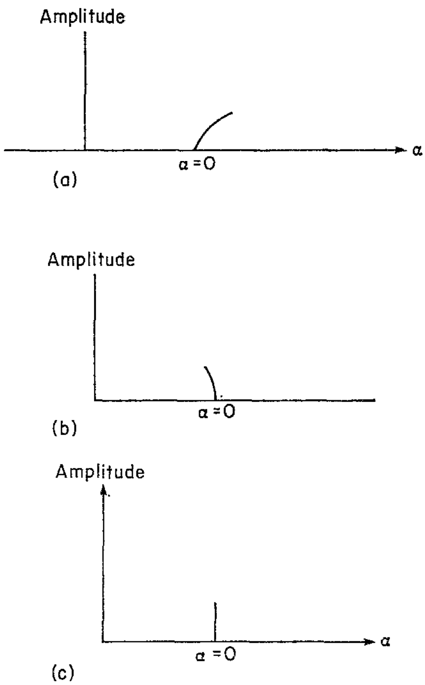

FIG. 1. Amplitude–parameter graph: (a) supercritical; (b) subcritical; (c) vertical.

Then, it was shown by Hopf that there are periodic orbits bifurcating from the zero solution. If one plots the amplitude–parameter graph (see Fig. 1), the three situations (a), (b), and (c) are all possible. Moreover, all three cases are of great physical interest. For example, Fig. 1a represents the first stage of transition to turbulance in fluids as postulated by Landau [31]; Fig. 1b represents an "inverted Hopf bifurcation" which often occurs for flows that exhibit an immediate transition behavior (see [38]); Fig. 1c represents a "degenerate Hopf bifurcation" which is related to the Liapunov center theorem (see [1]). The existence of Hopf bifurcation is a very elementary application of the implicit functions theorem. However, to determine the specific type of bifurcation depends upon certain analytic conditions which involve nonlinearities in the equation.

The purpose of this paper is to describe how, using the classical method of averaging (see [16, 18]) one can give conditions on the vector fields which ensure a supercritical or subcritical Hopf bifurcation. This is done in such a way that the theorems for FDE's (functional differential equations) and PDE's (partial differential equations) are essentially analogous to those of ODE's. The basic idea is to decompose the equation into three coupled equations which are equations for the amplitude r, the phase angle θ, and the stable part y. The decomposition is natural in the method of integral averaging. By means of a series of coordinate changes one then decouples the r equation up to a certain order in r. Examples from FDE's and PDE's are also included. Many of the transformations and techniques used in this paper are useful for problems involving a vector parameter α, as well as radial and angle coordinates (r, θ) which are vectors. This situation arises, for example, in the study of the bifurcation of an invariant torus from around a periodic orbit. See Lanford [32], Ruelle and Takens [40], and Sacker [41].

Since the appearance of Hopf's paper, there have been many papers related to similar problems, notably, Alexander and Yorke [1], Brunovsky [4], Bruslinskaya [5], Chafee [6, 7], Freedman [12], Friedrichs [13], Iooss [24], Josepg and Sattinger [26], Jost and Zehnder [27], Judovich [28], Kopell and Howard [30], Lanford [32], Marsden [35], Marsden and McCraken [36], McCracken [37], Ruelle and Takens [40], Sacker [41], Sattinger [42], Schmidt [43], Sotomayor [44], and Takens [46].

This paper is organized as follows. In Sections 2, 3 the method of averaging is described, without regard to any specific bifurcation problem. Although the exposition here is for ordinary differential equations, the mechanics of the averaging procedure carry over directly to infinite-dimensional systems (functional and partial differential equations). That is, many infinite-dimensional systems can be studied with averaging simply by rewriting them as an ordinary differential equation in a Banach space and proceeding formally from there. A more precise description of these ideas, as well as examples, is found in Sections 7–11.

In Section 4, the averaging method is applied to study the Hopf bifurcation in R^2, and in Section 5 these ideas are extended to the situation in R^n. Although a rather detailed study of the Hopf bifurcation is presented here, it should be noted that the techniques used have a much wider application. For example, the problem of bifurcation of an invariant torus from a periodic orbit can be treated in essentially the same manner, as is described at the end of Section 5. Thus the Hopf bifurcation (although an extremely interesting and important phenomenon) is presented here essentially as one illustration of the averaging technique.

In Section 6 the center manifold theorem is described. With this theorem one can give a rigorous proof of the existence of the periodic solutions formally obtained in the previous sections. This idea is especially important for infinite-dimensional systems.

The relation between infinite-dimensional systems and ordinary differential equations in a Banach space is explored in Sections 7, 8. The emphasis here is on writing such systems so that averaging can be applied as for finite-dimensional systems. Finally, in Sections 9–11 some specific examples of Hopf bifurcation are studied.

2. The Method of Averaging

Let us begin with a description of the averaging procedure for a class of ordinary differential (ODE's); in later sections this will be generalized and applied to solve bifurcation problems in both ODE's as well as functional differential equations (FDE's) and certain partial differential equations.

Consider first a two-dimensional system in polar coordinates (r, θ) of period 2π in θ, given by

$$\begin{aligned} \dot{r} &= \epsilon R_1(r, \theta, \alpha) + \epsilon^2 R_2(r, \theta, \alpha) + \cdots, \\ \dot{\theta} &= \omega + \epsilon W_1(r, \theta, \alpha) + \epsilon^2 W_2(r, \theta, \alpha) + \cdots, \end{aligned} \tag{2.1}$$

where ϵ and α are parameters, $\epsilon \in (-\epsilon_0, \epsilon_0) \subseteq R$, and $\omega \neq 0$ is constant. We assume this differential equation is sufficiently smooth for the following calculations to be performed; moreover, since only a finite number of coefficients of ϵ^k will be considered, it is sufficient that (2.1) represent a finite Taylor series, with a remainder term. We seek periodic solutions of (2.1), or integral manifolds in more general systems. In bifurcation problems, α is the bifurcation parameter while ϵ represents a scaling factor so that we need only consider r near a constant $r_0 > 0$; later, α will be chosen as a particular function of ϵ.

Now if each R_j is independent of θ, so $R_j(r, \theta, \alpha) = R_j(r, \alpha)$, in principle we are done, for the periodic solutions are precisely those circles $r = r_0$ satisfying

$$\epsilon R_1(r_0, \alpha) + \epsilon^2 R_2(r_0, \alpha) + \cdots = 0.$$

Thus, we strive to find new coordinates $(\bar{r}, \bar{\theta})$ for (2.1) in which enough of the R_j's are independent of $\bar{\theta}$. (In general, we can only hope a finite number of them will be, but this is sufficient.) In the Hopf bifurcation problem treated in this paper, it is necessary only to transform $r \to \bar{r}$, as the dependence of the W_j on θ is not important. For completeness however, in this section we transform $\theta \to \bar{\theta}$ as well, as this is important in considering bifurcation from a periodic orbit to a torus. These coordinate changes are constructed via integral averaging; here we present a description of them.

Suppose the coefficients of ϵ^j for $1 \leqslant j \leqslant k - 1$ are independent of θ, so that we have

$$\begin{aligned} \dot{r} &= \epsilon R_1(r, \alpha) + \cdots + \epsilon^{k-1} R_{k-1}(r, \alpha) + \epsilon^k R_k(r, \theta, \alpha) + O(\epsilon^{k+1}), \\ \dot{\theta} &= \omega + \epsilon W_1(r, \alpha) + \cdots + \epsilon^{k-1} W_{k-1}(r, \alpha) + \epsilon^k W_k(r, \theta, \alpha) + O(\epsilon^{k+1}). \end{aligned} \tag{2.2}$$

Consider a transformation of the form

$$\bar{r} = r + \epsilon^k u(r, \theta, \alpha), \qquad \bar{\theta} = \theta + \epsilon^k v(r, \theta, \alpha). \tag{2.3}$$

Clearly (2.3) brings (2.2) to the form

$$\dot{\bar{r}} = \epsilon R_1(\bar{r}, \alpha) + \cdots + \epsilon^{k-1} R_{k-1}(\bar{r}, \alpha) + \epsilon^k \bar{R}_k(\bar{r}, \bar{\theta}, \alpha) + O(\epsilon^{k+1}),$$
$$\dot{\bar{\theta}} = \omega + \epsilon W_1(\bar{r}, \alpha) + \cdots + \epsilon^{k-1} W_{k-1}(\bar{r}, \alpha) + \epsilon^k \bar{W}_k(\bar{r}, \bar{\theta}, \alpha) + O(\epsilon^{k+1}),$$

where

$$\bar{R}_k(\bar{r}, \bar{\theta}, \alpha) = R_k(\bar{r}, \bar{\theta}, \alpha) + \omega(\partial u/\partial \theta)(\bar{r}, \bar{\theta}, \alpha),$$
$$\bar{W}_k(\bar{r}, \bar{\theta}, \alpha) = W_k(\bar{r}, \bar{\theta}, \alpha) + \omega(\partial v/\partial \theta)(\bar{r}, \bar{\theta}, \alpha).$$

Thus u and v must be chosen to make $\bar{R}_k$ and $\bar{W}_k$ independent of $\bar{\theta}$; this choice is given in the following lemma.

LEMMA 2.1. *Consider the relation*

$$\bar{A}(r, \theta, \alpha) = A(r, \theta, \alpha) + \omega(\partial b/\partial \theta)(r, \theta, \alpha),$$

where A is given and all functions are 2π-periodic in θ. If b is chosen as

$$b(r, \theta, \alpha) = \frac{-1}{\omega} \int_0^\theta A(r, s, \alpha)\, ds + \frac{\theta}{2\pi\omega} \int_0^{2\pi} A(r, s, \alpha)\, ds$$

then $\bar{A}(r, \theta, \alpha) = \bar{A}(r, \alpha)$ is independent of θ. In fact, $\bar{A}$ is the mean value of A, so

$$\bar{A}(r, \alpha) = (1/2\pi) \int_0^{2\pi} A(r, \theta, \alpha)\, d\theta.$$

Proof. All that has to be checked is the easy fact that b is 2π-periodic in θ.

This lemma then guarantees the existence of a transformation (2.3) so that the coefficients of ϵ^j, $1 \leqslant j \leqslant k$, in (2.1) are independent of θ. Observe that the sequence of transformations averaging the $\epsilon, \epsilon^2, \ldots, \epsilon^k$ terms may be written as a single one,

$$\bar{r} = r + \epsilon u_1(r, \theta, \alpha) + \cdots + \epsilon^k u_k(r, \theta, \alpha),$$
$$\bar{\theta} = \theta + \epsilon v_1(r, \theta, \alpha) + \cdots + \epsilon^k v_k(r, \theta, \alpha).$$

The situation in dimension greater than two is somewhat more complicated. The equations here take the form

$$\begin{aligned} \dot{r} &= \epsilon R_1(r, \theta, y, \alpha) + \epsilon^2 R_2(r, \theta, y, \alpha) + \cdots, \\ \dot{\theta} &= \omega + \epsilon W_1(r, \theta, y, \alpha) + \epsilon^2 W_2(r, \theta, y, \alpha) + \cdots, \\ \dot{y} &= A_Q y + \epsilon Y_1(r, \theta, y, \alpha) + \epsilon^2 Y_2(r, \theta, y, \alpha) + \cdots, \end{aligned} \tag{2.4}$$

where (r, θ) is the rotational part of the differential equation, and $y \in R^q$ is the stable (or saddle point) part. It is thus assumed that the constant matrix A_Q

has no eigenvalues on the imaginary axis. This implies that for any solution $(r(t), \theta(t), y(t))$ of (2.4) bounded for all real t, we must have $y = O(\epsilon)$, since (if A_Q is stable)

$$y(t) = \epsilon \int_{-\infty}^{t} e^{A_Q(t-s)} [Y_1(r(s), \theta(s), y(s), \alpha) + O(\epsilon)]\, ds \tag{2.5}$$

with the analogous formula in case of a saddle point. It is not clear how to entirely eliminate the presence of θ and y in R_j and W_j; however, under very general conditions (but not always) we shall see it is possible to average (2.4) so that in the new coordinates,

$$\begin{aligned} \bar{R}_j(\bar{r}, \bar{\theta}, y, \alpha) &= R_{j0}(\bar{r}, \alpha) + O(|\, y\,|^m), \\ \bar{W}_j(\bar{r}, \bar{\theta}, y, \alpha) &= W_{j0}(\bar{r}, \alpha) + O(|\, y\,|^m), \end{aligned} \tag{2.6}$$

where $m = m(j)$ is given and $1 \leqslant j \leqslant k$. In view of (2.5), this is sufficient. In particular, this can always be done if A_Q is a stable matrix.

Assuming the coefficients of ϵ^j, $1 \leqslant j \leqslant k - 1$, are already averaged and so satisfy (2.6), let us describe the averaging of $R_k(r, \theta, y, \alpha)$; the situation for W_k is the same, so it is omitted. First expand R_k in powers of y,

$$R_k(r, \theta, y, \alpha) = \sum_{l=0}^{m-1} R_{kl}(r, \theta, \alpha)\, y^l + O(|\, y\,|^m),$$

so R_{kl} takes values in the vector space $\Lambda_q^{\,l}$ of symmetric l-linear maps from $R^q \times \cdots \times R^q = R^{ql}$ into R. Letting

$$\bar{r} = r + \epsilon^k u_k(r, \theta, y, \alpha) = r + \epsilon^k \sum_{l=0}^{m-1} u_{kl}(r, \theta, \alpha)\, y^l$$

it is seen that in the transformed variables $(\bar{r}, \theta, y)$, the coefficient of ϵ^k in $\dot{\bar{r}}$ is

$$\begin{aligned} \bar{R}_k(\bar{r}, \theta, y, \alpha) &= R_k(\bar{r}, \theta, y, \alpha) + \omega \frac{\partial u_k}{\partial \theta}(\bar{r}, \theta, y, \alpha) + \frac{\partial u_k}{\partial y}(\bar{r}, \theta, y, \alpha)\, A_Q y \\ &= \sum_{l=0}^{m-1} \left[R_{kl}(\bar{r}, \theta, \alpha) + \omega \frac{\partial u_{kl}}{\partial \theta}(\bar{r}, \theta, \alpha) + l u_{kl}(\bar{r}, \theta, \alpha)\, A_Q \right] y^l \\ &\quad + O(|\, y\,|^m). \end{aligned} \tag{2.7}$$

The notation $l u_{kl}(\bar{r}, \theta, \alpha)\, A_Q$ of (2.7) needs some explanation. Recall that u_{kl} (here we suppress the arguments $(\bar{r}, \theta, \alpha)$) is a symmetric l-linear map

$$u_{kl}\colon R^q \times \cdots \times R^q = R^{ql} \to R.$$

By $u_{kl}A_Q$ we mean that element of Λ_q^l given by

$$(u_{kl}A_Q)(a_1, ..., a_l) = (1/l) \sum_{i=1}^{l} u_{kl}(a_1, ..., A_Q a_i, ..., a_l).$$

If $l = 1$, u_{kl} is a linear functional on R^q and thus may be denoted by a row vector; in this case, $u_{kl}A_Q$ is the usual matrix multiplication. Observe also that the map

$$A_Q^l: \Lambda_q^l \to \Lambda_q^l \qquad \text{by} \qquad U \to UA_Q$$

is linear. Clearly A_Q^1 is just the adjoint of A_Q: $R^q \to R^q$.

Now, in order to obtain (2.6), it is necessary that

$$R_{k0}(\bar{r}, \theta, \alpha) + \omega \frac{\partial u_{k0}}{\partial \theta}(\bar{r}, \theta, \alpha) = \text{independent of } \theta, \tag{2.8}$$

$$R_{kl}(\bar{r}, \theta, \alpha) + \omega \frac{\partial u_{kl}}{\partial \theta}(\bar{r}, \theta, \alpha) + l u_{kl}(\bar{r}, \theta, \alpha) A_Q = 0 \qquad \text{for} \quad 1 \leqslant l \leqslant m - 1. \tag{2.9}$$

To get (2.8), simply choose u_{k0} as described in Lemma 2.1. The right-hand side of (2.8) is thus the average

$$(1/2\pi) \int_0^{2\pi} R_{k0}(\bar{r}, \theta, \alpha)\, d\theta = (1/2\pi) \int_0^{2\pi} R_k(\bar{r}, \theta, 0, \alpha)\, d\theta.$$

To obtain (2.9), observe that for each $(\bar{r}, \alpha)$, u_{kl} satisfies a linear inhomogeneous equation in θ, with periodic forcing term R_{kl}. A well-known result in differential equations asserts that (2.9) has a unique 2π-periodic solution if and only if the homogeneous equation

$$\omega(du/d\theta) + luA_Q = 0, \qquad u \in \Lambda_q^l$$

has no nontrivial 2π-periodic solutions; and this is true if and only if the linear map A_Q^l: $\Lambda_q^l \to \Lambda_q^l$ has no eigenvalues of the form $i\omega n/l$, for all integers n. This is certainly true for $l = 1$ since A_Q^1: $\Lambda_q^1 \to \Lambda_q^1$ is just right-matrix multiplication of row vectors by A_Q, and A_Q was assumed to have no pure imaginary eigenvalues. However, for $l > 1$, A_Q may have $i\omega n/l$ as an eigenvalue, and this motivates the following definition.

Definition. A $q \times q$ matrix M is called l-simple if the induced linear transformation M^l: $\Lambda_q^l \to \Lambda_q^l$ by $U \to UM$ has no eigenvalues of the form in/l for all integers n.

Thus (2.9) has a unique solution if and only if $(1/\omega) A_Q$ is l-simple. In Section 3, we shall give a necessary and sufficient criterion for a matrix to be l-simple; in particular, any stable matrix (all eigenvalues in the left half-plane) will be shown to be l-simple for all $l \geqslant 1$.

Here let us summarize the above as a theorem.

THEOREM 2.2. *Consider the differential equation*

$$\begin{aligned}\dot r &= \epsilon R_1(r,\theta,y,\alpha) + \epsilon^2 R_2(r,\theta,y,\alpha) + \cdots,\\ \dot\theta &= \omega + \epsilon W_1(r,\theta,y,\alpha) + \epsilon^2 W_2(r,\theta,y,\alpha) + \cdots,\\ \dot y &= A_Q y + \epsilon Y_1(r,\theta,y,\alpha) + \epsilon^2 Y_2(r,\theta,y,\alpha) + \cdots,\end{aligned}$$

for which the matrix $(1/\omega) A_Q$ *is l-simple for each* $1 \leqslant l \leqslant m-1$. *Then there exists a transformation*

$$\bar r = r + \epsilon^k u(r,\theta,y,\alpha),$$

where u is a polynomial in y of degree at most $m-1$, *such that in the new* $(\bar r,\theta,y)$ *coordinates, the term* R_k *becomes* $R_{k0}(\bar r,\alpha) + (|\,y\,|^m)$, *where*

$$R_{k0}(\bar r,\alpha) = (1/2\pi)\int_0^{2\pi} R_k(\bar r,\theta,0,\alpha)\,d\theta.$$

By means of a transformation

$$\bar\theta = \theta + \epsilon^k v(r,\theta,y,\alpha)$$

where v satisfies the same conditions as u, the term W_k *may be similarly averaged.*

We shall see that even for the simplest bifurcation problems, it is necessary to average not only θ terms, but also y terms as above. The problems associated with $(1/\omega) A_Q$ not being l-simple, however, do not arise in the generic case.

In studying the bifurcation from a periodic orbit to an invariant torus, one arrives at a system of the form (2.4), except that R_j, W_j, and Y_j are also 2π-periodic in t. Consider then

$$\begin{aligned}\dot r &= \epsilon R_1(r,\theta,y,t,\alpha) + \epsilon^2 R_2(r,\theta,y,t,\alpha) + \cdots,\\ \dot\theta &= \omega + \epsilon W_1(r,\theta,y,t,\alpha) + \epsilon^2 W_2(r,\theta,y,t,\alpha) + \cdots,\\ \dot y &= A_Q y + \epsilon Y_1(r,0,y,t,\alpha) + \epsilon^2 Y_2(r,0,y,t,\alpha) + \cdots,\end{aligned} \tag{2.10}$$

where ω, A_Q are constant as before, and all other terms are 2π-periodic in t. To average the term $\epsilon^k R_k(r,\theta,t,\alpha)\, y^l$, consider a transformation

$$\bar r = r + \epsilon^k u_{kl}(r,\theta,t,\alpha)\, y^l$$

of period 2π in both θ and t. The analogs of (2.8), (2.9) in this case are then

$$R_{ko} + \omega\frac{\partial u_{ko}}{\partial\theta} + \frac{\partial u_{ko}}{\partial t} = \text{independent of } (\theta, t), \tag{2.11}$$

$$R_{kl} + \omega\frac{\partial u_{kl}}{\partial\theta} + \frac{\partial u_{kl}}{\partial t} + l u_{kl} A_Q = 0, \qquad l \neq 0. \tag{2.12}$$

By expanding R_{ko} and u_{ko} in Fourier series

$$\begin{aligned} R_{ko}(r, \theta, t, \alpha) &= \sum R_{komn}(r, \alpha)\, e^{i(m\theta+nt)}, \\ u_{ko}(r, \theta, t, \alpha) &= \sum u_{komn}(r, \alpha)\, e^{i(m\theta+nt)}, \end{aligned} \tag{2.13}$$

one sees that (2.11) is equivalent to

$$u_{komn}(r, \alpha) = [i/(m\omega + n)]\, R_{komn}(r, \alpha), \qquad (m, n) \neq (0, 0),$$

which in general can be solved only if $\omega m + n \neq 0$ for all integers $(m, n) \neq (0, 0)$ occurring in the expansion (2.13) for R_{ko}. Very often this expansion is simply a trigonometric polynomial in 0, so this imposes only finitely many resonance conditions on ω. How such conditions arise in the bifurcation to a torus are described in Section 5.

To solve (2.12), again the Fourier expansion shows that one must assume a condition analogous to the l-simplicity of Theorem 2.2, namely, that the induced linear transformation $A_Q{}^l\colon \Lambda_q{}^l \to \Lambda_q{}^l$ have no eigenvalues of the form $i(m\omega + n)/l$ for integers (m, n) appearing in the Fourier expansion of R_{kl}. In practice, this restriction is generally of little consequence, as it appears only if terms of sufficiently high order must be averaged. Results of Section 3 show there is never any restriction when A_Q is a stable matrix.

3. Characterization of l-Simplicity

In this section, we prove the following result.

THEOREM 3.1. *Let M be a $q \times q$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_q$, and $M^l\colon \Lambda_q{}^l \to \Lambda_q{}^l$ the induced linear transformation described in Section 2. Then the eigenvalues of M^l are precisely*

$$\left\{ (1/l) \sum_{i=1}^{l} \lambda_{\alpha_i} \,\middle|\, 1 \leqslant \alpha_1 \leqslant \cdots \leqslant \alpha_l \leqslant q \right\}.$$

Two immediate consequences of this theorem are stated without proof.

COROLLARY 3.2. *The matrix M is l-simple if and only if*

$$\sum_{j=1}^{q} n_j \lambda_j \neq in$$

for all integers $n_j \geqslant 0$ with $\sum n_j = l$ and all integers n.

COROLLARY 3.3. *If all eigenvalues of M lie in the left half-plane, then M is l-simple for each $l \geqslant 1$.*

Proof of Theorem 3.1. It suffices to consider the case where M is diagonalizable, since any M can be approximated by a diagonalizable matrix, and the eigenvalues of M^l vary continuously with M.

Assume then there is a basis $\{e_1, ..., e_q\}$ such that $Me_i = \lambda_i e_i$ for each i. Letting $\mathscr{A}$ be the set of multiindices

$$\mathscr{A} = \{\alpha = (\alpha_1, ..., \alpha_l) \mid 1 \leqslant \alpha_1 \leqslant \cdots \leqslant \alpha_l \leqslant q\}$$

it is clear that any $U \in \Lambda_q{}^l$ is uniquely determined by the values

$$U(e_{\alpha_1}, ..., e_{\alpha_l}), \qquad \alpha \in \mathscr{A}.$$

For any $\beta \in \mathscr{A}$, let $U_\beta \in \Lambda_q{}^l$ be defined by

$$U_\beta(e_{\alpha_1}, ..., e_{\alpha_l}) = \begin{cases} 1 & \text{if } \beta = \alpha, \\ 0 & \text{if } \beta \neq \alpha. \end{cases} \tag{3.1}$$

Clearly, the U_β form a basis for $\Lambda_q{}^l$. Now we compute

$$\begin{aligned} (U_\beta M)(e_{\alpha_1}, ..., e_{\alpha_l}) &= \frac{1}{l} \sum_{i=1}^{l} U_\beta(e_{\alpha_1}, ..., Me_{\alpha_i}, ..., e_{\alpha_l}) \\ &= \left(\frac{1}{l} \sum_{i=1}^{l} \lambda_{\alpha_i}\right) U_\beta(e_{\alpha_1}, ..., e_{\alpha_l}) \\ &= \left(\frac{1}{l} \sum_{i=1}^{l} \lambda_{\beta_i}\right) U_\beta(e_{\alpha_1}, ..., e_{\alpha_l}), \end{aligned}$$

by (3.1). Hence U_β is an eigenvector of M^l, with eigenvalue $(1/l) \sum_{i=0}^{l} \lambda_{\beta_i}$. There are no other eigenvectors of M^l since the U_β form a basis.

Remark. In order to solve Eq. (2.11), we must assume

$$m\omega + n \neq 0 \tag{3.2}$$

for all $(m, n) \neq (0, 0)$ in the Fourier expansion of R_{ko}. In order to solve (2.12), we must assume

$$\sum_{j=1}^{q} n_j \mu_j \neq i(m\omega + n),$$

for all $n_j \geqslant 0, \sum n_j = l$, where $\{\mu_j\}$ are the eigenvalues of A_Q.

Remark. Professor Y. Bibikov has pointed out to us the possibility of dispensing with some of the above restrictions if one is willing to transform the y variable as well as (r, θ), namely, to consider transformations

$$\bar{y} = y + \epsilon^k u(r, \theta, y, \alpha).$$

Such ideas are briefly described by Pyartli [39]. In order to average (2.4), the l-simplicity restrictions are not necessary. Instead, one further decomposes $y = (y_1, y_2)$ into a stable and unstable part, corresponding to eigenvalues with real parts negative and positive. In the equation for $\dot{r}$, the terms involving y are *not* averaged; rather, in the equations for $\dot{y}_j$ one averages those terms which are independent of y_j. After sufficiently many averagings one sees as in (2.5) that in fact $y(t) = O(\epsilon^N)$. Some of the restrictions that arise in the time dependent case (2.10), specifically (3.2), are, however, essential and cannot be eliminated,

4. Hopf Bifurcation for an ODE in R^2

Our study of bifurcation begins with a discussion of the Hopf bifurcation for an autonomous ODE in the plane R^2. Consider such a system depending on a scalar parameter α near zero, such that the origin $x = 0$ is a fixed point for all α. To be specific, consider

$$\dot{x} = f(x, \alpha), \qquad x \in R^2, \qquad \alpha \in (-\alpha_0, \alpha_0);$$
$$f(0, \alpha) \equiv 0.$$

Assume the linearized equation about $x = 0$ is an exponentially stable spiral for $\alpha < 0$, a center when $\alpha = 0$ with eigenvalues $\pm i\omega_0 \neq 0$, and an unstable spiral when $\alpha > 0$. The eigenvalues of this equation are thus $\gamma(\alpha) \pm i\omega(\alpha)$ where $\gamma(0) = 0$, $\alpha\gamma(\alpha) > 0$ for $\alpha \neq 0$, and $\omega(0) = \omega_0$. We stipulate that $\gamma'(0) \neq 0$, and in fact, by using $\gamma(\alpha)$ instead of α as a bifurcation parameter we assume $\gamma(\alpha) \equiv \alpha$. The differential equation then takes the form

$$\dot{x} = A(\alpha)x + F(x, \alpha),$$
$$|F(x, \alpha)| = O(|x|^2),$$

where $A(\alpha)$ has eigenvalues $\alpha \pm i\omega(\alpha)$ with $\omega(0) = \omega_0$. By means of a linear coordinate change $x \to P(\alpha)x$, where $P(\alpha)$ is an appropriate 2×2 matrix, we may assume $A(\alpha)$ is in Jordan form

$$A(\alpha) = \begin{pmatrix} \alpha & -\omega(\alpha) \\ \omega(\alpha) & \alpha \end{pmatrix}. \tag{4.1}$$

Let us write

$$x = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \qquad F(x, \alpha) = \begin{pmatrix} F^1(x^1, x^2, \alpha) \\ F^2(x^1, x^2, \alpha) \end{pmatrix}.$$

Expanding in a Taylor series yields

$$\dot{x}^1 = \alpha x^1 - \omega(\alpha)\, x^2 + \sum_{j=2}^{\infty} B_j^{\,1}(x^1, x^2, \alpha),$$

$$\dot{x}^2 = \omega(\alpha)\, x^1 + \alpha x^2 + \sum_{j=2}^{\infty} B_j^{\,2}(x^1, x^2, \alpha),$$

$$B_j^{\,i} = \text{homogeneous polynomial of order } j \text{ in } (x^1, x^2).$$

As usual it is sufficient that this expansion be only a finite series. Passing to polar coordinates $(x^1, x^2) = (r \cos \theta, r \sin \theta)$ gives

$$\begin{aligned} \dot{r} &= \alpha r + r^2 C_3(\theta, \alpha) + r^3 C_4(\theta, \alpha) + \cdots, \\ \dot{\theta} &= \omega(\alpha) + r D_3(\theta, \alpha) + r^2 D_4(\theta, \alpha) + \cdots, \end{aligned} \tag{4.2}$$

where

$$C_j(\theta, \alpha) = (\cos \theta)\, B_{j-1}^1(\cos \theta, \sin \theta, \alpha) + (\sin \theta)\, B_{j-1}^2(\cos \theta, \sin \theta, \alpha)$$

$$D_j(\theta, \alpha) = (\cos \theta)\, B_{j-1}^2(\cos \theta, \sin \theta, \alpha) - (\sin \theta)\, B_{j-1}^1(\cos \theta, \sin \theta, \alpha).$$

Observe that C_j and D_j are homogeneous polynomials of degree j in $(\cos \theta, \sin \theta)$. We seek, for $\alpha \to 0$, periodic solutions of (4.2) with $r \to 0$. Scale r and α by replacing

$$r \to \epsilon r, \qquad \alpha \to \epsilon \alpha,$$

where the new r is to be considered near a constant $r_0 > 0$, to be determined later; we shall also later specify α as a function of ϵ, but it is not clear how to do this yet. After scaling, (4.2) becomes

$$\begin{aligned} \dot{r} &= \epsilon[\alpha r + r^2 C_3(\theta, \epsilon\alpha)] + \epsilon^2 r^3 C_4(\theta, \epsilon\alpha) + \cdots, \\ \dot{\theta} &= \omega_0 + \epsilon[\alpha \omega'(0) + r D_3(\theta, \epsilon\alpha)] + \cdots. \end{aligned} \tag{4.3}$$

Although this is not quite in the form (2.1) (since C_j and D_j depend on ϵ) it is clear that the averaging procedure still works.

Let us now work through the averaging of (4.3). The generic situation will be completely determined by averaging the ϵ and ϵ^2 terms in $\dot{r}$, and, in fact, there is no need to average the terms in $\dot{\theta}$. Thus the coordinate change

$$\bar{r} = r + \epsilon u_1(r, \theta, \alpha, \epsilon) + \epsilon^2 u_2(r, \theta, \alpha, \epsilon) \tag{4.4}$$

is considered. The argument ϵ appears in u_1 and u_2 since it appears in the coefficients of the expansion in (4.3). Note that the inverse of (4.4) is

$$r = \bar{r} - \epsilon u_1(\bar{r}, \theta, \alpha, \epsilon) + O(\epsilon^2).$$

Substitution into (4.3) yields

$$\begin{aligned}
\dot{\bar{r}} &= \epsilon[\alpha r + r^2 C_3(\theta, \epsilon\alpha)] + \epsilon^2 r^3 C_4(\theta, \epsilon\alpha) \\
&\quad + \epsilon^2 \frac{\partial u_1}{\partial r}(r, \theta, \alpha, \epsilon)[\alpha r + r^2 C_3(\theta, \epsilon\alpha)] \\
&\quad + \epsilon \frac{\partial u_1}{\partial \theta}(r, \theta, \alpha, \epsilon)[\omega_0 + \epsilon\alpha\omega'(0) + \epsilon r D_3(\theta, \epsilon\alpha)] \\
&\quad + \epsilon^2 \frac{\partial u_2}{\partial \theta}(r, \theta, \alpha, \epsilon)\,\omega_0 + O(\epsilon^3) \\
&= \epsilon\left[\alpha r + r^2 C_3(\theta, \epsilon\alpha) + \frac{\partial u_1}{\partial \theta}(r, \theta, \alpha, \epsilon)\,\omega_0\right] \\
&\quad + \epsilon^2\left[r^3 C_4(\theta, \epsilon\alpha) + \frac{\partial u_1}{\partial r}(r, \theta, \alpha, \epsilon)(\alpha r + r^2 C_3(\theta, \epsilon\alpha))\right. \\
&\quad \left. + \frac{\partial u_1}{\partial r}(r, \theta, \alpha, \epsilon)(\alpha\omega'(0) + r D_3(\theta, \epsilon\alpha)) + \frac{\partial u_2}{\partial \theta}(r, \theta, \alpha, \epsilon)\,\omega_0\right] \\
&= \epsilon\left[\alpha\bar{r} + \bar{r}^2 C_3(\theta, \epsilon\alpha) + \frac{\partial u_1}{\partial \theta}(\bar{r}, \theta, \alpha, \epsilon)\,\omega_0\right] \\
&\quad + \epsilon^2\left[\bar{r}^3 C_4(\theta, \epsilon\alpha) + \frac{\partial u_1}{\partial r}(\bar{r}, \theta, \alpha, \epsilon)(\alpha\bar{r} + \bar{r}^2 C_3(\theta, \epsilon\alpha))\right. \\
&\quad + \frac{\partial u_1}{\partial \theta}(\bar{r}, \theta, \alpha, \epsilon)(\alpha\omega'(0) + \bar{r} D_3(\theta, \epsilon\alpha)) \\
&\quad - u_1(\bar{r}, \theta, \alpha, \epsilon)\left(\alpha + 2\bar{r} C_3(\theta, \epsilon\alpha) + \frac{\partial^2 u_1}{\partial r\,\partial \theta}(\bar{r}, \theta, \alpha, \epsilon)\,\omega_0\right) \\
&\quad \left. + \frac{\partial u_2}{\partial \theta}(\bar{r}, \theta, \alpha, \epsilon)\,\omega_0\right] + O(\epsilon^3).
\end{aligned}$$

Following Lemma 2.1, u_1 is given by

$$u_1(r, \theta, \alpha, \epsilon) = -(r^2/\omega_0) \int_0^\theta C_3(s, \epsilon\alpha)\, ds$$

since C_3, being a homogeneous trigonometric polynomial of degree 3, has mean value zero. The coefficient of ϵ is thus $\alpha\bar{r}$. Next, u_2 is chosen as in the lemma; it is not necessary to determine u_2 explicitly since the coefficient of ϵ^2 is the mean value

$$\begin{aligned}
\text{mean}&\left[\bar{r}^3 C_4 + \left(\frac{\partial u_1}{\partial r}\right)(\alpha\bar{r} + \bar{r}^2 C_3) + \left(\frac{\partial u_1}{\partial \theta}\right)(\alpha\omega'(0) + \bar{r} D_3)\right. \\
&\quad \left. - (u_1)\left(\alpha + 2\bar{r} C_3 + \frac{\partial^2 u_1}{\partial r\,\partial \theta}\,\omega_0\right)\right] \\
&= \text{mean}\left[\bar{r}^3 C_4 - \frac{\bar{r}^3}{\omega_0} C_3 D_4\right] \stackrel{\text{def}}{=} \bar{r}^3 K.
\end{aligned}$$

This is summarized as a theorem.

THEOREM 4.1. *Consider the differential equation*

$$\begin{aligned} \dot r &= \epsilon[\alpha r + r^2 C_3(\theta, \epsilon\alpha)] + \epsilon^2 r^3 C_4(\theta, \epsilon\alpha) + O(\epsilon^3), \\ \dot\theta &= \omega_0 + \epsilon[\alpha\omega'(0) + r D_3(\theta, \epsilon\alpha)] + O(\epsilon^2) \end{aligned} \tag{4.5}$$

arising from the Hopf bifurcation problem in R^2 *described above, and the scaling* $r \to \epsilon r$, $\alpha \to \epsilon\alpha$. *Then there exists a coordinate change*

$$\bar r = r + \epsilon u_1(r, \theta, \alpha, \epsilon) + \epsilon^2 u_2(r, \theta, \alpha, \epsilon)$$

transforming (4.5) *into the averaged system of the form*

$$\begin{aligned} \dot{\bar r} &= \epsilon\alpha\bar r + \epsilon^2 \bar r^3 K + O(\epsilon^3), \\ \dot\theta &= \omega_0 + \epsilon[\alpha\omega'(0) + r D_3(\theta, \epsilon\alpha)] + O(\epsilon^2), \end{aligned} \tag{4.6}$$

where K *is the constant*

$$K = (1/2\pi) \int_0^{2\pi} C_4(\theta, 0) - (1/\omega_0)\, C_3(\theta, 0)\, D_3(\theta, 0)\, d\theta. \tag{4.7}$$

The generic case occurs when $K \neq 0$; for definiteness suppose $K < 0$. This suggests the choice $\alpha = \epsilon$, for then (4.6) becomes (dropping the bars)

$$\begin{aligned} \dot r &= \epsilon^2(r + r^3 K) + O(\epsilon^3), \\ \dot\theta &= \omega_0 + O(\epsilon), \end{aligned}$$

so a periodic solution for r near

$$r_0 = (-K)^{-1/2}$$

seems likely. This is, indeed, the case although we must verify that all periodic solutions are obtained in this manner (i.e., none are lost in scaling).

To see this, consider any periodic solution of the unscaled equation (4.2) bifurcating from $r = 0$, $\alpha = 0$. At some point (r_1, θ_1) of the solution $\dot r$ must vanish, and after scaling by $\epsilon = r_1/r_0$, $\dot r$ vanishes at (r_0, θ_1). Thus

$$0 = \epsilon\alpha r_0 + \epsilon^2 r_0^3 K + O(\epsilon^3) = \epsilon r_0(\alpha - \epsilon + O(\epsilon^2))$$

so that $\alpha = \epsilon + O(\epsilon^2)$. Consider an annulus of the form

$$\mathscr{A} \begin{cases} (1-\gamma)\, r_0 \leqslant r \leqslant (1+\gamma)\, r_0, \\ \gamma \to 0 \quad \text{as} \quad \epsilon \to 0. \end{cases} \tag{4.8}$$

We see that for an appropriate $\gamma = \gamma(\epsilon)$, $\mathscr{A}$ must be positively invariant since

$$r = (1+\gamma) r_0 \Rightarrow \dot{r} = \epsilon^2 (1+\gamma) r_0 \left(\frac{\alpha}{\epsilon} - (1+\gamma)^2 + O(\epsilon)\right) < 0,$$

$$r = (1-\gamma) r_0 \Rightarrow \dot{r} = \epsilon^2 (1-\gamma) r_0 \left(\frac{\alpha}{\epsilon} - (1-\gamma)^2 + O(\epsilon)\right) > 0,$$

hence the periodic solution lies entirely in $\mathscr{A}$. This then implies the following theorem.

THEOREM 4.2. *Let the constant K defined in* (4.7) *satisfy $K < 0$, and let $r_0 = (-K)^{-1/2}$. Then all periodic solutions of the original bifurcation problem*

$$\dot{x} = f(x, \alpha), \qquad x \in R^2$$

bifurcating from the origin $r = 0$, $\alpha = 0$ may be obtained by scaling $r \to \epsilon r$, $\alpha \to \epsilon\alpha$ and averaging as above to obtain

$$\dot{\bar{r}} = \epsilon\alpha\bar{r} + \epsilon^2 \bar{r}^3 K + O(\epsilon^3),$$
$$\dot{\theta} = \omega_0 + O(\epsilon),$$

then letting $\alpha = \epsilon$ and considering $\bar{r}$ near r_0. About each such solution, there is a positively invariant annulus $\mathscr{A}$ as in (4.8).

If $K > 0$, the same result holds except now $r_0 = K^{-1/2}$, $\alpha = -\epsilon$, and $\mathscr{A}$ is negatively invariant. Thus in either case, we have in the unscaled variables

$$\alpha = -(\operatorname{sgn} K)\, \epsilon^2,$$
$$r \sim |K|^{-1/2} \epsilon.$$

In the critical case $K = 0$ nothing more can be said until more terms of (4.5) are averaged. In order to study this situation, the following lemma is useful.

LEMMA 4.3. *Consider the system* (4.3), *but expanded in powers of ϵ*

$$\begin{aligned} \dot{r} &= \epsilon[\alpha r + r^2 C_3(\theta)] + \epsilon^2 [r^3 C_4(\theta) + \alpha r^2 C_3^1(\theta)] + \cdots, \\ \dot{\theta} &= \omega_0 + \epsilon[\alpha\omega'(0) + r D_3(\theta)] + \cdots, \end{aligned} \tag{4.9}$$

where $C_j(\theta) = C_j(\theta, 0)$, $C_j^1(\theta) = (\partial/\partial\alpha)\, C_j(\theta, 0)$, etc., and let the coefficients of $\epsilon, \ldots, \epsilon^k$ in $\dot{r}$ be averaged by a series of coordinate changes

$$\bar{r} = r + \epsilon^j u_j(r, \theta, \alpha), \qquad 1 \leqslant j \leqslant k.$$

Then the following properties hold for the averaged system:

(1) *For all m, the coefficient* $R_m(r, \theta, \alpha)$ *of* ϵ^m *in* $\dot{r}$ *is a polynomial in* (r, α) *of the form*

$$R_m(r, \theta, \alpha) = \sum_{j=0}^{m+1} r^{m+1-j}\alpha^j P_{mj}(\theta).$$

(2) *For all m, the coefficient* $W_m(r, \theta, \alpha)$ *of* ϵ^m *in* $\dot{\theta}$ *is a polynomial in* (r, α) *of the form*

$$W_m(r, \theta, \alpha) = \sum_{j=0}^{m} r^{m-j}\alpha^j V_{mj}(\theta).$$

(3) *The terms* $P_{mj}(\theta)$ *and* $V_{mj}(\theta)$ *are polynomials in* $(\cos\theta, \sin\theta)$ *and satisfy*

$$P_{mj}(\theta + \pi) = (-1)^{m+j} P_{mj}(\theta),$$
$$V_{mj}(\theta + \pi) = (-1)^{m+j} V_{mj}(\theta).$$

That is, only the terms $\cos^p\theta \sin^q\theta$ *appear, where* $p + q = m + j \pmod 2$.

Proof. Observe that 1 and 2 simply say that under the reverse scaling $r \to (1/\epsilon)r$, $\alpha \to (1/\epsilon)\alpha$, that ϵ is absent from the differential equation. Condition 3 says that under the substitution $\alpha \to -\alpha$, $\epsilon \to -\epsilon$, $\theta \to \theta + \pi$, the equation remains unchanged. It is thus clear that the unaveraged equation (4.9) (see also (4.2), (4.3)) satisfies these three properties. We induct on k, so assume the $\epsilon, \ldots, \epsilon^{k-1}$ terms have been averaged, and 1, 2, and 3 hold for all terms. Clearly, then, upon averaging the ϵ^k term, u_k must have the same form as the coefficient R_k in the sense that

$$\begin{aligned} u_k(r, \theta, \alpha) &= \sum_{j=0}^{k+1} r^{k+1-j}\alpha^j v_{kj}(\theta), \\ v_{kj}(\theta + \pi) &= (-1)^{k+j}\, v_{kj}(\theta), \\ v_{kj}(\theta) &= \text{polynomial in } (\cos\theta, \sin\theta). \end{aligned} \tag{4.10}$$

This is because Lemma 2.1 shows that u_k is obtained basically by integrating R_k with respect to θ; in particular, each $v_{kj}(\theta)$ (like $P_{kj}(\theta)$) involves only terms $\cos^p\theta \sin^q\theta$ with $p + q = k + j \pmod 2$. The coordinate change

$$\bar{r} = r + \epsilon^k u_k(r, \theta, \alpha) = r + \epsilon^k \sum_{j=0}^{k+1} r^{k+1-j}\alpha^j v_{kj}(\theta) \tag{4.11}$$

has the property that under the reverse scaling, ϵ is absent; this property thus carries over when the ϵ^k term is averaged, that is, the averaged equation satisfies 1 and 2. We also see that from (4.10), the transformation (4.11) remains unchanged under the substitution $\alpha \to -\alpha$, $\epsilon \to -\epsilon$, $0 \to \theta + \pi$; thus 3 also holds. This then proves the lemma.

Lemma 4.3 implies that when (4.9) has been averaged, it must have the form (for r bounded)

$$\begin{aligned}\dot{r} &= \epsilon\alpha r + \epsilon^2[r^3K_2 + O(\alpha)] + O(\epsilon^3\alpha) \\ &\quad + \epsilon^4[r^5K_4 + O(\alpha)] + \cdots + \epsilon^{2p}[r^{2p+1}K_{2p} + O(\alpha)] \\ &\quad + O(\epsilon^{2p+1}), \\ \dot{\theta} &= \omega_0 + O(\epsilon),\end{aligned}$$

where K_2, K_4,..., K_{2p} are computable constants and $k = 2p$. The generic case $K_2 \neq 0$ was analyzed above, so here assume

$$K_2 = \cdots = K_{2p-2} = 0, \qquad K \overset{\text{def}}{=} K_{2p} \neq 0.$$

Thus we are considering

$$\begin{aligned}\dot{r} &= \epsilon\alpha r + O(\epsilon^2\alpha) + \epsilon^{2p}r^{2p+1}K + O(\epsilon^{2p+1}), \\ \dot{\theta} &= \omega_0 + O(\epsilon).\end{aligned} \tag{4.12}$$

For the same reasons as in the generic case, the choice

$$\alpha = \begin{cases} \epsilon^{2p-1}, & K < 0, \\ -\epsilon^{2p-1}, & K > 0, \end{cases} \tag{4.13}$$

is made, and we work near $r_0 = |K|^{-1/2p}$. This leads immediately to the natural generalization of Theorem 4.2.

THEOREM 4.4. *Let the averaging procedure be performed on the terms in* $\dot{r}$ in (4.3) *until* (4.12) *is obtained for some* $p \geqslant 1$, $K \neq 0$. *Then the conclusions of Theorem* 4.2 *hold with the following changes*:

(1) *Choose*

$$\begin{aligned}\alpha &= -(\operatorname{sgn} K)\,\epsilon^{2p-1} \qquad (\textit{scaled}), \\ r_0 &= |K|^{-1/2p}.\end{aligned}$$

(2) *The annulus* $\mathscr{A}$ *is positively invariant if* $K < 0$ *and negatively invariant if* $K > 0$, *as before.*

(3) *In the unscaled variables,*

$$\begin{aligned}\alpha &= -(\operatorname{sgn} K)\,\epsilon^{2p}, \\ r &\sim |K|^{-1/2p}\epsilon.\end{aligned}$$

We omit any further justification of this theorem.

Of course Theorems 4.2 and 4.4 do not establish the existence of periodic solutions, but merely estimate the region where they may be found. In Section 6, existence of such solutions will be proved.

Let us close this section by briefly examining van der Pol's equation

$$\ddot{x} - \epsilon(1 - x^2)\dot{x} + x = 0 \tag{4.14}$$

from the point of view of averaging; we wish to compare this with the equations obtained in the Hopf bifurcation. It is known that (4.14) has a unique periodic solution for all ϵ, stable when $\epsilon > 0$, and with amplitude near 2 for small ϵ. Only ϵ near zero is considered here. In polar coordinates $(x, \dot{x}) = (r \cos \theta, r \sin \theta)$, (4.14) becomes

$$\begin{aligned} \dot{r} &= r(1 - r^2 \cos^2 \theta) \sin^2 \theta, \\ \dot{\theta} &= -1 + \epsilon(1 - r^2 \cos^2 \theta) \cos \theta \sin \theta. \end{aligned}$$

Upon averaging via a transformation

$$\bar{r} = r + \epsilon u(r, \theta),$$

we obtain

$$\begin{aligned} \dot{\bar{r}} &= \epsilon((\bar{r}/2) - (\bar{r}^3/8)) + O(\epsilon^2), \\ \dot{\theta} &= -1 + O(\epsilon), \end{aligned} \tag{4.15}$$

because of the computation of the mean value

$$(1/2\pi) \int_0^{2\pi} r(1 - r^2 \cos^2 \theta) \sin^2 \theta \, d\theta = (r/2) - (r^3/8).$$

This suggests the existence of a periodic solution, for small ϵ, near $r = 2$, the unique positive root of $(r/2) - (r^3/8) = 0$. Iy is instructive to compare (4.15) with the normal forms listed below, obtained for the Hopf bifurcation by Theorems 4.2 and 4.4.

$$\begin{aligned} \dot{r} &= \epsilon^2(r + r^3K) + O(\epsilon^3) && \text{(generic case, } K < 0\text{)}, \\ \dot{r} &= \epsilon^2(-r + r^3K) + O(\epsilon^3) && \text{(generic case, } K > 0\text{)}, \\ \dot{r} &= \epsilon^{2p}(r + r^{2p+1}K) + O(\epsilon^{2p+1}) && (p \geqslant 2, K < 0), \\ \dot{r} &= \epsilon^{2p}(-r + r^{2p+1}K) + O(\epsilon^{2p+1}) && (p \geqslant 2, K > 0), \\ \dot{\theta} &= \omega_0 + O(\epsilon) && \text{(all cases)}. \end{aligned}$$

5. Hopf Bifurcation in Higher Dimensions and More General Bifurcations

Our object in this section is first to carry over the results in Section 4 to ODE's in R_p $(n \geqslant 3)$ and then study more general systems such as bifurcation to an invariant torus. This will pave the way to considering infinite-dimensional evolution systems (FDE's and PDE's) in Section 7.

Consider an ODE in coordinates $(x, y) \in R^2 \times R^{n-2} = R^n$, depending on the parameter α, such that the origin $(x, y) = (0, 0)$ is a fixed point for all α, with linearized equation

$$\begin{aligned} \dot{x} &= A_P(\alpha)\, x + \alpha E(\alpha)\, y, \\ \dot{y} &= \alpha H(\alpha)\, x + [A_Q + \alpha M(\alpha)]\, y, \end{aligned} \tag{5.1}$$

with $A_P(\alpha)$ the matrix in (4.1), and where A_Q has no pure imaginary eigenvalues. When $\alpha = 0$ these linear equations decouple, and it is seen we are considering the appropriate generalization of the situation in Section 4. Expand the nonlinear equation as follows:

$$\begin{aligned} \dot{x} &= B_0(y, \alpha) + B_1(y, \alpha)\, x + B_2(y, \alpha)\, x^2 + \cdots, \\ \dot{y} &= \Gamma_0(x, \alpha) + \Gamma_1(x, \alpha)\, y + \Gamma_2(x, \alpha)\, y^2 + \cdots. \end{aligned} \tag{5.2}$$

Then the origin is a fixed point of (5.2) with variational equation (5.1) if and only if

$$\begin{aligned} B_0(0, \alpha) &= 0, & \Gamma_0(0, \alpha) &= 0, \\ (\partial B_0/\partial y)(0, \alpha) &= \alpha E(\alpha), & (\partial \Gamma_0/\partial x)(0, \alpha) &= \alpha H(\alpha), \\ B_1(0, \alpha) &= A_P(\alpha), & \Gamma_1(0, \alpha) &= A_Q + \alpha M(\alpha), \end{aligned}$$

and this implies B_0, B_1, Γ_0, Γ_1 have the form

$$\begin{aligned} B_0(y, \alpha) &= \alpha E(\alpha) y + F(y, \alpha) y^2, \\ B_1(y, \alpha) &= A_P(\alpha) + G(y, \alpha) y, \\ \Gamma_0(x, \alpha) &= \alpha H(\alpha) x + J(x, \alpha)\, x^2, \\ \Gamma_1(x, \alpha) &= A_Q + \alpha M(\alpha) + N(x, \alpha) x, \end{aligned}$$

for some functions F, G, J, N. The differential equation (5.2) then takes the form

$$\begin{aligned} \dot{x} &= \alpha E(\alpha) y + F(y, \alpha)\, y^2 + A_P(\alpha) x + G(y, \alpha)\, xy \\ &\quad + B_2(y, \alpha)\, x^2 + B_3(y, \alpha)\, x^3 + \cdots, \\ \dot{y} &= \alpha H(\alpha) x + J(x, \alpha)\, x^2 + A_Q y + \alpha M(\alpha) y + N(x, \alpha)\, xy \\ &\quad + \Gamma_2(x, \alpha)\, y^2 + \Gamma_3(x, \alpha)\, y^3 + \cdots, \end{aligned}$$

which in polar coordinates $x = (r \cos\theta, r \sin\theta)$ becomes

$$\begin{aligned} \dot{r} &= [\alpha E_1(\theta, \alpha)\, y + F_1(\theta, y, \alpha)\, y^2] + r[\alpha + G_2(\theta, y, \alpha)\, y] \\ &\quad + r^2 C_3(\theta, y, \alpha) + r^3 C_4(\theta, y, \alpha) + \cdots, \\ \dot{\theta} &= \frac{1}{r}\, [\alpha E_1^*(\theta, \alpha)\, y + F_1^*(\theta, y, \alpha)\, y^2] + [\omega(\alpha) + G_2^*(\theta, y, \alpha)\, y] \\ &\quad + r D_3(\theta, y, \alpha) + r^2 D_4(\theta, y, \alpha) + \cdots, \\ \dot{y} &= \text{as above but with } x = (r \cos\theta, r \sin\theta). \end{aligned} \tag{5.3}$$

The notation is such that E_1, E_1^*, F_1, etc., are computed from $E, F, \ldots$, just as C_j and D_j are computed from B_{j-1} as in Section 4. Moreover, the subscript j in (5.3) (such as on E_1, E_1^*, F_1, C_3, etc.) means the indicated function is homogeneous of degree j in $(\cos\theta, \sin\theta)$. Scale (5.3) by

$$r \to \epsilon r, \qquad y \to \epsilon y, \qquad \alpha \to \epsilon\alpha$$

to get

$$\begin{aligned}
\dot{r} &= \epsilon[\alpha r + r^2 C_3(\theta, \epsilon y, \epsilon\alpha) + \alpha E_1(\theta, \epsilon\alpha)\, y \\
&\quad + F_1(\theta, \epsilon y, \epsilon\alpha)\, y^2 + r G_2(\theta, \epsilon y, \epsilon\alpha)\, y] + \epsilon^2 r^3 C_4(\theta, \epsilon y, \epsilon\alpha) + O(\epsilon^3), \\
\dot{\theta} &= \omega_0 + \epsilon \Big[\alpha\omega'(0) + r D_3(\theta, \epsilon y, \epsilon\alpha) + \frac{\alpha}{r} E_1^*(\theta, \epsilon\alpha)\, y \\
&\quad + \frac{1}{r} F_1^*(\theta, \epsilon y, \epsilon\alpha)\, y^2 + G_2^*(\theta, \epsilon y, \epsilon\alpha)\, y\Big] + O(\epsilon^2), \qquad (5.4) \\
\dot{y} &= A_Q y + \epsilon[\alpha H(\epsilon\alpha)\, x + J(\epsilon x, \epsilon\alpha)\, x^2 + \alpha M(\epsilon\alpha)\, y \\
&\quad + N(\epsilon x, \epsilon\alpha)\, xy + \Gamma_2(\epsilon x, \epsilon\alpha)\, y^2] + O(\epsilon^2),
\end{aligned}$$

and we are ready to average. The generic case ought to be determined by averaging the ϵ, ϵy, and ϵ^2 terms in (5.4), as we anticipate $y = O(\epsilon)$. Thus the associated transformation has the form

$$\bar{r} = r + \epsilon u_1(r, \theta, \alpha, \epsilon) + \epsilon w(r, \theta, \alpha, \epsilon)\, y + \epsilon^2 u_2(r, \theta, \alpha, \epsilon)$$

with the inverse satisfying

$$r = \bar{r} - \epsilon u_1(\bar{r}, \theta, \alpha, \epsilon) + O(\epsilon\, |\, y\, |) + O(\epsilon^2).$$

Substituting into (5.4) yields

$$\begin{aligned}
\dot{\bar{r}} &= \epsilon[\alpha r + r^2 C_3] + \epsilon[\alpha E_1 + r G_2]\, y + \epsilon^2 r^3 C_4 \\
&\quad + \epsilon^2(\partial u_1/\partial r)(r, \theta, \alpha, \epsilon)[\alpha r + r^2 C_3] \\
&\quad + \epsilon(\partial u_1/\partial\theta)(r, \theta, \alpha, \epsilon)[\omega_0 + \epsilon\alpha\omega'(0) + \epsilon r D_3] \\
&\quad + \epsilon(\partial w/\partial\theta)(r, \theta, \alpha, \epsilon)\, \omega_0 y + \epsilon w(r, \theta, \alpha, \epsilon)\, A_Q y \\
&\quad + \epsilon^2 w(r, \theta, \alpha, \epsilon)[\alpha H x + J x^2] \\
&\quad + \epsilon^2(\partial u_2/\partial\theta)(r, \theta, \alpha, \epsilon)\, \omega_0 + O(\epsilon\, |\, y\, |^2) + O(\epsilon^2\, |\, y\, |) + O(\epsilon^3) \\
&= \epsilon[\alpha\bar{r} + \bar{r}^2 C_3 + (\partial u_1/\partial\theta)(\bar{r}, \theta, \alpha, \epsilon)\, \omega_0] \\
&\quad + \epsilon[\alpha E_1 + \bar{r} G_2 + (\partial w/\partial\theta)(\bar{r}, \theta, \alpha, \epsilon)\, \omega_0 + w(\bar{r}, \theta, \alpha, \epsilon)\, A_Q]\, y \\
&\quad + \epsilon^2[\bar{r}^3 C_4 + (\partial u_1/\partial r)(\bar{r}, \theta, \alpha, \epsilon)(\alpha\bar{r} + \bar{r}^2 C_3) \\
&\quad + (\partial u_1/\partial\theta)(\bar{r}, \theta, \alpha, \epsilon)(\alpha\omega'(0) + \bar{r} D_3) \\
&\quad + w(\bar{r}, \theta, \alpha, \epsilon)(\alpha H\bar{x} + J\bar{x}^2) \\
&\quad - u_1(\bar{r}, \theta, \alpha, \epsilon)(\alpha + 2\bar{r} C_3 + (\partial^2 u_1/\partial r\, \partial\theta)(\bar{r}, \theta, \alpha, \epsilon)\, \omega_0) \\
&\quad + (\partial u_2/\partial\theta)(\bar{r}, \theta, \alpha, \epsilon)\, \omega_0] + O(\epsilon\, |\, y\, |^2) + O(\epsilon^2\, |\, y\, |) + O(\epsilon^3),
\end{aligned}$$

where C_3, E_1,... are evaluated at $(\theta, 0, \epsilon\alpha)$ and $\bar{x} = (\bar{r}\cos\theta, \bar{r}\sin\theta)$. As before, the coefficient of ϵ (when $y = 0$) is averaged by letting

$$u_1(r, \theta, \alpha, \epsilon) = (-r^2/\omega_0) \int_0^\theta C_3(s, 0, \epsilon\alpha)\, ds;$$

this coefficient then becomes the mean value $\alpha\bar{r}$. To average the coefficient of ϵy, neglect the lower order term αE_1 and let $w(r, \theta, \alpha, \epsilon)$ be the unique 2π-periodic row vector solution of

$$rG_2(\theta, 0, \epsilon\alpha) + (\partial w/\partial\theta)(r, \theta, \alpha, \epsilon)\,\omega_0 + w(r, \theta, \alpha, \epsilon)\, A_Q = 0. \tag{5.5}$$

Finally, average the coefficient of ϵ^2 (with $y = 0$) by choosing u_2 so as to obtain the mean value

$$\begin{aligned}\text{mean}&\left[\bar{r}^3 C_4 + \left(\frac{\partial u_1}{\partial r}\right)(\alpha\bar{r} + \bar{r}^3 C_3) + \left(\frac{\partial u_1}{\partial\theta}\right)(\alpha\omega'(0) + \bar{r}D_3) + w(\alpha H\bar{x} + J\bar{x}^2)\right]\\ &= \bar{r}^3 \,\text{mean}\left[C_4 - \frac{1}{\omega_0} C_3 D_3\right] + \text{mean}[wJ\bar{x}^2] + O(\alpha).\end{aligned}$$

Observe this is *not* the same quantity obtained in Section 4, due to the additional term mean$(wJ\bar{x}^2)$. Thus even in the generic case, the terms involving y cannot be ignored. It is clear from (5.5) that

$$w(\bar{r}, \theta, \alpha, \epsilon) = \bar{r}w^*(\theta, \alpha, \epsilon)$$

for some w^*, and so, since $\bar{x} = (\bar{r}\cos\theta, \bar{r}\sin\theta)$, we have

$$\text{mean}(wJ\bar{x}^2) = \bar{r}^3 K^{**}$$

for some constant K^{**}. This gives the analog of Theorem 4.1.

THEOREM 5.1. *Consider the differential equation*

$$\begin{aligned}\dot{r} &= \epsilon[\alpha r + r^2 C_3(\theta, \epsilon y, \epsilon\alpha) + \alpha E_1(\theta, \epsilon\alpha)\, y\\ &\quad + F_1(\theta, \epsilon y, \epsilon\alpha)\, y^2 + rG_2(\theta, \epsilon y, \epsilon\alpha)\, y] + \epsilon^2 r^3 C_4(\theta, \epsilon y, \epsilon\alpha) + O(\epsilon^3),\\ \dot{\theta} &= \omega_0 + \epsilon\left[\alpha\omega'(0) + rD_3(\theta, \epsilon y, \epsilon\alpha) + \frac{\alpha}{r} E_1^*(\theta, \epsilon\alpha)\, y\right.\\ &\quad \left. + \frac{1}{r} F_1^*(\theta, \epsilon y, \epsilon\alpha)\, y^2 + G_2^*(\theta, \epsilon y, \epsilon\alpha)\, y\right] + O(\epsilon^2),\\ \dot{y} &= A_Q y + \epsilon[\alpha H(\epsilon\alpha)\, x + J(\epsilon x, \epsilon\alpha)\, x^2 + \alpha M(\epsilon\alpha)\, y\\ &\quad + N(\epsilon x, \epsilon\alpha)\, xy + \Gamma_2(\epsilon x, \epsilon\alpha)\, y^2] + O(\epsilon^2),\end{aligned} \tag{5.6}$$

arising from the Hopf bifurcation problem in R^n described above in Section 5, *and the scaling $r \to \epsilon r$, $y \to \epsilon y$, $\alpha \to \epsilon\alpha$. Then there exists a coordinate change*

$$\bar{r} = r + \epsilon u_1(r, \theta, \alpha, \epsilon) + \epsilon w(r, \theta, \alpha, \epsilon) y + \epsilon^2 u_2(r, \theta, \alpha, \epsilon)$$

transforming (5.6) *into the averaged system of the form*

$$\begin{aligned} \dot{\bar{r}} &= \epsilon\alpha\bar{r} + \epsilon^2\bar{r}^3 K + O(\epsilon \mid y \mid^2) + O(\epsilon^2 \mid y \mid) + O(\epsilon^3), \\ \dot{\theta} &= \omega_0 + O(\epsilon), \\ \dot{y} &= A_Q y + O(\epsilon), \end{aligned}$$

where K is the constant

$$\begin{aligned} K &= K^* + K^{**}, \\ K^* &= \frac{1}{2\pi}\int_0^{2\pi} C_4(\theta, 0, 0) - \frac{1}{\omega_0} C_3(\theta, 0, 0)\, D_3(\theta, 0, 0)\, d\theta, \\ K^{**} &= \frac{1}{2\pi}\int_0^{2\pi} w^*(\theta)\, J(0, 0)(\cos\theta, \sin\theta)^2\, d\theta, \end{aligned}$$

where $w^*(\theta)$ is the unique 2π-periodic solution of

$$G_2(\theta, 0, 0) + w^{*\prime}(\theta)\,\omega_0 + w^*(\theta)\, A_Q = 0. \tag{5.7}$$

We recall that for each (x, α), $J(x, \alpha)$ is a bilinear form in the x-space R^2, taking values in the y-space; in the theorem $J(0, 0)$ acts on the point $(\cos\theta, \sin\theta) \in R^2$. Also note that the easiest way of solving (5.7) and computing K^{**} may be to expand G_2 and w^* in Fourier series; see for example, Wright's equation in Section 9. Observe the following interesting fact: the property $K \neq 0$ depends only on the differential equation *at* $\alpha = 0$ and not on the particular parameterization passing through this equation, since the formulas for K^* and K^{**} do not involve derivatives of terms with respect to α. We are assuming $\operatorname{Re}\lambda(\alpha) = \alpha$ for the eigenvalues $\lambda(\alpha)$ of $A_P(\alpha)$, but if more generally, we have $\operatorname{Re}\lambda'(0) = \nu \neq 0$, then a generic bifurcation will still occur if $K \neq 0$. In this case the averaged equation for $\dot{r}$ would be

$$\begin{aligned} \dot{r} &= \operatorname{Re}(\lambda(\epsilon\alpha))r + \epsilon^2 r^3 K + O(\epsilon^3) \\ &= \epsilon\alpha\nu r + \epsilon^2 r^3 K + O(\epsilon^3) \\ &= \epsilon^2(\pm\nu r + r^3 K) + O(\epsilon^3), \qquad \pm = -\operatorname{sgn}(\nu K), \end{aligned}$$

where $\alpha = -\operatorname{sgn}(\nu K)\epsilon$. Thus, if $K \neq 0$ for the equation at $\alpha = 0$, then a generic bifurcation should occur if $\operatorname{Re}\lambda'(0) \neq 0$. We recall that the assumption here that $\operatorname{Re}\lambda(\alpha) = \alpha$ is simply for convenience. *As long* $\operatorname{Re}\lambda'(0) \neq 0$, *all of the averaging techniques described here apply.*

Higher order terms of (5.6) can be averaged provided A_Q is l-simple for appropriate l. In the following lemma, the terms $\epsilon, \epsilon^2, \ldots, \epsilon^{2p}$ are averaged. The proof involves essentially the same arguments as in Section 4, especially Lemma 4.3, so it is omitted.

LEMMA 5.2. *Let A_Q be l-simple for all $1 \leqslant l \leqslant 2p-1$. Then the coefficients of $\epsilon^i y^j$, $1 \leqslant i \leqslant 2p$, $0 \leqslant j \leqslant 2p-i$ in the expression for $\dot{r}$ in* (5.6) *may be averaged with a transformation of the form*

$$\bar{r} = r + \sum_{i=1}^{2p} \epsilon^i u_i(r, \theta, y, \alpha, \epsilon),$$

$$u_i = \text{polynomial in } y \text{ of degree } (2p-i).$$

The resulting averaged equations have the form

$$\begin{aligned}
\dot{\bar{r}} &= \epsilon\alpha r + O(\epsilon^2\alpha) + \epsilon^2\bar{r}^3 K_2 + \epsilon^4\bar{r}^5 K_4 + \cdots + \epsilon^{2p}\bar{r}^{2p+1}K_{2p} \\
&\quad + \sum_{i=1}^{2p} O(\epsilon^i \,|\, y\,|^{2p-i+1}) + O(\epsilon^{2p+1}) \\
\dot{\theta} &= \omega_0 + O(\epsilon), \\
\dot{y} &= A_Q y + O(\epsilon),
\end{aligned}$$

where $K_2, K_4, \ldots, K_{2p}$ are computable constants.

In the nongeneric case $K_2 = 0$; if $K_2 = \cdots = K_{2p-2} = 0$, $K \overset{\text{def}}{=} K_{2p} \neq 0$ then we are considering (dropping the bars)

$$\begin{aligned}
\dot{r} &= \epsilon\alpha r + O(\epsilon^2\alpha) + \epsilon^{2p} r^{2p+1} K + \sum_{i=1}^{2p} O(\epsilon^i \,|\, y\,|^{2p-i+1}) + O(\epsilon^{2p+1}), \\
\dot{\theta} &= \omega_0 + O(\epsilon), \\
\dot{y} &= A_Q y + \epsilon Y(r, \theta, y, \alpha, \epsilon),
\end{aligned} \tag{5.8}$$

for some function Y.

If $(r(t), \theta(t), y(t))$ is any solution bounded on $(-\infty, \infty)$ (e.g., a periodic solution) then the equation for y can be written in integrated form

$$y(t) = \epsilon \int_{-\infty}^{t} e^{A_Q(t-s)}\, Y(r(s), \theta(s), y(s), \alpha, \epsilon)\, ds \tag{5.9}$$

if A_Q is a stable matrix. (If A_Q is hyperbolic then two integrals are needed, one from $-\infty$ to t, the other from t to $+\infty$.) Temporarily choose ϵ as the supremum of $|r(t)| + |y(t)|$ (unscaled), so that upon scaling, $|r(t)| + |y(t)|$ has 1 as its supremum. But from (5.9), we have for appropriate constants $q > 0$, $\Omega > 0$

$$|y(t)| \leqslant |\epsilon| \int_{-\infty}^{t} e^{-q(t-s)}\, ds \sup_{|r|+|y|\leqslant 1} |Y(r, \theta, y, \alpha, \epsilon)| \leqslant \Omega |\epsilon|,$$

so that $|r(t)| \geqslant 1 - 2\Omega|\epsilon|$ for some t. Thus as long as we scale so that r remains in a bounded region, as $\epsilon \to 0$, we have (uniformly) $y = O(\epsilon)$. In particular, in (5.8), $\sum_{i=1}^{2p} O(\epsilon^i |y|^{2p-i+1}) = O(\epsilon^{2p+1})$, and this justifies the choice of $\alpha = -(\operatorname{sgn} K)\,\epsilon^{2p-1}$. Finally, the annulus $\mathscr{A}^*$ surrounding a periodic solution is given by

$$\mathscr{A}^* \begin{cases} (1-\gamma)\, r_0 \leqslant r \leqslant (1+\gamma)\, r_0, & \gamma \to 0 \text{ as } \epsilon \to 0, \\ |y| \leqslant \Omega |\epsilon|, & \Omega = \text{constant}. \end{cases} \tag{5.10}$$

Observe that $\mathscr{A}^*$ may *not* be invariant as before; if $K < 0$ but A_Q is not stable, for example, then solutions will enter $\mathscr{A}^*$ along the boundary $r = (1 \pm \gamma)\, r_0$, but may leave along $|y| = \Omega|\epsilon|$. We obtain then the following theorem.

THEOREM 5.3. *Let* (5.6) *be averaged as in Lemma* 5.2, *and assume the form* (5.8) *for some* $K \neq 0$. *Then all periodic solutions of the unscaled equation* (5.3) *bifurcating from* $r = 0$, $y = 0$, $\alpha = 0$ *may be obtained by letting* $\alpha = -(\operatorname{sgn} K)\,\epsilon^{2p-1}$ *in* (5.8) *to give*

$$\begin{aligned} \dot r &= \epsilon^{2p}(\pm r + r^{2p+1}K) + O(\epsilon^{2p+1}), \qquad \pm = -\operatorname{sgn} K, \\ \dot\theta &= \omega_0 + O(\epsilon), \\ \dot y &= A_Q y + O(\epsilon), \end{aligned} \tag{5.11}$$

and by considering r *near* $r_0 = |K|^{-1/2p}$ *and* y *near zero. The annulus* $\mathscr{A}^*$ *in* (5.10) *is positively invariant if* $K < 0$ *and* A_Q *is stable.*

Let us now briefly describe the phenomenon of bifurcation of an invariant torus from a periodic orbit. Suppose in R^n an autonomous differential equation $\dot x = f(x, \alpha)$ has for $\alpha = 0$ a nonconstant periodic solution $p(t)$, which is nondegenerate, that is, the characteristic multiplier $\mu = 1$ is simple. It is well known that for $|\alpha|$ small there is a unique periodic solution $p(t, \alpha)$, smooth in (t, α), with $p(t, 0) = p(t)$. We may assume also by rescaling the time that $p(t, \alpha)$ has period 2π. In an appropriate coordinate system around the periodic orbit, the autonomous equation may be rewritten as a nonautonomous equation

$$\dot x_1 = f_1(t, x_1, \alpha), \tag{5.12}$$

where $x_1 \in R^{n+1}$, f_1 is 2π-periodic in t, and $f_1(t, 0, \alpha) \equiv 0$. The solution $x_1 = 0$

corresponds in this coordinate system to the periodic orbit $p(t, \alpha)$. We may write (5.12) as

$$\begin{aligned} \dot{x}_1 &= A(\alpha)\, x_1 + F(x_1, t, \alpha), \\ |F(x_1, t, \alpha)| &= O(|x_1|^2). \end{aligned} \tag{5.13}$$

The linearized equation $\dot{x}_1 = A(\alpha)\, x_1$ here has been made autonomous by a linear transformation, using Floquet theory. Assume the same conditions as before on the eigenvalues of $A(\alpha)$. In particular, the eigenvalues $\pm i\omega_0$ at $\alpha = 0$ correspond to characteristic multipliers $\mu = e^{\pm 2\pi i\omega_0}$ of the original periodic orbit. The averaging of (5.13) then proceeds exactly as for the Hopf bifurcation, but with the appropriate modifications described in Section 2. In particular, in order to average the terms $\epsilon r^2 C_3(\theta, t)$ and $\epsilon^2 r^3 C_4(0, t)$ (as in (5.6)) it is necessary to assume $m\omega_0 + n \neq 0$ for $(m, n) \neq (0, 0)$, and $|m| \leqslant 4$. That is, one must assume the critical characteristic multipliers $\mu = e^{\pm 2\pi i\omega_0}$ satisfy

$$\mu^N \neq 1, \qquad N = 1, 2, 3, 4.$$

Under these conditions, averaging and scaling gives rise to a normal form as in (5.11), where $p = 1$, K is constant, and the higher order terms are periodic in t. If $K \neq 0$, one may expect a two-dimensional invariant manifold near the torus

$$r = |K|^{-1/2}, \qquad y = 0,\ \theta,\ t = \text{arbitrary},$$

where (θ, t) are the coorinates on the torus. This is indeed the case and follows from standard results on invariant manifolds.

6. Existence of the Bifurcating Solutions

We have yet to actually prove that the system (5.11) obtained by scaling and averaging (5.2), (5.3) has a periodic solution bifurcating from the fixed point $r = 0$, $y = 0$; this will now be shown. Briefly, the system (5.11) possesses an invariant manifold Σ, the center manifold, given by $y = y(r, \theta, \epsilon)$. It is defined near $r = 0$, $\epsilon = 0$, and passes through the origin for each ϵ, so that $y(0, \theta, \epsilon) = 0$. All orbits which stay near the origin for all $t \in (-\infty, \infty)$ lie on the center manifold; in particular, all periodic solutions lie on Σ. By substituting $y = y(r, \theta, \epsilon)$ into the differential equation (5.11) the search for periodic solutions has been reduced to a two-dimensional problem, as the equations now involve only (r, θ).

Strictly speaking, it is not necessary to use the center manifold, as it is not difficult to prove the existence of periodic solutions of (5.11) directly. Its advantage lies in the fact that many infinite-dimensional systems (such as functional and partial differential equations) have center manifolds. Obtaining periodic solutions of such systems directly may be very difficult; by looking only on the center manifold, however, the problem is reduced to a two-dimensional ODE.

To obtain the center manifold, augment the original system (5.2) by considering the parameter α as a state variable satisfying $\dot{\alpha} = 0$. The linearized equation about the origin $(x, y, \alpha) = (0, 0, 0)$ is then

$$\begin{aligned} \dot{x} &= A_P(0)\, x, \qquad A_P(0) = \begin{pmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{pmatrix}, \\ \dot{y} &= A_Q y, \\ \dot{\alpha} &= 0. \end{aligned}$$

Since no eigenvalues of A_Q are purely imaginary, the center manifold theorem guarantees the existence of a smooth invariant three-dimensional manifold Σ, passing through $(x, y, \alpha) = (0, 0, 0)$ and tangent to the (x, α) space. In polar coordinates $x = (r \cos \theta,\ r \sin \theta)$ then Σ has the form

$$\Sigma\colon y = y^*(r, \theta, \alpha),$$

$$0 = y^*(0, \theta, 0) = \frac{\partial y^*}{\partial r}(0, \theta, 0) = \frac{\partial y^*}{\partial \alpha}(0, \theta, 0),$$

for some function y^*. Since the fixed point $r = 0$, $y = 0$ must lie on Σ for all α we also have $y^*(0, \theta, \alpha) = 0$, so that

$$\begin{aligned} y^*(r, \theta, \alpha) &= r z^*(r, \theta, \alpha), \\ z^*(0, \theta, 0) &= 0 \end{aligned}$$

for some smooth z^*. After scaling $r \to \epsilon r$, $y \to \epsilon y$, $\alpha \to \epsilon\alpha$ and averaging by $r \to r + O(\epsilon)$, and setting $\alpha = \pm\epsilon^{2p-1}$, the equation for Σ takes the form

$$y = r z^*(\epsilon r + O(\epsilon^2), \theta, \pm\epsilon^{2p}) + O(\epsilon) \stackrel{\text{def}}{=} y(r, \theta, \epsilon).$$

Note that $y(r, \theta, \epsilon) = O(\epsilon)$ uniformly, from (6.1). Thus substitution of $y = y(r, \theta, \epsilon)$ into (5.11) yields the differential equation on Σ

$$\begin{aligned} \dot{r} &= \epsilon^{2p}(\pm r + r^{2p+1}K) + O(\epsilon^{2p+1}), \qquad \pm = \operatorname{sgn} K, \\ \dot{\theta} &= \omega_0 + O(\epsilon). \end{aligned} \tag{6.2}$$

The above reduction of the bifurcation problem to Σ carries over for more general (infinite-dimensional) systems whenever the following hold:

(1) there exists a smooth invariant manifold Σ given by $y = y^*(r, \theta, \alpha)$ through the origin $(r, y, \alpha) = (0, 0, 0)$ and tangent to the (r, θ, α)-space;

(2) each point of Σ lies on a trajectory of the differential equation, that is, the differential equation induces a smooth flow on Σ; and

(3) all orbits lying near the origin for all time $t \in (-\infty, \infty)$ lie on Σ.

Quite generally, if the equations generate a semigroup $T(t, x, \alpha)$ such that the map $(x, \alpha) \to T(t, x, \alpha)$ is smooth for each fixed t, then a center manifold with these properties exists. This is the case for retarded functional differential equations and for many classes of PDE's such as certain nonlinear parabolic equations, and the Navier–Stokes equation.

It is now very easy to obtain periodic solutions for the two-dimensional system (6.2).

LEMMA 6.1. *For r restricted to a sufficiently large bounded region, the system (6.2) has a unique periodic solution $r(t, \epsilon)$, $\theta(t, \epsilon)$ for small $|\epsilon| \neq 0$. As $\epsilon \to 0$,*

$$r(t, \epsilon) \to r_0 = |K|^{-1/2p} \text{ uniformly,}$$

$$\tau(\epsilon) \overset{\text{def}}{=} \text{period of the solution} \to 2\pi/\omega_0 .$$

The solution, restricted to Σ, is stable when $K < 0$ and unstable when $K > 0$.

THEOREM 6.2. *Let the Hopf bifurcation problem (5.6) be averaged, and the substitution $\alpha = -(\operatorname{sgn} K)\,\epsilon^{2p-1}$ made, as described in Theorem 5.3, so that*

$$\begin{aligned} \dot{r} &= \epsilon^{2p}(\pm r + r^{2p+1}K) + O(\epsilon^{2p+1}), \qquad \pm = -\operatorname{sgn} K, \\ \dot{\theta} &= \omega_0 + O(\epsilon), \\ \dot{y} &= A_Q y + O(\epsilon), \end{aligned}$$

is obtained. Then in the original (unaveraged and unscaled) equation, there is a unique periodic solution bifurcating from the origin, either for $\alpha > 0$ (when $K < 0$) or $\alpha < 0$ (when $K > 0$). More precisely, in the original coordinates (x, y, α) with $x = (r \cos \theta, r \sin \theta)$ the solution has the form

$$\begin{aligned} r(t, \epsilon) &= \epsilon r_0 + O(\epsilon^2), \qquad r_0 = |K|^{-1/2p}, \\ \theta(t, \epsilon) &= \omega_0 t + O(\epsilon), \\ y(t, \epsilon) &= O(\epsilon^2), \\ \tau(\epsilon) &= \text{period of the solution} = (2\pi/\omega_0) + O(\epsilon), \end{aligned}$$

where ϵ is related to the bifurcation parameter α by

$$\alpha = -(\operatorname{sgn} K)\,\epsilon^{2p}.$$

The solutions obtained for ϵ and $-\epsilon$ are identical and only differ by a time translation. The solution is stable if and only if all eigenvalues of A_Q lie in the left half-plane and $K < 0$.

In case the eigenvalues $\lambda(\alpha)$, $\overline{\lambda(\alpha)}$ *of the* (r, θ)*-subspace of the linearized equation satisfy merely* $\operatorname{Re} \lambda'(0) \neq 0$ *rather than* $\operatorname{Re} \lambda(\alpha) \equiv \alpha$, *then all of the above carries over with the obvious modifications.*

Only Lemma 6.1 will be proved, as the theorem is a simple application of this lemma to the scaled and averaged system considered in Theorem 5.3.

Proof of Lemma 6.1. Consider the solution $r(t)$, $\theta(t)$ of (6.2) for ϵ fixed, with initial conditions

$$r(0) = \rho, \qquad \theta(0) = 0,$$

and define $R(\rho, \epsilon) = r(\tau)$ where $\theta(\tau) = 2\tau$. Periodic solutions then are given by solving

$$R(\rho, \epsilon) = \rho, \tag{6.3}$$

and $\tau = \tau(\rho, \epsilon)$ is the period. From the form of (6.2) it is clear that

$$\begin{aligned} \tau(\rho, \epsilon) &= (2\pi/\omega_0) + O(\epsilon), \\ R(\rho, \epsilon) &= \rho + \epsilon^{2p}(2\pi/\omega_0)(\pm\rho + \rho^{2p+1}K) + O(\epsilon^{2p+1}). \end{aligned}$$

Thus (6.3) reduces to solving

$$\begin{aligned} S(\rho, \epsilon) &\overset{\text{def}}{=} \epsilon^{-2p}(R(\rho, \epsilon) - \rho) \\ &= (2\pi/\omega_0)(\pm\rho + \rho^{2p+1}K) + O(\epsilon) \\ &= 0. \end{aligned}$$

It is immediate that $S(r_0, 0) = 0$, $(\partial S/\partial\rho_0)(r_0, 0) \neq 0$. Hence by the implicit function theorem there is a unique zero $\rho = \rho(\epsilon)$ of S, with $\rho(0) = r_0$. This establishes the existence of the periodic solution. The assertions about stability are a consequence of the existence of the annulus $\mathscr{A}$ in Theorem 4.3; alternatively one may observe that

$$\begin{aligned} K < 0 &\Rightarrow (\partial S/\partial r_0)(\rho(\epsilon), \epsilon) < 0 \Rightarrow \text{stability (in } \Sigma), \\ K > 0 &\Rightarrow (\partial S/\partial r_0)(\rho(\epsilon), \epsilon) > 0 \Rightarrow \text{instability.} \end{aligned}$$

7. Infinite-Dimensional Systems

Here we show how the averaging procedure and its application to bifurcation carries over to certain classes of infinite-dimensional evolution equations, such as functional and certain partial differential equations. Assume the equation can be written abstractly as

$$\begin{aligned} \dot{z} = f(z, \alpha) &= A(\alpha)\, z + F(z, \alpha), \\ F(z, \alpha) &= O(|z|^2), \end{aligned} \tag{7.1}$$

where z and $\dot{z}$ lie in (generally different) Banach spaces. Specifically, assume

$$f: X_1 \times (-\alpha_0, \alpha_0) \to X_2$$

is sufficiently smooth, where X_1 and X_2 are Banach spaces with X_1 continuously and densely contained in X_2. For example, if (7.1) represents a parabolic equation of the form

$$\begin{aligned} \partial z/\partial t &= g(x, z, \partial z/\partial x, \partial^2 z/\partial x^2, \alpha), \\ x \in \Omega &= \text{smooth, open, bounded set in } R^n, \\ z &= 0 \quad \text{for} \quad x \in \partial\Omega, \end{aligned}$$

then choices for (X_1, X_2) would possibly be $(H_0{}^2(\Omega), L^2(\Omega))$ or $(C_0^{2+\alpha}(\Omega), C^\alpha(\Omega))$.

We also assume, as before,

$$\begin{aligned} f(0, \alpha) &= 0, \\ (\partial f/\partial z)(0, \alpha) &= A(\alpha) \\ &= \text{bounded linear operator from } X_1 \text{ to } X_2, \end{aligned}$$

so that the linearized equation about the origin $z = 0$, when $\alpha = 0$, is

$$\dot{z} = A(0)z.$$

By considering X_1 as a subset of X_2, then $A(0)$ (or more generally $A(\alpha)$) may be regarded as an unbounded closed operator from X_2 into itself, with domain $X_1 \subseteq X_2$. Typically, the space X_1 (when not considered a subset of X_2) will be eldowed with the graph norm of $A(0)$. We have the spectral decomposition

$$X_2 = P \oplus Q,$$

where P is the two-dimensional eigenspace of $A(0)$ corresponding to simple eigenvalues $\pm i\omega_0 \neq 0$ and the spectrum of $A_Q(0)$ ($= A(0)$ restricted to Q) is assumed to lie a positive distance δ from the imaginary axis. In fact, we assume the decomposition

$$\begin{aligned} X_2 &= P(\alpha) \oplus Q(\alpha), \\ \dim P(\alpha) &= 2 \end{aligned}$$

holds for all α near 0, the positive distance δ holds uniformly, and the eigenvalues of $A(\alpha)$ restricted to $P(\alpha)$ are $\alpha \pm i\omega(\alpha)$. Since the eigenspace P lies in $X_1 \subseteq X_2$ the decomposition restricts to X_1

$$X_1 = P \oplus (Q \cap X_1) \stackrel{\text{def}}{=} P \oplus Q_1.$$

It is now clear that (7.1) can be decomposed by writing

$$\begin{aligned} z &= x + y \in P \oplus Q_1 = X_1, \\ \dot{z} &= x + y \in P \oplus Q = X_2, \\ x &= (r \sin \theta, r \cos \theta). \end{aligned}$$

Assuming all of the above conditions on the spectrum of $A(\alpha)$, and the resulting decomposition, the only other assumption needed is the existence of a center manifold

$$\begin{aligned} \Sigma\colon y &= y^*(r, \theta, \alpha) \in X_1, \\ 0 &= y^*(0, \theta, \alpha) = (\partial y^*/\partial r)(0, \theta, 0), \end{aligned}$$

through the origin $r = 0$, and tangent to the (r, θ, α) space. It is important that y^* be a smooth map taking values in X_1, since after averaging, we must substitute $y = y^*(r, \theta, \alpha)$ into the right-hand side of (7.1). As described in Section 6, a smooth flow is induced on Σ, and all periodic orbits bifurcating from the origin lie on Σ. The basic assumption on (7.1) necessary for Σ to exist is that the nonlinear semigroup $T(t, z, \alpha)$ in X_1 generated by (7.1) be smooth in (z, α) for each fixed t.

With the above setup, all there remains to do is to rigorously justify the formal averaging procedure applied to (7.1). Recall from Theorem 5.1 the coordinate change

$$\bar{r} = r + \epsilon u_1(r, \theta, \alpha, \epsilon) + \epsilon w(r, \theta, \alpha, \epsilon) y + \epsilon^2 u_2(r, \theta, \alpha, \epsilon) \tag{7.2}$$

used to compute the first order constant K. (For simplicity we consider only this case, as averaging of higher order terms is similar.) Here u_1 and u_2 are scalar valued, while w takes values in the dual $Q_1{}^*$ of the y-space Q_1. In fact $w(r, \theta, \alpha, \epsilon) = rw^*(\theta)$ whereas in (5.7), w^* is the unique 2π-periodic solution of

$$G_2(\theta, 0, 0) + w^{*\prime}(\theta)\, \omega_0 + w^*(\theta)\, A_Q = 0. \tag{7.3}$$

Let us consider more carefully the meaning of Eq. (7.3) and trasformation (7.2) in the infinite-dimensional space. Now G_2 arises as a coefficient of y in the differential equation involving $\dot{r}$ (after decomposing); hence $G_2(\theta, 0, 0)$ for each θ is a linear functional acting on $y \in Q_1$. In particular, writing G_2 as a Fourier series yields

$$G_2(\theta, 0, 0) = \sum_{n=-\infty}^{\infty} g_n e^{in\theta},$$

$$g_n \in Q_1{}^*, \qquad \sum_{n=-\infty}^{\infty} | g_n |_{Q_1{}^*}^2 < \infty.$$

By expanding $w^*(\theta)$ as a Fourier series

$$w^*(\theta) = \sum_{n=-\infty}^{\infty} w_n e^{in\theta}, \tag{7.4}$$

inserting this into Eq. (7.3) and equating coefficients, we arrive at

$$w_n = -g_n(A_Q + in\omega_0)^{-1}. \tag{7.5}$$

Since A_Q is a bounded operator from Q_1 into Q, then $(A_Q + in\omega_0)^{-1}$ is bounded from Q into Q_1. This implies in particular

$$w_n \in Q^* \qquad | w_n |_{Q^*} \leqslant (\text{const.})\, n^{-1} \, | g_n |_{Q_1^*} .$$

Thus by defining w^* by (7.4) and (7.5), both w^* and its derivative $w^{*\prime}$ are square integrable functions taking values in Q^*. This is stronger than saying they take values in Q_1^*, and is due to the presence of the smoothing operator $(A_Q + i\omega_0)^{-1}$ in (7.5). This last observation is important since when Eq. (7.1) is rewritten in terms of the new averaged coordinates $(\bar{r}, \theta, y)$, it is seen that the functional $w(r, \theta, \alpha, \epsilon)$ acts on $\dot{y} \in Q$. Indeed, this is what happens when (7.2) is differentiated with respect to time.

One thus concludes that the form of the equation, that is, $z \in X_1$, and $\dot{z} \in X_2$, is preserved under any sequence of averaging transformations (7.2), and (of course) scaling; the equation $y = y(r, \theta, \epsilon)$ describing the center manifold in scaled, averaged coordinates may be substituted into the infinite-dimensional system, to reduce the problem to the two-dimensional case, as before.

8. Functional Differential Equations

In order to average retarded functional differential equations, care must be taken as to how the equation is interpreted as an (abstract) ordinary differential equation in a Banach space, as in (7.1). Consider the RFDE

$$\dot{z}(t) = f(z_t, \alpha), \tag{8.1}$$

where the notation of Hale is followed. In particular, assume $z \in R^n$, and z_t is the function defined by

$$z_t(\theta) = z(t + \theta), \qquad -r \leqslant \theta \leqslant 0;$$

thus

$$z_t \in C \overset{\text{def}}{=} C([-r, 0], R^n),$$

where r is fixed. The phase space of (8.1) is thus the Banach space C, and $f = f(\phi, \alpha)$ satisfies

$$f: C \times (-\alpha_0, \alpha_0) \to R^n,$$
$$f(0, \alpha) = 0.$$

Equation (8.1) can be solved forward in time by specifying an initial condition $z_0 = \phi \in C$ at time $t = 0$.

Some remarks on notation are in order here. As there is some overlap in the notation for FDE's and that of the previous seven sections, it will be necessary to adopt several new conventions for the integral averaging. For example, we use θ from now on to denote the argument of z_t, so that $\theta \in [-r, 0]$, and not the angle variable as before. The symbol ζ henceforth denotes the angle in polar coordinates (r, ζ).

Write (8.1) as

$$\begin{aligned} \dot{z}(t) &= L(\alpha)\, z_t + F(z_t, \alpha), \\ L(\alpha) &= (\partial f/\partial\phi)(0, \alpha), \\ |F(\phi, \alpha)| &= O(|\phi|^2), \end{aligned} \tag{8.2}$$

so that the linearized equation at the origin is, for each α,

$$\dot{z}(t) = L(\alpha)\, z_t .$$

Here $L(\alpha)$ is a linear functional on C but takes values in R^n; it thus has the Stieltjes integral representation

$$L(\alpha)\,\phi = \int_{-r}^{0} d_\theta \eta(\alpha, \theta)\, \phi(\theta),$$

where the $n \times n$ matrix $\eta(\alpha, \theta)$ is of bounded variation in $\theta \in [-r, 0]$, and smooth in α when considered as a $BV[-r, 0]$ valued function.

Equations (8.1), (8.2) certainly do not fit the framework of Section 7, in particular (7.1). The left-hand side $\dot{z}(t)$ lies in R^n and this cannot be considered a phase space of (8.2). The clue to writing (8.2) as an abstract ODE comes from the variation of constants formula for retarded equations.

First scale $z \to \epsilon z$, $\alpha \to \epsilon\alpha$ so that (8.2) takes the form (with a different F)

$$\begin{aligned} \dot{z}(t) &= Lz_t + \epsilon F(z_t, \alpha, \epsilon), \\ L &= L(0). \end{aligned} \tag{8.3}$$

The linear equation $\dot{z}(t) = Lz_t$ at $\epsilon = 0$ generates a strongly continuous semigroup $T(t)$ of bounded linear operators on C, with infinitesimal generator A given by

$$A\phi = (d/d\theta)\,\phi,$$
$$\phi \in C^1 \cap \{\phi \mid \dot{\phi}(0) = L\phi\} = \text{domain of } A.$$

In integrated form, (8.3) becomes

$$z_t = T(t)\, z_0 + \epsilon \int_0^t T(t-s)\, X_0 F(z_s, \alpha, \epsilon)\, ds \tag{8.4}$$

where $X_0 = X_0(\theta)$ is given by

$$X_0(\theta) = \begin{cases} I = n \times n \text{ identity matrix}, & \theta = 0, \\ 0, & -r \leqslant \theta < 0. \end{cases}$$

Strictly speaking X_0 does not belong to C because of the discontinuity at $\theta = 0$; nevertheless X_0 can serve as the initial condition for the linear equation, so the semigroup can act on it to produce $T(t)\, X_0$. Equation (8.4) must be interpreted as an equality for each $\theta \in [-r, 0]$, but one may informally think of it as an equation in C. If (8.4) is differentiated with respect to t. we obtain the *formal* expression

$$(d/dt)\, z_t = Az_t + \epsilon X_0 F(z_t, \alpha, \epsilon). \tag{8.5}$$

As it stands, (8.5) does not make sense, for two reasons:

(1) In general z_t does not belong to the domain of A; it is certainly C^1 (at least for $t \geqslant r$) but may not satisfy $\dot{z}(t) = Lz_t$.

(2) The nonlinear term is a multiple of X_0, hence does not belong to C.

We shall show that (8.5) does make sense if interpreted correctly. Both of the above problems can be remedied at once if we think of z_t as belonging to

$$z_t \in C^1 = \{\phi \in C \mid \phi \text{ is of class } C^1\}$$

and extend the domain of A to all of C^1.

To extend the domain of A, consider the formula for A^{-1}; to solve $A\phi = \psi$ for ϕ we have

$$\phi(\theta) = \phi(0) + \int_0^\theta \psi(s)\, ds \tag{8.6}$$

with $\phi(0)$ determined by

$$\begin{aligned} \psi(0) &= L\left[\phi(0) + \int_0^\theta \psi(s)\, ds\right] \\ &= \left[\int_{-r}^0 d\eta(\theta)\right] \phi(0) + \int_{-r}^0 d\eta(\theta) \int_0^\theta \psi(s)\, ds. \end{aligned} \tag{8.7}$$

As long as

$$\det \left[\int_{-r}^0 d\eta(\theta)\right] \neq 0$$

then $\phi(0)$ is uniquely determined, hence A has a bounded inverse. What is crucial is that formulas (8.6), (8.7) are defined even for $\psi = X_0$. In this case, ϕ is the constant (matrix valued) function

$$\phi = A^{-1}X_0 = \left[\int_{-r}^{0} d\eta(\theta)\right]^{-1}. \tag{8.8}$$

This means that the domain of A has been extended to include all constant functions, provided we let A take values in

$$BC = C \oplus \langle X_0 \rangle,$$

the space of all functions continuous on $-r \leqslant \theta < 0$, with a jump discontinuity at $\theta = 0$. It is easy to see that A is well defined on C^1, since any C^1 function ϕ can be uniquely written as

$$\begin{aligned} \phi &= \phi^1 + \phi^2, \\ \dot{\phi}^1(0) &= L\phi^1, \\ \phi^2 &= \text{constant function}. \end{aligned}$$

Indeed, to attain this decomposition it is enough to let ϕ^2 be defined by

$$\dot{\phi}(0) = L(\phi - \phi^2) = L\phi - \left[\int_{-r}^{0} d\eta(\theta)\right] \phi^2.$$

Once this is done, we have A defined on all of C^1 by

$$A\phi = A\phi^1 + A\phi^2 = \dot{\phi} + X_0 \left[\int_{-r}^{0} d\eta(\theta)\right] \phi^2$$

from (8.8), which implies that

$$A\phi = \dot{\phi} + X_0[L\phi - \dot{\phi}(0)] \tag{8.9}$$

for all $\phi \in C^1$. Observe here that we may now drop the restriction that A^{-1} exist, and *let* (8.9) *define A in its extended domain* C^1. With A interpreted in this extended sense, we claim in the following theorem that (8.5) holds for solutions of the RFDE (8.2), as long as $t \geqslant r$.

THEOREM 8.1. *Consider the retarded functional differential equation*

$$\dot{z}(t) = Lz_t + \epsilon F(z_t, \alpha, \epsilon), \tag{8.10}$$

where

$$L\phi = \int_{-r}^{0} d\eta(\theta)\, \phi(\theta).$$

Let the operator A map C^1 into $BC = C \oplus \langle X_0 \rangle$ by

$$A\phi = \dot{\phi} + X_0[L\phi - \dot{\phi}(0)].$$

Then any solution of (8.10) *for $t \geqslant t_0$, satisfies*

$$(d/dt)\, z_t = Az_t + \epsilon X_0 F(z_t, \alpha, \epsilon) \tag{8.11}$$

as long as $t \geqslant t_0 + r$ (or, in fact, as long as $z_t \in C^1$).

The proof of this is immediate, so it is omitted.

With this interpretation of the retarded equation (8.10) as the abstract ODE (8.11), all that needs to be done before averaging is to decompose BC as $P \oplus Q$ with the two-dimensional eigenspace $P \subseteq C^1$ and complement Q, according to the spectrum of A. This is done by Hale and we review the main points in preparation for the examples of later sections.

The spectrum of A is determined, as with ODE's, by exponential solutions of $\dot{z}(t) = Lz_t$. In particular, all spectral values λ are isolated and of finite multiplicity, and are determined by solving the characteristic equation

$$\det \left[\lambda I - \int_{-r}^{0} d\eta(\theta)\, e^{\lambda\theta}\right] = 0.$$

The elements of the eigenspaces are all exponential polynomials. A basis $\Phi = (\phi^1, \phi^2)$ for P is chosen, as well as a dual basis $\Psi = \operatorname{col}(\psi^1, \psi^2)$ for the adjoint equation, where

$$\psi^j \in C^* \overset{\text{def}}{=} C([0, r], R^n) \qquad \text{(row vectors)}.$$

With the bilinear form on $C^* \times C$ defined by

$$(\psi, \phi) = \psi(0)\,\phi(0) - \int_{-r}^{0} \int_{0}^{\theta} \psi(s - \theta)\, d\eta(\theta)\, \phi(s)\, ds$$

we assume $(\Psi, \Phi) = I$. The projections ϕ^P and ϕ^Q of any $\phi \in BC$ onto P and Q are given by

$$\phi^P = \Phi(\Psi, \phi),$$
$$\phi^Q = \phi - \phi^P.$$

Relative to the basis Φ, the operator A restricted to P may be represented by a matrix A_P defined by

$$A\Phi = (d/d\theta)\Phi = \Phi A_P; \tag{8.12}$$

we let A_Q denote A restricted to Q. Thus, for each t we may decompose z_t as

$$\begin{aligned} z_t &= z_t^P + z_t^Q \\ &= \Phi(\Psi, z_t) + z_t^Q \\ &\overset{\text{def}}{=} \Phi x(t) + y_t, \end{aligned}$$

where $x(t)$ lies in R^2 and y_t takes values in Q, but does not necessarily satisfy $y_t(\theta) = y(t+\theta)$. With this decomposition (8.11) becomes

$$\begin{aligned} \dot{x}(t) &= A_P x(t) + \epsilon\Psi(0) F(\Phi x(t) + y_t, \alpha, \epsilon), \\ (d/dt)\, y_t &= A_Q y_t + X_0{}^Q F(\Phi x(t) + y_t, \alpha, \epsilon). \end{aligned}$$

By writing x in polar coordinates (r, ζ), we may now average over the angle ζ.

9. Wright's Equation

We study the Hopf bifurcation for Wright's equation

$$\dot{z}(t) = -az(t-1)[1 + z(t)] \tag{9.1}$$

with the real parameter a. This equation has been studied by many authors; in particular, it is known that the origin $z = 0$ is stable when $0 < a < \pi/2$ and unstable when $a > \pi/2$. Moreover, for all $a > \pi/2$, (9.1) always has a periodic solution. Such solutions have been shown to exist by means of topological fixed point theorem; however, this method sheds no light on qualitative behavior, such as amplitude, period, or stability. Using integral averaging we shall study the local behavior of periodic solutions near the bifurcation points

$$a_N = (-1)^N[(\pi/2) + N\pi], \qquad z = 0 \tag{9.2}$$

for each nonnegative integer N. In all cases the constant K turns out to be nonzero, so a generic bifurcation occurs. In fact, our computations reveal the following.

THEOREM 9.1. *Consider Wright's equation* (9.1). *At each bifurcation point* $a = a_N$ *in* (9.2), *a generic Hopf bifurcation occurs from* $z = 0$. *The bifurcating solution has the form*

$$z(t) = \left[\frac{20(a - a_N)}{3a_N - 1}\right]^{1/2} \cos a_N t + O(a - a_N)$$

for a *near* a_N. *Thus bifurcation occurs to the right of* a_N *for* N *even, and the left*

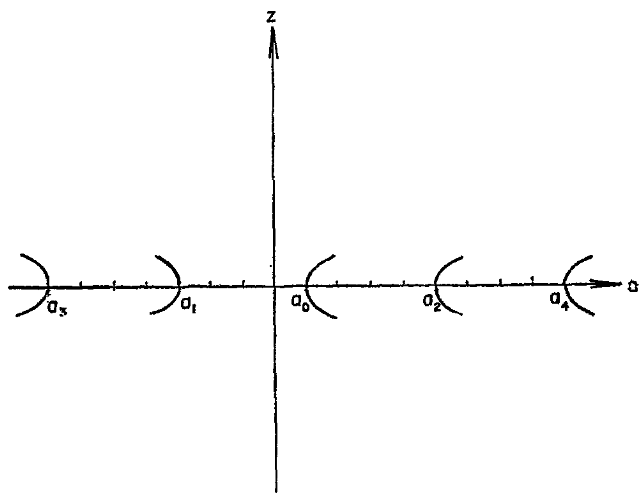

FIGURE 9.1

for N odd, in accordance with Fig. 9.1. *Only the solution at* $N = 0$ *is stable, and it then has the form*

$$z(t) = \left[\frac{40(a - (\pi/2))}{3\pi - 2}\right]^{1/2} \cos\left(\frac{\pi t}{2}\right) + O\left(a - \frac{\pi}{2}\right),$$

with the constant

$$[40/(3\pi - 2)]^{1/2} \cong 2 \cdot 3210701.$$

The proof of this theorem consists in writing (9.1) as an abstract ODE, and then averaging. The characteristic equation near $z = 0$ is

$$\lambda + ae^{-\lambda} = 0 \tag{9.3}$$

and at $a = a_N$ has imaginary roots

$$\lambda = \pm ib_N,$$
$$b_N = |a_N| = (\pi/2) + N\pi.$$

For a near a_N, there is a unique pair of conjugate roots $\lambda(a)$, $\overline{\lambda(a)}$ near $\pm ib_N$ with $\lambda(a_N) = ib_N$. This follows from the implicit function theorem; moreover, differentiating (9.3) shows

$$\nu = \operatorname{Re} \lambda'(a_N) = a_N/(1 + a_N^2) \neq 0. \tag{9.4}$$

Let us scale $z \to \epsilon z$ and set $a = a_N$. We compute the constant K of Theorem 5.1, and as noted before, K depends only on the differential equation at $a = a_N$.

Consider then

$$\dot{z}(t) = -a_N z(t-1) - \epsilon a_N z(t-1)\, z(t).$$

Hale gives a basis

$$\Phi(\theta) = (\cos b_N\theta, \sin b_N\theta)$$

for the eigenspace corresponding to $\lambda = \pm i b_N$, and a dual basis

$$\Psi(\theta) = \frac{2}{1+b_N^2}\begin{pmatrix}\cos b_N\theta - b_N \sin b_N\theta \\ \sin b_N\theta + b_N \cos b_N\theta\end{pmatrix}$$

relative to the bilinear form

$$(\psi, \phi) = \psi(0)\,\phi(0) - a_N \int_{-1}^{0} \psi(\theta+1)\,\phi(\theta)\, d\theta.$$

Thus $(\Psi, \Phi) = I$, the identity matrix. Upon decomposing as in Section 8 we obtain the equations

$$\begin{aligned} \dot{x}(t) &= A_P x(t) + \epsilon\Psi(0) f(\Phi x(t) + y_t), \\ (d/dt)\, y_t &= A_Q y_t + \epsilon X_0{}^Q f(\Phi x(t) + y_t), \\ f(\phi) &= -a_N \phi(-1)\,\phi(0). \end{aligned} \tag{9.5}$$

The defining condition (8.12) for A_P implies

$$A_P = \begin{pmatrix} 0 & b_N \\ -b_N & 0 \end{pmatrix},$$

and we observe

$$\begin{aligned} A\phi &= \dot{\phi} - X_0[a_N\phi(-1) + \dot{\phi}(0)], \qquad \phi \in C^1, \\ A_Q &= A \text{ restricted to } Q = \{\phi \in C^1 \mid (\Psi, \phi) = 0\}, \\ X_0{}^Q &= X_0 - X_0{}^P = X_0 - \Phi\Psi(0). \end{aligned}$$

In polar coordinates (9.5) becomes

$$\begin{aligned} \dot{r} &= \frac{2\epsilon}{1+b_N^2}(\cos\zeta + b_N \sin\zeta)\, f(\Phi x + y_t), \\ \dot{\zeta} &= -b_N + \frac{2\epsilon}{r(1+b_N^2)}(b_N\cos\zeta - \sin\zeta)\, f(\Phi x + y_t), \\ \frac{d}{dt}\, y_t &= \text{as before}, \\ x &= \operatorname{col}(r\cos\zeta, r\sin\zeta), \\ f(\Phi x + y_t) &= -a_N((-1)^{N+1} r \sin\zeta + y_t(-1))(r\cos\zeta + y_t(0)). \end{aligned}$$

Therefore, with the notation of Theorem 5.1, and in particular (5.6), we have

$$\begin{aligned}
C_3(\zeta, 0, 0) &= \frac{2b_N}{1 + b_N^2}(\cos\zeta + b_N \sin\zeta)\sin\zeta\cos\zeta, \\
C_4(\zeta, 0, 0) &= 0, \\
D_3(\zeta, 0, 0) &= \frac{2b_N}{1 + b_N^2}(b_N \cos\zeta - \sin\zeta)\sin\zeta\cos\zeta, \\
\omega_0 &= -b_N, \\
G_2(\zeta, 0, 0)\phi &= \frac{2b_N}{1 + b_N^2}(\cos\zeta + b_N \sin\zeta) \\
&\quad \times (\phi(0)\sin\zeta + (-1)^{N+1}\phi(-1)\cos\zeta), \\
J_2(0, 0)(\cos\zeta, \sin\zeta)^2 &= b_N(\cos\zeta\sin\zeta)\, X_0^Q.
\end{aligned}$$

Direct calculation yields

$$K^* = (1/2\pi b_N)\int_0^{2\pi} C_3 D_3 \, d\zeta = 0.$$

Writing G_2 as

$$G_2(\zeta, 0, 0) = \sum g_n e^{in\zeta}$$

we solve Eq. (5.7)

$$G_2(\zeta) - w^{*\prime}(\zeta)\, b_N + w^*(\zeta)\, A_Q = 0$$

to yield

$$w^*(\zeta) = \sum g_n (inb_N - A_Q)^{-1} e^{in\zeta}.$$

From Theorem 5.1 we then have

$$\begin{aligned}
K^{**} &= \frac{b_N}{2\pi}\sum g_n(inb_N - A_Q)^{-1} X_0^Q \int_0^{2\pi} e^{in\zeta}\cos\zeta\sin\zeta\, d\zeta \\
&= \frac{ib_N}{4}[g_2(2ib_N - A_Q)^{-1} + g_{-2}(2ib_N + A_Q)^{-1}]\, X_0^Q.
\end{aligned}$$

Since $X_0^Q = X_0 - \Phi\Psi(0)$ and $G_2(\zeta, 0, 0)$ are real, we have $g_{-2} = \bar{g}_2$, which implies

$$K^{**} = (-b_N/2)\,\mathrm{Im}\, g_2(2ib_N - A_Q)^{-1} X_0^Q.$$

First let us determine the linear functional g_2. By writing $\cos\zeta\sin\zeta$ in terms of $e^{\pm i\zeta}$ it is easy to see

$$g_2\phi = \frac{b_N}{2(1 + b_N^2)}(1 - ib_N)(-\phi(0)\, i + (-1)^{N+1}\phi(-1))$$

for all $\phi \in C$. Next we determine

$$\phi = (2ib_N - A_Q)^{-1}X_0^Q.$$

To calculate, more generally,

$$\phi = (2ib_N - A)^{-1}\psi$$

we must solve

$$\dot{\phi}(\theta) = 2ib_N\phi(\theta) - \psi(\theta)$$

subject to the boundary condition

$$\dot{\phi}(0) = -a_N\phi(-1).$$

The solution is easily obtained as

$$\phi(\theta) = e^{2ib_N\theta}\phi(0) - \int_0^\theta e^{2ib_N(\theta-s)}\psi(s)\,ds,$$

$$\phi(0) = [2ib_N + a_N e^{-2ib_N}]^{-1}\left[\psi(0) + a_N\int_0^{-1} e^{-2ib_N(1+s)}\psi(s)\,ds\right]$$

$$= [(\pi/2) + N\pi]^{-1}\,(2i + (-1)^{N+1})^{-1}\left[\psi(0) + a_N\int_0^{-1} e^{-2ib_N(1+s)}\psi(s)\,ds\right].$$

For $\psi = X_0$ we then have

$$\begin{aligned}
\phi(0) &= \left(\frac{\pi}{2} + N\pi\right)^{-1}(2i + (-1)^{N+1})^{-1},\\
\phi(-1) &= -\phi(0),\\
g_2\phi &= g_2(2ib_N - A)^{-1}X_0\\
&= \frac{b_N}{2(1 + b_N^2)}(1 - ib_N)(-1 + (-1)^N)\\
&\quad\times\left(\frac{\pi}{2} + N\right)^{-1}(2i + (-1)^{N+1})^{-1},
\end{aligned}$$

$$\frac{-b_N}{2}\operatorname{Im}[g_2(2ib_N - A)^{-1}X_0] = -\frac{b_N^2[3b_N + (-1)^{N-1}]}{20(1 + b_N^2)[(\pi/2) + N\pi]}.$$

The calculation of $(2ib_N - A)^{-1}X_0^P$ uses the fact that $X_0^P = \Phi\Psi(0)$, and that A is represented by the matrix A_P relative to the basis Φ of P. Thus

$$\begin{aligned}
(2ib_N - A)^{-1}X_0^P &= \Phi(2ib_N - A_P)^{-1}\Psi(0)\\
&= \Phi\begin{pmatrix}2ib_N & b_N\\ b_N & 2ib_N\end{pmatrix}^{-1}\begin{pmatrix}1\\ b_N\end{pmatrix}\frac{2}{1 + b_N^2}\\
&= \frac{-2}{3(1 + b_N^2)}\,\Phi\begin{pmatrix}(2i/b_N) + 1\\ (-1/b_N) + 2i\end{pmatrix}
\end{aligned}$$

and this yields

$$g_2(2ib_N - A)^{-1} X_0^P = \frac{b_N}{2(1+b_N^2)}(1+ib_N)\left(\frac{-2}{3(1+b_N^2)}\right)\left(i+\frac{1}{b_N}\right)$$
$$- \frac{b_N}{2}\operatorname{Im}[g_2(2ib_N - A)^{-1} X_0^P] = 0.$$

We therefore finally obtain

$$K = K^{**} = -\frac{b_N^2[3b_N + (-1)^{N+1}]}{20(1+b_N^2)[(\pi/2)+N\pi]} = -\frac{b_N[3b_N+(-1)^{N+1}]}{20(1+b_N^2)}.$$

Thus the averaged equation has the normal form

$$\dot{r} = \epsilon\alpha\nu r + \epsilon^2 r^3 K + O(\epsilon^3),$$
$$\dot{\zeta} = -b_N + O(\epsilon),$$
$$a - a_N = \epsilon\alpha,$$

where ν is as in (9.4). The form of the bifurcating solution (in unscaled coordinates now)

$$r \sim |\nu/K|^{1/2}\epsilon, \qquad a - a_N = -\operatorname{sgn}(\nu K)\,\epsilon^2$$

is clear, and this implies the theorem.

10. Equations with First Integrals

Consider the equation

$$z(t) = a\int_{t-L_1-L_2}^{t-L_1} g(z(s))\,ds, \tag{10.1}$$

where $L_1 > 0$ and $L_2 > 0$ are fixed constants, a is a parameter, and g is a known smooth function mapping the reals into the reals. Observe that (10.1) can be written as a functional differential equation:

$$\dot{z}(t) = ag(z(t-L_1)) - ag(z(t-L_1-L_2)) \tag{10.2}$$

with a first integral

$$I(\varphi) = \varphi(0) - \int_{-L_1-L_2}^{-L_1} g(\varphi(\theta))\,d\theta, \qquad \varphi \in C[-L_1-L_2, 0].$$

For appropriate choices of g, L_1, and L_2, there is an equilibrium $c = c(a)$ obtained by solving

$$c - aL_2 g(c) = 0.$$

The linearized equation about the equilibrium $c = c(a)$ has the following characteristic equation

$$\lambda - ag'(c(a))[e^{-L_1\lambda} - e^{-(L_1+L_2)\lambda}] = 0.$$

Note that zero is always an eigenvalue because of the first integral $I(\varphi)$. In addition, we may have a pair of conjugate eigenvalues $\lambda(a)$ and $\bar{\lambda}(a)$ crossing the imaginary axis with nonzero speed at $a = a_0$.

In general, one may expect a Hopf bifurcation occurring on each integral surface

$$I(\varphi) = \text{constant},$$

where I is near zero. Since the integral surface

$$I(\varphi) = 0$$

is of interest in applications, we consider only the bifurcation in that surface.

To do the integral averaging on the integral surface, we first decompose

$$z_t - c = \Phi x + \Phi_0 \sigma + y_t,$$

where

$$\begin{aligned} \Phi(\theta) &= (\cos \mu\theta, \sin \mu\theta), \qquad \mu = \operatorname{Im} \lambda(a_0), \\ \Phi_0(\theta) &\equiv 1 \end{aligned}$$

are the bases for the appropriate eigenspaces, $x \in R^2$, $\sigma \in R^1$, and y_t is in the complementary subspace. As in the case of Wright's equation, we obtain the following equations after scaling,

$$\begin{aligned} \dot{x}(t) &= A_P x(t) + \epsilon \Psi(0)\, N(x(t), \sigma(t), y_t, \epsilon), \\ \sigma(t) &= \epsilon \Psi_0(0)\, N(x(t), \sigma(t), y_t, \epsilon), \\ (d/dt)\, y_t &= A_Q(y_t) + \epsilon X_0{}^Q N(x(t), \sigma(t), y_t, \epsilon), \end{aligned} \tag{10.3}$$

where $\Psi(\theta)$, $-(L_1 + L_2) \leqslant \theta \leqslant 0$, is the basis for the adjoint eigenspace (see Section 9), $\Psi_0(0)$ is some fixed constant, N is the nonlinearity, A_Q is the infinitesimal generator on the complementary subspace, $X_0{}^Q$ is as in Section 9 and

$$A_P = \begin{pmatrix} 0 & \mu \\ -\mu & 0 \end{pmatrix}.$$

At the bifurcation,

$$\begin{aligned} &I(c + \Phi x + \Phi_0 \sigma + y_t) \\ &\qquad = [1 - aL_2 g'(c(a))]\sigma + O(|x|^2 + \sigma^2 + |y_t|^2). \end{aligned}$$

So, after scaling we may solve $I = 0$ to get

$$\rho = \epsilon S(x, y_t, \epsilon).$$

Now, by eliminating σ from Eqs. (10.3) we obtain two equations involving x and y_t only. Thus proceeding as in Section 9 we obtain the bifurcations. Details are omitted.

11. Diffusion Equations

Consider the scalar parabolic equation

$$u_t = u_{xx}, \qquad t \geqslant 0, \quad 0 < x < 1 \tag{11.1}$$

with nonlinear boundary conditions

$$u_x(0, t) = 0 \tag{11.2}$$

$$u_x(1, t) = ag(u(0, t), u(1, t)), \tag{11.3}$$

where a is a parameter and

$$g(u, v) = \alpha u + \beta v + O(u^2 + v^2).$$

For simplicity, we assume $\alpha = 1$ and $\beta = 0$. A necessary and sufficient condition for $\lambda = i\mu$, $\mu > 0$, to be an eigenvalue is ($\sigma = (\mu/2)^{1/2}$)

$$2\sigma \sinh \sigma \cos \sigma = a_0,$$
$$2\sigma \cosh \sigma \sin \sigma = -a_0.$$

It is also not difficult to see that for appropriate parameters the eigenvalues cross the imaginary axis with nonzero speed as a passes through a_0.

Scaling (11.1), (11.2), and (11.3), we obtain

$$\begin{aligned} u_t &= u_{xx}, \\ u_x(0, t) &= 0, \\ u_x(1, t) &= a_0 u(0, t) + \epsilon N(u(0, t), u(1, t), \epsilon), \end{aligned} \tag{11.4}$$

where N is the nonlinearity.

In order to write (11.4) as an ordinary differential equation, we consider the operator

$$\tilde{A}\colon H^2(0, 1) \cap \tilde{H} \to L^2(0, 1),$$
$$u \to (d^2/dx^2)u,$$

where $L^2(0, 1)$ is the Lebesgue space, $H^2(0, 1)$ is the standard Sobolev space, and $\tilde{H}$ consists of functions such that

$$(d/dx)u\,|_{x=0} = 0, \qquad (d/dx)u\,|_{x=1} = a_0 u\,|_{x=0}\,.$$

Note that $H^2(0, 1) \cap \tilde{H}$ is well defined by the trace theorem and is in fact a subspace of codimension two. The operator $\tilde{A}$ is in fact a linear isomorphism with $\tilde{A}^{-1}$ given by

$$\begin{aligned} &\tilde{A}^{-1}: L^2(0, 1) \to H^2(0, 1) \cap \tilde{H}, \\ &\qquad u \to (1/a_0) \int_0^1 u(x)\, dx + \int_0^x \int_0^t u(s)\, ds\, dt. \end{aligned} \tag{11.5}$$

Because of the nonlinear boundary condition (11.3), the solution $u(\cdot, t)$ is in general not in $H^2(0, 1) \cap \tilde{H}$. However, the solution $u(\cdot, t) \in H^2(0, 1)$. Therefore, we extend the definition of $\tilde{A}$ to $H^2(0, 1)$ by permitting the extended map A to take values in a larger space. Specifically, observe that (11.5) is defined for the Dirac measures $\delta(x)$ and $\delta(x - 1)$, and in fact

$$\begin{aligned} \tilde{A}^{-1}(\delta(x)) &= (1/a_0) + x, \\ \tilde{A}^{-1}(\delta(x - 1)) &= (1/a_0). \end{aligned} \tag{11.6}$$

Now, for $u \in H^2(0, 1)$ we have

$$u(x) = \tilde{u}(x) + \left(\frac{1}{a_0} + x\right) u_x(0) + \frac{1}{a_0}(a_0 u(0) - u_x(1)),$$

where $\tilde{u}(x) \in H^2(0, 1) \cap \tilde{H}$ and

$$\tilde{u}(x) = u(x) - x u_x(0) + \frac{1}{a_0}[u_x(1) - u_x(0) - a_0 u(0)].$$

Motivated by (11.6), we define

$$\begin{aligned} &A: H^2(0, 1) \to L^2(0, 1) \oplus \langle \delta(x), \delta(x - 1)\rangle, \\ &\qquad u(x) \to u_{xx}(x) + u_x(0)\, \delta(x) + [a_0 u(0) - u_x(0)]\, \delta(x - 1), \end{aligned}$$

where $\langle \delta(x), \delta(x - 1)\rangle$ denotes the vector space generated by $\delta(x)$ and $\delta(x - 1)$.

Using the definition of A, we may rewrite (11.4) as an ordinary differential equation:

$$\begin{aligned} du/dt &= Au - u_x(0)\, \delta(x) - [a_0 u(0) - u_x(1)]\, \delta(x - 1) \\ &= Au + \epsilon N(u(0, t), u(1, t), \epsilon)\, \delta(x - 1) \\ &= f(u, \epsilon), \end{aligned}$$

where $f: H^2(0, 1) \times (-\epsilon_0, \epsilon_0) \to L^2(0, 1) \oplus \langle \delta(x), \delta(x - 1)\rangle$. Now, we may proceed to discuss the bifurcations as before.

12. Notes and Remarks

Section 2. Changes of coordinates as (2.3) are given in the textbook of Hale [16] and a survey paper of Volosov [47]. For more references on integral averaging, see the references in Hale [15, 16] and Volosov [47].

In the literature, Eq. (2.4) is usually averaged by eliminating the variables r and θ in the equation for $\dot{y}$ rather than eliminating the variable y in the equations for $\dot{r}$ and $\dot{\theta}$, as was done here. For example, in the book of Lefschetz [33] such an approach is to be used to study stability properties in which no bifurcation parameter appears (see also the papers of Hale [19] and Hausrath [[22]).

Section 3. The resonance condition of Corollary 3.2 for the eigenvalues has been used by a number of authors, notably, Sternberg [45], Hartman [21], and Lefschetz [33]. If the resonance condition is not satisfied, then it is impossible to average the equations as required since certain terms cannot be eliminated. In principle, one can eliminate these terms by a homeomorphic (but not diffeomorphic) change of coordinates. In practice, these changes of coordinates are not necessary as one could analyze the problem even with the presence of these terms.

Section 5. The existence of the invariant forms described at the end of this section follows from Chapter 18 of Hale [15].

Section 6. Apparently, the center manifold technique was first used by Chafee [6, 7] and also by Ruelle and Takens [39] (see also [32]). For PDE's, it was used by Marsden [35] and McCracken and Marsden [36]. Dorroh and Marsden [11] showed that for a large class of PDE, Properties (1), (2), and (3) hold.

Section 8. For the theory of functional differential equations, see Hale [17]. The space $BC = C \oplus \langle X_0 \rangle$ was used by Hale [19] in studying stability properties of equilibria of neutral FDE in critical cases.

Section 9. Equation (9.1) has been studied by many authors, notably, Wright [48] and Kakutani and Markus [29]. Jones [25] first showed that there are periodic solutions of (9.1) for $a > \pi/2$.

Section 10. Equation (10.1) is model of population growth introduced by Cooke and Yorke [10]. Hale [20] showed that for generic g's there is a Hopf bifurcation. Greenberg [14] studied the existence of periodic solutions for large values of a.

Section 11. The problem (11.1), (11.2), and (11.3) is a simplified model of a diffusion equation with nonlinear boundary conditions, occurring in the study of the interaction and production of two enzymes (see [3]). The actual system is:

$$\begin{aligned} u_t &= u_{xx} - Q^2 u, \\ v_t &= v_{xx} - Q^2 v, \\ u_x(0, t) &= -PQf(v(0, t)), \\ u_x(1, t) &= 0, \\ v_x(0, t) &= 0, \\ v_x(1, t) &= PQg(u(1, t)), \end{aligned}$$

where P and Q are constants and f and g are nonlinearities.

For parabolic equations in R^n with nonlinear boundary conditions a similar procedure may be followed. Consider, for example,

$$\begin{aligned} u_t &= \Delta u \qquad \text{in } \Omega \subseteq R^n, \\ \partial u/\partial n &= \xi(u) \qquad \text{on } \partial\Omega, \end{aligned}$$

where ξ is a nonlinear function which may depend on values of u even in Ω. Define

$$A: H^1(\Omega) \to H^1(\Omega)^*,$$

$$u \to \left[v \to -\int_\Omega \nabla u \cdot \nabla v\right],$$

where $H^1(\Omega)^*$ is the dual space of $H^1(\Omega)$. Formally, if $\langle\ ,\ \rangle$ represents the duality between $H^1(\Omega)^*$ and $H^1(\Omega)$ we have

$$\begin{aligned} \langle Au, v\rangle &= -\int_{\partial\Omega} (\partial u/\partial n)\, v + \int_\Omega (\Delta u)\, v \\ &= -\int_{\partial\Omega} \xi(u)\, v + \int u_t v. \end{aligned}$$

Assume that the nonlinear functional

$$\xi: H^1(\Omega) \to H^{-1/2}(\partial\Omega)$$

and define

$$N: H^1(\Omega) \to H^1(\Omega)^*,$$

$$\langle N(u), v\rangle = \int_{\partial\Omega} \xi(u)\, v.$$

Thus we obtain for the differential equation

$$\int_\Omega u_t v = \langle Au + N(u), v\rangle, \qquad v \in H^1(\Omega).$$

If we identify the function u_t with the element of $H^1(\Omega)^*$ defined by $v \to \int_\Omega u_t v$, $v \in H^1(\Omega)$, the differential equation may be written

$$du/dt = Au + N(u),$$

where $u \in H^1(\Omega)$, $du/dt \in H^1(\Omega)^*$. For more details see, for example, Lions and Magenes [34].

REFERENCES

1. J. C. ALEXANDER AND J. A. YORKE, Global bifurcation of periodic orbits, *Amer. J. Math.*, to appear.
2. A. ANDRONOV AND C. E. CHAIKIN, "Theory of Oscillations," Princeton Univ. Press, Princeton, N.J., 1949.
3. D. G. ARONSON, to appear.
4. P. BRUNOVSKY, On one parameter families of diffeomorphisms, *Comment. Math. Univ. Carolinae* **11** (1970), 559–582.
5. N. N. BRUSLINSKAYA, Qualitative integration of a system of n differential equations in a region containing a singular point and a limit cycle, *Dokl. Akad. Nauk SSSR* **139** (1961), 9–12.
6. N. N. CHAFEE, The bifurcation of one or more closed orbits from an equilibrium points of an autonomous differential systems, *J. Differential Equations* **4** (1968), 661–679.
7. N. N. CHAFEE, A bifurcation problem for a functional differential equation of finitely retarded type, *J. Math. Anal. Appl.* **35** (1971), 312–348.
8. S. N. CHOW, J. MALLET-PARET, AND J. K. HALE, Applications of generic bifurcation, I, *Arch. Rational Mech. Anal.* **59** (1975), 159–188.
9. S. N. CHOW, J. MALLET-PARET, AND J. K. HALE, Applications of generic bifurcation, II, *Arch. Rat. Mech. Anal.* **62** (1976), 209–235.
10. K. COOKE AND J. A. YORKE, Some equations modelling growth processes and gonorrhea epidemics, *Math. Biosci.* **16** (1973), 75–101.
11. J. R. DORROH AND J. E. MARSDEN, Smoothness of nonlinear semigroups, *J. Functional Analysis*, to appear.
12. H. I. FREEDMAN, On a bifurcation theorem of Hopf and Friedrichs, to appear.
13. K. O. FRIEDRICHS, "Advanced Ordinary Differential Equations," Gordon and Breach, New York, 1965.
14. J. M. GREENBERG, Periodic solutions to a population equation. Preprint.
15. J. K. HALE, "Oscillations in Nonlinear Systems," McGraw–Hill, New York, 1963.
16. J. K. HALE, "Ordinary Differential Equations," McGraw–Hill, New York, 1969.
17. J. K. HALE, "Theory of Functional Differential Equations," Springer–Verlag, Berlin, 1977.
18. J. K. HALE, Integral manifolds of perturbed differential systems, *Ann. Math.* **73** (1961), 496–531.
19. J. K. HALE, Critical cases for neutral functional differential equations, *J. Differential Equations* **10** (1971), 59–82.
20. J. K. HALE, Behavior near constant solutions of functional differential equations, *J. Differential Equations* **15** (1974), 278–294.
21. P. HARTMAN, "Ordinary Differential Equations," Wiley, New York, 1964.
22. A. R. HAUSRATH, Stability in the critical case of pure imaginary roots for neutral functional differential equations, *J. Differential Equations* **13** (1973), 329–357.
23. E. HOPF, Abzweigung einer periodischen Lösung von einer stationaren Losung eines Differential systems, *Ber. Math. Phys., Kl. Sachs Akad. Wiss. Leipzig* **94** (1942), 1–22.

24. G. Iooss, Existence et stabilité de la solution périodique secondaire intervenant dans les problèmes d'évolution du type Navier-Stokes, *Arch. Rational Mech. Anal.* **46** (1972), 301–329.
25. G. S. Jones, The existence of periodic solutions of $f'(x) = -\alpha f(x-1)[1+f(x)]$, *J. Math. Anal. Appl.* **5** (1962), 435–450.
26. D. D. Joseph and D. H. Sattinger, Bifurcating time periodic solutions and their stability, *Arch. Rational Mech. Anal.* **45** (1972), 79–109.
27. R. Jost and E. Zehneder, A generalization of the Hopf bifurcation theorem, *Helv. Phys. Acta* **45** (1972), 258–276.
28. V. I. Judovich, The onset of auto-oscillations in a fluid, *PMM* **35** (1971), 638–655.
29. S. Kakutani and L. Markus, On the nonlinear difference-differential equation $y'(t) = [A - By(t-\tau)]\, y(t)$, *Contrib. Theory Nonlinear Oscillations* **4** (1958), 1–18.
30. N. Kopell and L. N. Howard, Bifurcations under nongeneric conditions, *Advances in Math.* **13** (1974), 274–283.
31. L. D. Landau and E. M. Lifshitz, "Fluid Mechanics," Addison–Wesley, Reading, Mass., 1959.
32. O. E. Lanford, "Bifurcation of Periodic Solutions into Invariant Tori: The Work of Ruelle and Takens," Springer–Verlag Lecture Notes 322, pp. 159–192, Springer–Verlag, Berlin, 1972.
33. S. Lefschetz, "Differential Equations: Geometric Theory," Interscience, New York, 1963.
34. J. L. Lions and E. Magenes, "Nonhomogeneous Boundary Value Problems and Applications," Vol. 1, Springer–Verlag, Berlin, 1972.
35. J. E. Marsden, The Hopf bifurcation for nonlinear semigroups, *Bull. Amer. Math. Soc.* **79** (1973), 537–541.
36. J. E. Marsden and M. F. McCracken, The Hopf bifurcation and its applications, Springer–Verlag, New York, 1976.
37. M. F. McCracken, Computation of stability for the Hopf bifurcation theorem, to appear.
38. J. B. McLaughlin and P. C. Martin, "The Transition to Turbulence in a Statically Stressed Fluid," preprint, Dept. of Physics, Harvard University.
39. A. S. Pyartli, Birth of complex invariant manifolds close to a singular point of a parametrically dependent vector field, *Functional Anal. Appl.* **6** (1972), 339–340.
40. D. Ruelle and F. Takens, On the nature of turbulence, *Comm. Math. Phys.* **20** (1971), 167–192.
41. R. J. Sacker, A new approach to the perturbation theory of invariant surfaces, *Comm. Pure Appl. Math.* **18** (1965), 717–732.
42. D. H. Sattinger, "Topics in Stability and Bifurcation Theory," Springer–Verlag Lecture Notes, 309, Springer–Verlag, Berlin, 1973.
43. D. S. Schmidt, Hopf's bifurcation theorem and the center theorem of Liapunov, *J. Math. Anal. Appl.*, to appear.
44. J. Sotomayer, Generic one-parameter families of vectors fields, *Bull. Amer. Math. Soc.* **74** (1968), 722–726.
45. S. Sternberg, Local contraction and a theorem of Poincaré, *Amer. J. Math.* **79** (1957), 809–824.
46. F. Takens, Unfolding of certain singularities of vector fields: generalized Hopf bifurcations, *J. Differential Equations* **14** (1973), 476–493.
47. V. M. Volosov, Averaging in systems of ordinary differential equations, *Russian Math. Surveys* **17** (1962), 1–126.
48. E. M. Wright, A nonlinear difference-differential equation, *J. Reine Angew. Math.* **194** (1955), 66–87.

Printed in Belgium