1
Two Examples

In this chapter, two examples, one semisimple and the other not, will be treated from an elementary point of view. The purpose of the treatment is to motivate the concerns and themes of the remainder of the book. The semisimple example is the nonlinear center; when unfolded, this becomes the Hopf bifurcation. The nonsemisimple example is the generic double-zero eigenvalue, which unfolds to the Takens–Bogdanov bifurcation.

Although the discussion in this chapter is generally elementary, some topics are meant as previews of subjects to be discussed in detail later. In these cases the treatment here may appear somewhat sketchy, but it is not necessary to understand every point completely at this stage.

1.1 The (Single) Nonlinear Center

Consider the system of differential equations

$$\dot{u} = Au + Q(u) + C(u), \tag{1.1.1}$$

where $u = (x, y)$ (regarded as a column vector),

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{1.1.2}$$

and Q and C are quadratic and cubic terms, respectively. (We understand the words quadratic and cubic to mean *homogeneous* quadratic and cubic. That is, Q contains only quadratic terms, and C only cubic terms. It is understood here, and below, that any vector written with parentheses and

commas is actually a column vector; if we need a row vector it will be written as a row matrix, with square brackets and no commas.) The system (1.1.1) has a rest point at the origin, and the linear terms have eigenvalues $\pm i$. If these eigenvalues were in the left half-plane, the origin would be asymptotically stable; the rest point would be what is called a *sink*, or an *attracting rest point*, meaning that all orbits beginning near the origin approach the origin as $t \to \infty$. If one or both were in the right half-plane, the rest point would be unstable, either a *saddle* or a *source*. But none of these is the case, because the eigenvalues are on the imaginary axis. If the system were linear (that is, if the nonlinear terms were zero), the pure imaginary eigenvalues would imply that the origin was neutrally stable and surrounded by periodic solutions, a configuration called a *linear center*. But this also is not the case, because we do not assume that the nonlinear terms vanish. In fact, the stability or instability of the rest point at the origin for the system (1.1.1) cannot be determined from the linear terms alone, but will depend on the higher-order terms. This rest point is called a *nonlinear center*, or more precisely the *two-dimensional cubic nonlinear center* (since $u \in \mathbb{R}^2$ and (1.1.1) contains no terms of higher order than cubic). We propose the following problem: to introduce new coordinates $v = (\xi, \eta)$ into (1.1.1) so as to bring the quadratic and cubic terms into the simplest possible form, in the hope that this will facilitate determining the stability of the origin. This strategy is based on the fact that changing coordinates cannot affect the stability of the origin: If orbits of (1.1.1) beginning near the rest point approach the rest point as $t \to \infty$, this will be remain true no matter what coordinates are used. (We will, in fact, use only coordinate changes that map the origin to itself, so the words "rest point" can be replaced by "origin" in the last sentence.)

1.1.1. Remark. In elementary differential equations courses one studies the *linear spring*, or *harmonic oscillator*, governed by the equation

$$\ddot{x} + x = 0.$$

Introducing $y = \dot{x}$, this can be written as the system of differential equations

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

In polar coordinates ($x = r\cos\theta$, $y = r\sin\theta$) this becomes

$$\dot{r} = 0,$$
$$\dot{\theta} = -1.$$

The negative sign in the last equation is somewhat awkward, and can be avoided by interchanging the roles of x and y, so that we begin with

$$\ddot{y} + y = 0$$

and set $x = \dot{y}$ to obtain

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

or

$$\dot{r} = 0,$$
$$\dot{\theta} = 1.$$

This fits the pattern of (1.1.1) with vanishing nonlinear terms. The full system (1.1.1) can arise as a model of a *nonlinear spring* in which the restoring forces are not exactly proportional to the extension.

The Normal Form for the Nonlinear Center

The simplification will be carried out in two stages, using an approach called *format 1a* in later chapters. The first stage is to perform a change of variables of the form

$$u = v + q(v), \tag{1.1.3}$$

where $q : \mathbb{R}^2 \to \mathbb{R}^2$ is homogeneous quadratic. Since this transformation has no constant term, it does not move the origin; the origin is still the rest point. Since "$u = v$ up to first order" (since q is strictly quadratic), the linear term Au of the differential equation (1.1.1) should be unchanged in the new coordinates (that is, it should become Av). Therefore, the new equations should have the form

$$\dot{v} = Av + \widehat{Q}(v) + \widehat{C}(v) + \cdots, \tag{1.1.4}$$

where the dots denote higher-order terms introduced by the transformation. This can be proved, and an expression for $\widehat{Q}(v)$ can be obtained, by differentiating (1.1.3) with respect to time and using (1.1.4) to obtain, through quadratic terms,

$$\dot{u} = \dot{v} + q'(v)\dot{v} = Av + \left(\widehat{Q}(v) + q'(v)Av\right) + \cdots,$$

and comparing this with

$$\dot{u} = A(v + q(v)) + Q(v + q(v)) + \cdots = Av + (Aq(v) + Q(v)) + \cdots,$$

which results from substituting (1.1.3) into (1.1.1). The result is

$$\widehat{Q}(v) = Q(v) + Aq(v) - q'(v)Av. \tag{1.1.5}$$

Here q' denotes the matrix of partial derivatives of q. (If this derivation is not convincing, a more rigorous approach will be given in Lemma 4.3.1.) If we define the *homological operator* $\pounds$ acting on mappings $q : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$(\pounds q)(u) = q'(u)Au - Aq(u), \tag{1.1.6}$$

then (1.1.5) can be written

$$\pounds q = Q - \widehat{Q}. \tag{1.1.7}$$

Equation (1.1.7) is called a *homological equation*. We have not indicated the variables in (1.1.7); it is irrelevant whether calculations using this equation are performed using the letter u or v, as long as the correct variables are used in (1.1.4). Written out in greater detail, with $u = (x, y)$, $q(u) = (f(x, y), g(x, y))$, and A as in (1.1.2), the $\pounds$ operator (1.1.6) takes the form

$$\pounds \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} -yf_x + xf_y + g \\ -yg_x + xg_y - f \end{bmatrix}. \tag{1.1.8}$$

1.1.2. Remark. The omission of the terms indicated by $\cdots$ in (1.1.4) has an important consequence: Conclusions drawn from this equation (or any later modifications of this equation) are valid only in a neighborhood of the origin. Near the origin, the terms of low degree are the dominant terms, and it is reasonable to hope that these terms determine the behavior of the system. When x and y are large, on the other hand, terms of high degree dominate terms of lower degree, and it would not be reasonable to expect the truncated system to determine the behavior. This is the reason for the name *local* dynamical systems: We deal only with behavior local to the origin. Examples of local behavior include the stability or instability of the rest point at the origin, and the bifurcation of a limit cycle from the origin, both of which will be studied in this section.

1.1.3. Remark. Any mapping of $\mathbb{R}^2 \to \mathbb{R}^2$, such as q or Q, can be regarded in several ways:

1. as a *transformation* from one "copy" of $\mathbb{R}^2$ to another;
2. as a *deformation* within the same "copy" of $\mathbb{R}^2$ (for instance, a stretching of a rubber sheet); or
3. as a *vector field* on $\mathbb{R}^2$.

In our situation, Q is naturally regarded as a vector field and q as a transformation. Notice that $\pounds$ simply acts on a quadratic mapping $\mathbb{R}^2 \to \mathbb{R}^2$ and produces another such mapping. The fact that q in (1.1.7) is thought of as a transformation, while Q and $\widehat{Q}$ are viewed as vector fields, is of no consequence. There is a vector space $\mathcal{V}_1^2$ whose elements are quadratic maps of $\mathbb{R}^2 \to \mathbb{R}^2$, and $\pounds$ is an operator on this vector space; that is, $\pounds : \mathcal{V}_1^2 \to \mathcal{V}_1^2$. We will refer to $\mathcal{V}_1^2$ as the "space of quadratic vector fields on $\mathbb{R}^2$." It is six-dimensional, and a basis for it is given in (1.1.9) below. The notation is that $\mathcal{V}_j^n$ is the space of vector fields on $\mathbb{R}^n$ of degree $j + 1$; see Definition 4.1.1.

1.1.4. Remark. Throughout this book we will use the British monetary symbol $\pounds$ (which can be read "pounds") for the homological operator. The symbol $\pounds$ is chosen because it resembles the letter L (which stands for *Lie operator*, to be introduced later, of which $\pounds$ is a special case). For those who have seen a little homological algebra, the *homological equation* (1.1.7) can be pictured as follows. Let $\mathcal{V}_j^n$

be the space of homogeneous vector fields of degree $j+1$ on $\mathbb{R}^n$, let $\pounds$ be the homological operator defined by an $n \times n$ matrix A, and consider the (very short) chain complex

$$0 \to \mathcal{V}_j^n \to \mathcal{V}_j^n \to 0,$$

where the only nontrivial map is $\pounds : \mathcal{V}_j^n \to \mathcal{V}_j^n$. The homology space $\mathcal{V}_j^n / \operatorname{im} \pounds$ of this chain complex is an abstract version of the normal form space for degree $j+1$, in the sense that it is isomorphic to any normal form style (or complement of $\operatorname{im} \pounds$) but does not, by itself, select a style. This remark is entirely trivial and does not contribute to the study of normal forms, but it does justify the terminology. The connection between normal forms and homological algebra becomes more serious in connection with hypernormalization; see the Notes and References to Section 4.10.

We now ask whether it is possible to choose q in (1.1.7) so as to eliminate the quadratic terms entirely, that is, to make $\widehat{Q} = 0$. According to (1.1.7), this requires finding q such that $\pounds q = Q$. This is possible, for arbitrary Q, if and only if $\pounds : \mathcal{V}_1^2 \to \mathcal{V}_1^2$ is onto, where $\mathcal{V}_1^2$ is the space of quadratic vector fields mentioned in Remark 1.1.3. Since $\pounds$ is linear, and the space of quadratic vector fields on $\mathbb{R}^2$ is 6-dimensional, $\pounds$ can be represented as a 6×6 matrix. With respect to the basis

$$\begin{bmatrix} x^2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} xy \\ 0 \end{bmatrix}, \quad \begin{bmatrix} y^2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ x^2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ xy \end{bmatrix}, \quad \begin{bmatrix} 0 \\ y^2 \end{bmatrix} \tag{1.1.9}$$

for the space of (homogeneous) quadratic vector fields, the matrix of $\pounds$ is easily (but tediously) calculated to be

$$\pounds = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ -2 & 0 & 2 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & -2 & 0 & 2 \\ 0 & 0 & -1 & 0 & -1 & 0 \end{bmatrix}. \tag{1.1.10}$$

Further tedious processes of linear algebra, such as row reduction to echelon form, can be used to verify that this matrix is indeed invertible, showing that $\pounds$ is onto and the quadratic term can be eliminated. (A much easier way of seeing this will be explained later in this section.) At this stage our system has been simplified to the form (1.1.4) with $\widehat{Q}(v) = 0$. Notice that we have not calculated the expression for $\widehat{C}$. This expression will be quite complicated, and will depend on Q and q in addition to C. (In fact, we have not even calculated q, only shown that it exists.) At this point a crucial decision must be made: Do we wish to carry out the *computations* necessary to reduce any actual *specific* system of the form (1.1.1) to simplest form, or do we wish only to *describe* the simplest form to which any *arbitrary* system of the form (1.1.1) can be reduced? We call the latter problem

the *description problem* for the normal form of (1.1.1), and the former the *computation problem*. Only the description problem will be addressed in this introductory chapter. For this purpose we need not calculate q or $\widehat{C}$, but merely treat $\widehat{C}$ as an arbitrary cubic vector field. In fact, we do not even need to keep track of the coordinate changes that we use (which is of course essential for the computation problem). So we can replace v by u, drop the hat on $\widehat{C}$, and express our results so far in the form

$$\dot{u} = Au + C(u) + \cdots . \tag{1.1.11}$$

That is, the general equation of form (1.1.1) can be put into form (1.1.11) by changing variables and changing the cubic terms C.

The change in C that has been made so far was uncontrolled, and does not constitute a simplification of the cubic terms. The next step (the second of the two stages mentioned above) is to make a change of variables in (1.1.11) targeted at simplifying C. To that end we introduce new variables, again temporarily called $v = (\xi, \eta)$, by

$$u = v + c(v),$$

where c is cubic. This time the linear and (absent) quadratic terms in (1.1.11) are unchanged, while the cubic term C is replaced by $\widehat{C}$ with

$$\pounds c = C - \widehat{C}, \tag{1.1.12}$$

and the higher-order terms (represented by dots) are modified in an uncontrollable way. The operator $\pounds$ is the same as in (1.1.6), except that now it is taken as mapping the vector space $\mathcal{V}_2^2$ of (homogeneous) cubic vector fields on $\mathbb{R}^2$ into itself. With respect to the basis

$$\begin{bmatrix} x^3 \\ 0 \end{bmatrix}, \begin{bmatrix} x^2y \\ 0 \end{bmatrix}, \begin{bmatrix} xy^2 \\ 0 \end{bmatrix}, \begin{bmatrix} y^3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x^3 \end{bmatrix}, \begin{bmatrix} 0 \\ x^2y \end{bmatrix}, \begin{bmatrix} 0 \\ xy^2 \end{bmatrix}, \begin{bmatrix} 0 \\ y^3 \end{bmatrix} \tag{1.1.13}$$

the matrix of $\pounds$ is

$$\pounds = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -3 & 0 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -3 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & -2 & 0 & 3 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \end{bmatrix}. \tag{1.1.14}$$

The reduced row echelon form of this matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.1.15}$$

The two rows of zeros at the bottom shows that this matrix has a two-dimensional kernel (or null space), so $\mathcal{L}$ acting on cubic vector fields is not invertible, and the cubic terms cannot be entirely eliminated.

We have now arrived at an issue so crucial to the subject of normal forms that it will occupy all of Chapter 2 as well as large parts of Chapters 3 and 4. This issue is, how do we handle the homological equation (1.1.12) when $\mathcal{L}$ is not invertible? Remember that our goal is to simplify C. That means we want to choose $\widehat{C}$ to be as simple as possible, subject to the condition that (1.1.12) must be solvable for c. In other words, $C - \widehat{C}$ must lie in the image (or range) of $\mathcal{L}$. Briefly (and this will be developed at length later in the book), the image $\operatorname{im}\mathcal{L}$ is a subspace of the space $\mathcal{V}_2^2$ of cubic vector fields; if we select a complement to this image and let $\widehat{C}$ be the projection of C into this complement, then $C - \widehat{C}$ will belong to the image, and (1.1.12) will be solvable. The problem, then, comes down to selecting a complement to $\operatorname{im}\mathcal{L}$. We call this the choice of a normal form *style* , and it is here (as the idea of "style" suggests) that one's notion of simplicity comes into play.

In the present example, these issues can be handled fairly easily, since there is only one normal form style (called the *semisimple normal form*) that is accepted by everyone as the best. Recall that according to (1.1.15), $\dim\ker\mathcal{L} = 2$. Since $\dim\operatorname{im}\mathcal{L} + \dim\ker\mathcal{L} = \dim\mathcal{V}_2^2 = 8$, it follows that $\dim\operatorname{im}\mathcal{L} = 6$. It is therefore plausible, based on dimension count alone, that $\ker\mathcal{L}$ is a subspace of $\mathcal{V}_2^2$ complementary to $\operatorname{im}\mathcal{L}$:

$$\mathcal{V}_2^2 = \operatorname{im}\mathcal{L} \oplus \ker\mathcal{L}. \tag{1.1.16}$$

It will be shown (in Theorem 2.1.3 and Lemma 4.5.2) that this conjecture is true when the matrix A is semisimple, that is, diagonalizable over the complex numbers; this is the case for the matrix A given by (1.1.2). The validity of (1.1.16) can be verified (by tedious but elementary methods) in the current instance, by the following calculations: First, from the row reduced matrix (1.1.15), determine that $\ker\mathcal{L}$ is spanned by $(1,0,1,0,0,1,0,1)$ and $(0,1,0,1,-1,0,-1,0)$. Next, check that these are linearly independent of the columns of (1.1.14), which span $\operatorname{im}\mathcal{L}$. This can be done, for instance, by appending these vectors as additional columns in (1.1.14) and repeating the row reduction.

Having established that (1.1.16) is true, it follows that $\widehat{C}$ may be taken to lie in $\ker \mathcal{L}$. The vectors in $\mathbb{R}^8$ that we have just calculated as a basis for $\ker \mathcal{L}$ are, of course, coordinate vectors with respect to the basis (1.1.13) for $\mathcal{V}_2^2$, and correspond to the vector fields

$$\left(x^2+y^2\right)\begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad \left(x^2+y^2\right)\begin{bmatrix} -y \\ x \end{bmatrix}. \tag{1.1.17}$$

Thus, we may take

$$\widehat{C}(\xi,\eta) = \alpha\left(\xi^2+\eta^2\right)\begin{bmatrix} \xi \\ \eta \end{bmatrix} + \beta\left(\xi^2+\eta^2\right)\begin{bmatrix} -\eta \\ \xi \end{bmatrix}.$$

Since we are solving only the description problem, we may rename the variables as x and y, and state our result as follows: The normal form for (1.1.1), calculated to degree 3, is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \alpha\left(x^2+y^2\right)\begin{bmatrix} x \\ y \end{bmatrix} + \beta\left(x^2+y^2\right)\begin{bmatrix} -y \\ x \end{bmatrix} + \cdots . \tag{1.1.18}$$

Stability and Sufficiency of Jets

Now let us see whether this normal form calculation enables us to solve the stability problem for (1.1.1). For a preliminary investigation it is convenient to drop the terms of (1.1.18) that are represented by dots; this is called passing to the *truncated system*, or, more precisely, taking the *3-jet* of (1.1.18). (A *k-jet* is simply a Taylor polynomial of degree k, or the truncation of a power series after the kth term.) Introducing polar coordinates ($x = r\cos\theta$, $y = r\sin\theta$) brings the truncated system into the form

$$\begin{aligned} \dot{r} &= \alpha r^3, \\ \dot{\theta} &= 1 + \beta r^2. \end{aligned} \tag{1.1.19}$$

It is easy to understand the behavior of (1.1.19): If $\alpha < 0$, the origin is asymptotically stable (because r is decreasing); if $\alpha > 0$, the origin is unstable; and if $\alpha = 0$, the origin is neutrally stable, surrounded by periodic orbits. The period of these orbits is not usually independent of the amplitude, as it would be in a linear system, but increases with amplitude (a *soft center*) if $\beta < 0$ and decreases (*hard center*) if $\beta > 0$. The terms "hard and soft center" are borrowed from the more familiar "hard and soft spring," and have meaning even when $\alpha \neq 0$, provided that the "period" is understood as the time for one complete rotation of the angle.

This discussion of (1.1.19) is not conclusive for (1.1.18), but instead raises the important question of *sufficiency of jets*, which in this case takes the following form: Is the 3-jet of (1.1.18) sufficient to determine the stability of the system? The answer in this case is fairly easy to guess (details will be given below): If $\alpha < 0$ the origin of (1.1.18) is asymptotically stable; if

$\alpha > 0$, it is unstable; and if $\alpha = 0$, the stability of the origin cannot be determined, because it depends on the higher-order (dotted) terms. Thus, we say that *the 3-jet is sufficient to determine stability* if $\alpha \neq 0$, insufficient if $\alpha = 0$. In addition, if $\alpha < 0$, we will say that *the stability of the origin is 3-determined*. It is important to understand that the sufficiency of a jet is never absolute, but is relative to the question being asked. For example, if $\beta > 0$, the origin of (1.1.18) is hard; if $\beta < 0$, it is soft; and if $\beta = 0$, the 3-jet is insufficient to determine hardness or softness; therefore the 3-jet may be sufficient to determine stability but not sufficient to determine hardness (if $\alpha \neq 0$ but $\beta = 0$), or the reverse.

Here are the necessary details to show that if $\alpha < 0$, the origin of (1.1.18) is asymptotically stable. By the monomial division theorem (see Section A.2, especially Theorem A.2.16 and Remark A.2.17), the first equation of (1.1.18) may be written in full as $\dot{x} = -y + (\alpha x - \beta y)(x^2 + y^2) + f_0(x,y)x^4 + f_1(x,y)x^3y + f_2(x,y)x^2y^2 + f_3(x,y)xy^3 + f_4(x,y)y^4$, where $f_0, \ldots, f_4$ are unknown but smooth. Consider the term $f_1(x,y)x^3y$, and let M_1 be the maximum of $|f_1(x,y)|$ on the unit disk $r \leq 1$; then $|f_1(x,y)x^3y| \leq M_1 r^4 |\cos^3\theta \sin\theta| \leq M_1 r^4$. The other unknown terms, and those of $\dot{y}$, may be treated similarly. Since $\dot{r} = \dot{x}\cos\theta + \dot{y}\sin\theta$, we conclude after a little calculation that $\dot{r} = \alpha r^3 + g(r,\theta)$ with $|g(r,\theta)| \leq Mr^4$ on the unit disk, for some $M > 0$. Since $\alpha < 0$, it follows that $\dot{r} \leq (\alpha + Mr)r^3$ is negative for $0 < r < |\alpha|/M$, and the origin is asymptotically stable. (The final step requires a reference to the technical definition of asymptotic stability and an argument from the theory of Lyapunov functions, or simply a reliance on the intuitive meaning of asymptotic stability. Strictly speaking it is $r^2 = x^2 + y^2$, rather than r, that is the Lyapunov function, because r^2 is smooth at the origin.)

Finally, it should again be mentioned that if it is required to find the stability or instability of the origin in a *specific* system of the form (1.1.1), then it is necessary to reduce *that system* to normal form so that the sign of α can be determined. For this it is not sufficient to have only the *description* (1.1.18) of the normal form, but also the full *computation* of the normal form. This requires working out the details of several steps that we were able to omit in solving the description problem. These steps are best handled by recursive algorithms to be given in Appendices C and D.

A Complete Solution of the Description Problem

Up to this point, our discussion of the nonlinear center has followed the pattern of most elementary presentations of the subject. But there are many shortcomings to this approach. For example, consider our discussion of the homological equation (1.1.7) for the quadratic terms. In order to show that $\widehat{Q}$ could be taken to be zero, we showed that $\pounds$ is invertible by writing down its matrix and performing row operations. We promised

at the time that there was an easier way. It will be shown in Chapter 4 (Lemma 4.5.2) that if A is semisimple (that is, diagonalizable when the use of complex numbers is allowed), then the operator $\pounds$ defined by (1.1.6) is also semisimple, and its eigenvalues, when acting on homogeneous vector fields of a given degree, can be computed from those of A. In the present problem, the eigenvalues of $\pounds$ acting on vector fields of degree m turn out to be the numbers $(m_1 - m_2 \pm 1)i$, where m_1 and m_2 are nonnegative integers with $m_1 + m_2 = m$. Thus, we can compute the eigenvalues in the quadratic terms by writing down the table

m_1	m_2	$(m_1 - m_2 - 1)i$	$(m_1 - m_2 + 1)i$
2	0	i	$3i$
1	1	$-i$	i
0	2	$-3i$	$-i$

Since none of the eigenvalues are zero, $\pounds$ is invertible on the quadratic terms (that is, on the space of homogeneous quadratic vector fields), and we can take $\widehat{Q} = 0$.

The table of eigenvalues for the cubic terms is as follows:

m_1	m_2	$(m_1 - m_2 - 1)i$	$(m_1 - m_2 + 1)i$
3	0	$2i$	$4i$
2	1	0	$2i$
1	2	$-2i$	0
0	3	$-4i$	$-2i$

Since $\pounds$ is diagonalizable and the eigenvalue zero occurs twice, there are two linearly independent eigenvectors of $\pounds$ associated with the eigenvalue zero. That is, the dimension of the kernel of $\pounds$ is 2; it follows that there will be two parameters (α and β) in the cubic terms of the normal form. The eigenvalue table, by itself, is not sufficient to complete the computation of the normal form; it is still necessary to calculate the kernel in order to determine the explicit form of the terms in which α and β appear. But there is a way to find the kernel without first finding the large matrix (1.1.14).

The computation of this kernel can be viewed as a problem in partial differential equations, instead of one in linear algebra. According to (1.1.8), a vector field (f, g) lies in $\ker \pounds$ precisely when

$$\begin{aligned} -yf_x + xf_y &= -g, \\ -yg_x + xg_y &= f. \end{aligned} \tag{1.1.20}$$

In polar coordinates,

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = \frac{\partial}{\partial \theta},$$

so (1.1.20) can be written as $f_\theta = -g$, $g_\theta = f$. Together these imply $f_{\theta\theta} + f = 0$, which is the familiar equation of a spring, with solution $f(r, \theta) = A(r)\cos\theta + B(r)\sin\theta$; since $g = f_\theta$ it follows that $g(r, \theta) =$

$A(r)\sin\theta - B(r)\cos\theta$. Returning to rectangular coordinates,

$$f(x,y) = \frac{A(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}x + \frac{B(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}y, \tag{1.1.21}$$
$$g(x,y) = \frac{A(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}y - \frac{B(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}x.$$

To obtain solutions that are homogeneous cubic polynomials, it is necessary to take $A = a(x^2+y^2)^{3/2}$ and $B = b(x^2+y^2)^{3/2}$, for constants a and b. This gives two linearly independent cubic vector fields (f,g) in $\ker \mathcal{L}$, which is the correct number according to our eigenvalue table, and once again yields the normal form (1.1.18).

One of the advantages of (1.1.21) is that we can deduce from it not only the cubic terms of the normal form, but the terms of any order. But before doing this, it is worthwhile to reconsider the original system (1.1.1) that we have been studying, which contains no terms beyond cubic. Notice that terms of higher order are introduced at once in (1.1.4), and these remain present (with modifications) throughout all subsequent steps. When a particular feature that we wish to study (such as the stability of the origin) is not 3-determined, it is natural to normalize to a higher order. But it is pointless to do this if the original system (1.1.1) was obtained by dropping the higher-order terms of a more complete system. Many equations studied in applications arise in exactly this way, and then the dotted terms in (1.1.4) cannot be considered to be correct. In order to make sure that normalizing the higher-order terms is meaningful, let us replace (1.1.1) with a system of the form

$$\dot{x} = -y + f(x,y), \tag{1.1.22}$$
$$\dot{y} = x + g(x,y),$$

where f and g are smooth (infinitely differentiable) functions whose power series begin with quadratic terms:

$$f(x,y) \sim f_2(x,y) + f_3(x,y) + f_4(x,y) + \cdots, \tag{1.1.23}$$
$$g(x,y) \sim g_2(x,y) + g_3(x,y) + g_4(x,y) + \cdots,$$

where f_j and g_j are homogeneous of degree j. (Warning: In Chapter 4, for technical reasons, the indices will denote *one less than the degree*, but for now we will use the degree.)

1.1.5. Remark. The power series (or Taylor series) of a smooth function f need not converge (hence the symbol $\sim$, rather than $=$), and even when it does converge, it need not converge to f. This comes about because of the existence of *flat functions*, also called *exponentially small* or *transcendentally small* functions, which vanish at the origin together with all of their derivatives, although the function itself is not identically zero. (The basic example in one variable is

$f(x) = e^{-(1/x)^2}$ for $x \neq 0$, $f(0) = 0$.) Any two smooth functions having the same Taylor series differ by a flat function. In particular, if the Taylor series converges, the function itself may differ from the sum of its Taylor series by a flat function. Since the full Taylor series does not adequately represent a smooth function, the most important fact is that the Taylor series is asymptotic. That is, f differs from its k-jet $j^k f$ (the truncation of its Taylor series at degree k) by a smooth function of the order of the first omitted term:

$$|f(x) - (j^k f)(x)| \leq c|x|^{k+1}$$

in the case of functions of a single variable, with analogous formulas (with $|\cdot|$ replaced by $\|\cdot\|$) if x or f is a vector. It is sometimes of interest to know that every formal power series is the Taylor series of some smooth function, that is, that the mapping from smooth functions to power series is onto. This is called the Borel–Ritt theorem, proved in Appendix A (Theorem A.3.2).

Replacing (1.1.1) by (1.1.22) and (1.1.23) is a natural generalization that does not affect any of our previous calculations. It is not hard to extend these calculations to show that for any given positive integer $k \geq 2$, the vector fields (f_j, g_j) for $j = 2, 3, \ldots, k$ can be brought into a normal form characterized by the condition that $\pounds(f_j, g_j) = 0$. (Several ways of doing this will be developed in Chapter 4.) It is not, in general, possible to bring all terms into normal form, only a finite number of them.

1.1.6. Remark. Since we will soon find it useful to speak of the "normal form to all orders," in spite of its purported impossibility, it is worthwhile to point out the subtleties involved here. First, it is standard practice in mathematics to speak of mathematical objects as existing if they are uniquely determined, even if they are not calculated. Thus, we regard the decimal digits of π as existing, and think of π as an infinitely long decimal. (Constructivists are entitled to their own opinions.) Now, there is clearly no obstacle in principle to the computation of the terms of the normal form to any order, and therefore all of the terms exist in the mathematical sense. Therefore, it is possible to speak of the formal power series that would be obtained if the system were put into normal form to all orders. Of course, this series may not converge, and is subject to the limitations discussed in Remark 1.1.5. Even if the normal form to all orders converges, the sequence of transformations that normalize the successive terms need not converge. These difficulties can be overcome (in principle) by careful use of the Borel–Ritt theorem to show the existence of smooth functions filling the required roles, but since these smooth functions are not computable, invocation of the Borel–Ritt theorem is not actually helpful here. So when we do speak of the normal form to all orders, it should be kept in mind that in the end this must be truncated at some degree k, and then the original system can be transformed into a system that agrees with the normal form up to that degree.

Now we can turn to (1.1.21) to solve the description problem for the normal form of (1.1.22) to order k. It is not possible to choose A and B in (1.1.21) so as to obtain a homogeneous polynomial of even order; therefore all even terms in the normal form must vanish. Terms of odd order arise by taking A and B to be constants multiplied by odd powers of $\sqrt{x^2+y^2}$, and they have the same form as the cubic terms except that they contain a higher power of $\left(x^2+y^2\right)$. That is, (1.1.18) becomes

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} + \alpha_1\left(x^2+y^2\right)\begin{bmatrix}x\\ y\end{bmatrix} + \beta_1\left(x^2+y^2\right)\begin{bmatrix}-y\\ x\end{bmatrix} \tag{1.1.24}$$
$$+ \alpha_2\left(x^2+y^2\right)^2\begin{bmatrix}x\\ y\end{bmatrix} + \beta_2\left(x^2+y^2\right)^2\begin{bmatrix}-y\\ x\end{bmatrix} + \cdots .$$

When normalized to degree $2k+1$, truncated (taking the $(2k+1)$-jet), and expressed in polar coordinates, the result is

$$\begin{aligned}\dot{r} &= \alpha_1 r^3 + \alpha_2 r^5 + \cdots + \alpha_k r^{2k+1},\\ \dot{\theta} &= 1 + \beta_1 r^2 + \beta_2 r^4 + \cdots + \beta_k r^{2k}.\end{aligned} \tag{1.1.25}$$

Now it is easy to extend the previously given stability criterion: If $\alpha_1 < 0$, the origin is stable and the stability is 3-determined; if $\alpha_1 = 0$ and $\alpha_2 < 0$, it is stable and the stability is 5-determined, and so forth. Of course, if the first nonzero α_j is positive, the origin is unstable. If all of the α_j through $j = k$ are zero, the stability is not $(2k+1)$-determined, and the normalization must be continued to higher k. There remains the possibility that absolutely all α_j are zero (in the sense of Remark 1.1.6); in that case the stability of the origin is said to be *not finitely determined*.

1.1.7. Remark. In order to show that in this case the origin may be either asymptotically stable or unstable (and need not be neutrally stable), one may take f and g in (1.1.22) to be flat functions. Then all terms of (1.1.23) are zero, the system is already in normal form to all orders, and in polar coordinates $\dot{r} = h(r,\theta)$ where h is flat in r. Since h may be either positive or negative, the origin may be asymptotically stable or unstable. The existence of problems that are not finitely determined places an ultimate limitation upon all methods that will be presented in this book.

The Normal Form Module; Invariants and Equivariants

Equation (1.1.25) gives a complete solution of the description problem for the normal form of a nonlinear center to any order k, but does not yet give a complete description of the mathematical structure of the normal form. To do this, it is necessary to pass to the *normal form to all orders*, bearing in mind that (according to Remark 1.1.6) this is only a convenient manner of speaking, and that all "actual" normal forms are truncated at some order k. It turns out that the set of differential equations that have a given

linear term and are in normal form to all orders possesses the structure of a module over a ring.

1.1.8. Remark. Most of what we do will be clear without a technical knowledge of ring and module theory. Very briefly, a *field* is a set of objects that can be added, subtracted, multiplied, and divided, with the ordinary rules of algebra holding. The rational numbers, the real numbers, and the complex numbers form fields. A *ring* (by which we always mean a commutative ring with identity) is a similar structure that permits addition, subtraction, and multiplication, but not necessarily division. Examples are the ring of integers, the ring of polynomials in a specified number of variables, and the ring of formal power series ("infinitely long polynomials") in a specified number of variables. See Appendix A for more information about rings. Just as a vector space is a set that admits addition of vectors and multiplication of vectors by scalars taken from some field, a *module* is a set that admits addition of its elements, and multiplication of its elements by elements of a specified ring. The most important difference between a vector space and a module is that a vector space spanned by a finite number of elements always has a basis (linearly dependent elements can be eliminated from the spanning set). This is not always true for finitely generated modules (because the elimination process would require division in the ring), and in general there will be relations called *syzygies* among the generators of a module. Examples of this phenomenon will appear in Sections 4.5 and 4.7; see also Appendix B.

If the normal form (1.1.24) is extended to all orders and rearranged, it may be written (formally, that is, without regard to convergence) as

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} + \varphi\left(x^2+y^2\right)\begin{bmatrix}x\\ y\end{bmatrix} + \psi\left(x^2+y^2\right)\begin{bmatrix}-y\\ x\end{bmatrix},$$

where φ and ψ are formal power series in $s = x^2 + y^2$ beginning with the linear term (in s):

$$\begin{aligned}\varphi(s) &= \alpha_1 s + \alpha_2 s^2 + \alpha_3 s^3 + \cdots,\\ \psi(s) &= \beta_1 s + \beta_2 s^2 + \beta_3 s^3 + \cdots.\end{aligned}$$

It is convenient to allow constant terms in these series even though they do not appear in the normal form (they will appear later in the unfolding), so that

$$\begin{aligned}\varphi(s) &= \alpha_0 + \alpha_1 s + \alpha_2 s^2 + \cdots,\\ \psi(s) &= \beta_0 + \beta_1 s + \beta_2 s^2 + \cdots.\end{aligned}$$

The collection of formal power series in $s = x^2 + y^2$ forms a ring. (Such formal power series can be added, subtracted, and multiplied, but not always

divided.) The vector fields

$$\varphi(s)\begin{bmatrix} x \\ y \end{bmatrix} + \psi(s)\begin{bmatrix} -y \\ x \end{bmatrix} \tag{1.1.26}$$

appearing in the normal form are linear combinations of the basic vector fields (x, y) and $(-y, x)$ with coefficients taken from the ring (in the same way that vectors in the plane are linear combinations of two unit vectors with coefficients taken from the field of real numbers). Thus, the set of vector fields forms a module over the ring, with generators (x, y) and $(-y, x)$. (This is a particularly simple module, in that there are no syzygies among the generators.)

This example illustrates what is, in principle, the best possible solution of the description problem for a normal form. In the simplest cases such a description has the following ingredients:

1. A collection of functions $I_1, \ldots, I_p$ of the coordinates. These functions are called the *basic invariants*, for a reason to be explained shortly. In the present example $p = 1$ and $I_1 = x^2 + y^2$.
2. The ring of formal power series in $I_1, \ldots, I_p$. This is called the *ring of (formal) invariants*.
3. A collection of vector fields $v_1, \ldots, v_q$ called the *basic equivariants*. In our example $q = 2$, $v_1 = (x, y)$, and $v_2 = (-y, x)$.
4. The *normal form module*, or *module of equivariants*, which is the set of linear combinations of the basic equivariants, with coefficients taken from the ring of invariants.

In more complicated cases, the same pattern holds, except that there may be relations among the basic invariants (so that the ring of invariants is a quotient ring of a power series ring by an ideal of relations) and also relations among the basic equivariants. These issues will be dealt with as they arise.

It remains to explain the terms *invariant* and *equivariant*. The function $I = x^2 + y^2$ is invariant under rotation, in the sense that if the vector (x_1, y_1), thought of as an arrow from the origin, is rotated around the origin to give (x_2, y_2), then $I(x_1, y_1) = I(x_2, y_2)$. The vector fields $v_1 = (x, y)$ and $v_2 = (-y, x)$ also appear to be invariant under rotation, *when thought of as whole vector fields*. That is, if the vectors of the vector field are drawn on a sheet of paper and the paper is rotated, the vector field will appear the same. But this "invariance" is different from the invariance of I, because if (x_1, y_1) rotates to (x_2, y_2), then $v_i(x_1, y_1) \neq v_i(x_2, y_2)$. Instead, the vector $v_i(x_1, y_1)$ *rotates into* $v_i(x_2, y_2)$. This type of invariance is technically called equivariance.

But why are rotations involved at all? This comes from the fact that the solutions of the *linear part* of our original equation (1.1.1) are rotations.

Remember that normalizing a system of differential equations means normalizing it *relative to its linear part*. Our entire discussion has been based on the operator $\pounds$, defined in (1.1.6) in terms of the matrix A that gives the linear part of (1.1.1). It will turn out that for each A there is a group of transformations e^{At} under which the associated normal form is equivariant (in the present case, the group of rotations), and the vector fields equivariant under this group form a module over the scalar invariants for the same group. Details will have to await the development of sufficient machinery (in Chapters 3 and 4).

Unfolding the Nonlinear Center

There are two reasons for introducing what are called *unfoldings* of a dynamical system such as (1.1.1):

1. *Imperfection in modeling.* The given differential equation may be an imperfect mathematical model of a real system, that is, the model may have been derived using certain simplifying assumptions (such as an absence of friction). If the predictions of the model do not exactly match the behavior of the real system, but it is not feasible to improve the model (perhaps because the neglected quantities are too small to measure or too difficult to model), a natural step is to consider the behavior of *all* differential equations that are close to the original model, to see whether one of these will predict the observed behavior. This leads directly to the notion of unfolding.

2. *Bifurcation theory.* The given differential equation may have been obtained by fixing the value of one or more parameters in a more general equation. In this case one may be interested in studying how the behavior of the solutions changes as the parameters are varied. Varying the parameters produces systems close to the original one, which again belong to the unfolding of the given system. This is most important when the original values of the parameters have been taken to be *critical* values (or a *bifurcation point* in parameter space) at which some important transition in behavior takes place.

It is important to understand that there are many ways in which one system of differential equations can be considered "close" to another, and not all of these are covered by the idea of an unfolding. In particular, if a model contains fewer state variables than the actual system, this imperfection cannot be corrected by passing to the unfolding.

In order to illustrate these ideas, which are developed more completely in Chapter 6, we will briefly and sketchily treat the unfolding of the nonlinear center and the resultant bifurcation, known as the *Hopf bifurcation*. We begin by adding certain arbitrary terms, multiplied by a small parameter

ε, to (1.1.1) to obtain

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + Q(x,y) + C(x,y) + \varepsilon \left\{ \begin{bmatrix} p \\ q \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}. \tag{1.1.27}$$

The added terms represent the constant and linear terms of an arbitrary smooth function of (x, y), and the symbol $\equiv$ has been used in place of $=$ to indicate that we are omitting (or "calculating modulo") terms of degree 4 or higher in x and y, terms of degree 2 or higher in ε, and also products of ε times terms of degree 2 or higher in x and y. This represents a (somewhat arbitrary) choice of a (somewhat asymmetrical in x, y, and ε) jet, which we hope will be a sufficient jet to determine the behavior of the system. (Of course, the choice is not actually arbitrary, but is based on the knowledge that it gives useful results. But for exploratory purposes such a choice can be tried arbitrarily.)

Thus, *every* system that is close to (1.1.1) can be written in the form (1.1.27), up to the jet that we have just described. Our task is to simplify this system as much as possible, determine its behavior (that is, determine the behavior of the truncated system), and then study whether this behavior is the same as that of the full system (with the deleted terms restored). There are six arbitrary quantities in (1.1.27), namely, p, q, a, b, c, and d. In the course of simplifying the system these will be reduced to two arbitrary quantities, the *unfolding parameters*, which in this instance have the physical meaning of damping and frequency modulation.

1.1.9. Remark. Until Chapter 6, where a more precise definition of *asymptotic unfolding* will be given, we will use the word *unfolding* in a somewhat loose sense, to mean a system obtained by adding an arbitrary perturbation, taking a specified jet, and simplifying as much as possible. The arbitrary parameters that remain will be called the unfolding parameters, and the number of unfolding parameters is called the *codimension*. For those who are familiar with the usual technical definitions, we point out that our unfoldings are not necessarily miniversal unfoldings with respect to topological equivalence, and our codimensions are not necessarily the true codimensions in this sense. In fact, most dynamical systems do not have true miniversal unfoldings, because their codimension is infinite. In our approach this will be reflected in the facts that the number of unfolding parameters usually increases when the order of the jet is increased, and that in most cases no jet is sufficient to determine the complete behavior of the full system (up to topological equivalence). Nevertheless, the unfoldings that we calculate using jets of specified order can be sufficient to determine certain aspects of the behavior of the full system, such as (in the present example) the existence and stability of a Hopf bifurcation.

The first step in simplifying (1.1.27) is to perform the changes of variables needed to place the unperturbed ($\varepsilon = 0$) terms into normal form (1.1.18). These changes of variables may modify the perturbation terms, but since these are arbitrary anyway at this point, we need not be concerned. The result is

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} \equiv \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} + \alpha\left(x^2+y^2\right)\begin{bmatrix}x\\ y\end{bmatrix} + \beta\left(x^2+y^2\right)\begin{bmatrix}-y\\ x\end{bmatrix} \tag{1.1.28}$$
$$+ \varepsilon\left\{\begin{bmatrix}p\\ q\end{bmatrix} + \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}\right\}.$$

The presence of the small constant term $\varepsilon(p, q)$ means that there is no longer a rest point at the origin, but in fact there is still a rest point close to the origin, and it may be moved back to the origin by the small shift of coordinates

$$x = \xi - \varepsilon q,$$
$$y = \eta + \varepsilon p.$$

In fact, this transformation, followed by renaming (ξ, η) as (x, y), eliminates the constant term and reduces our equation to

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} \equiv \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} + \alpha\left(x^2+y^2\right)\begin{bmatrix}x\\ y\end{bmatrix} + \beta\left(x^2+y^2\right)\begin{bmatrix}-y\\ x\end{bmatrix} \tag{1.1.29}$$
$$+ \varepsilon\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}.$$

(The quantities a, b, c, d are again modified by this change of variables, as are higher-order terms that are deleted. But, since a, b, c, d are still arbitrary, we do not care. The general rule is that once the terms of a given order have been simplified, we must be careful to preserve them, but until then it does not matter. Of course, to solve the *computation problem* it would be necessary to record all the changes so that they can be implemented for a system with specific numerical coefficients, but again we are only solving the *description problem* at the moment.)

The next step is to simplify the linear terms in the perturbation. Putting $u = (x, y)$, (1.1.29) may be written as

$$\dot{u} \equiv Au + f(u) + \varepsilon Bu,$$

with

$$B = \begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

To this equation, we will apply a transformation of the form

$$u = (I + \varepsilon T)v,$$

with T initially unspecified; at the end we will choose T to make the transformed equation as simple as possible. Using the fact that the in-

verse transformation can be expressed as $v = (I - \varepsilon T + \cdots)u$ (where the omitted terms are of higher order in ε), the transformed equation is seen to be

$$\dot{v} \equiv Av + f(v) + \varepsilon \widehat{B} v,$$

where $\widehat{B} = B + AT - TA$. (Near-identity linear transformations of this type will be studied in great detail in Chapter 3.) If we define a new homological operator $\pounds'$, acting on 2×2 matrices, by

$$\pounds' Q = [Q, A] = QA - AQ,$$

then the relation between T and $\widehat{B}$ may be expressed as the homological equation

$$\pounds' T = B - \widehat{B}, \tag{1.1.30}$$

which is similar in structure to (1.1.7). In fact, the similarity is more than superficial: $\pounds'$ is the same as the operator $\pounds$ defined in (1.1.6), applied to linear vector fields represented by their matrices. More precisely, if Q is a matrix and Qu is the linear vector field that it defines, then $\pounds(Qu) = (\pounds' Q)u$.

Because the structure of this problem is the same as that of the normal form problem, the work necessary to describe the simplest form for $\widehat{B}$ has already been done: $\widehat{B}$ must belong to $\ker \pounds'$, or equivalently, $\widehat{B}v$ must belong to $\ker \pounds$, which means that $\widehat{B}v$ must be equivariant under the group of rotations. Thus, returning (as we always do after changing variables) to the original notation (replacing v by $u = (x, y)$ and $\widehat{B}$ by B), the normalized linear vector field Bu must be formed from the basic equivariant vector fields (x, y) and $(-y, x)$ by linear combinations using invariants as coefficients (see (1.1.26)). But the only invariants that will give linear vector fields are constants. So (with the conventional choice of signs, so that δ may be called the *damping*) we must have $Bu = -\delta(x, y) + \nu(-y, x)$, that is,

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -\delta & -\nu \\ \nu & -\delta \end{bmatrix}.$$

Thus, (1.1.29) finally simplifies to

$$\begin{aligned} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \alpha \left(x^2 + y^2\right) \begin{bmatrix} x \\ y \end{bmatrix} + \beta \left(x^2 + y^2\right) \begin{bmatrix} -y \\ x \end{bmatrix} \\ + \varepsilon \begin{bmatrix} -\delta & -\nu \\ \nu & -\delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \end{aligned} \tag{1.1.31}$$

containing (as promised) only the two unfolding parameters δ and ν. To determine the physical meaning of these parameters, we put (1.1.31) into

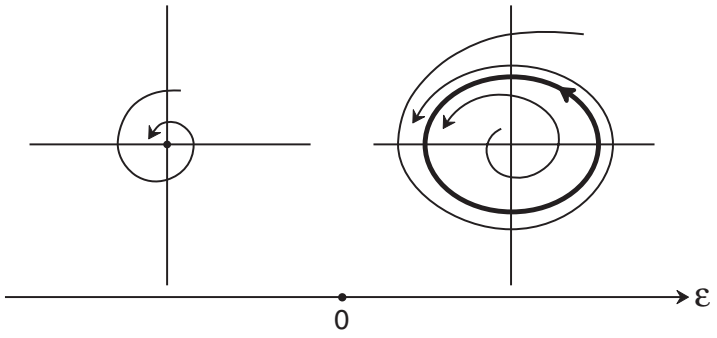

Figure 1.1. Hopf bifurcation with $\alpha < 0$ and $\delta < 0$.

polar coordinates to obtain (compare (1.1.19))

$$\begin{aligned} \dot{r} &= -\varepsilon\delta r + \alpha r^3, \\ \dot{\theta} &= 1 + \varepsilon\nu + \beta r^2. \end{aligned} \tag{1.1.32}$$

Thus, δ represents linear damping ($\delta > 0$ stabilizes the origin, regardless of the sign of α that determines nonlinear damping or excitation), while ν represents frequency modulation.

1.1.10. Remark. By scaling time (introducing a new independent variable τ by $\tau = (1+\nu)t$), it is possible to obtain equations having the same form as (1.1.32) with ν absent. Then δ is the only remaining unfolding parameter. Because of this, the Hopf bifurcation is usually classified as a codimension-one bifurcation. In fact, as the following discussion shows, δ controls the topological features of the bifurcation and ν affects only the period of the periodic orbit. Scaling time does not fit within the framework of normal form transformations, which involve only the dependent variables.

Observe that, if $\varepsilon\delta/\alpha$ is positive, there is a circle of radius

$$r^* = \sqrt{\frac{\varepsilon\delta}{\alpha}}$$

on which $\dot{r} = 0$; see Figure 1.1. This represents an invariant circle, or *limit cycle*, for (1.1.32). In particular, suppose $\alpha < 0$, so that the origin of the unperturbed ($\varepsilon = 0$) system is stable. Suppose $\delta < 0$ is fixed, and let ε be gradually increased from zero. Then the radius r^* of the limit cycle gradually increases from zero, as if the origin were blowing a smoke ring. The origin becomes unstable (because $\delta < 0$), but outside the limit cycle r is still decreasing (because $\alpha < 0$). Therefore, the limit cycle itself is asymptotically orbitally stable, all solutions of (1.1.32) approaching the limit cycle either from the inside or the outside. This process of "blowing a smoke ring" is the famous *Hopf bifurcation*.

1.1.11. Remark. When we say "let ε be gradually increased from zero," we do not mean that ε increases gradually *with time*. Instead we mean to consider a number of distinct systems of the form (1.1.32), in each of which ε is constant, but ε increases from zero as we pass from one system to the next. The question of what happens in a single system in which ε is varied with time is a much more difficult question, referred to in the literature as the problem of a "slowly varying bifurcation parameter" (or sometimes as "dynamic bifurcation," although this phrase usually has a different meaning). It will not be addressed here.

Of course, the discussion we have given so far does not prove the existence of a Hopf bifurcation in the original system. Equation (1.1.32) is a truncated system (because of the $\equiv$ in (1.1.31)). In order to prove the existence of the Hopf bifurcation, it is necessary to show that the (asymmetrical) jet used to define $\equiv$ is sufficient to determine the behavior that we have described. This will be done in Section 6.5. Because our analysis is local (Remark 1.1.2) we can expect the limit cycle to exist only for small ε, with r^* small, and we can expect the stability of the limit cycle to be only local. Solutions beginning at a fair distance from the origin may tend to other attractors that do not appear in (1.1.32).

Notes and References

Introductory treatments of normal form theory using matrices such as (1.1.10) are contained in Guckenheimer and Holmes [52], Wiggins [111], and Arrowsmith and Place [7]. Another useful introductory book is Kuznetsov [67], although the normal form work in this book is carried out on an ad hoc basis for each example and is not separated from the rest of the analysis. An extensive treatment of normal forms for (mostly) two-dimensional systems is given in Chow, Li, and Wang [29]. The partial differential equations approach, as in (1.1.20), is included in Elphick et al. [41]. Further references for normal forms will be found in Chapter 4 below, and for unfoldings in Chapter 6.

1.2 The Nonsemisimple Double-Zero Eigenvalue

The system to be studied in this section is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + Q(x, y) + \cdots , \tag{1.2.1}$$

where Q is the homogeneous quadratic term.

Normal Form Styles for the Nonsemisimple Double Zero

To simplify the quadratic term, the discussion in the previous section beginning with (1.1.3) down to (1.1.7) applies word for word, except that now

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \tag{1.2.2}$$

which entails, by way of (1.1.6), that

$$\pounds \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} yf_x - g \\ yg_x \end{bmatrix}. \tag{1.2.3}$$

But the solution of the homological equation

$$\pounds q = Q - \widehat{Q} \tag{1.2.4}$$

goes rather differently, because A is not semisimple (diagonalizable). The normal form cannot be taken to belong to $\ker \pounds$, because this does not form a complement to $\operatorname{im} \pounds$.

With respect to the basis (1.1.9), the matrix of $\pounds$ on the quadratic terms is

$$\pounds = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 2 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{1.2.5}$$

It is easy to see by inspection that the third and sixth columns of this matrix are linear combinations of the other four, which are linearly independent and therefore form a basis for the image of $\pounds$. Expressing these four columns as quadratic vector fields (after dividing the first column by two) yields

$$\begin{bmatrix} x^2 \\ -2xy \end{bmatrix}, \quad \begin{bmatrix} xy \\ 0 \end{bmatrix}, \quad \begin{bmatrix} y^2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -xy \\ y^2 \end{bmatrix}.$$

Any two quadratic vector fields linearly independent of these may be chosen as a basis for a complement to $\operatorname{im} \pounds$, and therefore as the basis for a possible *style* for the quadratic normal form. Two simple choices are

$$\begin{bmatrix} 0 \\ x^2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ xy \end{bmatrix} \tag{1.2.6}$$

and

$$\begin{bmatrix} x^2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ x^2 \end{bmatrix}. \tag{1.2.7}$$

With the first choice (1.2.6), the normal form becomes

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} + \alpha\begin{bmatrix}0\\ x^2\end{bmatrix} + \beta\begin{bmatrix}0\\ xy\end{bmatrix} + \cdots . \tag{1.2.8}$$

This normal form style will be called a *simplified normal form*. One of its distinguishing features is that it contains nonzero entries only in rows corresponding to the bottom rows of Jordan blocks of A (assumed to be in Jordan normal form).

One drawback to this ad hoc manner of selecting a normal form style (or a complement to $\operatorname{im}\mathcal{L}$) is that it has to be done again from scratch in each new degree. Thus, if we wanted to find the normal form for cubic terms, we would write down an 8×8 matrix, similar to (1.1.14), calculate its image, and select a complement. There would be no relationship between this complement and the one found for the quadratic terms, and hence little likelihood that the resulting normal form would have a nice mathematical description, such as a module structure involving invariants and equivariants. There has to be a better way.

In Section 4.6, it will be shown that a complement to $\operatorname{im}\mathcal{L}$ is always given by $\ker\mathcal{L}^*$, where

$$(\mathcal{L}^*q)(u) = q'(u)A^*(u) - A^*q(u), \tag{1.2.9}$$

A^* being the adjoint (or conjugate transpose) of A. (When A is real, as here, this is simply the transpose.) The normal form style resulting from this choice of complement will be called the *inner product normal form* (because it is related to a special inner product on vector fields, explained in Section 4.6). In our case

$$\mathcal{L}^*\begin{bmatrix}f(x,y)\\ g(x,y)\end{bmatrix} = \begin{bmatrix}xf_y\\ xg_y - f\end{bmatrix}, \tag{1.2.10}$$

so the inner product normal form (to any degree) consists of polynomial vector fields (of the required degree) satisfying the partial differential equations

$$\begin{aligned} xf_y(x,y) &= 0,\\ xg_y &= f. \end{aligned} \tag{1.2.11}$$

Calling such vector fields *equivariants*, and allowing formal power series (rather than only polynomials), it is not hard to see that the equivariants form a module over the *invariants*, or (scalar) formal power series $\varphi(x,y)$ satisfying $x\varphi_x(x,y) = 0$. (We can, of course, drop the x factor and write simply $\varphi_x(x,y) = 0$, and similarly in the first, but not the second, line of (1.2.11). We have written the longer form to emphasize the recurrence of the operator $x\partial/\partial y$, which is differentiation along the flow of the linear vector field $A^*(x,y)$. This flow defines the group with respect to which the invariants and equivariants are, respectively, invariant and equivariant.) In

Section 6.6, it will be seen that the only basic invariant is $I_1 = x$, and the basic equivariants are $v_1 = (0, 1)$ and $v_2 = (x, y)$; the normal form is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \sum_{i=1}^{\infty} \left(\alpha_i x^i \begin{bmatrix} x \\ y \end{bmatrix} + \beta_i x^{i+1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right). \tag{1.2.12}$$

The inner product normal form is more complicated than the simplified normal form (1.2.8), having nonzero entries outside the bottom row, but it has stronger mathematical properties. In particular, the simplified normal form is not equivariant under the flow of A^*.

A third normal form style, to be developed in Section 4.8, is what we call the sl(2) normal form (because it involves the "special linear algebra" of 2×2 matrices with zero trace). In the present problem it coincides with the inner product normal form, but in general it is different. (Once again, in appearance it is usually more complicated even than the inner product normal form, but has additional advantageous structural features.) The development of the sl(2) normal form depends upon a substantial amount of mathematical theory. This theory, and the sl(2) normal form itself, is presented in the starred sections, which may be omitted without loss of continuity.

Unfolding the Nonsemisimple Double Zero

We begin with the simplified normal form (1.2.8), and consider a system to be "close" to (1.2.8) if it is a perturbation of (1.2.8), obtained by adding terms multiplied by a perturbation parameter ε. Calculating modulo cubic terms in x and y, quadratic terms in ε, and products that are linear in ε and quadratic in x and y, the "arbitrary perturbation" of (1.2.8) is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix} + \varepsilon \left\{ \begin{bmatrix} p \\ q \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}. \tag{1.2.13}$$

Our goal is to reduce the number of arbitrary parameters p, q, a, b, c, d from six to two (in the generic case) or three (in all cases) by performing a series of coordinate changes. As usual, each coordinate change will be written as a change from (x, y) to (ξ, η), after which the notation reverts to (x, y). The first such change is

$$y = \eta + \varepsilon k,$$

with x unchanged; we call this transformation a *primary shift*, and it carries (1.2.13) into the following form:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix} + \varepsilon \left\{ \begin{bmatrix} p + k \\ q \end{bmatrix} + \begin{bmatrix} a & b \\ c & d + k\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}.$$

Choosing $k = -p$ removes p from the constant term. Renaming $d + k\beta$ as d yields the 5-parameter expression

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix} + \varepsilon \left\{ \begin{bmatrix} 0 \\ q \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}.$$

Next, by the same methods used to obtain (1.1.31), a linear transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} = (I + \varepsilon T) \begin{bmatrix} \xi \\ \eta \end{bmatrix}$$

can be found that reduces the system further to the three-parameter form

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix} + \varepsilon \left\{ \begin{bmatrix} 0 \\ q \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}, \tag{1.2.14}$$

where c and d are again allowed to differ from their original values. The next, and final, transformation does not fit into the usual normal-form framework, but instead is an example of "hypernormalization" or "normalization beyond the normal form." We call it a *secondary shift*:

$$x = \xi + \varepsilon h$$

carries (1.2.14) into

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix} + \varepsilon \left\{ \begin{bmatrix} 0 \\ q \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c + 2\alpha h & d + \beta h \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\}. \tag{1.2.15}$$

Observe that unlike the primary shift, the secondary shift has no effect on the constant term of the perturbation. It would have been useless to incorporate this shift into the first step, where our goal was to simplify the constant term. But now the secondary shift is useful. It has the effect of "injecting" the quantities $2\alpha h$ and βh into the bottom row of the matrix for the perturbed linear terms, *without disrupting the simplification of this matrix that has already been achieved* (that is, the a and b entries are still zero). It remains to choose h, which we do in one of the following ways: If $\alpha \neq 0$, choosing $h = -c/2\alpha$ eliminates the parameter c; if $\beta \neq 0$, choosing $h = -a/\beta$ eliminates the parameter d; if both α and β are nonzero, we can choose between eliminating c and d; if both α and β are zero, the secondary shift has no effect and we must retain both c and d.

Let us now assume that $\alpha \neq 0$; this is a generic condition on the unperturbed quadratic term. In this case, we can eliminate c (and once again modify d), and the resulting system can be rearranged as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ \varepsilon q \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & \varepsilon d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta y^2 \end{bmatrix}. \tag{1.2.16}$$

It is convenient to introduce new independent small unfolding parameters $\mu_1 = \varepsilon q$, $\mu_2 = \varepsilon d$, and write this unfolded system as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ \mu_1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & \mu_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta y^2 \end{bmatrix}. \tag{1.2.17}$$

On the other hand, if $\beta \neq 0$, we can achieve instead the unfolding

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ \mu_1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ \mu_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix}, \tag{1.2.18}$$

or (if we wish to permit arbitrary α and β with no condition imposed) we must accept the codimension-three unfolding given by

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ \mu_1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ \mu_2 & \mu_3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha x^2 + \beta xy \end{bmatrix}. \tag{1.2.19}$$

The first of these unfoldings, (1.2.17), is the most commonly studied; as the two parameters μ_1 and μ_2 are varied (in a neighborhood of $(0,0)$), saddle-node, Hopf, and homoclinic bifurcations are discovered. This is called the *Takens–Bogdanov bifurcation* and will be discussed in greater detail in Section 6.6, along with the question of the sufficiency of the jet indicated by $\equiv$. This is perhaps the best understood of all local dynamical systems, in the sense that there exists a complete proof that this jet is sufficient for all topological features of the motion.

Notes and References

The example studied in this section is treated in all of the references listed at the end of Section 1.1. Further treatment of this example, and further references, will be found in Sections 5.4 and 6.6.

http://www.springer.com/978-0-387-95464-6

Normal Forms and Unfoldings for Local Dynamical Systems

Murdock, J.

2003, XX, 500 p., Hardcover
ISBN: 978-0-387-95464-6