# Chapter 2
# Stability and Normal Forms

In this chapter our concern is with a system of ordinary differential equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ in a neighborhood of a point $\mathbf{x}_0$ at which $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$. Early investigations into the nature of solutions of the system of differential equations in a neighborhood of such a point were made in the late 19th century by A. M. Lyapunov ([114, 115]) and H. Poincaré ([143]). Lyapunov developed two methods for investigating the stability of $\mathbf{x}_0$. The so-called First Method involves transformation of the system to *normal form*; the Second or Direct Method involves the use of what are now termed *Lyapunov functions*. In the first section of this chapter we present several of the principal theorems of Lyapunov's Direct Method. Since smoothness of $\mathbf{f}$ is not necessary for these results, we do not assume it in this section. The second and third sections are devoted to the basics of the theory of normal forms.

## 2.1 Lyapunov's Second Method

Let $\Omega$ be an open subset of $\mathbb{R}^n$, $\mathbf{x}_0$ a point of $\Omega$, and $\mathbf{f} : \Omega \to \mathbb{R}^n$ continuous and such that solutions of initial value problems associated with the autonomous system of differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{2.1}$$

are unique. For $\mathbf{x}_1 \in \Omega$, we let $\mathbf{x}_1(t)$ denote the unique solution of (2.1) that satisfies $\mathbf{x}(0) = \mathbf{x}_1$, on its maximal interval of existence $J_1$; this is the *trajectory through* $\mathbf{x}_1$. The point set $\{\mathbf{x}_1(t) : t \in J_1\}$ is the *orbit* of (or through) $\mathbf{x}_1$. If $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$, then $\mathbf{x}(t) \equiv \mathbf{x}_0$ solves (2.1) uniquely, the orbit through $\mathbf{x}_0$ is just $\{\mathbf{x}_0\}$, and $\mathbf{x}_0$ is termed an *equilibrium* or *rest point* of the system, or a *singularity* or *singular point* (particularly when we view $\mathbf{f}$ as a vector field on $\Omega$; see Remark 3.2.4(b)). Any orbit is topologically a point, a circle (a *closed orbit* or a *cycle*), or a line (see [44] or [140]). The decomposition of $\Omega$, the *phase space* of (2.1), into the union of disjoint orbits is the *phase portrait* of (2.1). In part (a) of the following definition we have incorporated into the definition the result from the theory of differential equations

V.G. Romanovski, D.S. Shafer, *The Center and Cyclicity Problems*,
DOI 10.1007/978-0-8176-4727-8_2,

that if $\mathbf{x}_1(t)$ is confined to a compact set for all nonnegative $t$ in its maximal interval of existence, then that interval must contain the half-line $[0,\infty)$, so that existence of $\mathbf{x}_1(t)$ for all nonnegative $t$ need not be assumed in advance. The same consideration applies to part (b) of the definition.

**Definition 2.1.1.**
(a) An equilibrium $\mathbf{x}_0$ of (2.1) is *stable* if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathbf{x}_1$ satisfies $|\mathbf{x}_1 - \mathbf{x}_0| < \delta$ then $|\mathbf{x}_1(t) - \mathbf{x}_0| < \varepsilon$ for all $t \geq 0$.
(b) An equilibrium $\mathbf{x}_0$ of (2.1) is *asymptotically stable* if it is stable and if there exists $\delta_1 > 0$ such that if $\mathbf{x}_1$ satisfies $|\mathbf{x}_1 - \mathbf{x}_0| < \delta_1$, then $\lim_{t\to\infty} \mathbf{x}_1(t) = \mathbf{x}_0$.
(c) An equilibrium of (2.1) is *unstable* if it is not stable.

Note that it is possible that the trajectory of every point in a neighborhood of an equilibrium $\mathbf{x}_0$ tends to $\mathbf{x}_0$ in forward time, yet in every neighborhood there exist a point whose forward trajectory travels a uniformly large distance away from $\mathbf{x}_0$ before returning to limit on $\mathbf{x}_0$ (Exercise 2.1). This is the reason for the specific requirement in point (b) that $\mathbf{x}_0$ be stable. What we have called *stability* and *asymptotic stability* are sometimes referred to in the literature as *positive stability* and *positive asymptotic stability*, and the equilibrium is then said to be *negatively stable* or *negatively asymptotically stable* for (2.1) if it is respectively positively stable or positively asymptotically stable for $\dot{\mathbf{x}} = -\mathbf{f}(\mathbf{x})$.

The problem of interest in this section is to obtain a means of determining when an equilibrium of system (2.1) is stable, asymptotically stable, or unstable without actually solving or estimating solutions of the system, particularly when linear estimates fail. In a system of differential equations that models a mechanical system, the total energy function typically holds the answer: if total energy strictly decreases along positive trajectories near the equilibrium, then it is asymptotically stable. The concept of Lyapunov function generalizes this idea.

Since without loss of generality we may assume that the equilibrium is located at the origin of $\mathbb{R}^n$, we henceforth assume that $\Omega$ is a neighborhood of $\mathbf{0} \in \mathbb{R}^n$ and that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. To understand the meaning of the quantity $\dot{W}$ in point (b) of the following definition, note that if $\mathbf{x}(t)$ is a trajectory of system (2.1) (corresponding to some initial condition $\mathbf{x}(0) = \mathbf{x}_0$), then the expression $w(t) = W(\mathbf{x}(t))$ defines a differentiable function from a punctured neighborhood of $0 \in \mathbb{R}$ into $\mathbb{R}$, and by the chain rule its derivative is $\mathrm{d}W(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x})$. The expression in (b) can thus be understood as giving the instantaneous rate of change (at $\mathbf{x}$) of the function $W$ along the unique trajectory of (2.1) which is at $\mathbf{x}$ at time zero.

**Definition 2.1.2.** Let $U$ be an open neighborhood of $\mathbf{0} \in \mathbb{R}^n$ and let $W : U \to \mathbb{R}$ be a continuous function that is $C^1$ on $U \setminus \{\mathbf{0}\}$.
(a) $W$ is *positive definite* if $W(\mathbf{0}) = 0$ and $W(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$.
(b) $W$ is a *Lyapunov function* for system (2.1) if it is positive definite and if the function $\dot{W} : U \setminus \{\mathbf{0}\} \to \mathbb{R} : \mathbf{x} \mapsto \mathrm{d}W(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x})$ is nonpositive.
(c) $W$ is a *strict Lyapunov function* for system (2.1) if it is positive definite and if $\dot{W}$ is negative.

**Theorem 2.1.3.** *Let $\Omega$ be an open neighborhood of $\mathbf{0} \in \mathbb{R}^n$, and let $\mathbf{0}$ be an equilibrium for system* (2.1) *on $\Omega$.*

1. *If there exists a Lyapunov function for system* (2.1) *on a neighborhood $U$ of $\mathbf{0}$, then $\mathbf{0}$ is stable.*
2. *If there exists a strict Lyapunov function for system* (2.1) *on a neighborhood $U$ of $\mathbf{0}$, then $\mathbf{0}$ is asymptotically stable.*

*Proof.* Suppose there exists a Lyapunov function $W$ defined on a neighborhood $U$ of $\mathbf{0}$, and let $\varepsilon > 0$ be given. Decrease $\varepsilon$ if necessary so that $\{\mathbf{x} : |\mathbf{x}| \le \varepsilon\} \subset U$, and let $S = \{\mathbf{x} : |\mathbf{x}| = \varepsilon\}$. By continuity of $W$, compactness of $S$, and the fact that $W$ is positive definite, it is clear that $m := \min\{W(\mathbf{x}) : \mathbf{x} \in S\}$ is finite and positive, and that there exists a positive number $\delta < \varepsilon$ such that $M := \max\{W(\mathbf{x}) : |\mathbf{x}| \le \delta\} < m$. We claim that $\delta$ is as required. For fix $\mathbf{x}_1$ such that $0 < |\mathbf{x}_1| \le \delta$, and as usual let $\mathbf{x}_1(t)$ denote the trajectory through $\mathbf{x}_1$. If, contrary to what we wish to establish, there exists a positive value of $t$ for which $|\mathbf{x}_1(t)| = \varepsilon$, then there exists a smallest such value $T$. Then for $0 \le t \le T$, $\mathbf{x}_1(t)$ is in $U \setminus \{\mathbf{0}\}$, hence $v(t) := W(\mathbf{x}_1(t))$ is defined and smooth, and $v'(t) = \dot{W}(\mathbf{x}_1(t)) \le 0$, so that $v(0) \ge v(T)$, in contradiction to the fact that $v(0) \le M < m \le v(T)$.

For our proof of the second point we recall the notion of the omega limit set $\omega(\mathbf{x})$ of a point $\mathbf{x}$ for which $\mathbf{x}(t)$ is defined for all $t \ge 0$: $\omega(\mathbf{x})$ is the set of all points $\mathbf{y}$ for which there exists a sequence $t_1 < t_2 < \cdots$ of numbers such that $t_n \to \infty$ and $\mathbf{x}(t_n) \to \mathbf{y}$ as $n \to \infty$. Note that $\omega(\mathbf{x})$ is closed and invariant under the flow of system (2.1) (see [44] or [140]).

Suppose there exists a strict Lyapunov function $W$ defined on a neighborhood $U$ of $\mathbf{0}$. Since $W$ is also a Lyapunov function, by point (1) $\mathbf{0}$ is stable. Choose $\varepsilon > 0$ so small that $\{\mathbf{x} : |\mathbf{x}| \le \varepsilon\} \subset U$, and let $\delta$ be such that if $|\mathbf{x}| < \delta$, then $\mathbf{x}(t)$ exists for all $t \ge 0$ and satisfies $|\mathbf{x}(t)| < \varepsilon/2$, hence the omega limit set $\omega(\mathbf{x})$ is defined and is a nonempty, compact, connected set ([44, 140]). Note that if $0 < |\mathbf{x}_1| < \delta$, then for all $t \ge 0$, $W(\mathbf{x}_1(t))$ exists and $\dot{W}(\mathbf{x}_1(t)) < 0$.

Fix any $\mathbf{x}_1$ such that $0 < |\mathbf{x}_1| < \delta$. If $\mathbf{a}, \mathbf{b} \in \omega(\mathbf{x}_1)$, then $W(\mathbf{a}) = W(\mathbf{b})$. Indeed, there exist sequences $t_n$ and $s_n$, $t_n \to \infty$ and $s_n \ge 0$, such that both $\mathbf{x}_1(t_n) \to \mathbf{a}$ and $\mathbf{x}_1(t_n + s_n) \to \mathbf{b}$. Since $W$ is continuous and strictly decreases on every trajectory in $\Omega \setminus \{\mathbf{0}\}$, $W(\mathbf{a}) \ge W(\mathbf{b})$. Similarly, $W(\mathbf{b}) \ge W(\mathbf{a})$.

Again, if $\mathbf{a} \in \omega(\mathbf{x}_1)$, then $\mathbf{a}$ must be an equilibrium. For if $|\mathbf{a}| \ne \mathbf{0}$ and $\mathbf{f}(\mathbf{a}) \ne \mathbf{0}$, then $\mathbf{a}(t) \in \Omega \setminus \{\mathbf{0}\}$ when it is defined, and for sufficiently small $\tau > 0$ the time-$\tau$ image $\mathbf{a}(\tau)$ of $\mathbf{a}$ satisfies $\mathbf{a}(\tau) \ne \mathbf{a}$. But if $\mathbf{a} \in \omega(\mathbf{x}_1)$, then $|\mathbf{a}| \le \varepsilon/2 < \varepsilon$, so that $W(\mathbf{a}(\tau)) < W(\mathbf{a})$. Yet $\mathbf{a}(\tau) \in \omega(\mathbf{x}_1)$, hence $W(\mathbf{a}(\tau)) = W(\mathbf{a})$, a contradiction.

Finally, $\mathbf{0}$ is the only equilibrium in $\{\mathbf{x} : |\mathbf{x}| < \delta\}$. For given any $\mathbf{x}$ satisfying $0 < |\mathbf{x}| < \delta$, for sufficiently small $\tau > 0$, $W(\mathbf{x}(\tau))$ is defined and $W(\mathbf{x}(\tau)) < W(\mathbf{x})$, hence $\mathbf{x}(\tau) \ne \mathbf{x}$.

In short, $\omega(\mathbf{x}_1) = \{\mathbf{0}\}$, so $\mathbf{x}_1(t) \to \mathbf{0}$ as $t \to \infty$, as required. □

The following slight generalization of part (2) of Theorem 2.1.3 is sometimes useful (see Example 2.1.6, but also Exercise 2.4).

**Theorem 2.1.4.** *Let $\Omega$ be an open neighborhood of $\mathbf{0} \in \mathbb{R}^n$, and let $\mathbf{0}$ be an equilibrium for system* (2.1) *on $\Omega$. Let $C$ be a smooth curve in $\Omega$ to which $\mathbf{f}$ is nowhere* *tangent except at* $\mathbf{0}$*. If there exists a Lyapunov function* $W$ *for system* (2.1) *on a neighborhood* $U$ *of* $\mathbf{0}$ *such that* $\dot{W}$ *is negative on* $\Omega \setminus C$*, then* $\mathbf{0}$ *is asymptotically stable.*

*Proof.* The proof of part (2) of Theorem 2.1.3 goes through unchanged, since it continues to be true that $W$ is strictly decreasing on every trajectory in $\Omega \setminus \{\mathbf{0}\}$. □

An instability theorem analogous to Theorem 2.1.3 is the following.

**Theorem 2.1.5.** *Let* $\Omega$ *be an open neighborhood of* $\mathbf{0} \in \mathbb{R}^n$*, and let* $\mathbf{0}$ *be an equilibrium for system* (2.1) *on* $\Omega$*.*

1. *If there exists a positive definite function* $W$ *on a neighborhood* $U$ *of* $\mathbf{0}$ *such that* $\dot{W} > 0$ *on* $U \setminus \{\mathbf{0}\}$*, then* $\mathbf{0}$ *is unstable.*
2. *If there exists a function* $W$ *on a neighborhood* $U$ *of* $\mathbf{0}$ *such that* $W(\mathbf{0}) = 0$*,* $\dot{W}$ *is positive definite, and* $W$ *takes a positive value in every neighborhood of* $\mathbf{x}_0$*, then* $\mathbf{x}_0$ *is unstable.*

*Proof.* The proof is left as Exercise 2.5. □

**Example 2.1.6.** Consider a quadratic system on $\mathbb{R}^2$ with an equilibrium at which the linear part has one eigenvalue negative and the other eigenvalue zero. (To say that the system is *quadratic* means that the right-hand sides in (2.1) are polynomials, the maximum of whose degrees is two.) By translating the equilibrium to the origin, performing a linear change of coordinates, and rescaling time, we may place the system in the form

$$\begin{aligned} \dot{x} &= -x + ax^2 + bxy + cy^2 \\ \dot{y} &= \qquad dx^2 + exy + fy^2 . \end{aligned}$$

We will use a Lyapunov function to show that the equilibrium is stable if $f = 0$ and $ce < 0$, or if $c = e = f = 0$. To do so, consider any function $W$ of the form

$$W(x,y) = (Ax^2 + By^2)/2. \tag{2.2}$$

Then

$$\dot{W}(x,y) = -Ax^2 + Aax^3 + (Ab + Bd)x^2y + (Ac + Be)xy^2 + Bfy^3 .$$

Choosing $A = |e|$ and $B = |c|$ if $ce < 0$ and $A = B = 1$ if $c = e = 0$, $W$ is positive definite. When $f = 0$, $\dot{W}$ becomes $\dot{W}(x,y) = -x^2(A - Aax - (Ab + Bd)y)$, hence is nonpositive on a neighborhood of the origin, implying stability. We note that any equilibrium of a planar system at which the linear part has exactly one eigenvalue zero is a node, a (topological) saddle, or a saddle-node (see, for example, Theorem 65, §21 of [12]), so that the equilibrium in question must actually be a stable node, hence be asymptotically stable, a fact that the Lyapunov function does not reveal if only Theorem 2.1.3 is used, but one that is shown by Theorem 2.1.4 when $c \neq 0$ (but not when $c = 0$).

**Example 2.1.7.** Consider a quadratic system on $\mathbb{R}^2$ with an equilibrium at the origin at which the linear part has purely imaginary eigenvalues. By a linear change of coordinates and time rescaling (Exercise 2.7), we may write the system in the form

$$\begin{aligned}\dot{x} &= -y + ax^2 + bxy + cy^2 \\ \dot{y} &= \quad x + dx^2 + exy + fy^2 .\end{aligned} \tag{2.3}$$

If we look for a Lyapunov function in the form (2.2), then

$$\dot{W}(x,y) = (B-A)xy + Aax^3 + (Ab+Bd)x^2y + (Ac+Be)xy^2 + Bfy^3 ,$$

which is nonpositive in a neighborhood of the origin if and only if $A = B$ and

$$a = f = 0, \quad b = -d, \quad c = -e \tag{2.4}$$

(in which case $\dot{W} \equiv 0$). Therefore the origin of system (2.3) is stable if the coefficients satisfy (2.4). The sufficient conditions (2.4) that we have found for stability of the origin of system (2.3) in fact are not necessary, however, but arose simply as a result of our specific choice of the form (2.2) of our trial function $W$. Indeed, consider system (2.3) with $a = f = 1$ and $c = d = e = 0$, that is,

$$\dot{x} = -y + x^2 + bxy, \quad \dot{y} = x + y^2 , \tag{2.5}$$

and in place of (2.2) the far more complicated trial function

$$\begin{aligned}W(x,y) &= x^2 + y^2 \\ &\quad + \frac{2\,(2+b)}{3}x^3 - 2x^2y + 2xy^2 - \frac{4}{3}y^3 \\ &\quad - \frac{\left(26 - 30b^2 - 9b^3\right)}{6\,(4+3b)}x^4 + \frac{4}{3\,(4+3b)}x^3y \\ &\quad - \frac{2\,(1+3b)}{3\,(4+3b)}x^2y^2 + \frac{4\,(1+b)\,(1+3b)}{3\,(4+3b)}xy^3 - \frac{\left(26+36b+9b^2\right)}{6\,(4+3b)}y^4 ,\end{aligned}$$

for which

$$\dot{W}(x,y) = \frac{2(26+30b+9b^2)}{12+9b}x^4 - \frac{4\left(13+13b+3b^2\right)y^4}{12+9b} + o(|x^4+y^4|) .$$

When $\frac{-13-\sqrt{13}}{6} < b < \frac{-13+\sqrt{13}}{6}$, $W$ is a strict Lyapunov function, hence for such $b$ the origin is an asymptotically stable singular point of system (2.5).

Example 2.1.7 illustrates that even for a system that is apparently as simple as system (2.3) it is a difficult problem to find in the space of parameters $\{a,b,c,d,e,f\}$ the subsets corresponding to systems with a stable, unstable, or asymptotically stable singularity at the origin. The theorems of this section do not provide a procedure for resolving this problem. We will study this situation in detail in Chapter 3.

## 2.2 Real Normal Forms

Suppose $\mathbf{x}_0$ is a regular point of system (2.1), that is, a point $\mathbf{x}_0$ at which $\mathbf{f}(\mathbf{x}_0) \neq \mathbf{0}$, and that $\mathbf{f}$ is $C^\infty$ (or real analytic). The Flowbox Theorem (see, for example, §1.7.3 of [44]) states that there is a $C^\infty$ (respectively, real analytic) change of coordinates $\mathbf{x} = \mathbf{H}(\mathbf{y})$ in a neighborhood of $\mathbf{x}_0$ so that with respect to the new coordinates system (2.1) becomes

$$\dot{y}_1 = 1 \qquad \text{and} \qquad \dot{y}_j = 0 \qquad \text{for} \quad 2 \leq j \leq n. \tag{2.6}$$

The Flowbox Theorem confirms the intuitively obvious answers to two questions about regular points, those of *identity* and of *structural stability* or *bifurcation*: regardless of the infinitesimal generator $\mathbf{f}$, the phase portrait of system (2.1) in a neighborhood of a regular point $\mathbf{x}_0$ is topologically equivalent to that of the parallel flow of system (2.6) at any point, and when $\mathbf{f}$ is regarded as an element of any "reasonably" topologized set $V$ of vector fields, there is a neighborhood $N$ of $\mathbf{f}$ in $V$ such that for any $\tilde{\mathbf{f}}$ in $N$, the flows of $\mathbf{f}$ and $\tilde{\mathbf{f}}$ (or we often say $\mathbf{f}$ and $\tilde{\mathbf{f}}$ themselves) are topologically equivalent in a neighborhood of $\mathbf{x}_0$. (Two phase portraits, or the systems of differential equations or vector fields that generate them, are said to be *topologically equivalent* if there is a homeomorphism carrying the orbits of one onto the orbits of the other, preserving the direction of flow along the orbits, but not necessarily their actual parametrizations, say as solutions of the differential equation.) In short, the phase portrait in a neighborhood of a regular point is known (up to diffeomorphism), and there is no bifurcation under sufficiently small perturbation.

Now suppose that $\mathbf{x}_0$ is an equilibrium of system (2.1). As always we assume, by applying a translation of coordinates if necessary, that the equilibrium is located at the origin and that $\mathbf{f}$ is defined and is sufficiently smooth on some open neighborhood of the origin. The questions of identity and of structural stability or bifurcation have equally simple answers in the case that the real parts of the eigenvalues of the linear part $A := d\mathbf{f}(\mathbf{0})$ of $\mathbf{f}$ are nonzero, in which case the equilibrium is called *hyperbolic*. For the Hartman–Grobman Theorem (see for example [44]) states that, in such a case, in a neighborhood of the origin the local flows $\phi(t,\mathbf{x})$ generated by (2.1) and $\psi(t,\mathbf{x})$ generated by the linear system

$$\dot{\mathbf{x}} = A\mathbf{x} \qquad (A = d\mathbf{f}(\mathbf{0})) \tag{2.7}$$

are topologically conjugate: there is a homeomorphism $\mathbf{H}$ of a neighborhood of the origin onto its image such that $\mathbf{H}(\phi(t,\mathbf{x})) = \psi(t,\mathbf{H}(\mathbf{x}))$. (Here, for each $\mathbf{x}_0$, $\phi(t,\mathbf{x}_0)$ (respectively, $\psi(t,\mathbf{x}_0)$) is the unique solution $\mathbf{x}(t)$ of (2.1) (respectively, of (2.7)) satisfying the initial condition $\mathbf{x}(0) = \mathbf{x}_0$.) Because the homeomorphism $\mathbf{H}$ thus carries trajectories of system (2.1) onto those of system (2.7), preserving their sense (in fact, their parametrization), it is a fortiori a topological equivalence between the full system (2.1) and the linear system (2.7) in a neighborhood of $\mathbf{0}$, the latter of which is explicitly known. Furthermore, if $\mathbf{f}$ is an element of a set $V$ of vector fields that is topologized in such a way that eigenvalues of linear parts depend continuously on the vector fields, then system (2.1) is structurally stable as well.

The real situation of interest then is that of an isolated equilibrium of system (2.1) that is nonhyperbolic. The identity problem is most fully resolved in dimension two. If there is a characteristic direction of approach to the equilibrium, then apart from certain exceptional cases, the topological type of the equilibrium can be found by means of a finite sequence of "blow-ups" and is determined by a finite initial segment of the Taylor series expansion of $\mathbf{f}$ at $\mathbf{0}$ (see [65]). In higher dimensions, results on the topological type of degenerate equilibria are less complete.

Normal form theory enters in when we wish to understand the bifurcations of system (2.1) in a neighborhood of a nonhyperbolic equilibrium. Supposing that we have solved the identity problem (although in actual practice an initial normal form computation may be done in order to simplify the identity problem for the original system), we wish to know what phase portraits are possible in a neighborhood of $\mathbf{0}$, for any vector field in a neighborhood of $\mathbf{f}$ in some particular family $V$ of vector fields, with a given topology. The idea of normal form theory quite simply is to perform a change of coordinates $\mathbf{x} = \mathbf{H}(\mathbf{y})$, or a succession of coordinate transformations, so as to place the original system (2.1) into a form most amenable to study. Typically, this means eliminating as many terms as possible from an initial segment of the power series expansion of $\mathbf{f}$ at the origin.

It is useful to compare this idea to an application of the Hartman–Grobman Theorem. Although the homeomorphism $\mathbf{H}$ guaranteed by the Hartman–Grobman Theorem can rightly be regarded as a change of coordinates in a neighborhood of the origin, in that it is invertible, it would be incorrect to say that system (2.1) has been transformed into system (2.7) by the change of coordinates $\mathbf{H}$, since $\mathbf{H}$ could fail to be smooth. In fact, for some choices of $\mathbf{f}$ in (2.1) it is impossible to choose $\mathbf{H}$ to be smooth. An example for which this happens is instructive and will lay the groundwork for the general approach in the nonhyperbolic case.

**Example 2.2.1.** Consider the linear system

$$\begin{aligned} \dot{x}_1 &= 2x_1 \\ \dot{x}_2 &= \phantom{2}x_2, \end{aligned} \tag{2.8}$$

which has a hyperbolic equilibrium at the origin, and the general quadratic system with the same linear part,

$$\begin{aligned} \dot{x}_1 &= 2x_1 + \phantom{'}ax_1^2 + \phantom{'}bx_1x_2 + \phantom{'}cx_2^2 \\ \dot{x}_2 &= \phantom{2}x_2 + a'x_1^2 + b'x_1x_2 + c'x_2^2 \,. \end{aligned} \tag{2.9}$$

We make a $C^2$ change of coordinates $\mathbf{x} = \mathbf{H}(\mathbf{y}) = \mathbf{h}^{(0)}(\mathbf{y}) + \mathbf{h}^{(1)}(\mathbf{y}) + \mathbf{h}^{[2]}(\mathbf{y})$, where $\mathbf{h}^{(0)}(\mathbf{y})$ denotes the constant terms, $\mathbf{h}^{(1)}(\mathbf{y})$ the linear terms in $y_1$ and $y_2$, and $\mathbf{h}^{[2]}(\mathbf{y})$ *all* the remaining terms. Our goal is to eliminate all the quadratic terms in (2.9). To keep the equilibrium situated at the origin, we choose $\mathbf{h}^{(0)}(\mathbf{y}) = \mathbf{0}$, and to maintain the same linear part, which is already in "simplest" form, namely Jordan normal form, we choose $\mathbf{h}^{(1)}(\mathbf{y}) = \mathbf{y}$. Thus the change of coordinates is

$$\mathbf{x} = \mathbf{y} + \mathbf{h}^{[2]}(\mathbf{y}),$$

whence

$$\dot{\mathbf{x}} = \dot{\mathbf{y}} + d\mathbf{h}^{[2]}(\mathbf{y})\dot{\mathbf{y}} = (\mathbf{Id} + d\mathbf{h}^{[2]}(\mathbf{y}))\dot{\mathbf{y}}, \tag{2.10}$$

where for the $n$-dimensional vector function $\mathbf{u}$ we denote by $d\mathbf{u}$ the Jacobian matrix

$$\begin{pmatrix} \frac{\partial u_1}{\partial y_1} & \cdots & \frac{\partial u_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_n}{\partial y_1} & \cdots & \frac{\partial u_n}{\partial y_n} \end{pmatrix}.$$

For $\mathbf{y}$ sufficiently close to $\mathbf{0}$ the geometric series $\sum_{k=0}^{\infty}(-d\mathbf{h}^{[2]}(\mathbf{y}))^k$ converges in the real vector space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ with the uniform norm $\|\mathbf{T}\| = \max\{|\mathbf{T}\mathbf{x}| : |\mathbf{x}| \le 1\}$. Therefore, the linear transformation $\mathbf{Id} + d\mathbf{h}^{[2]}(\mathbf{y})$ is invertible and $(\mathbf{Id} + d\mathbf{h}^{[2]}(\mathbf{y}))^{-1} = \mathbf{Id} - d\mathbf{h}^{[2]}(\mathbf{y}) + \cdots$, so that (2.10) yields

$$\dot{\mathbf{y}} = (\mathbf{Id} + d\mathbf{h}^{[2]}(\mathbf{y}))^{-1}\dot{\mathbf{x}} = (\mathbf{Id} - d\mathbf{h}^{[2]}(\mathbf{y}) + \cdots)\dot{\mathbf{x}}. \tag{2.11}$$

Writing $\mathbf{h}^{[2]}$ as

$$\mathbf{h}^{[2]}(\mathbf{y}) = \begin{pmatrix} a_{20}y_1^2 + a_{11}y_1y_2 + a_{02}y_2^2 + \cdots \\ b_{20}y_1^2 + b_{11}y_1y_2 + b_{02}y_2^2 + \cdots \end{pmatrix},$$

we have

$$d\mathbf{h}^{[2]}(\mathbf{y}) = \begin{pmatrix} 2a_{20}y_1 + a_{11}y_2 + \cdots & a_{11}y_1 + 2a_{02}y_2 + \cdots \\ 2b_{20}y_1 + b_{11}y_2 + \cdots & b_{11}y_1 + 2b_{02}y_2 + \cdots \end{pmatrix},$$

so that (2.11) is

$$\begin{aligned}
\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} &= \begin{pmatrix} 1 - 2a_{20}y_1 - a_{11}y_2 - \cdots & -a_{11}y_1 - 2a_{02}y_2 - \cdots \\ -2b_{20}y_1 - b_{11}y_2 - \cdots & 1 - b_{11}y_1 - 2b_{02}y_2 - \cdots \end{pmatrix} \\
&\quad \times \begin{pmatrix} 2(y_1 + a_{20}y_1^2 + a_{11}y_1y_2 + a_{02}y_2^2 + \cdots) \\ \qquad + a(y_1 + \cdots)^2 + b(y_1 + \cdots)(y_2 + \cdots) + c(y_2 + \cdots)^2 \\ y_2 + b_{20}y_1^2 + b_{11}y_1y_2 + b_{02}y_2^2 + \cdots \\ \qquad + a'(y_1 + \cdots)^2 + b'(y_1 + \cdots)(y_2 + \cdots) + c'(y_2 + \cdots)^2 \end{pmatrix} \\
&= \begin{pmatrix} 2y_1 + (a - 2a_{20})y_1^2 + (b - a_{11})y_1y_2 + cy_2^2 + \cdots \\ y_2 + (a' - 3b_{20})y_1^2 + (b' - 2b_{11})y_1y_2 + (c' + b_{02})y_2^2 + \cdots \end{pmatrix}.
\end{aligned}$$

From this last expression it is clear that suitable choices of the coefficients $a_{ij}$ and $b_{ij}$ exist to eliminate all the quadratic terms in the transformed system except for the $y_2^2$ term in $\dot{y}_1$. In short, no matter what our choice of the transformation $\mathbf{H}(\mathbf{y})$, if $\mathbf{H}$ is twice differentiable, then system (2.9) can be simplified to

$$\begin{aligned} \dot{y}_1 &= 2y_1 + cy_2^2 \\ \dot{y}_2 &= y_2 \end{aligned} \tag{2.12}$$

through order two, but the quadratic term in the first component cannot be removed; (2.12) is the *normal form* for system (2.9) through order two.

We now turn our attention to the general situation in which system (2.1) has an isolated equilibrium at the origin, not necessarily hyperbolic. We will need the following notation and terminology. As in Chapter 1, for $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$, $\mathbf{x}^\alpha$ denotes $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. We let $\mathscr{H}_s$ denote the vector space of functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ each of whose components is a homogeneous polynomial function of degree $s$; elements of $\mathscr{H}_s$ will be termed vector homogeneous functions. If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, $\mathbf{e}_j = (0, \dots, 0, \overset{j}{1}, 0, \dots, 0)^T$, then a basis for $\mathscr{H}_s$ is the collection of vector homogeneous functions

$$\mathbf{v}_{j,\alpha} = \mathbf{x}^\alpha \mathbf{e}_j \tag{2.13}$$

for all $j$ such that $1 \le j \le n$ and all $\alpha$ such that $|\alpha| = s$ (this is the product of a monomial and a vector). For example, the first three vectors listed in the basis for $\mathscr{H}_2$ given in Example 2.2.2 below are $\mathbf{v}_{1,(2,0)}$, $\mathbf{v}_{1,(1,1)}$, and $\mathbf{v}_{1,(0,2)}$. Thus $\mathscr{H}_s$ has dimension $N = nC(s+n-1, s)$ (see Exercise 2.6).

Assuming that $\mathbf{f}$ is $C^2$, we expand $\mathbf{f}$ in a Taylor series so as to write (2.1) as

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{f}^{(2)}(\mathbf{x}) + \mathbf{R}(\mathbf{x}), \tag{2.14}$$

where $\mathbf{f}^{(2)} \in \mathscr{H}_2$ and the remainder satisfies the condition that $|\mathbf{R}(\mathbf{x})|/|\mathbf{x}|^2 \to 0$ as $|\mathbf{x}| \to 0$. We assume that $A$ has been placed in a standard form by a preliminary linear transformation, so that in practical situations the nonlinear terms may be different from what they were in the system as originally encountered.

Applying the reasoning of Example 2.2.1, we make a coordinate transformation of the form

$$\mathbf{x} = \mathbf{H}(\mathbf{y}) = \mathbf{y} + \mathbf{h}^{(2)}(\mathbf{y}), \tag{2.15}$$

where in this case $\mathbf{h}^{(2)} \in \mathscr{H}_2$, that is, each of the $n$ components of $\mathbf{h}^{(2)}(\mathbf{y})$ is a homogeneous quadratic polynomial in $\mathbf{x}$. Since $d\mathbf{H}(\mathbf{0}) = \mathbf{Id}$ is invertible, the Inverse Function Theorem guarantees that $\mathbf{H}$ has an analytic inverse on a neighborhood of $\mathbf{0}$. Using (2.15) in the right-hand side of (2.14) and inserting that in turn into the analogue of (2.11) (that is, into (2.11) with $\mathbf{h}^{[2]}$ replaced by $\mathbf{h}^{(2)}$) yields

$$\dot{\mathbf{y}} = A\mathbf{y} + A\mathbf{h}^{(2)}(\mathbf{y}) + \mathbf{f}^{(2)}(\mathbf{y}) - d\mathbf{h}^{(2)}(\mathbf{y})A\mathbf{y} + \mathbf{R}_2(\mathbf{y}), \tag{2.16}$$

where the remainder satisfies the condition $|\mathbf{R}_2(\mathbf{y})|/|\mathbf{y}|^2 \to 0$ as $|\mathbf{y}| \to 0$. The quadratic terms can be eliminated from (2.16) if and only if $\mathbf{h}^{(2)}(\mathbf{y})$ can be chosen so that

$$\mathscr{L}\mathbf{h}^{(2)}(\mathbf{y}) = d\mathbf{h}^{(2)}(\mathbf{y})A\mathbf{y} - A\mathbf{h}^{(2)}(\mathbf{y}) = \mathbf{f}^{(2)}(\mathbf{y}), \tag{2.17}$$

where $\mathscr{L}$, the so-called *homological operator*, is the linear operator on $\mathscr{H}_2$ defined by

$$\mathscr{L} : \mathbf{p}(\mathbf{y}) \mapsto d\mathbf{p}(\mathbf{y})A\mathbf{y} - A\mathbf{p}(\mathbf{y}). \tag{2.18}$$

In other words, all quadratic terms can be eliminated from (2.16) if and only if $\mathscr{L}$ maps $\mathscr{H}_2$ *onto* itself. If $\mathscr{L}$ is not onto, then because $\mathscr{H}_2$ is finite-dimensional, it decomposes as a direct sum $\mathscr{H}_2 = \text{Image}(\mathscr{L}) \oplus \mathscr{K}_2$, where Image($\mathscr{L}$) denotes the image of $\mathscr{L}$ in $\mathscr{H}_2$, although the complementary subspace $\mathscr{K}_2$ is not unique. The quadratic terms in (2.16) that can be eliminated by a $C^2$ change of coordinates are precisely those that lie in Image($\mathscr{L}$). Those that remain have a form dependent on the choice of the complementary subspace $\mathscr{K}_2$.

**Example 2.2.2.** Let us reconsider the system (2.9) of Example 2.2.1 in this context. The basis (2.13) for $\mathscr{H}_2$ is

$$\left\{ \begin{pmatrix} x_1^2 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1x_2 \\ 0 \end{pmatrix}, \begin{pmatrix} x_2^2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ x_1^2 \end{pmatrix}, \begin{pmatrix} 0 \\ x_1x_2 \end{pmatrix}, \begin{pmatrix} 0 \\ x_2^2 \end{pmatrix} \right\},$$

which we order by the order in which we have listed the basis elements, and which for ease of exposition we label $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$, $\mathbf{u}_5$, and $\mathbf{u}_6$. A straightforward computation based on the definition (2.18) of the homological operator $\mathscr{L}$ yields

$$\mathscr{L}\begin{pmatrix} y_1^{\alpha_1}y_2^{\alpha_2} \\ 0 \end{pmatrix} = (2\alpha_1 + \alpha_2 - 2)\begin{pmatrix} y_1^{\alpha_1}y_2^{\alpha_2} \\ 0 \end{pmatrix} \tag{2.19a}$$

and

$$\mathscr{L}\begin{pmatrix} 0 \\ y_1^{\alpha_1}y_2^{\alpha_2} \end{pmatrix} = (2\alpha_1 + \alpha_2 - 1)\begin{pmatrix} 0 \\ y_1^{\alpha_1}y_2^{\alpha_2} \end{pmatrix}, \tag{2.19b}$$

so that each basis vector is an eigenvector. The eigenvalues are, in the order of the basis vectors to which they correspond, 2, 1, 0, 3, 2, and 1. Image($\mathscr{L}$) is thus the five-dimensional subspace of $\mathscr{H}_2$ spanned by the basis vectors other than $\mathbf{u}_3$, and a natural complement to Image($\mathscr{L}$) is Span$\{\mathbf{u}_3\}$, corresponding precisely to the quadratic term in (2.12).

Returning to the general situation, beginning with (2.1), written in the form (2.14), we compute the operator $\mathscr{L}$ of (2.18), choose a complement $\mathscr{K}_2$ to Image($\mathscr{L}$) in $\mathscr{H}_2$, and decompose $\mathbf{f}^{(2)}$ as $\mathbf{f}^{(2)} = (\mathbf{f}^{(2)})_0 + \tilde{\mathbf{f}}^{(2)} \in \text{Image}(\mathscr{L}) \oplus \mathscr{K}_2$. Then for any $\mathbf{h}^{(2)} \in \mathscr{H}_2$ satisfying $\mathscr{L}\mathbf{h}^{(2)} = (\mathbf{f}^{(2)})_0$, by (2.16) the change of coordinates (2.15) reduces (2.1) (which is the same as (2.14)) to

$$\dot{\mathbf{y}} = A\mathbf{y} + \tilde{\mathbf{f}}^{(2)}(\mathbf{y}) + \tilde{\mathbf{R}}_2(\mathbf{y}), \tag{2.20}$$

where, to repeat, $\tilde{\mathbf{f}}^{(2)} \in \mathscr{K}_2$, and $|\tilde{\mathbf{R}}_2(\mathbf{y})|/|\mathbf{y}|^2 \to 0$ as $|\mathbf{y}| \to 0$. This is the normal form for (2.1) through order two: the quadratic terms have been simplified as much as possible.

Turning to the cubic terms, and assuming one more degree of differentiability, let us return to $\mathbf{x}$ for the current coordinates and write (2.20) as

$$\dot{\mathbf{x}} = A\mathbf{x} + \tilde{\mathbf{f}}^{(2)}(\mathbf{x}) + \mathbf{f}^{(3)}(\mathbf{x}) + \mathbf{R}(\mathbf{x}), \tag{2.21}$$

where $\mathbf{R}$ denotes a new remainder term, satisfying the condition $|\mathbf{R}(\mathbf{x})|/|\mathbf{x}|^3 \to 0$ as $|\mathbf{x}| \to 0$. A change of coordinates that will leave the constant, linear, and quadratic terms in (2.21) unchanged is one of the form

$$\mathbf{x} = \mathbf{y} + \mathbf{h}^{(3)}(\mathbf{y}), \tag{2.22}$$

where $\mathbf{h}^{(3)} \in \mathscr{H}_3$, that is, each of the $n$ components of $\mathbf{h}^{(3)}(\mathbf{y})$ is a homogeneous cubic polynomial function in $\mathbf{y}$. Using (2.22) in the right-hand side of (2.21) and inserting that in turn into the analogue of (2.11) (that is, into (2.11) with $\mathbf{h}^{[2]}$ replaced by $\mathbf{h}^{(3)}$) yields

$$\dot{\mathbf{y}} = A\mathbf{y} + \tilde{\mathbf{f}}^{(2)}(\mathbf{y}) + A\mathbf{h}^{(2)}(\mathbf{y}) + \mathbf{f}^{(3)}(\mathbf{y}) - d\mathbf{h}^{(3)}(\mathbf{y})A\mathbf{y} + \mathbf{R}_3(\mathbf{y}), \tag{2.23}$$

where the remainder satisfies the condition $|\mathbf{R}_3(\mathbf{y})|/|\mathbf{y}|^2 \to 0$ as $|\mathbf{y}| \to 0$. The cubic terms can be eliminated from (2.23) if and only if $\mathbf{h}^{(3)}(\mathbf{y})$ can be chosen so that

$$\mathscr{L}\mathbf{h}^{(3)}(\mathbf{y}) = d\mathbf{h}^{(3)}(\mathbf{y})A\mathbf{y} - A\mathbf{h}^{(3)}(\mathbf{y}) = \mathbf{f}^{(3)}(\mathbf{y}). \tag{2.24}$$

Comparing (2.17) and (2.24), the reader can see that the condition for the elimination of all cubic terms is exactly the same as that for the elimination of all quadratic terms. The homological operator $\mathscr{L}$ is again defined by (2.18), except that it now operates on the vector space $\mathscr{H}_3$ of functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ all of whose components are homogeneous *cubic* polynomial functions. If $\mathscr{L}$ does not map onto $\mathscr{H}_3$, then exactly as for $\mathscr{H}_2$ when $\mathscr{L}$ does not map onto $\mathscr{H}_2$, $\mathscr{H}_3$ decomposes as a direct sum $\mathscr{H}_3 = \text{Image}(\mathscr{L}) \oplus \mathscr{K}_3$, although again the complementary subspace $\mathscr{K}_3$ is not unique. Once we have chosen a complement $\mathscr{K}_3$, $\mathbf{f}^{(3)}$ decomposes as $\mathbf{f}^{(3)} = (\mathbf{f}^{(3)})_0 + \tilde{\mathbf{f}}^{(3)} \in \text{Image}(\mathscr{L}) \oplus \mathscr{K}_3$. Then for any $\mathbf{h}^{(3)} \in \mathscr{H}_3$ satisfying $\mathscr{L}\mathbf{h}^{(3)} = (\mathbf{f}^{(3)})_0$, by (2.23) the change of coordinates (2.22) reduces (2.1) (which is the same as (2.21)) to

$$\dot{\mathbf{y}} = A\mathbf{y} + \tilde{\mathbf{f}}^{(2)}(\mathbf{y}) + \tilde{\mathbf{f}}^{(3)}(\mathbf{y}) + \tilde{\mathbf{R}}_3(\mathbf{y}), \tag{2.25}$$

where $\tilde{\mathbf{f}}^{(3)} \in \mathscr{K}_3$, and $|\tilde{\mathbf{R}}_3(\mathbf{y})|/|\mathbf{y}|^3 \to 0$ as $|\mathbf{y}| \to 0$. This is the normal form for (2.1) through order three: the quadratic and cubic terms have been simplified as much as possible.

It is apparent that the pattern continues through all orders, as long as $\mathbf{f}$ is sufficiently differentiable. Noting that a composition of transformations of the form $\mathbf{x} = \mathbf{y} + \mathbf{p}(\mathbf{y})$, where each component of $\mathbf{p}(\mathbf{y})$ is a polynomial function, is a transformation of the same form and has an analytic inverse on a neighborhood of $\mathbf{0}$, we have the following theorem.

**Theorem 2.2.3.** *Let* $\mathbf{f}$ *be defined and* $C^r$ *on a neighborhood of* $\mathbf{0}$ *in* $\mathbb{R}^n$ *and satisfy* $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. *Let* $A = d\mathbf{f}(\mathbf{0})$. *For* $2 \le k \le r$, *let* $\mathscr{H}_k$ *denote the vector space of functions from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *all of whose components are homogeneous polynomial functions of degree* $k$, *let* $\mathscr{L}$ *denote the linear operator on* $\mathscr{H}_k$ *(the "homological operator") defined by* $\mathscr{L}\mathbf{p}(\mathbf{y}) = d\mathbf{p}(\mathbf{y})A\mathbf{y} - A\mathbf{p}(\mathbf{y})$, *and let* $\mathscr{K}_k$ *be any complement to* $\text{Image}(\mathscr{L})$ *in* $\mathscr{H}_k$, *so that* $\mathscr{H}_k = \text{Image}(\mathscr{L}) \oplus \mathscr{K}_k$. *Then there is a polynomial change of coordinates* $\mathbf{x} = \mathbf{H}(\mathbf{y}) = \mathbf{y} + \mathbf{p}(\mathbf{y})$ *such that in the new coordinates system* (2.1) *is*

$$\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{f}^{(2)}(\mathbf{y}) + \cdots + \mathbf{f}^{(r)}(\mathbf{y}) + \mathbf{R}(\mathbf{y}), \tag{2.26}$$

*where for* $2 \le k \le r$, $\mathbf{f}^{(k)} \in \mathscr{H}_k$, *and the remainder* $\mathbf{R}$ *satisfies* $|\mathbf{R}(\mathbf{y})|/|\mathbf{y}|^r \to 0$ *as* $|\mathbf{y}| \to 0$.

**Definition 2.2.4.** In the context of Theorem 2.2.3, expression (2.26) is a *normal form through order r* for system (2.1).

## 2.3 Analytic and Formal Normal Forms

In this section we study in detail the homological operator $\mathscr{L}$ and normal forms of the system

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{X}(\mathbf{x}), \tag{2.27}$$

where now $\mathbf{x} \in \mathbb{C}^n$, $A$ is a possibly complex $n \times n$ matrix, and each component $X_k(\mathbf{x})$ of $\mathbf{X}$, $1 \le k \le n$, is a formal or convergent power series, possibly with complex coefficients, that contains no constant or linear terms. Our treatment mainly follows the lines of [19].

To see why this is of importance, even when our primary interest is in real systems, recall that the underlying assumption in the discussion leading up to Theorem 2.2.3, which was explicitly stated in the first sentence following (2.14), was that the linear terms in the right-hand side of (2.1), that is, the $n \times n$ matrix $A$ in (2.14), had already been placed in some standard form by a preliminary linear transformation. To elaborate on this point, typically at the beginning of an investigation of system (2.1), expressed as (2.14), the matrix $A$ has no particularly special form. From linear algebra we know that there exists a nonsingular $n \times n$ matrix $S$ such that the similiarity transformation $SAS^{-1} = J$ produces the Jordan normal form $J$ of $A$. If we use the matrix $S$ to make the linear coordinate transformation

$$\mathbf{y} = S\mathbf{x} \tag{2.28}$$

of phase space, then in the new coordinates (2.14) becomes

$$\dot{\mathbf{y}} = J\mathbf{y} + S\mathbf{f}^{(2)}(S^{-1}\mathbf{y}) + S\mathbf{R}(S^{-1}\mathbf{y}). \tag{2.29}$$

Although the original system (2.1) or (2.14) is real, the matrices $J$ and $S$ can be complex, hence system (2.29) can be complex as well. Thus even if we are primarily interested in studying real systems, it is nevertheless fruitful to investigate normal forms of complex systems (2.27). Since we are working with systems whose right-hand sides are power series, we will also allow formal rather than convergent series as well.

Whereas previously $\mathscr{H}_s$ denoted the vector space of functions from $\mathbb{R}^n$ to $\mathbb{R}^n$, all of whose components are homogeneous polynomial functions of degree $s$, we now let $\mathscr{H}_s$ denote the vector space of functions from $\mathbb{C}^n$ to $\mathbb{C}^n$, all of whose components are homogeneous polynomial functions of degree $s$. The collection $\mathbf{v}_{j,\alpha}$ of (2.13) remains a basis of $\mathscr{H}_s$. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $\kappa = (\kappa_1, \dots, \kappa_n) \in \mathbb{C}^n$ we will let $(\alpha, \kappa)$ denote the scalar product

$$(\alpha, \kappa) = \sum_{j=1}^{n} \alpha_j \kappa_j \,.$$

**Lemma 2.3.1.** *Let $A$ be an $n \times n$ matrix with eigenvalues $\kappa_1, \dots, \kappa_n$, and let $\mathscr{L}$ be the corresponding homological operator on $\mathscr{H}_s$, that is, the linear operator on $\mathscr{H}_s$ defined by*

$$\mathscr{L}\mathbf{p}(\mathbf{y}) = \mathrm{d}\mathbf{p}(\mathbf{y})A\mathbf{y} - A\mathbf{p}(\mathbf{y}) \,. \tag{2.30}$$

*Let $\kappa = (\kappa_1, \dots, \kappa_n)$. Then the eigenvalues $\lambda_j$, $i = j, \dots, N$, of $\mathscr{L}$ are*

$$\lambda_j = (\alpha, \kappa) - \kappa_m,$$

*where $m$ ranges over $\{1, \dots, n\} \subset \mathbb{N}$ and $\alpha$ ranges over $\{\beta \in \mathbb{N}_0^n : |\beta| = s\}$.*

*Proof.* For ease of exposition, just for this paragraph let $\mathscr{T}$ denote the linear transformation of $\mathbb{C}^n$ whose matrix representative with respect to the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of $\mathbb{C}^n$ is $A$. There exists a nonsingular $n \times n$ matrix $S$ such that $J = SAS^{-1}$ is the lower-triangular Jordan form of $A$ (omitted entries are zero),

$$J = \begin{pmatrix} \kappa_1 & & & & & \\ \sigma_2 & \kappa_2 & & & & \\ & \sigma_3 & \kappa_3 & & & \\ & & \sigma_4 & \kappa_4 & & \\ & & & \ddots & \ddots & \\ & & & & \sigma_n & \kappa_n \end{pmatrix},$$

where $\kappa_1$ through $\kappa_n$ are the eigenvalues of $A$ (repeated eigenvalues grouped together) and $\sigma_j \in \{0, 1\}$ for $2 \le j \le n$. We will compute the eigenvalues of $\mathscr{L}$ by finding its $N \times N$ matrix representative $L$ with respect to a basis of $\mathscr{H}_s$ corresponding to the new basis of $\mathbb{C}^n$ in which $J$ is the matrix of $\mathscr{T}$. This corresponds to the change of coordinates $\mathbf{x} = S\mathbf{y}$ in $\mathbb{C}^n$.

In Exercise 2.11 the reader is led through a derivation of the fact that with respect to the new coordinates the expression for $\mathscr{L}$ changes to (2.30) with $A$ replaced by $J$, that is,

$$\mathscr{L}\mathbf{p}(\mathbf{y}) = \mathrm{d}\mathbf{p}(\mathbf{y})J\mathbf{y} - J\mathbf{p}(\mathbf{y}) \,. \tag{2.31}$$

Column $(j, \alpha)$ of $L$ is the coordinate vector of the image under $\mathscr{L}$ of the basis vector $\mathbf{v}_{j,\alpha}$ of (2.13). Computing directly from (2.31), $\mathscr{L}\mathbf{v}_{j,\alpha}(\mathbf{y})$ is (omitted entries are zeros)

$$
\begin{aligned}
\mathscr{L}\mathbf{v}_{j,\alpha}(\mathbf{y}) &= \begin{pmatrix} \\ \alpha_1 y_1^{\alpha_1-1} y_2^{\alpha_2}\cdots y_n^{\alpha_n} \;\cdots\; \alpha_n y_1^{\alpha_1} y_2^{\alpha_2}\cdots y_n^{\alpha_n-1} \\ \\ \end{pmatrix} \begin{pmatrix} \kappa_1 y_1 \\ \kappa_2 y_2 + \sigma_2 y_1 \\ \kappa_3 y_3 + \sigma_3 y_2 \\ \vdots \\ \kappa_n y_n + \sigma_n y_{n-1} \end{pmatrix} \\
&\qquad - \begin{pmatrix} \\ \kappa_j y_1^{\alpha_1} y_2^{\alpha_2}\cdots y_n^{\alpha_n} \\ \sigma_{j+1} y_1^{\alpha_1} y_2^{\alpha_2}\cdots y_n^{\alpha_n} \\ \\ \end{pmatrix} \\
&= \left[ ((\alpha,\kappa)-\kappa_j))\mathbf{y}^{\alpha} + \sum_{i=2}^{n} \sigma_i \alpha_i y_1^{\alpha_1}\cdots y_{i-1}^{\alpha_{i-1}+1} y_i^{\alpha_i-1}\cdots y_n^{\alpha_n} \right] \mathbf{e}_j \\
&\qquad + \sigma_{j+1}\mathbf{y}^{\alpha}\mathbf{e}_{j+1} \\
&= ((\alpha,\kappa)-\kappa_j))\mathbf{v}_{j,\alpha}(\mathbf{y}) + \sum_{i=2}^{n} \sigma_i \alpha_i \mathbf{v}_{j,(\alpha_1,\ldots,\alpha_{i-1}+1,\alpha_i-1,\ldots,\alpha_n)}(\mathbf{y}) \\
&\qquad + \sigma_{j+1}\mathbf{v}_{j+1,\alpha}(\mathbf{y})\,.
\end{aligned}
$$

If the basis of $\mathscr{H}_s$ is ordered so that $\mathbf{v}_{r,\beta}$ precedes $\mathbf{v}_{s,\gamma}$ if and only if the first nonzero entry (reading left to right) in the row vector $(r-s, \beta-\gamma)$ is negative, then the basis vector $\mathbf{v}_{j,\alpha}$ precedes all the remaining vectors in the expression for $\mathscr{L}\mathbf{v}_{j,\alpha}$. This implies that the corresponding $N \times N$ matrix $L$ for $\mathscr{L}$ is lower triangular, and has the numbers $(\alpha,\kappa)-\kappa_m$ ($|\alpha| = s$, $1 \le m \le n$) on the main diagonal. □

The order of the basis of $\mathscr{H}_s$ referred to in the proof of Lemma 2.3.1 is the lexicographic order. See Exercise 2.12.

We say that our original system (2.27) under consideration is *formally equivalent* to a like system

$$\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{Y}(\mathbf{y}) \tag{2.32}$$

if there is a change of variables

$$\mathbf{x} = \mathbf{H}(\mathbf{y}) = \mathbf{y} + \mathbf{h}(\mathbf{y}) \tag{2.33}$$

that transforms (2.27) into (2.32), where the coordinate functions of $\mathbf{Y}$ and $\mathbf{h}$, $Y_j$ and $h_j$, $j = 1,\ldots,n$, are formal power series. (Of course, in this context it is natural to allow the coordinate functions $X_j$ of $\mathbf{X}$ to be merely formal series as well.) If all $Y_j$ and $h_j$ are convergent power series (and all $X_j$ are as well), then by the Inverse Function Theorem the transformation (2.33) has an analytic inverse on a neighborhood of $\mathbf{0}$ and we say that (2.27) and (2.32) are *analytically equivalent*. (See the paragraph following Corollary 6.1.3 for comments on the convergence of power series of several variables.)

We alert the reader to the fact that, as indicated by the notation introduced in (2.33), $\mathbf{h}$ stands for just the terms of order at least two in the equivalence transformation $\mathbf{H}$ between the two systems.

Lemma 2.3.1 yields the following theorem.

**Theorem 2.3.2.** *Let $\kappa_1, \dots, \kappa_n$ be the eigenvalues of the $n \times n$ matrix $A$ in* (2.27) *and* (2.32)*, set $\kappa = (\kappa_1, \dots, \kappa_n)$, and suppose that*

$$(\alpha, \kappa) - \kappa_m \neq 0 \tag{2.34}$$

*for all $m \in \{1, \dots, n\}$ and for all $\alpha \in \mathbb{N}_0^n$ for which $|\alpha| \geq 2$. Then systems* (2.27) *and* (2.32) *are formally equivalent for all* $\mathbf{X}$ *and* $\mathbf{Y}$*, and the equivalence transformation* (2.33) *is uniquely determined by* $\mathbf{X}$ *and* $\mathbf{Y}$.

*Proof.* Differentiating (2.33) with respect to $t$ and applying (2.27) and (2.32) yields the condition

$$d\mathbf{h}(\mathbf{y})A\mathbf{y} - A\mathbf{h}(\mathbf{y}) = \mathbf{X}(\mathbf{y} + \mathbf{h}(\mathbf{y})) - d\mathbf{h}(\mathbf{y})\mathbf{Y}(\mathbf{y}) - \mathbf{Y}(\mathbf{y}) \tag{2.35}$$

that $\mathbf{h}$ must satisfy. We determine $\mathbf{h}$ by a recursive process like that leading up to Theorem 2.2.3.

Decomposing $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{h}$ as the sum of their homogeneous parts,

$$\mathbf{X} = \sum_{s=2}^{\infty} \mathbf{X}^{(s)}, \qquad \mathbf{Y} = \sum_{s=2}^{\infty} \mathbf{Y}^{(s)}, \qquad \mathbf{h} = \sum_{s=2}^{\infty} \mathbf{h}^{(s)}, \tag{2.36}$$

where $\mathbf{X}^{(s)}, \mathbf{Y}^{(s)}, \mathbf{h}^{(s)} \in \mathscr{H}_s$, (2.35) decomposes into the infinite sequence of equations

$$\mathscr{L}(\mathbf{h}^{(s)}) = \mathbf{g}^{(s)}(\mathbf{h}^{(2)}, \dots, \mathbf{h}^{(s-1)}, \mathbf{Y}^{(2)}, \dots, \mathbf{Y}^{(s-1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(s)}) - \mathbf{Y}^{(s)}, \tag{2.37}$$

for $s = 2, 3, \dots$, where $\mathbf{g}^{(s)}$ denotes the function that is obtained after the substitution into $\mathbf{X}(\mathbf{y} + \mathbf{h}(\mathbf{y})) - d\mathbf{h}(\mathbf{y})\mathbf{Y}(\mathbf{y})$ of the expression $\mathbf{y} + \sum_{i=1}^{s} \mathbf{h}^{(i)}$ in the place of $\mathbf{y} + \mathbf{h}(\mathbf{y})$ and the expression $\sum_{i=1}^{s} \mathbf{Y}^{(i)}(\mathbf{y})$ in the place of $\mathbf{Y}(\mathbf{y})$, and maintaining only terms that are of order $s$. For $s = 2$, the right-hand side of (2.37) is to be understood to stand for $\mathbf{X}^{(2)}(\mathbf{y}) - \mathbf{Y}^{(2)}(\mathbf{y})$, which is known. For $s > 2$, the right-hand side of (2.37) is known if $\mathbf{h}^{(2)}, \dots, \mathbf{h}^{(s-1)}$ have already been computed. By the hypothesis (2.34) and Lemma 2.3.1 the operator $\mathscr{L}$ is invertible. Thus for any $s \geq 2$ there is a unique solution $\mathbf{h}^{(s)}$ to (2.37). Therefore a unique solution $\mathbf{h}(\mathbf{y})$ of (2.35) is determined recursively. □

Choosing $\mathbf{Y} = \mathbf{0}$ in (2.32) yields the following corollary and motivates the definition that follows it.

**Corollary 2.3.3.** *If condition* (2.34) *holds, then system* (2.27) *is formally equivalent to its linear approximation $\dot{\mathbf{y}} = A\mathbf{y}$. The (possibly formal) coordinate transformation that transforms* (2.27) *into $\dot{\mathbf{y}} = A\mathbf{y}$ is unique.*

**Definition 2.3.4.** System $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{X}(\mathbf{x})$ is *linearizable* if there is an analytic normalizing transformation $\mathbf{x} = \mathbf{y} + \mathbf{h}(\mathbf{y})$ that places it in the normal form $\dot{\mathbf{y}} = A\mathbf{y}$.

Both the linear system $\dot{\mathbf{y}} = A\mathbf{y}$ and the linearizing transformation that produces it are referred to as a "linearization" of the system $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{X}(\mathbf{x})$. We will see later (Corollary 4.2.3) that, at least when $A$ is diagonal and $n = 2$ (the only case of practical interest for us), the existence of a merely formal linearization implies the existence of a convergent linearization.

As we saw in Example 2.2.2, when (2.34) does not hold some equations in (2.37) might not have a solution. This means that in such a case we might not be able to transform system (2.27) into a linear system by even a formal transformation (2.33). The best we are sure to be able to do is to transform (2.27) into a system in which all terms that correspond to pairs $(m, \alpha)$ for which (2.34) holds have been eliminated. However, terms corresponding to those pairs $(m, \alpha)$ for which (2.34) fails might be impossible to eliminate. These troublesome terms have a special name.

**Definition 2.3.5.** Let $\kappa_1, \ldots, \kappa_n$ be the eigenvalues of the matrix $A$ in (2.27), ordered according to the choice of a Jordan normal form $J$ of $A$, and let $\kappa = (\kappa_1, \ldots, \kappa_n)$. Suppose $m \in \{1, \ldots, n\}$ and $\alpha \in \mathbb{N}_0^n$, $|\alpha| = \alpha_1 + \cdots + \alpha_n \geq 2$, are such that

$$(\alpha, \kappa) - \kappa_m = 0\,.$$

Then $m$ and $\alpha$ are called a *resonant pair*, the corresponding coefficient $X_m^{(\alpha)}$ of the monomial $\mathbf{x}^\alpha$ in the $m$th component of $\mathbf{X}$ is called a *resonant coefficient*, and the corresponding term is called a *resonant term* of $\mathbf{X}$. Index and multi-index pairs, terms, and coefficients that are not resonant are called *nonresonant*.

A "normal form" for system (2.27) should be a form that is as simple as possible. The first step in the simplification process is to apply (2.28) to change the linear part $A$ in (2.27) into its Jordan normal form. We will assume that this preliminary step has already been taken, so we begin with (2.27) in the form

$$\dot{\mathbf{x}} = J\mathbf{x} + \mathbf{X}(\mathbf{x}), \tag{2.38}$$

where $J$ is a lower-triangular Jordan matrix. (Note that the following definition is based on this supposition.) The simplest form that we are sure to be able to obtain is one in which all nonresonant terms are zero, so we will take this as the meaning of the term "normal form" from now on.

**Definition 2.3.6.** A *normal form* for system (2.27) is a system (2.38) in which every nonresonant coefficient is equal to zero. A *normalizing transformation* for system (2.27) is any (possibly formal) change of variables (2.33) that transforms (2.27) into a normal form; it is called *distinguished* if for each resonant pair $m$ and $\alpha$, the corresponding coefficient $h_m^{(\alpha)}$ is zero, in which case the resulting normal form is likewise termed *distinguished*.

Two remarks about this definition are in order. The first is that it is more restrictive than the definition of normal form through order $k$ for a smooth function, Definition 2.2.4, since it requires that every nonresonant term be eliminated. In Example 2.2.2, for instance, $\mathrm{Span}\{c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5 + c_6\mathbf{u}_6\}$ for any fixed choice of the constants $c_j$ is an acceptable complement to $\mathrm{Image}(\mathscr{L})$, hence

$$\begin{aligned} \dot{x} &= 2x + c(c_1x^2 + c_2xy + y^2) \\ \dot{y} &= \ \ y + c(c_4x^2 + c_5xy + c_6y^2) \end{aligned} \tag{2.39}$$

is a normal form through order two according to Definition 2.2.4. But equations (2.19) show that the single resonant term is the $y^2$ term in the first component, so that (2.39) does not give a normal form according to Definition 2.3.6 unless the $c_j$ are all chosen to be zero.

The second remark concerning Definition 2.3.6 is almost the reverse of the first: although a normal form is the simplest form that we are sure to be able to obtain in general, for a particular system it might not be the absolute simplest. In other words, the fact that a coefficient is resonant does not mean that it must be (or remain) nonzero under every normalization: a normalizing transformation that eliminates all the nonresonant terms could very well eliminate some resonant terms as well. For example, the condition in Corollary 2.3.3 is sufficent for linearizability, but it is by no means necessary. In Chapter 4 we will treat the question of the possibility of removing all resonant as well as nonresonant terms under normalization.

**Remark 2.3.7.** Suppose the system $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{X}(\mathbf{x})$ is transformed into the normal form $\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{Y}(\mathbf{y})$ by the normalizing transformation $\mathbf{x} = \mathbf{y} + \mathbf{h}(\mathbf{y})$, and let $\lambda$ be any nonzero number. It is an immediate consequence of Definitions 2.3.5 and 2.3.6 that $\dot{\mathbf{y}} = \lambda A\mathbf{y} + \lambda\mathbf{Y}(\mathbf{y})$ is a normal form for $\dot{\mathbf{x}} = \lambda A\mathbf{x} + \lambda\mathbf{X}(\mathbf{y})$, and inspection of (2.35) shows that the same transformation $\mathbf{x} = \mathbf{y} + \mathbf{h}(\mathbf{y})$ is a normalizing transformation between the scaled systems.

Henceforth we will use the following notation. For any multi-index $\alpha$, the coefficient of the monomial $\mathbf{x}^\alpha$ in the $m$th component $X_m$ of $\mathbf{X}$ will be denoted $X_m^{(\alpha)}$. Thus for example in system (2.9), $X_1^{((2,0))} = a$ and $X_2^{((1,1))} = b'$, although when $\alpha$ is given explicitly, by slight abuse of notation we will write simply $X_m^{(\alpha_1,\ldots,\alpha_n)}$ instead of $X_m^{((\alpha_1,\ldots,\alpha_n))}$. Hence, for example, we write $X_1^{(2,0)} = a$ instead of $X_1^{((2,0))} = a$. We will use the same notational convention for $\mathbf{Y}$ and $\mathbf{h}$.

Every system is at least formally equivalent to a normal form, and as the proof of the following theorem shows, there is some freedom in choosing it, although that freedom disappears if we restrict ourselves to distinguished normalizing transformations. Theorem 2.3.11 has more to say about this.

**Theorem 2.3.8.** *Any system* (2.38) *is formally equivalent to a normal form (which need not be unique). The normalizing transformation can be chosen to be distinguished.*

*Proof.* Since the linear part is already in simplest form, we look for a change of coordinates of the form (2.33) that transforms system (2.38) into $\dot{\mathbf{y}} = J\mathbf{y} + \mathbf{Y}(\mathbf{y})$ in which all nonresonant coefficients are zero. Writing $\mathbf{h}(\mathbf{y}) = \sum_{s=2}^{\infty} \mathbf{h}^{(s)}(\mathbf{y})$, for each $s$ the function $\mathbf{h}^{(s)}$ must satisfy equation (2.37), arising from (2.35) by the process described immediately below (2.37). We have shown in the proof of Lemma 2.3.1 that the matrix of the operator $\mathscr{L}$ on the left-hand side of (2.37) is lower triangular with the eigenvalues $(\alpha,\kappa)-\kappa_m$ on the main diagonal. Therefore any coefficient $h_m^{(\alpha)}$ of $\mathbf{h}^{(s)}$ is determined by the equation

$$[(\alpha,\kappa)-\kappa_m]h_m^{(\alpha)} = g_m^{(\alpha)} - Y_m^{(\alpha)}, \tag{2.40}$$

where $g_m^{(\alpha)}$ is a known expression depending on the coefficients of $\mathbf{h}^{(i)}$ satisfying $j<s$. Suppose that for $i=2,3,\ldots,s-1$, the homogeneous terms $h^{(j)}$ and $Y^{(j)}$ have been determined . Then for any $m\in\{1,\ldots,n\}$ and any multi-index $\alpha$ with $|\alpha|=s$, if the pair $m$ and $\alpha$ is nonresonant, that is, if $(\alpha,\kappa)-\kappa_m\neq 0$, then we choose $Y_m^{\alpha}=0$ so that $\mathbf{Y}$ will be a normal form, and choose $h_m^{(\alpha)}$ as uniquely determined by equation (2.40). If $(\alpha,\kappa)-\kappa_m=0$, then we may choose $h_m^{(\alpha)}$ arbitrarily (and in particular, the choice $h_m^{(\alpha)}=0$ every time yields a distinguished transformation), but the resonant coefficient $Y_m^{(\alpha)}$ must be chosen to be $g_m^{(\alpha)}$, $Y_m^{(\alpha)}=g_m^{(\alpha)}$. The process can be started because for $s=2$ the right-hand side of (2.40) is $X_m^{(\alpha)}-Y_m^{(\alpha)}$. Thus formal series for a normal form and a normalizing transformation, distinguished or not, as we decide, are obtained. □

For simplicity, from now on we will assume that the matrix $J$ is diagonal, that is, that $\sigma_k=0$ for $k=2,\ldots,n$. (All applications of normal form theory in this book will be confined to systems that meet this condition.) Then the $m$th component on the right-hand side of (2.35) is obtaining by expanding $X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))$ in powers of $\mathbf{y}$; we let $\{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)}$ denote the coefficient of $\mathbf{y}^{(\alpha)}$ in this expansion. Using this fact and Exercise 2.14, the coefficient $g_m^{(\alpha)}$ in (2.40) is given by the expression

$$g_m^{(\alpha)} = \{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)} - \sum_{j=1}^{n}\sum_{\substack{2\le|\beta|\le|\alpha|-1\\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} \beta_j h_m^{(\beta)} Y_j^{(\alpha-\beta+e_j)}, \tag{2.41}$$

where again $\{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)}$ denotes the coefficient of $\mathbf{y}^{\alpha}$ obtained after expanding $X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))$ in powers of $\mathbf{y}$, and $e_j=(0,\ldots,0,\overset{j}{1},0,\ldots,0)\in\mathbb{N}_0^n$. Note that for $|\alpha|=2$ the sum over $\beta$ is empty, so that $g_m^{(\alpha)}=\{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)}$, which reduces to $g_m^{(\alpha)}=X_m^{(\alpha)}$, since $\mathbf{X}$ and $\mathbf{h}$ begin with quadratic terms. For $|\alpha|>2$, $|\beta|<|\alpha|$ and $|\alpha-\beta+e_j|<|\alpha|$ ensure that $g_m^{(\alpha)}$ is uniquely determined by (2.41).

The proof of Theorem 2.3.8 and formula (2.41) yield the normalization procedure that is displayed in Table 2.1 on page 75 for system (2.38) in the case that the matrix $J$ is diagonal.

**Example 2.3.9.** Fix any $C^\infty$ system (2.1) with an equilibrium at $\mathbf{0}$ whose linear part is the same as that of system (2.8):

$$\begin{aligned}\dot{x}_1 &= 2x_1 + ax_1^2 + bx_1x_2 + cx_2^2 + \cdots\\ \dot{x}_2 &= \phantom{2}x_2 + a'x_1^2 + b'x_1x_2 + c'x_2^2 + \cdots .\end{aligned} \tag{2.42}$$

The resonant coefficients are determined by the equations

**Normal Form Algorithm**

**Input:**

system $\dot{\mathbf{x}} = J\mathbf{x} + \mathbf{X}(\mathbf{x})$,
$J$ diagonal with eigenvalues $\kappa_1, \ldots, \kappa_n$
$\kappa = (\kappa_1, \ldots, \kappa_n)$
an integer $k > 1$

**Output:**

a normal form $\dot{\mathbf{y}} = J\mathbf{y} + \mathbf{Y}(\mathbf{y}) + o(|\mathbf{y}|^k)$ up to order $k$
a distinguished transformation $\mathbf{x} = \mathbf{H}(\mathbf{y}) = \mathbf{y} + \mathbf{h}(\mathbf{y}) + o(|\mathbf{y}|^k)$ up to order $k$

**Procedure:**

$\mathbf{h}(\mathbf{y}) := \mathbf{0}$; $\mathbf{Y}(\mathbf{y}) := \mathbf{0}$
FOR $s = 2$ TO $s = k$ DO
  FOR $m = 1$ TO $m = n$ DO
    compute $\mathbf{X}_m(\mathbf{y} + \mathbf{h}(\mathbf{y}))$ through order $s$
  FOR $|\alpha| = s$ DO
    FOR $m = 1$ TO $m = n$ DO
      compute $g_m^{(\alpha)}$ using (2.41)
      IF
      $(\alpha, \kappa) - \kappa_m \neq 0$
      THEN
      $h_m^{(\alpha)} := \dfrac{g_m^{(\alpha)}}{(\alpha, \kappa) - \kappa_m}$
      $h_m(\mathbf{y}) := h_m(\mathbf{y}) + h_m^{(\alpha)} \mathbf{y}^\alpha$
      ELSE
      $Y_m^{(\alpha)} := g_m^{(\alpha)}$
      $Y_m(\mathbf{y}) := Y_m(\mathbf{y}) + Y_m^{(\alpha)} \mathbf{y}^\alpha$
$\dot{\mathbf{y}} := J\mathbf{y} + \mathbf{Y}(\mathbf{y}) + o(|\mathbf{y}|^k)$
$\mathbf{H} := \mathbf{y} + \mathbf{h}(\mathbf{y}) + o(|\mathbf{y}|^k)$

**Table 2.1** Normal Form Algorithm

$$
\begin{aligned}
(\alpha, \kappa) - 2 &= 2\alpha_1 + \alpha_2 - 2 = 0 \\
(\alpha, \kappa) - 1 &= 2\alpha_1 + \alpha_2 - 1 = 0 .
\end{aligned}
$$

When $|\alpha| = 2$, the first equation has the unique solution $(\alpha_1, \alpha_2) = (0, 2)$ and the second equation has no solution; for $|\alpha| \geq 3$, neither equation has a solution. Thus by Definition 2.3.6, for any $k \in \mathbb{N}_0$, the normal form through order $k$ is

$$
\begin{aligned}
\dot{y}_1 &= 2y_1 + Y_1^{(0,2)} y_2^2 + o(|\mathbf{y}|^k) \\
\dot{y}_2 &= \quad y_2 + o(|\mathbf{y}|^k) .
\end{aligned}
$$

As remarked immediately after equation (2.41), for $|\alpha| = 2$, $g_m^{(\alpha)} = X_m^{(\alpha)}(\mathbf{y})$, so we know that in fact $Y_1^{(0,2)} = c$. If we ignore the question of convergence, then (2.42) is formally equivalent to

$$\begin{aligned}
\dot{y}_1 &= 2y_1 + cy_2^2 \\
\dot{y}_2 &= \ \ y_2 .
\end{aligned}$$

In Exercise 2.15 the reader is asked to use the Normal Form Algorithm to find the normalizing transformation $\mathbf{H}$ through order two.

**Example 2.3.10.** Let us change the sign of the coefficient of $x_2$ in the second equation of system (2.42) and consider the resulting system:

$$\begin{aligned}
\dot{x}_1 &= 2x_1 + \ ax_1^2 + \ bx_1x_2 + \ cx_2^2 + \cdots \\
\dot{x}_2 &= -x_2 + a'x_1^2 + b'x_1x_2 + c'x_2^2 + \cdots .
\end{aligned} \tag{2.43}$$

The normal form of (2.43) is drastically different from the normal form of (2.42). For now the resonant coefficients are determined by the equations

$$\begin{aligned}
(\alpha,\kappa) - 2 &= 2\alpha_1 - \alpha_2 - 2 = 0 \\
(\alpha,\kappa) + 1 &= 2\alpha_1 - \alpha_2 + 1 = 0 .
\end{aligned}$$

Solutions of the first equation that correspond to $|\alpha| \geq 2$ are the pairs $(k, 2k-2)$, $k \in \mathbb{N}_0$, $k \geq 2$; solutions of the second equation that correspond to $|\alpha| \geq 2$ are the pairs $(k, 2k+1)$, $k \in \mathbb{N}_0$. By Definition 2.3.6, the normal form of (2.43) is

$$\begin{aligned}
\dot{y}_1 &= 2y_1 + y_1 \sum_{k=1}^{\infty} Y_1^{(k+1,2k)} (y_1 y_2^2)^k , \\
\dot{y}_2 &= -y_2 + y_2 \sum_{k=1}^{\infty} Y_2^{(k,2k+1)} (y_1 y_2^2)^k .
\end{aligned} \tag{2.44}$$

In Exercise 2.16 the reader is asked to use the Normal Form Algorithm to find the resonant coefficients $Y_1^{(2,2)}$ and $Y_2^{(1,3)}$ and the normalizing transformation $\mathbf{H}$ through order three.

**Theorem 2.3.11.** *Let system* (2.38)*, with $J$ diagonal, be given. There is a unique normal form that can be obtained from system* (2.38) *by means of a distinguished normalizing transformation, which we call the distinguished normal form for* (2.38)*, and the distinguished normalizing transformation that produces it is unique. The resonant coefficients of the distinguished normal form are given by the formula*

$$Y_m^{(\alpha)} = \{X_m(\mathbf{y} + \mathbf{h}(\mathbf{y}))\}^{(\alpha)} , \tag{2.45}$$

*where $\{X_m(\mathbf{y} + \mathbf{h}(\mathbf{y}))\}^{(\alpha)}$ denotes the coefficient of $\mathbf{y}^\alpha$ obtained after expanding $X_m(\mathbf{y} + \mathbf{h}(\mathbf{y}))$ in powers of $\mathbf{y}$.*

*Proof.* In the inductive process in the proof of Theorem 2.3.8, at each step the choice of $h_m^{(\alpha)}$ is already uniquely determined if $m$ and $\alpha$ are a nonresonant pair, while if they are resonant, then $h_m^{(\alpha)}$ must be chosen to be zero so that the transformation will be distinguished. Consider now the choice of $Y_m^{(\alpha)}$ at any step. If $m$ and $\alpha$ are a nonresonant pair then of course $Y_m^{(\alpha)}$ must be chosen to be zero. If $m$ and $\alpha$ are a resonant pair, so that $(\alpha,\kappa)-\kappa_m=0$, then by (2.40) and (2.41)

$$Y_m^{(\alpha)} = g_m^{(\alpha)} = \{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)} - \sum_{j=1}^{n} \sum_{\substack{2\le|\beta|\le|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} \beta_j h_m^{(\beta)} Y_j^{(\alpha-\beta+e_j)}, \tag{2.46}$$

so again $Y_m^{(\alpha)}$ is uniquely determined at this stage.

At the start of the process, when $|\alpha|=2$, as noted immediately below (2.41) the sum in (2.46) is empty; if $m$ and $\alpha$ are a resonant pair, then $Y_m^{(\alpha)}$ is determined as $Y_m^{(\alpha)} = \{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)} = X_m^{(\alpha)}$, since $\mathbf{X}$ and $\mathbf{h}$ begin with quadratic terms. To obtain formula (2.45), consider the sum in (2.46). Any particular coefficient $Y_j^{(\alpha-\beta+e_j)}$ can be nonzero only if $j$ and $\alpha-\beta+e_j$ are a resonant pair, that is, only if $(\alpha-\beta+e_j,\kappa)-\kappa_j=0$, which by linearity and the equality $(e_j,\kappa)=\kappa_j$ holds if and only if $(\alpha-\beta,\kappa)=0$. That is,

$$Y_j^{(\alpha-\beta+e_j)} \neq 0 \quad \text{implies} \quad (\alpha-\beta,\kappa)=0\,. \tag{2.47}$$

Thus if in the sum $j$ and $\beta$ are such that $Y_j^{(\alpha-\beta+e_j)}\neq 0$, then

$$(\beta,\kappa)-\kappa_m = (\beta-\alpha,\kappa)+(\alpha,\kappa)-\kappa_m = 0$$

by (2.47) and the assumption that $m$ and $\alpha$ are a resonant pair. Since $\mathbf{h}$ is distinguished this means that $h_m^{(\beta)}=0$. Thus every term in the sum in (2.46) is zero, and the theorem is established. □

We close this section with a theorem that gives a criterion for convergence of normalizing transformations that applies to all the situations of interest in this book. Recall that a series $v(\mathbf{z})=\sum_\alpha v^{(\alpha)}\mathbf{z}^\alpha$ is said to *majorize* a series $u(\mathbf{z})=\sum_\alpha u^{(\alpha)}\mathbf{z}^\alpha$, and $v(\mathbf{z})$ is called a *majorant* of $u(\mathbf{z})$, denoted $u(\mathbf{z})\prec v(\mathbf{z})$, if $|u^{(\alpha)}|\le v^{(\alpha)}$ for all $\alpha\in\mathbb{N}_0^n$. If a convergent series $v(\mathbf{z})$ majorizes a series $u(\mathbf{z})$, then $u(\mathbf{z})$ converges on some neighborhood of $\mathbf{0}$. By way of notation, for any series $f(\mathbf{z})=\sum_\alpha f^{(\alpha)}\mathbf{z}^\alpha$ we denote by $f^\natural(\mathbf{z})$ its trivial majorant, the series that is obtained by replacing each coefficient of $f$ by its modulus: $f^\natural(\mathbf{z}) := \sum_\alpha |f^{(\alpha)}|\mathbf{z}^\alpha$. Note that in the following lemma $f(\mathbf{x})$ begins with terms of order at least two.

**Lemma 2.3.12.** *Suppose the series $f(\mathbf{x})=\sum_{\alpha:|\alpha|\ge 2} f^{(\alpha)}\mathbf{x}^\alpha$ converges on a neighborhood of $\mathbf{0}\in\mathbb{C}^n$. Then there exist positive real numbers $a$ and $b$ such that*

$$f^{\natural}(\mathbf{x}) \prec \frac{a\left(\sum_{j=1}^{n} x_j\right)^2}{1 - b\sum_{j=1}^{n} x_j}. \tag{2.48}$$

*Proof.* There exist positive real numbers $a_0$ and $b_0$ such that $f(\mathbf{x})$ converges on a neighborhood of $M := \{\mathbf{x} : |x_j| \le b_0 \text{ for } 1 \le j \le n\}$, and $|f(\mathbf{x})| \le a_0$ for $\mathbf{x} \in M$. Then the Cauchy Inequalities state that

$$|f^{(\alpha)}| \le \frac{a_0}{b_0^{\alpha_1} \cdots b_0^{\alpha_n}} \quad \text{for all} \quad \alpha \in \mathbb{N}_0^n. \tag{2.49}$$

From the identity $\sum_{|\alpha|\ge 0} y_1^{\alpha_1} \cdots y_n^{\alpha_n} = \prod_{j=1}^{n} \left[\sum_{s=0}^{\infty} y_j^s\right]$ it follows that

$$\begin{aligned}\sum_{|\alpha|\ge 0} \frac{a_0}{b_0^{\alpha_1} \cdots b_0^{\alpha_n}} x_1^{\alpha_1} \cdots x_n^{\alpha_n} &= a_0 \sum_{|\alpha|\ge 0} \left(\frac{x_1}{b_0}\right)^{\alpha_1} \cdots \left(\frac{x_n}{b_0}\right)^{\alpha_n} = a_0 \prod_{j=1}^{n} \left[\sum_{s=0}^{\infty} \left(\frac{x_j}{b_0}\right)^s\right] \\ &= a_0 \prod_{j=1}^{n} \left(1 - \frac{x_j}{b_0}\right)^{-1}\end{aligned} \tag{2.50}$$

holds on $\mathrm{Int}(M)$. By the definition of majorization, (2.49) and (2.50) yield

$$f^{\natural}(\mathbf{x}) \prec a_0 \prod_{j=1}^{n} \left(1 - \frac{x_j}{b_0}\right)^{-1}. \tag{2.51}$$

It follows readily from the fact that for all $n, k \in \mathbb{N}$, $(n+k)/(1+k) \le n$, and the series expansion of $(1+x)^{-n}$ about $0 \in \mathbb{C}$ that for any $n \in \mathbb{N}$,

$$(1+x)^{-n} \prec \frac{1}{1-nx}. \tag{2.52}$$

It is also readily verified that

$$\prod_{j=1}^{n} \left(1 - \frac{x_j}{b_0}\right)^{-1} \prec \left(1 - \frac{1}{b_0} \sum_{j=1}^{n} x_j\right)^{-n}. \tag{2.53}$$

Thus applying (2.52) with $\mathbf{x}$ replaced by $-\frac{1}{b_0}\sum_{j=1}^{n} x_j$, and using the fact that for $c \in \mathbb{R}^+$, $1/(1+cu) \prec 1/(1-cu)$, (2.53) yields

$$\prod_{j=1}^{n} \left(1 - \frac{x_j}{b_0}\right)^{-1} \prec \frac{1}{1 - \frac{n}{b_0}\sum_{j=1}^{n} x_j}. \tag{2.54}$$

Combining (2.51) and (2.54) yields

$$f^{\natural}(\mathbf{x}) \prec \frac{a_0}{1 - \frac{n}{b_0}\sum_{j=1}^{n} x_j}.$$

But since $f$ has no constant or linear terms, the constant and linear terms in the right-hand side may be removed, yielding finally

$$f^{\natural}(\mathbf{x}) \prec a_0 \left[ \frac{1}{1 - \frac{n}{b_0}\sum_{j=1}^{n} x_j} - 1 - \frac{n}{b_0}\sum_{j=1}^{n} x_j \right],$$

so that (2.48) holds with $a = (n/b_0)^2 a_0$ and $b = n/b_0$. □

**Theorem 2.3.13.** *Let $\kappa_1, \dots, \kappa_n$ be the eigenvalues of the matrix $J$ in* (2.38) *and set $\kappa = (\kappa_1, \dots, \kappa_n)$. Suppose $\mathbf{X}$ is analytic, that is, that each component $X_m$ is given by a convergent power series, and that for each resonant coefficient $Y_j^{(\alpha)}$ in the distinguished normal form $\mathbf{Y}$ of $\mathbf{X}$, $\alpha \in \mathbb{N}^n$ (that is, every entry in the multi-index $\alpha$ is positive). Suppose further that there exist positive constants $d$ and $\varepsilon$ such that the following conditions hold:*
*(a) for all $\alpha \in \mathbb{N}_0^n$ and all $m \in \{1, \dots, n\}$ such that $(\alpha, \kappa) - \kappa_m \neq 0$,*

$$|(\alpha, \kappa) - \kappa_m| \geq \varepsilon\,; \tag{2.55}$$

*(b) for all $\alpha$ and $\beta$ in $\mathbb{N}_0^n$ for which $2 \leq |\beta| \leq |\alpha| - 1$, $\alpha - \beta + e_m \in \mathbb{N}_0^n$ for all $m \in \{1, \dots, n\}$, and*

$$(\alpha - \beta, \kappa) = 0, \tag{2.56}$$

*the following inequality holds:*

$$\left| \sum_{j=1}^{n} \beta_j Y_j^{(\alpha - \beta + e_j)} \right| \leq d\,|(\beta, \kappa)| \sum_{j=1}^{n} \left| Y_j^{(\alpha - \beta + e_j)} \right| . \tag{2.57}$$

*Then the distinguished normalizing transformation $\mathbf{x} = \mathbf{H}(\mathbf{y})$ is analytic as well, that is, each component $h_m(\mathbf{y})$ of $\mathbf{h}$ is given by a convergent power series, so that system* (2.38) *is analytically equivalent to its normal form.*

*Proof.* Suppose that a particular pair $m$ and $\alpha$ correspond to a nonresonant term. Then $Y_m^{(\alpha)} = 0$ and by (2.40) and (2.41)

$$\begin{aligned} |h_m^{(\alpha)}| &\leq \tfrac{1}{\varepsilon} \left| \{X_m(\mathbf{y} + \mathbf{h}(\mathbf{y}))\}^{(\alpha)} \right| \\ &\quad + \frac{1}{|(\alpha, \kappa) - \kappa_m|} \left| \sum_{j=1}^{n} \sum_{\substack{2 \leq |\beta| \leq |\alpha| - 1 \\ \alpha - \beta + e_j \in \mathbb{N}_0^n}} \beta_j h_m^{(\beta)} Y_j^{(\alpha - \beta + e_j)} \right| . \end{aligned} \tag{2.58}$$

For any nonzero term $Y_j^{(\alpha - \beta + e_j)}$ in (2.58), by (2.47)

$$(\alpha, \kappa) - \kappa_m = (\alpha - \beta, \kappa) + (\beta, \kappa) - \kappa_m = (\beta, \kappa) - \kappa_m\,,$$

so by hypothesis (a) $|(\beta, \kappa) - \kappa_m| = |(\alpha, \kappa) - \kappa_m| \geq \varepsilon$. We adopt the convention that $Y_j^{(\gamma)} = 0$ if $\gamma \notin \mathbb{N}_0^n$, so that we can reverse the order of summation in (2.58) and apply hypothesis (b). Thus the second term in (2.58) is bounded above by

$$\begin{aligned}
\frac{1}{|(\alpha,\kappa)-\kappa_m|} & \sum_{2\le|\beta|\le|\alpha|-1} |h_m^{(\beta)}| \left| \sum_{j=1}^{n} \beta_i Y_j^{(\alpha-\beta+e_j)} \right| \\
& \le \frac{1}{|(\alpha,\kappa)-\kappa_m|} \sum_{2\le|\beta|\le|\alpha|-1} |h_m^{(\beta)}| d |(\alpha,\kappa)-\kappa_m+\kappa_m| \sum_{j=1}^{n} |Y_j^{(\alpha-\beta+e_j)}| \\
& \le d\left(1+\frac{|\kappa_m|}{\varepsilon}\right) \sum_{j=1}^{n} \sum_{\substack{2\le|\beta|\le|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} |h_m^{(\beta)}||Y_j^{(\alpha-\beta+e_j)}| .
\end{aligned}$$

Applying this to (2.58) gives

$$|h_m^{(\alpha)}| \le \left| \frac{1}{\varepsilon} \{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)} \right| + d_0 \sum_{j=1}^{n} \sum_{\substack{2\le|\beta|\le|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} |h_m^{(\beta)}||Y_j^{(\alpha-\beta+e_j)}| , \tag{2.59}$$

where $d_0 = \max_{1\le m\le n} d(1+|\kappa_m|/\varepsilon)$. Thus

$$h_m^{\natural}(\mathbf{y}) \prec \tfrac{1}{\varepsilon} \sum_{|\alpha|\ge 2} |\{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)}| \mathbf{y}^{\alpha} + d_0 \sum_{|\alpha|\ge 2} \sum_{j=1}^{n} \sum_{\substack{2\le|\beta|\le|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} |h_m^{(\beta)}||Y_j^{(\alpha-\beta+e_j)}| \mathbf{y}^{\alpha} . \tag{2.60}$$

Clearly

$$\sum_{|\alpha|\ge 2} |\{X_m(\mathbf{y}+\mathbf{h}(\mathbf{y}))\}^{(\alpha)}| \mathbf{y}^{\alpha} \prec \sum_{|\alpha|\ge 2} X_m^{\natural}(\mathbf{y}+\mathbf{h}^{\natural}(\mathbf{y})) . \tag{2.61}$$

Turning to the term to the right of the plus sign in (2.60), for convenience index elements of $\mathbb{N}^n$ as $\{\gamma_r\}_{r=1}^{\infty}$. Recalling our convention that $Y^{(\gamma)} = 0$ if $\gamma \notin \mathbb{N}_0^n$,

$$\begin{aligned}
& \sum_{|\alpha|\ge 2} \sum_{j=1}^{n} \sum_{\substack{2\le|\beta|\le|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} |h_m^{(\beta)}||Y_j^{(\alpha-\beta+e_j)}| \mathbf{y}^{\alpha} \\
& \qquad = \sum_{r=1}^{\infty} \left[ |h_m^{(\beta_r)}| \mathbf{y}^{\beta_r} \sum_{j=1}^{n} \sum_{s=1}^{\infty} |Y_j^{(\alpha_s-\beta_r+e_j)}| \mathbf{y}^{\alpha_s-\beta_r} \right] .
\end{aligned} \tag{2.62}$$

For any fixed multi-index $\beta_r$, consider the sum

$$\sum_{j=1}^{n} \sum_{s=1}^{\infty} |Y_j^{(\alpha_s-\beta_r+e_j)}| \mathbf{y}^{\alpha_s-\beta_r} . \tag{2.63}$$

The number $Y_j^{(\alpha_s-\beta_r+e_j)}$ is nonzero only if $j$ and $\alpha_s-\beta_r+e_j$ form a resonant pair. The same term, times the same power of $\mathbf{y}$, will occur in the sum (2.63) corresponding to the multi-index $\beta_{r'}$ if and only if there exists a multi-index $\alpha_{s'}$ that satisfies $\alpha_{s'}-\beta_{r'} = \alpha_s-\beta_r$, which is true if and only if $\alpha_s-\beta_r+\beta_{r'} \in \mathbb{N}_0^n$. Writing $\alpha_s = (\alpha_s^1,\dots,\alpha_s^n)$ and $\beta_r = (\beta_r^1,\dots,\beta_r^n)$, then by the fact that $\alpha_s-\beta_r+e_j$ is part of a resonant pair and the hypothesis that no entry in the multi-index of a resonant pair is zero, if $k \neq j$, then $\alpha_s^k - \beta_r^k \geq 1$, while $\alpha_s^j - \beta_r^j \geq 0$. Thus $\alpha_s - \beta_r + \beta_{r'} \in \mathbb{N}_0^n$, so we conclude that the expression in (2.63) is the same for all multi-indices $\beta_{r'}$, and may be written

$$\sum_{j=1}^{n} \sum_{(\gamma,j)\ \text{resonant}} |Y_j^{(\gamma)}| \mathbf{y}^{\gamma - e_j},$$

hence (2.62) is

$$\begin{aligned}
&\sum_{|\alpha|\geq 2} \sum_{j=1}^{n} \sum_{\substack{2\leq|\beta|\leq|\alpha|-1 \\ \alpha-\beta+e_j\in\mathbb{N}_0^n}} \left|h_m^{(\beta)}\right| \left|Y_j^{(\alpha-\beta+e_j)}\right| \mathbf{y}^{\alpha} \\
&\quad \prec \left( \sum_{|\beta|\geq 2} \left|h_m^{(\beta)}\right| \mathbf{y}^{\beta} \right) \left( \sum_{j=1}^{n} \sum_{(\gamma,j)\ \text{resonant}} \left|Y_j^{(\gamma)}\right| \mathbf{y}^{\gamma-e_j} \right) \\
&\quad = h_m^{\natural}(\mathbf{y}) \sum_{j=1}^{n} \sum_{(\gamma,j)\ \text{resonant}} y_j^{-1} \left|Y_j^{(\gamma)}\right| \mathbf{y}^{\gamma} \\
&\quad = h_m^{\natural}(\mathbf{y}) \sum_{j=1}^{n} y_j^{-1} Y_j^{\natural}(\mathbf{y}),
\end{aligned} \tag{2.64}$$

which is well-defined even at $\mathbf{y} = \mathbf{0}$ by the hypothesis on resonant pairs. Thus applying (2.61) and (2.64) to (2.60) yields

$$h_m^{\natural}(\mathbf{y}) \prec \tfrac{1}{\varepsilon} X_m^{\natural}(\mathbf{y} + \mathbf{h}^{\natural}(\mathbf{y})) + h_m^{\natural}(\mathbf{y}) \sum_{j=1}^{n} y_j^{-1} Y_j^{\natural}(\mathbf{y}). \tag{2.65}$$

Multiplying (2.45) by $\mathbf{y}^{\alpha}$ and summing it over all $\alpha$ for which $|\alpha| \geq 2$, from (2.61) we obtain $Y_m^{\natural}(\mathbf{y}) \prec X_m^{\natural}(\mathbf{y} + \mathbf{h}^{\natural}(\mathbf{y}))$. Summing this latter expression and expression (2.65) over $m$, and recalling that for a distinguished transformation $h_m^{(\alpha)} = 0$ for all nonresonant pairs, we obtain the existence of a real constant $c_1 > 0$ such that the following majorizing relation holds:

$$\sum_{m=1}^{n} Y_m^{\natural}(\mathbf{y}) + \sum_{m=1}^{n} h_m^{\natural}(\mathbf{y}) \prec c_1 \sum_{m=1}^{n} X_m^{\natural}(\mathbf{y} + \mathbf{h}^{\natural}(\mathbf{y})) + c_1 \sum_{m=1}^{n} h_m^{\natural}(\mathbf{y}) \sum_{j=1}^{n} \frac{Y_j^{\natural}(\mathbf{y})}{y_j}. \tag{2.66}$$

By Lemma 2.3.12

$$X_m^{\natural}(\mathbf{y} + \mathbf{h}^{\natural}(\mathbf{y})) \prec \frac{a\left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n} h_j^{\natural}(\mathbf{y})\right)^2}{1 - b\left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n} h_j^{\natural}(\mathbf{y})\right)}, \tag{2.67}$$

so that (2.66) becomes

$$\sum_{m=1}^{n} Y_m^{\natural}(\mathbf{y}) + \sum_{m=1}^{n} h_m^{\natural}(\mathbf{y}) \prec \frac{c_1 a \left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n} h_j^{\natural}(\mathbf{y})\right)^2}{1 - b\left(\sum_{j=1}^{n} y_j + \sum_{j=1}^{n} h_j^{\natural}(\mathbf{y})\right)} + c_1 \sum_{m=1}^{n} h_m^{\natural}(\mathbf{y}) \sum_{j=1}^{n} y_j^{-1} Y_j^{\natural}(\mathbf{y}). \tag{2.68}$$

To prove the theorem it suffices to show that the series

$$S(\mathbf{y}) = \sum_{m=1}^{n} Y_m^{\natural}(\mathbf{y}) + \sum_{m=1}^{n} h_m^{\natural}(\mathbf{y})$$

converges at some nonzero value of $\mathbf{y}$. We will show convergence at $\mathbf{y} = (\eta, \ldots, \eta)$ for some positive value of the real variable $\eta$. To do so, we consider the real series $S(\eta, \ldots, \eta)$. Since $\mathbf{Y}$ and $\mathbf{h}$ begin with quadratic or higher terms we can write $S(\eta, \ldots, \eta) = \eta U(\eta)$ for a real series $U(\eta) = \sum_{k=1}^{\infty} u_k \eta^k$, where the coefficients $u_k$ are nonnegative. By the definition of $S$ and $U$, $U(\eta)$ clearly satisfies $\sum_{m=1}^{n} h_m^{\natural}(\eta, \ldots, \eta) \prec \eta U(\eta)$ and $\sum_{m=1}^{n} Y_m^{\natural}(\eta, \ldots, \eta) \prec \eta U(\eta)$, which when applied to (2.68) yield

$$\eta U(\eta) = S(\eta, \ldots, \eta) \prec \frac{c_1 a (n\eta + \eta U(\eta))^2}{1 - b(n\eta + \eta U(\eta))} + c_1 \eta U^2(\eta)$$

so that

$$U(\eta) \prec c_1 U^2(\eta) + \frac{c_1 a \eta \, (n + U(\eta))^2}{1 - b\eta \, (n + U(\eta))}. \tag{2.69}$$

Consider the real analytic function $F$ defined on a neighborhood of $(0,0) \in \mathbb{R}^2$ by

$$F(x,y) = c_1 x^2 + \frac{c_1 a y (n+x)^2}{1 - b y (n+x)}.$$

Using the geometric series to expand the second term, it is clear that

$$F(x,y) = F_0(y) + F_1(y)x + F_2(y)x^2 + \cdots,$$

where $F_j(0) = 0$ for $j \neq 2$ and $F_j(y) \geq 0$ if $y \geq 0$. Thus for any sequence of real constants $r_1, r_2, r_3, \ldots$,

$$F\left(\sum_{k=1}^{\infty} r_k y^k, y\right) = \delta_1 y + \delta_2(r_1) y^2 + \delta_3(r_1, r_2) y^3 + \cdots, \tag{2.70}$$

where $\delta_1 \geq 0$ and for $k \geq 2$ $\delta_k$ is a polynomial in $r_1, r_2, \ldots, r_{k-1}$ with nonnegative coefficients. Thus

$$0 \leq a_j \leq b_j \text{ for } j = 1, \ldots, k-1 \quad \text{implies} \quad \delta_k(a_1, \ldots, a_{k-1}) \leq \delta_k(b_1, \ldots, b_{k-1}). \tag{2.71}$$

By the Implicit Function Theorem there is a unique real analytic function $w = w(y)$ defined on a neighborhood of 0 in $\mathbb{R}$ such that $w(0) = 0$ and $x - F(x,y) = 0$ if and only if $x = w(y)$; write $w(y) = \sum_{k=1}^{\infty} w_k y^k$. By (2.70) the coefficients $w_k$ satisfy

$$w_k = \delta_k(w_1, \dots, w_{k-1}) \quad \text{for} \quad k \geq 2, \tag{2.72}$$

and by (2.69) the coefficients $u_k$ satisfy

$$u_k \leq \delta_k(u_1, \dots, u_{k-1}) \quad \text{for} \quad k \geq 2. \tag{2.73}$$

A simple computation shows that $u_1 = c_1 a n^2 = w_1$, hence (2.71), (2.72), and (2.73) imply by mathematical induction that $u_k \leq w_k$ for $k \geq 1$. Thus $U(\eta) \prec w(\eta)$, implying the convergence of $U$ on a neighborhood of 0 in $\mathbb{R}$, which implies convergence of $S$, hence of $\mathbf{h}$ and $\mathbf{Y}$, on a neighborhood of 0 in $\mathbb{C}^n$. □

## 2.4 Notes and Complements

Our concern in this chapter has been with vector fields in a neighborhood of an equilibrium. In Section 2.1 we presented a few of Lyapunov's classical results on the stability problem. Other theorems of this sort are available in the literature. The reader is referred to the monograph of La Salle ([109]) for generalizations of the classical Lyapunov theory.

The idea of placing system (2.1) in some sort of normal form in preparation for a more general study, or in order to take advantage of a general property of the collections of systems under consideration, has wide applicability. To cite just one example, a class of systems of differential equations that has received much attention is the set of quadratic systems in the plane, which are those of the form

$$\begin{aligned} \dot{x} &= a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \\ \dot{y} &= b_{00} + b_{10}x + b_{01}y + b_{20}x^2 + b_{11}xy + b_{02}y^2 . \end{aligned} \tag{2.74}$$

Using special properties of such systems, it is possible by a sequence of coordinate transformations and time rescalings to place any such system that has a cycle in its phase portrait into the special form

$$\begin{aligned} \dot{x} &= \delta x - y + \ell x^2 + mxy + ny^2 \\ \dot{y} &= x(1 + ax + by) , \end{aligned} \tag{2.75}$$

thereby simplifying the expression, reducing the number of parameters, and making one parameter ($\delta$) into a rotation parameter on the portion of the plane in which a cycle can exist. (See §12 of [202].)

Turning to the problem of computing normal forms, if we wish to know the actual coefficients in the normal form in terms of the *original* coefficients, whether numerical or symbolic, we must keep track of all coefficients exactly at each successive step of the sequence of transformations leading to the normal form. Hand computation can quickly become infeasible. Computer algebra approaches to the actual computation of normal forms is treated in the literature; the reader is referred to [71, 139, 146]. Sample code for the algorithm in Table 2.1 is in the Appendix.

We have noted that unless $\mathscr{L}$ maps $\mathscr{H}_s$ onto itself, the subspace $\mathscr{K}_s$ complementary to Image($\mathscr{L}$) is not unique. It is reasonable to attempt to make a uniform, or at least systematic, choice of the subspace $\mathscr{K}_s$, $s \geq 2$. Such a systematic choice is termed in [139] a normal form *style*, and the reader is directed there for a full discussion.

For more exhaustive treatments of the theory of normal forms the reader can consult, for example, the references [15, 16, 19, 24, 25, 115, 139, 142, 180, 181].

## Exercises

2.1 Show that the trajectory of every point in a neighborhood of the equilibrium $\mathbf{x}_0 = (1,0)$ of the system

$$\begin{aligned}\dot{x} &= x - (x+y)\sqrt{x^2+y^2} + xy\\ \dot{y} &= y + (x-y)\sqrt{x^2+y^2} - x^2\end{aligned}$$

on $\mathbb{R}^2$ tends to $\mathbf{x}_0$ in forward time, yet in every neighborhood of $\mathbf{x}_0$ there exists a point whose forward trajectory travels distance at least 1 away from $\mathbf{x}_0$ before returning to limit on $\mathbf{x}_0$.
Hint. Change to polar coordinates.

2.2 Consider the general second-order linear homogeneous ordinary differential equation in one dependent variable in standard form,

$$\ddot{x} + f(x)\dot{x} + g(x) = 0\,, \tag{2.76}$$

generally called a *Liénard equation*.

a. Show that if there exists $\varepsilon > 0$ such that $xg(x) > 0$ whenever $0 < |x| < \varepsilon$, then the function $W(x,y) = \frac{1}{2}y^2 + \int_0^x g(u)\,du$ is positive definite on a neighborhood of $\mathbf{0} \in \mathbb{R}^2$.
b. Assuming the truth of the condition on $g(x)$ in part (a), use the function $W$ to show that the equilibrium of the system

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -g(x_1) - f(x_1)x_2,\end{aligned}$$

which is equivalent to (2.76), is stable if $f(x) \equiv 0$ and is asymptotically stable if $f(x) > 0$ for $0 < |x| < \varepsilon$.

2.3 Transform system (2.76) into the equivalent Liénard form

$$\begin{aligned}\dot{x}_1 &= x_2 - F(x_1)\\ \dot{x}_2 &= -g(x_1)\,,\end{aligned} \tag{2.77}$$

where $F(x)=\int_0^x f(u)du$. Formulate conditions in terms of $g(x)$ and $F(x)$ so that the equilibrium of (2.77) is stable, asymptotically stable, or unstable.

2.4 Construct a counterexample to Theorem 2.1.4 if tangency to $C\setminus\{\mathbf{0}\}$ is allowed, even if at only countably many points.

2.5 Prove Theorem 2.1.5.

2.6 Show that the dimension of the vector space $\mathscr{H}_s$ of functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ (or from $\mathbb{C}^n$ to $\mathbb{C}^n$), all of whose components are homogeneous polynomial functions of degree $s$, is $nC(s+n-1,s)=n(s+n-1)!/(s!(n-1)!)$.
Hint. Think in terms of distributing $p$ identical objects into $q$ different boxes, which is the same as selecting with repetition $p$ objects from $q$ types of objects.

2.7 Show that if the eigenvalues of the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ are $\pm i\beta$ ($\beta\in\mathbb{R}$), then by a linear transformation the system

$$\begin{aligned}\dot{x}&=ax+by\\ \dot{y}&=cx+dy\end{aligned}$$

can be brought to the form

$$\begin{aligned}\dot{x}&=\quad\beta y\\ \dot{y}&=-\beta x\,.\end{aligned}$$

2.8 For system (2.1) on $\mathbb{R}^2$, suppose $\mathbf{f}(\mathbf{0})=\mathbf{0}$ and $A=d\mathbf{f}(\mathbf{0})$ has exactly one zero eigenvalue. Show that for each $k\in\mathbb{N}$, $k\geq 2$, each vector in the ordered basis of $\mathscr{H}_k$ of Example 2.2.2 is an eigenvector of $\mathscr{L}$, and that $\mathscr{L}$ has zero as an eigenvalue of multiplicity two, with corresponding eigenvectors $\begin{pmatrix} xy^{k-1} \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ y^k \end{pmatrix}$, where $x$ and $y$ denote the usual coordinates on $\mathbb{R}^2$. Write down the normal form for (2.1) through order $k$. (Assume from the outset that $A$ has been placed in Jordan normal form, that is, diagonalized.)

2.9 For system (2.1) on $\mathbb{R}^2$, suppose $\mathbf{f}(\mathbf{0})=\mathbf{0}$ and $A=d\mathbf{f}(\mathbf{0})$ has both eigenvalues zero but is not itself the zero transformation, hence has upper-triangular Jordan normal form $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. (This form for the linear part is traditional in this context.)

a. Show that the matrix of $\mathscr{L}:\mathscr{H}_2\to\mathscr{H}_2$ with respect to the ordered basis of Example 2.2.2 is

$$L=\begin{pmatrix} 0&0&0&-1&0&0\\ 2&0&0&0&-1&0\\ 0&1&0&0&0&-1\\ 0&0&0&0&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&1&0 \end{pmatrix}.$$

b. Show that the rank of $L$ (the dimension of its column space) is four, hence that $\dim\operatorname{Image}(\mathscr{L})=4$.

c. Show that none of the basis vectors $\begin{pmatrix} x^2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ xy \end{pmatrix}$, and $\begin{pmatrix} 0 \\ x^2 \end{pmatrix}$ lies in Image($\mathscr{L}$).
Hint. Work in coordinates, for example replacing the second vector in the list by its coordinate vector, the standard basis vector $e_5$ of $\mathbb{R}^6$.

d. Explain why two possible normal forms of $\mathbf{f}$ through order two are

$$\text{(i)}\quad \begin{aligned}\dot{x} &= y + O(3)\\ \dot{y} &= ax^2 + bxy + O(3)\end{aligned} \qquad \text{and} \qquad \text{(ii)}\quad \begin{aligned}\dot{x} &= y + ax^2 + O(3)\\ \dot{y} &= bx^2 + O(3)\,.\end{aligned}$$

Remark: This is the "Bogdanov–Takens Singularity." Form (i) in part (d) is the normal form of Bogdanov ([20]); form (ii) is that of Takens ([187]).

2.10 In the same situation as that of Exercise 2.9, show that for all $k \in \mathbb{N}$, $k \geq 3$, $\dim(\mathrm{Image}(\mathscr{L})) = \dim(\mathscr{H}_k) - 2$ and that $\mathscr{K}_k = \mathrm{Span}\left\{\binom{x^k}{0}, \binom{0}{x^k}\right\}$ is one choice of $\mathscr{K}_k$. Hence a normal form for (2.1) through order $r$ is

$$\begin{aligned}\dot{x} &= y + a_2x^2 + a_3x^3 + \cdots + a_rx^r + O(r+1)\\ \dot{y} &= \phantom{y + {}} b_2x^2 + b_3x^3 + \cdots + b_rx^r + O(r+1)\,.\end{aligned}$$

2.11 Show that if $S$ is a nonsingular $n \times n$ matrix such that $J = SAS^{-1}$, then under the coordinate change $\mathbf{x} = S\mathbf{y}$, expression (2.30) for $\mathscr{L}$ is transformed into (2.31), as follows.

a. Writing $\mathscr{S} : \mathbb{C}^n \to \mathbb{C}^n : \mathbf{x} \mapsto \mathbf{y} = S\mathbf{x}$ and for $\mathbf{h} \in \mathscr{H}_s$ letting $\mathbf{u} = \mathscr{S}^{-1} \circ \mathbf{h} \circ \mathscr{S}$, use the fact that $d\mathscr{S}(\mathbf{y})z = Sz$ to show that $\mathscr{L}\mathbf{h}(\mathbf{x}) = Sd\mathbf{u}(\mathbf{y})J\mathbf{y} - AS\mathbf{u}(\mathbf{y})$.

b. Use the fact that $\mathscr{L}\mathbf{u} = \mathscr{S}^{-1} \circ \mathscr{L}\mathbf{h} \circ \mathscr{S}$ and the result of part (a) to obtain (2.31).

2.12 a. Order the basis vectors of Example 2.2.2 according to the lexicographic order of the proof of Lemma 2.3.1.

b. Rework part (a) of Exercise 2.9 using lexicographic order. (Merely use the result of Exercise 2.9 without additional computation.)

c. By inspection of the diagonal entries of the matrix in part (b), determine the eigenvalues of $\mathscr{L}$.

d. Verify by direct computation of the numbers $(\kappa, \alpha) - \kappa_j$ that the eigenvalues of $\mathscr{L}$ are all zero.

2.13 [Referenced in Section 3.2.] Let $\kappa_1, \ldots, \kappa_n$ be the eigenvalues of the matrix $A$ in display (2.27).

a. Suppose $n = 2$. Show that if $\kappa_1 \neq 0$, then a necessary condition that there be a resonant term in either component of $\mathbf{X}$ is that $\kappa_2/\kappa_1$ be a rational number. Similarly if $\kappa_2 \neq 0$. (If the ratio of the eigenvalues is $p/q$, $\mathrm{GCD}(p,q) = 1$, then the resonance is called a $p : q$ resonance.)

b. Show that the analogous statement is not true when $n \geq 3$.

2.14 Derive the rightmost expression in (2.41) in the following three steps.

a. Use the expansions of $\mathbf{h}$ and $\mathbf{Y}$ analogous to (2.36) to determine that the vector homogeneous function of degree $s$ in $d\mathbf{h}(\mathbf{y})\mathbf{Y}(\mathbf{y})$ is a sum of $s-2$ products of the form $d\mathbf{h}^{(k)}(\mathbf{y})\mathbf{Y}^{(\ell)}(\mathbf{y})$.

b. Show that for each $r \in \{2, \ldots, s-1\}$, the $m$th component of the corresponding product in part (a) is

$$\sum_{j=1}^{n}\left(\left[\sum_{|\beta|=r}\beta_j h_m^{(\beta)} y_1^{\beta_1}\cdots y_j^{\beta_j-1}\cdots y_n^{\beta_n}\right]\left[\sum_{|\gamma|=s-r+1} Y_j^{(\gamma)}\mathbf{y}^{\gamma}\right]\right).$$

c. For any $\alpha$ such that $|\alpha| = s$, use the expression in part (b) to find the coefficient of $\mathbf{y}^{\alpha}$ in the $m$th component of $d\mathbf{h}(\mathbf{y})\mathbf{Y}(\mathbf{y})$, thereby obtaining the negative of the rightmost expression in (2.41).

2.15 Apply the Normal Form Algorithm displayed in Table 2.1 on page 75 to show that the normalizing transformation $\mathbf{H}$ through order two in Example 2.2.2 is

$$\mathbf{H}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{2}ay_1^2 + by_1y_2 \\ \frac{1}{3}a'y_1^2 + \frac{1}{2}b'y_1y_2 + c'y_2^2 \end{pmatrix} + \cdots.$$

2.16 Apply the Normal Form Algorithm displayed in Table 2.1 on page 75 to compute the coefficients $Y_1^{(2,2)}$ and $Y_2^{(1,3)}$ of the normal form (2.44) of Example 2.3.10 and the normalizing transformation $\mathbf{H}(\mathbf{y})$ up to order three.

2.17 In system (2.38) let $J$ be a diagonal $2\times 2$ matrix with eigenvalues $\kappa_1 = p$ and $\kappa_2 = -q$, where $p, q \in \mathbb{N}$ and $\mathrm{GCD}(p,q) = 1$. Show that the normal form of system (2.38) is

$$\begin{aligned} \dot{y}_1 &= \quad py_1 + y_1Y_1(w) \\ \dot{y}_2 &= -qy_2 + y_2Y_2(w), \end{aligned}$$

where $Y_1$ and $Y_2$ are formal power series without constant terms and $w = y_1^q y_2^p$.

2.18 For system (2.1) on $\mathbb{R}^n$, suppose that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and that the linear part $A = d\mathbf{f}(\mathbf{0})$ of $f$ at $\mathbf{0}$ is diagonalizable. For $k \in \mathbb{N}$, $k \geq 2$, choose as a basis of $\mathscr{H}_k$ the analogue of the basis of Example 2.2.2, consisting of functions of the form

$$\mathbf{h}_{\alpha,j} : \mathbb{R}^n \to \mathbb{R}^n : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}\mathbf{e}_j,$$

where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$, $\alpha_1 + \alpha_2 + \cdots + \alpha_n = k$, $j \in \{1, 2, \ldots, n\}$, and for each such $j$, $\mathbf{e}_j$ is the $j$th standard basis vector of $\mathbb{R}^n$, as always. Show by direct computation using the definition (2.18) of $\mathscr{L}$ that each basis vector $\mathbf{h}_{\alpha,j}$ is an eigenvector of $\mathscr{L}$, with corresponding eigenvalue $\lambda_j - \sum_{i=1}^{n}\alpha_i\lambda_i$. (Note how Example 2.2.2 and Exercise 2.8 fit into this framework but Exercises 2.9 and 2.10 do not.) When $\lambda_j \neq 0$ for all $j$ the equilibrium is hyperbolic, and the non-resonance conditions

$$\lambda_j - \sum_{k=1}^{n}\alpha_k\lambda_k \neq 0 \quad \text{for all } 1 \leq j \leq n, \quad \text{for all } \alpha \in \mathbb{N}_0^n$$

are necessary conditions for the existence of a smooth linearization. Results on the existence of smooth linearizations are given in [183] (for the $C^\infty$ case) and in [180] (for the analytic case).

http://www.springer.com/978-0-8176-4726-1

The Center and Cyclicity Problems
A Computational Algebra Approach
Romanovski, V.; Shafer, D.
2009, XII, 330 p. 4 illus., Softcover
ISBN: 978-0-8176-4726-1
A product of Birkhäuser Basel