Chapter 2
Introduction to Functional Differential Equations

There are different types of functional differential equations (FDEs) arising from important applications: delay differential equations (DDEs) (also referred to as retarded FDEs [RFDEs]), neutral FDEs (NFDEs), and mixed FDEs (MFDEs). The classification depends on how the current change rate of the system state depends on the history (the historical status of the state only or the historical change rate and the historical status) or whether the current change rate of the system state depends on the future expectation of the system. Later we will also see that the delay involved may also depend on the system state, leading to DDEs with state-dependent delay.

2.1 Infinite Dynamical Systems Generated by Time Lags

In Newtonian mechanics, the system's state variable changes over time, and the law that governs the change of the system's state is normally described by an ordinary differential equation (ODE). Assuming that the function involved in this ODE is sufficiently smooth (locally Lipschitz, for example), the corresponding Cauchy initial value problem is well posed, and thus knowing the current status, one is able to reconstruct the history and predict the future of the system.

In many applications, a close look at the physical or biological background of the modeling system shows that the change rate of the system's current status often depends not only on the current state but also on the history of the system, see, for example, [50, 76, 198, 199]. This usually leads to so-called DDEs with the following prototype:

$$\dot{x}(t) = f(x(t), x(t-\tau)), \tag{2.1}$$

where $x(t)$ is the system's state at time t, $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a given mapping, and the time lag $\tau > 0$ is a constant.

Such an equation arises naturally, for example, from the population dynamics of a single-species structured population. In such an example, if $x(t)$ denotes the pop-

S. Guo and J. Wu, *Bifurcation Theory of Functional Differential Equations*, Applied Mathematical Sciences 184, DOI 10.1007/978-1-4614-6992-6_2,

ulation density of the mature/reproductive population, and if the maturation period is assumed to be a constant, then we have

$$f(x(t),x(t-\tau)) = -d_m x(t) + \mathrm{e}^{-d_i\tau} b(x(t-\tau)), \tag{2.2}$$

where d_m and d_i are the death rates of the mature and immature populations, respectively, and $b\colon \mathbb{R} \to \mathbb{R}$ is the birth rate. Death is instantaneous, so the term $-d_m x(t)$ is without delay. However, the rate into the mature population is the maturation rate (not the birth rate), that is, the birth rate at time τ, multiplied by the survival probability $\mathrm{e}^{-d_i\tau}$ during the maturation process.

Clearly, to specify a function $x(t)$ of $t \geq 0$ that satisfies (2.1) (called a solution of (2.1)), we must prescribe the history of x on $[-\tau,0]$. On the other hand, once the initial value data

$$\varphi : [-\tau,0] \to \mathbb{R}^n \tag{2.3}$$

is given as a continuous function and if $f\colon \mathbb{R}^n \times \mathbb{R}^n \ni (x,y) \to f(x,y) \in \mathbb{R}^n$ is continuous and locally Lipschitz with respect to the first state variable $x \in \mathbb{R}^n$, then (2.1) on $[0,\tau]$ becomes an ODE for which the initial value problem

$$\dot{x}(t) = f(x(t),\varphi(t-\tau)), \quad t \in [0,\tau],\ x(0) = \varphi(0), \tag{2.4}$$

is solvable. If such a solution exists on $[0,\tau]$, we can repeat the argument to the initial value problem

$$\begin{cases} \dot{x}(t) = f(x(t),\underbrace{x(t-\tau)}_{\text{given}}), \quad t \in [\tau,2\tau], \\ x(\tau) \text{ is given in the previous step,} \end{cases} \tag{2.5}$$

to obtain a solution on $[\tau,2\tau]$. This process may be continued to yield a solution of (2.1) subject to $x|_{[-\tau,0]} = \varphi$ given in (2.3).

Let $C_{n,\tau} = C([-\tau,0];\mathbb{R}^n)$ be the Banach space of continuous mappings from $[-\tau,0]$ to $\mathbb{R}^n$ equipped with the supremum norm

$$\|\phi\| = \sup_{-\tau\leq\theta\leq 0} |\phi(\theta)| \ \text{ for } \ \phi \in C_{n,\tau},$$

and if we define $x_t\colon C_{n,\tau} \to C_{n,\tau}$ by the segment of x on the interval $[t-\tau,t]$ translated back to the initial interval $[-\tau,0]$, namely,

$$x_t(\theta) = x(t+\theta), \quad \theta \in [-\tau,0], \tag{2.6}$$

then (2.1) subject to $x_0 = \varphi \in C_{n,\tau}$ gives a semiflow $[0,\infty] \ni t \mapsto x_t \in C_{n,\tau}$. This clearly shows that an appropriate state space of a DDE is $C_{n,\tau}$ and that a DDE gives an infinite-dimensional dynamical system on this phase space.

Many applications call for the study of asymptotic behaviors (as $t \to \infty$) of solutions of (2.1), and such a study seems to be very difficult due to the infinite-dimensionality of the phase space and the generated semiflow, even for a scalar

DDE (2.1) (that is, when $n = 1$). Even to restrict the study of the asymptotic behaviors of solutions near a specified solution is highly nontrivial. Take a steady state as an example. A vector $x^* \in \mathbb{R}^n$ is called an equilibrium of (2.1) if

$$f(x^*, x^*) = 0. \tag{2.7}$$

This vector gives a state $\hat{x}^* \in C_{n,\tau}$, which is a constant mapping on $[-\tau, 0]$ with the constant value $x^* \in \mathbb{R}^n$, and a solution of (2.1) with the initial value $\hat{x}^*$ is a constant function x: $[0,\infty) \to \mathbb{R}^n$ with the constant value x^*. Behaviors of solutions of (2.1) in a neighborhood of $\hat{x}^*$ may be determined by the zero solution of the linearization

$$\dot{x}(t) = D_x f(x^*, x^*) x(t) + D_y f(x^*, x^*) x(t - \tau) \tag{2.8}$$

with

$$D_x f(x^*, x^*) \stackrel{\text{def}}{=} \frac{\partial}{\partial x} f(x,y)\bigg|_{x=x^*, y=x^*},$$
$$D_y f(x^*, x^*) \stackrel{\text{def}}{=} \frac{\partial}{\partial y} f(x,y)\bigg|_{x=x^*, y=x^*}.$$

In the case $\tau > 0$, even when $n = 1$, the behaviors of solutions of (2.8) can be more complicated than any given linear system of ODEs, since (2.8) even when $n = 1$ may have infinitely many linearly independent solutions $\mathrm{e}^{\lambda t}$ with λ being given by the so-called characteristic equation

$$\lambda = D_x f(x^*, x^*) + D_y f(x^*, x^*) \mathrm{e}^{-\lambda \tau}. \tag{2.9}$$

In particular, the infinite-dimensionality of the problem (2.1) leads to a transcendental equation (rather than a polynomial), which can have multiple zeros on the imaginary axis, giving rise to complicated critical cases.

On the other hand, some special features (specially the eventual compactness of the solution semiflow) of DDEs ensure that the sequence of zeros of the characteristic equation on the imaginary axis (counting multiplicity, either algebraically or geometrically, as will be specified later) must be finite. This gives a finite-dimensional center manifold of system (2.1) in a neighborhood of the equilibrium state $\hat{x}^*$, so that the asymptotic behaviors of solutions of (2.1) in a neighborhood of $\hat{x}^*$ can be captured by the reduced system on the center manifold, and this reduced system is an ODE system even though its dimension can be high.

We aim to introduce systematically the approach that enables us to derive the specific form of the reduced ODE system on the center manifold, explicitly in terms of the original system (2.1). Some forms of system (2.1) from application problems come with a parameter, and since the asymptotic behaviors of solutions near a given equilibrium may change qualitatively when the parameter varies (the so-called bifurcation), our focus will be on how the center manifold and the reduced ODE system on the center manifold change when the parameter is varied.

We should mention the step-by-step method in solving (2.1) on $[0,\tau]$, $[\tau,2\tau]$, ... inductively, which, though effectively numerically, may not give useful qualitative information about asymptotic behaviors of solutions. This method is also not useful in solving the kind of DDE with distributed delay such as

$$\dot{x}(t) = \int_{-\tau}^{0} f(x(t), x(t+\theta))d\theta$$

or

$$\dot{x}(t) = f\left(x(t), \int_{-\tau}^{0} g(x(t+\theta))d\theta\right)$$

with g: $\mathbb{R}^n \to \mathbb{R}^n$. One should also mention that in case the change rate of $x(t)$ depends on the historical value of $\dot{x}(t+\theta)$ with $\theta \in [-\tau, 0]$, such as

$$\dot{x}(t) = c\dot{x}(t-\tau) + f(x(t), x(t-\tau)),$$

we encounter additional difficulties, which shall be discussed later.

2.2 The Framework for DDEs

2.2.1 Definitions

Assume that $\mathbb{R}^n$ is equipped with the Euclidean norm $|\cdot|$. For a given constant $\tau \geq 0$, $C_{n,\tau} \stackrel{\text{def}}{=} C([-\tau,0],\mathbb{R}^n)$ denotes the Banach space of continuous mappings from $[-\tau,0]$ into $\mathbb{R}^n$ equipped with the supremum norm $\|\phi\| = \sup_{-\tau\leq\theta\leq 0}|\phi(\theta)|$ for $\phi \in C_{n,\tau}$. Moreover, if $t_0 \in \mathbb{R}$, $A \geq 0$, and $x : [t_0-\tau, t_0+A] \to \mathbb{R}^n$ is a continuous mapping, then for every $t \in [t_0, t_0+A]$, $x_t \in C_{n,\tau}$ is defined by $x_t(\theta) = x(t+\theta)$ for $\theta \in [-\tau, 0]$.

If f: $C_{n,\tau} \to \mathbb{R}^n$ is a mapping, we say that the equation

$$\dot{x} = f(x_t) \tag{2.10}$$

is a retarded functional differential equation (RFDE), or a delay differential equation (DDE). A function x is said to be a solution of (2.10) on $[t_0, t_0+A)$ if there are $t_0 \in \mathbb{R}$ and $A > 0$ such that $x \in C([t_0-\tau, t_0+A), \mathbb{R}^n)$, and $x(t)$ is differentiable and satisfies (2.10) for all $t \in [t_0, t_0+A)$. If f is locally Lipschitz (i.e., for every $\varphi \in C_{n,\tau}$ there exist a neighborhood $U \subseteq C_{n,\tau}$ of φ and a constant L such that $\|f(\phi) - f(\psi)\| \leq L\|\phi - \psi\|$ for all ϕ, $\psi \in U$), then for each given initial condition $(t_0, \varphi) \in \mathbb{R} \times C_{n,\tau}$, system (2.10) has a unique mapping x^{φ}: $[t_0-\tau, \beta) \to \mathbb{R}^n$ such that $x^{\varphi}\mid_{[t_0-\tau,t_0]} = \varphi$, x^{φ} is continuous for all $t \geq t_0 - \tau$, is differentiable, and satisfies (2.10) for $t \in (t_0, \beta)$, the maximal interval of existence of the solution x^{φ}. Furthermore, if $\beta < \infty$, then there exists a sequence $t_k \to \beta^-$ such that $|x^{\varphi}(t_k)| \to \infty$ as $k \to \infty$. For further results on existence, uniqueness, continuation, and continuous dependence of solutions for DDEs, see, for example, [18, 30, 51, 70, 120, 144–147, 154, 206, 208, 300, 302].

System (2.10) includes the following DDE with distributed delay

$$\dot{x}(t) = \int_{-\tau}^{0} g(\theta, x(t+\theta))d\theta, \tag{2.11}$$

and the following DDE with discrete delay

$$\dot{x}(t) = h(x(t), x(t-\tau_1), \ldots, x(t-\tau_k)), \tag{2.12}$$

where $\tau = \max\{\tau_1, \ldots, \tau_k\}$, g: $[-\tau, 0] \times \mathbb{R}^n \to \mathbb{R}^n$, and h: $\mathbb{R}^n \times \cdots \times \mathbb{R}^n (= \mathbb{R}^{n(k+1)}) \to \mathbb{R}^n$ are continuous. In these cases, for $\varphi \in C_{n,\tau}$,

$$f(\varphi) = \int_{-\tau}^{0} g(\theta, \varphi(\theta))d\theta$$

and

$$f(\varphi) = h(\varphi(0), \varphi(-\tau_1), \ldots, \varphi(-\tau_k)),$$

respectively. It can be shown that if h is locally Lipschitz (in (2.12)), then so is f. Similarly, if for every $x \in \mathbb{R}^n$ there exist a neighborhood U of $x \in \mathbb{R}^n$ and a constant $L > 0$ such that $|g(\theta, z) - g(\theta, y)| \leq L|z - y|$ for all $\theta \in [-\tau, 0]$ and $z, y \in U$, then the corresponding f is locally Lipschitz.

2.2.2 An Operator Equation

Throughout this chapter, we always assume that f: $C_{n,\tau} \to \mathbb{R}^n$ is continuously differentiable. Without loss of generality, assume that $f(0) = 0$, that is, 0 is an equilibrium point of (2.10). Let L be the linearized operator of f at this equilibrium point. Then the linearization of system (2.10) at this equilibrium point is

$$\dot{x}(t) = Lx_t. \tag{2.13}$$

We will consider the above linear system with a general linear operator $L : C_{n,\tau} \to \mathbb{R}^n$. Such an operator is clearly locally Lipschitz. For $\varphi \in C_{n,\tau}$, let $x = x^{\varphi}$ be the unique solution of (2.13) satisfying $x_0^{\varphi} = \varphi$. Then we have $|x(t)| \leq |\varphi(0)| + \int_0^t |L| ||x_s|| ds$ for all $t \geq 0$, from which it follows that $||x_t|| \leq ||\varphi|| + \int_0^t |L| ||x_s|| ds$ for $t \geq 0$ and hence $||x_t|| \leq ||\varphi|| e^{|L|t}$ for $t \geq 0$. This implies that the solution is defined for all $t \geq 0$. Here we use $|L|$ to denote the operator norm of the bounded operator L.

Define the solution operators $T(t) : C_{n,\tau} \to C_{n,\tau}$ by the relation

$$(T(t)\varphi)(\theta) = x_t^{\varphi}(\theta) = x(t+\theta) \tag{2.14}$$

for $\varphi \in C_{n,\tau}$, $\theta \in [-\tau, 0]$, $t \geq 0$. Then (2.13) can be thought of as maps from $C_{n,\tau}$ to $C_{n,\tau}$. Moreover,

(i) $T(t)$ is bounded and linear for $t \geq 0$;
(ii) $T(0)\varphi = \varphi$ or $T(0) = \mathrm{Id}$;

(iii) $\lim_{t\to t_0^+} \|T(t)\varphi - T(t_0)\varphi\| = 0$ for $\varphi \in C_{n,\tau}$.

Note that the inverse of $T(t), t \geq 0$, does not necessarily exist. Therefore, $T(t), t \geq 0$, is a *strongly continuous semigroup*.

An *infinitesimal generator* of a semigroup $T(t)$ is defined by

$$\mathscr{A}\varphi = \lim_{t\to 0^+} \frac{T(t)\varphi - \varphi}{t} \quad \text{for} \quad \varphi \in C_{n,\tau}.$$

In the case of the linear system (2.13), the infinitesimal generator can be constructed as

$$(\mathscr{A}\varphi)(\theta) = \begin{cases} d\varphi/d\theta, & \text{if} \quad \theta \in [-\tau, 0), \\ L\varphi, & \text{if} \quad \theta = 0. \end{cases} \tag{2.15}$$

We can show that the domain of $\mathscr{A}$ is given by

$$\mathrm{dom}(\mathscr{A}) = \{\varphi : \phi \in C^1_{n,\tau}, \varphi'(0) = L\varphi\}.$$

Then $T(t)\varphi$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t)\varphi = \mathscr{A}T(t)\varphi,$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t)\varphi = \lim_{h\to 0} \frac{T(t+h)\varphi - T(t)\varphi}{h}.$$

We may enlarge the phase space $C_{n,\tau}$ in such a way that (2.10) can be written as an abstract ODE in a Banach space. To accomplish this, for a positive integer n, let BC_n be the set of all functions from $[-\tau, 0]$ to $\mathbb{R}^n$ that are uniformly continuous on $[-\tau, 0)$ and may have a possible jump discontinuity at 0. We also introduce $X_0 : [-\tau, 0] \to BL(\mathbb{R}^n)$ defined by

$$X_0(\theta) = \begin{cases} \mathrm{Id}_n, & \theta = 0 \\ 0, & \theta \in [-\tau, 0). \end{cases} \tag{2.16}$$

Then every $\psi \in BC_n$ can be expressed as $\psi = \varphi + X_0\xi$ with $\varphi \in C_{n,\tau}$ and $\xi \in \mathbb{R}^n$. Thus BC_n can be identified with $C_{n,\tau} \times \mathbb{R}^n$. Equipped with the norm $|\varphi + X_0\xi|_{BC_n} = \|\varphi\| + |\xi|$, BC_n is a Banach space. In BC_n, we consider an extension of the infinitesimal generator of $\{T(t)\}_{t\geq 0}$, still denoted by $\mathscr{A}$,

$$\mathscr{A} : C^1_{n,\tau} \ni \psi \mapsto \dot{\psi} + X_0[L\psi - \dot{\psi}(0)] \in BC_n,$$

where $\dot{\psi} = \frac{\mathrm{d}}{\mathrm{d}\theta}\psi$. Thus, the abstract ODE in BC_n associated with (2.10) can be rewritten in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \mathscr{A}x_t + X_0 F(x_t), \tag{2.17}$$

where $F(x_t) = f(x_t) - Lx_t$. For $\theta \in [-\tau, 0)$, (2.17) is just the trivial equation $\mathrm{d}u_t/\mathrm{d}t = \mathrm{d}u_t/\mathrm{d}\theta$; for $\theta = 0$, it is (2.10).

2.2.3 Spectrum of the Generator

If the linear operator L: $C_{n,\tau} \to \mathbb{R}^n$ defined in (2.13) is continuous, then by the Riesz representation theorem, there exists an $n \times n$ matrix-valued function $\eta : [-\tau, 0] \to \mathbb{R}^{n^2}$ whose elements are of bounded variation such that (see, for example, Hale and Verduyn Lunel [154] for more details)

$$L\varphi = \int_{-\tau}^{0} d\eta(\theta)\varphi(\theta), \qquad \varphi \in C_{n,\tau}. \tag{2.18}$$

For example, consider $x'(t) = -x(t) + bx(t-1)$. Let $\eta : [-1, 0] \to \mathbb{R}$ be given such that $\eta(\theta) = 0$ for all $\eta \in (-1, 0)$ and $\eta(0) = -1$ and $\eta(-1) = -b$. Then $\int_{-1}^{0} d\eta(\theta)\varphi(\theta) = -\varphi(0) + b\varphi(-1)$ for $\varphi \in C_{1,1}$.

In general, the spectrum of an operator may consist of three different types of points, namely, the residual spectrum, the continuous spectrum, and the point spectrum. Moreover, points of the point spectrum are called eigenvalues of this operator. It is interesting to see that the spectrum $\sigma(\mathscr{A})$ of $\mathscr{A}$ consists of only the point spectrum. This implies that $\sigma(\mathscr{A})$ consists of eigenvalues of $\mathscr{A}$ and that λ is in $\sigma(\mathscr{A})$ if and only if λ satisfies the characteristic equation

$$\det \Delta(\lambda) = 0, \tag{2.19}$$

where $\Delta(\lambda)$ is the characteristic matrix of (2.13) and is given by

$$\Delta(\lambda) = \lambda \mathrm{Id}_n - \int_{-\tau}^{0} \mathrm{e}^{\lambda\theta} d\eta(\theta). \tag{2.20}$$

Here and in what follows, Id_n is the $n \times n$ identity matrix. We will not use the subscript n if that does not cause confusion.

For any $\lambda \in \sigma(\mathscr{A})$, the generalized eigenspace $\mathscr{M}_\lambda(\mathscr{A})$ is finite-dimensional, and there exists an integer k such that $\mathscr{M}_\lambda(\mathscr{A}) = \mathrm{Ker}((\lambda \mathrm{Id} - \mathscr{A})^k)$ and we have the direct sum decomposition

$$C_{n,\tau} = \mathrm{Ker}((\lambda \mathrm{Id} - \mathscr{A})^k) \oplus \mathrm{Ran}((\lambda \mathrm{Id} - \mathscr{A})^k),$$

where $\mathrm{Ker}((\lambda \mathrm{Id} - \mathscr{A})^k)$ and $\mathrm{Ran}((\lambda \mathrm{Id} - \mathscr{A})^k)$ represent the kernel and image of $(\lambda \mathrm{Id} - \mathscr{A})^k$, respectively. Clearly, $\mathscr{A}\mathscr{M}_\lambda(\mathscr{A}) \subseteq \mathscr{M}_\lambda(\mathscr{A})$.

The dimension $\mathscr{M}_\lambda(\mathscr{A})$ is the same as the order of zero for $\det\Delta(\lambda) = 0$.

Let $d = \dim \mathscr{M}_\lambda(\mathscr{A})$, let $\varphi_1, \ldots, \varphi_d$ be a basis for $\mathscr{M}_\lambda(\mathscr{A})$, and let $\Phi_\lambda = (\varphi_1, \ldots, \varphi_d)$. Then there exists a $d \times d$ constant matrix B_λ such that $\mathscr{A}\Phi_\lambda = \Phi_\lambda B_\lambda$. Moreover, we have the following properties:

(i) the only eigenvalue of B_λ is λ;
(ii) $\Phi_\lambda(\theta) = \Phi_\lambda(0)\mathrm{e}^{B_\lambda \theta}$;
(iii) $T(t)\Phi_\lambda = \Phi_\lambda \mathrm{e}^{B_\lambda t}$.

Therefore, we have the following result.

Theorem 2.1 (Hale and Verduyn Lunel [154]). *Suppose Λ is a finite set $\{\lambda_1, \ldots, \lambda_p\}$ of eigenvalues of (2.13), and let $\Phi_\Lambda = (\Phi_{\lambda_1}, \ldots, \Phi_{\lambda_p})$ and $B_\Lambda = \mathrm{diag}(B_{\lambda_1}, \ldots, B_{\lambda_p})$, where Φ_{λ_j} is a basis for the generalized space of $\mathscr{A}$ associated with λ_j and B_{λ_j} is the matrix defined by $\mathscr{A}\Phi_{\lambda_j} = \Phi_{\lambda_j} B_{\lambda_j}$, $j = 1, 2, \ldots, p$. Then the only eigenvalue of B_{λ_j} is λ_j, and for every vector v of the same dimension as the space P_Λ spanned by Φ_Λ, the solution $T(t)\Phi_\Lambda v$ with initial value $\Phi_\Lambda v$ at $t = 0$ may be defined on $(-\infty, \infty)$ by the relations*

$$T(t)\Phi_\Lambda v = \Phi_\Lambda \mathrm{e}^{B_\Lambda t} v$$

and

$$\Phi_\Lambda(\theta) = \Phi_\Lambda(0)\mathrm{e}^{B_\Lambda \theta}, \quad -\tau \leq \theta \leq 0.$$

Furthermore, there exists a subspace Q_Λ of $C_{n,\tau}$ such that $T(t)Q_\Lambda \subseteq Q_\Lambda$ for all $t \geq 0$ and

$$C_{n,\tau} = P_\Lambda \oplus Q_\Lambda. \tag{2.21}$$

2.2.4 An Adjoint Operator

We now describe a formal adjoint operator associated with (2.15). Let $C^*_{n,\tau} = C([0,\tau]; \mathbb{R}^{n*})$ be the space of continuous functions from $[0,\tau]$ to $\mathbb{R}^{n*}$ with

$$\|\psi\| = \sup_{t \in [0,\tau]} |\psi(t)|$$

for $\psi \in C^*_{n,\tau}$, where $\mathbb{R}^{n*}$ is the space of n-dimensional real row vectors. The formal adjoint equation associated with the linear RFDE (2.13) is given by

$$\dot{y} = -\int_{-\tau}^{0} y(t-\theta) d\eta(\theta). \tag{2.22}$$

For $\psi \in C^*_{n,\tau}$, let y^ψ be the unique solution of (2.22) satisfying $y^\psi_0 = \psi$ (in this subsection, $y_t \in C^*_{n,\tau}$ is defined as $y_t(s) = y(t+s)$ for $s \in [0,\tau]$).

If we define

$$(T^*(t)\psi)(\theta) = y^\psi_t(\theta) = y(t+\theta) \tag{2.23}$$

for $\psi \in C^*_{n,\tau}$, $\theta \in [0,\tau]$, $t \le 0$, then (2.23) defines a strongly continuous semigroup with the infinitesimal generator

$$(\mathscr{A}^*\psi)(\xi) = \begin{cases} -\mathrm{d}\psi(\xi)/\mathrm{d}\xi, & \text{if} \quad \xi \in (0,\tau], \\ \int_{-\tau}^{0} \psi(-\theta)d\eta(\theta), & \text{if} \quad \xi = 0. \end{cases} \tag{2.24}$$

Note that although the formal infinitesimal generator for (2.23) is defined as

$$A^*\psi = \lim_{t\to 0^-} \frac{T(t)\psi - \psi}{t} \quad \text{for} \quad \varphi \in C_{n,\tau},$$

Hale [144], for convenience, takes $\mathscr{A}^* = -A^*$ in (2.24) as the formal adjoint to (2.15). This family of operators (2.23) satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} T^*(t)\psi = -\mathscr{A}^* T^*(t)\psi.$$

In addition, it is easy to obtain the following results.

Theorem 2.2. *The following hold:*

(i) λ is an eigenvalue of $\mathscr{A}$ if and only if $\overline{\lambda}$ is an eigenvalue of $\mathscr{A}^*$.
(ii) The dimensions of the eigenspaces of $\mathscr{A}$ and $\mathscr{A}^*$ are finite and equal.
(iii) The dimensions of the generalized eigenspaces of $\mathscr{A}$ and $\mathscr{A}^*$ are finite and equal.

2.2.5 A Bilinear Form

In contrast to $\mathbb{R}^n$, the space $C_{n,\tau}$ does not have a natural inner product associated with its norm. However, following Hale [144], one can introduce a substitute device that acts like an inner product in $C_{n,\tau}$. This is an approach that is often taken when a function space does not have a natural inner product associated with its norm. Throughout, we will be assuming the complexification of the spaces so that we can work with complex eigenvalues and eigenvectors.

Define two operators $\Pi\colon C^1(\mathbb{R};\mathbb{R}^n) \to C(\mathbb{R};\mathbb{R}^n)$ and $\Omega\colon C^1(\mathbb{R};\mathbb{R}^{n*}) \to C(\mathbb{R};\mathbb{R}^{n*})$ as follows:

$$(\Pi x)(t) = \dot{x}(t) - \int_{-\tau}^{0} d\eta(\theta)x(t+\theta)$$

and

$$(\Omega y)(t) = \dot{y}(t) + \int_{-\tau}^{0} y(t-\theta)d\eta(\theta).$$

Then we have

$$\bar{y}(t)(\Pi x)(t) + (\overline{\Omega y})(t)x(t) = \frac{\mathrm{d}}{\mathrm{d}t}\langle y,x\rangle(t),$$

where

$$\langle y,x\rangle(t) = \bar{y}(t)x(t) - \int_{-\tau}^{0}\int_{0}^{\theta} \bar{y}(t+\xi-\theta)d\eta(\theta)x(t+\xi)d\xi. \tag{2.25}$$

Thus, if $x \in C^1(\mathbb{R};\mathbb{R}^n)$ and $y \in C^1(\mathbb{R};\mathbb{R}^{n*})$ satisfy $\Pi x = 0$ and $\Omega y = 0$, then $\langle y,x\rangle(t)$ is constant, and one can set $t = 0$ in (2.25) to define the bilinear form

$$\langle \psi,\varphi\rangle = \overline{\psi}(0)\varphi(0) - \int_{-\tau}^{0}\int_{0}^{\theta} \overline{\psi}(\xi-\theta)d\eta(\theta)\varphi(\xi)d\xi, \ \ \psi \in C^*_{n,\tau}, \varphi \in C_{n,\tau}. \tag{2.26}$$

In terms of (2.15) and (2.24), we see that

$$\langle \psi, \mathscr{A}\varphi\rangle = \langle \mathscr{A}^*\psi, \varphi\rangle$$

for $\varphi \in C_{n,\tau}$ and $\psi \in C^*_{n,\tau}$.

Let Λ be a set of some eigenvalues of $\mathscr{A}$ satisfying $\overline{\lambda} \in \Lambda$ if $\lambda \in \Lambda$. Denote by P and P^* the generalized eigenspaces of $\mathscr{A}$ and $\mathscr{A}^*$ associated with Λ, respectively. It follows from Theorem 2.2 that $\dim P = \dim P^*$. If $\varphi_1, \varphi_2, \ldots, \varphi_m$ is a basis for P and $\psi_1, \psi_2, \ldots, \psi_m$ is a basis for P^*, then construct the matrices $\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_m)$ and $\Psi = (\psi_1, \psi_2, \ldots, \psi_m)^T$. Define the bilinear form between Ψ and Φ by

$$\langle \Psi, \Phi\rangle = \begin{bmatrix} \langle \psi_1, \varphi_1\rangle & \ldots & \langle \psi_1, \varphi_m\rangle \\ \vdots & \ddots & \vdots \\ \langle \psi_m, \varphi_1\rangle & \ldots & \langle \psi_m, \varphi_m\rangle \end{bmatrix}.$$

This matrix is nonsingular and can be chosen so that $\langle \Psi, \Phi\rangle = \mathrm{Id}_m$. In fact, if $\langle \Psi, \Phi\rangle$ is not the identity, then a change of coordinates can be performed by setting $K = \langle \Psi, \Phi\rangle^{-1}$ and $\tilde{\Psi} = K\Psi$. Then $\langle \tilde{\Psi}, \Phi\rangle = \langle K\Psi, \Phi\rangle = K\langle \Psi, \Phi\rangle = \mathrm{Id}_m$. The decomposition (2.21) of $C_{n,\tau}$ given by Theorem 2.21 may be written explicitly as

$$\varphi = \varphi_p + \varphi_q,$$

where $\varphi_p = \Phi_\Lambda \langle \Psi_\Lambda, \varphi\rangle \in P_\Lambda$, $\varphi_q \in Q_\Lambda = \{\phi : \langle \Psi_\Lambda, \phi\rangle = 0\}$.

Remark 2.1. The bilinear form in $C^*_{n,\tau} \times C_{n,\tau}$ given by (2.26) can be extended in a natural way to $C^*_{n,\tau} \times BC_n$ by setting $\langle \psi, X_0\rangle = \overline{\psi}(0)$. We defer to Sect. 2.3 for a discussion how this extended bilinear form allows us to cast a functional differential equation to a system defined on the spaces P and Q_Λ.

2.2.6 Neural Networks with Delay: A Case Study on Characteristic Equations

In this section, we provide a detailed case study for the characteristic equation of the linearization at the trivial equilibrium of a coupled network of neurons with delayed feedback. Such a network with feedback with different interneuron and intraneuron time lags arises naturally in biological neural populations and their hardware implementation, and such a network also provides a simple-looking delay differential system that can exhibit complicated dynamics due to the existence of multiple eigenvalues of the infinitesimal generator of the linearized system at a given equilibrium when the synaptic connections and signal transmission delays are in certain ranges.

2.2.6.1 General Additive Neural Networks with Delay

We first describe an artificial neural network consisting of electronic neurons (amplifiers) interconnected through a matrix of resistors. Here an electronic neuron, the building block of the network, consists of a nonlinear amplifier that transforms an input signal u_i into the output signal v_i, and the input impedance of the amplifier unit is described by the combination of a resistor ρ_i and a capacitor C_i. We assume that the input–output relation is completely characterized by a voltage amplification function $v_i = f_i(u_i)$. The synaptic connections of the network are represented by resistors R_{ij} that connect the output terminal of the amplifier j with the input part of the neuron i. In order for the network to function properly, the resistances R_{ij} must be able to take on negative values. This can be realized by supplying each amplifier with an inverting output line that produces the signal $-v_j$. The number of rows in the resistor matrix is doubled, and whenever a negative value of R_{ij} is needed, this is realized using an ordinary resistor that is connected to the inverting output line.

The time evolution of the signals of the network is described by the Kirchhoff's law. Namely, the strengths of the incoming and outgoing current at the amplifier input port must balance. Consequently, we arrive at

$$C_i \frac{\mathrm{d}u_i}{\mathrm{d}t} + \frac{u_i}{\rho_i} = \sum_{j=1}^{n} \frac{1}{R_{ij}}(v_j - u_i).$$

Let

$$\frac{1}{R_i} = \frac{1}{\rho_i} + \sum_{j=1}^{n} \frac{1}{R_{ij}}.$$

We get

$$C_i R_i \frac{\mathrm{d}u_i}{\mathrm{d}t} + u_i = \sum_{j=1}^{n} \frac{R_i}{R_{ij}} v_j.$$

In the above derivation of the model equation for an artificial neural network, we implicitly assumed that the neurons communicate and respond instantaneously. Consideration of the finite switching speed of amplifiers requires that the input–output relation be replaced by $v_i = f_i(u_i(t-\tau_i))$ with a positive constant $\tau_i > 0$, and thus we obtain the following system of delay differential equations (see also [168, 209, 252, 267, 278]):

$$C_i R_i \frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -u_i(t) + \sum_{j=1}^{n} \frac{R_i}{R_{ij}} f_j(u_j(t-\tau_j)), \qquad 1 \le i \le n.$$

In what follows, for the sake of simplicity, we assume that

$$C_i = C, \quad R_i = R, \quad 1 \leq i \leq n,$$

and thus all local relaxation times $C_iR_i = CR$ are the same. Rescaling the time delay with respect to the network's relaxation time and rescaling the synaptic connection by

$$x_i(t) = u_i(CRt), \quad r_j = \frac{\tau_j}{RC}, \quad w_{ij} = \frac{R}{R_{ij}},$$

we get

$$x_i'(t) = -x_i(t) + \sum_{j=1}^{n} w_{ij} f_j(x_j(t - r_j)).$$

It is now easy to observe that it is the relative size of the delay r_j that determines the dynamics and the computational performance of the network, and designing a network to operate more quickly will increase this relative size of the delay.

It is therefore important to examine the effect of signal delays on the network dynamics. An important issue that has been addressed in the literature is how signal delays change the stability of equilibria, causing nonlinear oscillations and inducing periodic solutions. It will be shown that increasing the delay is among many mechanisms to create a network that exhibits periodic oscillations. Obviously, whether delay can generate oscillation also depends on the network connection topology. We refer to the monographs [224, 304] and a book chapter [52] for discussions about the relevance of this type of artificial neural network for the study of biological neural populations. In particular, we emphasize the importance of temporal delays in the coupling between cells, since in many chemical and biological oscillators (cells coupled via membrane transport of ions), the time needed for transport or processing of chemical components or signals may be of considerable length.

2.2.6.2 Special Case: Two Neurons

We now consider the following system of two neurons:

$$\begin{cases} \dot{x}_1(t) = -x_1(t) + \beta f(x_1(t-\tau)) + a_{12} f(x_2(t-\tau_1)), \\ \dot{x}_2(t) = -x_2(t) + \beta f(x_2(t-\tau)) + a_{21} f(x_1(t-\tau_2)), \end{cases} \tag{2.27}$$

where $x_1(t)$ and $x_2(t)$ denote the activations of the two neurons, $\tau_i (i = 1,2)$ and τ denote the synaptic transmission delays, a_{12} and a_{21} are the synaptic coupling weights, $f : \mathbb{R} \to \mathbb{R}$ is the activation function. Throughout this subsection, we always assume that $\tau_1 + \tau_2 = 2\tau > 0$ and $f : \mathbb{R} \to \mathbb{R}$ is a C^1-smooth function with $f(0)=0$. Without loss of generality, we also assume that $\tau_1 \geq \tau_2$ and $f'(0) = 1$. Letting $x(t) = (x_1(t), x_2(t))^T$ and $x_t(\theta) = x(t+\theta)$ for $\theta \in [-\tau_1, 0]$, we can rewrite (2.27) as

$$\dot{x}(t) = Lx_t + F(x_t)$$

with

$$L\varphi = -\varphi(0) + B_1\varphi(-\tau_1) + B_2\varphi(-\tau_2) + B\varphi(-\tau)$$

and

$$F(\varphi) = \frac{f''(0)}{2}\begin{bmatrix} a_{11}\varphi_1^2(-\tau) + a_{12}\varphi_2^2(-\tau_1) \\ a_{21}\varphi_1^2(-\tau_2) + a_{22}\varphi_2^2(-\tau) \end{bmatrix}$$

$$+ \frac{f'''(0)}{6}\begin{bmatrix} a_{11}\varphi_1^3(-\tau) + a_{12}\varphi_2^3(-\tau_1) \\ a_{21}\varphi_1^3(-\tau_2) + a_{22}\varphi_2^3(-\tau) \end{bmatrix} + o(\|\varphi\|^3)$$

for $\varphi = (\varphi_1, \varphi_2)^T \in C_{2,\tau_1}$, where

$$B_1 = \begin{bmatrix} 0 & a_{12} \\ 0 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 \\ a_{21} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix}.$$

The linearized system of (2.27) can be written as

$$\dot{x} = Lx_t = \int_{-\tau_1}^{0} d\eta(\theta)x(t+\theta), \tag{2.28}$$

where the matrix function $\eta(\theta)$ is given by

$$\eta(\theta,\mu) = \begin{cases} B_1 + B + B_2 - \mathrm{Id}_n, & \theta = 0, \\ B_1 + B + B_2, & \theta \in [-\tau_2, 0), \\ B_1 + B, & \theta \in [-\tau, -\tau_2), \\ B_1, & \theta \in (-\tau_1, -\tau), \\ 0, & \theta = -\tau_1, \end{cases}$$

and $\delta(\theta)$ is the Dirac delta function. The formal adjoint equation associated with (2.28) is given by

$$\dot{y}(t) = y(t) - y(t+\tau_1)B_1 - y(t+\tau_2)B_2 - y(t+\tau)B.$$

The bilinear form is

$$\begin{aligned} \langle \psi, \varphi \rangle &= \overline{\psi}(0)\varphi(0) + \textstyle\int_{-\tau_1}^{0} \overline{\psi}(s+\tau_1)B_1\varphi(s)ds \\ &\quad + \textstyle\int_{-\tau_2}^{0} \overline{\psi}(s+\tau_2)B_2\varphi(s)ds + \int_{-\tau}^{0} \overline{\psi}(s+\tau)B\varphi(s)ds. \end{aligned} \tag{2.29}$$

The operators $\mathscr{A}$ and $\mathscr{A}^*$ are given by

$$(\mathscr{A}\varphi)(\theta) = \begin{cases} \frac{d\varphi(\theta)}{d\theta}, & \text{if} \quad \theta \in [-\tau_1, 0), \\ -\varphi(0) + B_1\varphi(-\tau_1) + B_2\varphi(-\tau_2) + B\varphi(-\tau), & \text{if} \quad \theta = 0, \end{cases}$$

and

$$(\mathscr{A}^*\psi)(\theta) = \begin{cases} -\frac{d\psi(\xi)}{d\xi}, & \text{if} \quad \xi \in (0, \tau_1], \\ -\psi(0) + \psi(\tau_1)B_1 + \psi(\tau_2)B_2 + \psi(\tau)B, & \text{if} \quad \xi = 0. \end{cases}$$

Moreover, φ is in $\mathrm{Ker}(\lambda \mathrm{Id} - \mathscr{A})$ if and only if $\varphi(\theta) = \mathrm{e}^{\lambda\theta} v$, $-\tau_1 \leq \theta \leq 0$, where v is a vector in $\mathbb{R}^2$ such that $\Delta(\lambda)v = 0$ and the characteristic matrix $\Delta(\lambda)$ is

$$\Delta(\lambda) = \begin{bmatrix} \lambda + 1 - \beta\mathrm{e}^{-\lambda\tau} & -a_{12}\mathrm{e}^{-\lambda\tau_1} \\ -a_{21}\mathrm{e}^{-\lambda\tau_2} & \lambda + 1 - \beta\mathrm{e}^{-\lambda\tau} \end{bmatrix}.$$

Thus, the characteristic equation is

$$\det\Delta(\lambda) = [\lambda + 1 - \beta\mathrm{e}^{-\lambda\tau}]^2 - a_{12}a_{21}\mathrm{e}^{-2\lambda\tau} = 0. \tag{2.30}$$

Also, ψ is in $\mathrm{Ker}(\lambda \mathrm{Id} - \mathscr{A}^*)$ if and only if $\psi(\xi) = \mathrm{e}^{\lambda\xi} u$, $0 \leq \xi \leq \tau_1$, where u is a vector in $\mathbb{R}^{2*}$ such that $u\Delta(-\lambda) = 0$.

Let $\gamma_\pm = \beta \pm \sqrt{a_{12}a_{21}}$, where $\sqrt{a_{12}a_{21}}$ is a real if $a_{12}a_{21} > 0$ and purely imaginary otherwise. Then, $\det\Delta(\lambda)$ can be decomposed as

$$\det\Delta(\lambda) = [\lambda + 1 - \gamma_+\mathrm{e}^{-\lambda\tau}][\lambda + 1 - \gamma_-\mathrm{e}^{-\lambda\tau}].$$

Thus, in order to investigate the distribution of zeros of $\det\Delta(\lambda)$, we first consider the distribution of zeros of the following function:

$$P_z(\lambda) = \lambda + 1 - z\mathrm{e}^{-\lambda\tau}, \tag{2.31}$$

where $z \in \mathbb{C}$. Define a parametric curve Σ with the parametric equations

$$\begin{cases} u(t) = \cos\tau t - t\sin\tau t, \\ v(t) = t\cos\tau t + \sin\tau t, \end{cases} \qquad t \in \mathbb{R}. \tag{2.32}$$

It is easy to see that the curve Σ is symmetric about the u-axis. Let $\theta(t) = v(t)/u(t)$. Then $\theta'(t) = u^{-2}(t)[1 + \tau + \tau t^2] > 0$ for all $t \in \mathbb{R}$ such that $u(t) \neq 0$. This implies that as t increases, the corresponding point $(u(t), v(t))$ on the curve Σ moves counterclockwise about the origin. Moreover, it follows from $u^2(t) + v^2(t) = 1 + t^2$ that $\Sigma^+ = \{(u(t), v(t)) : t \in \mathbb{R}^+\}$ is simple, i.e., it cannot intersect itself. Let $\{t_n\}_{n=0}^{+\infty}$ be the monotonic increasing sequence of the nonnegative zeros of $v(t)$, and $c_n = u(t_n)$ for all $n \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$. Obviously, we have $t_0 = 0$ and $t_n \in ((2n-1)\pi/(2\tau), n\pi/\tau)$ for all $n \in \mathbb{N}$. Therefore, the curve Σ intersects with the u-axis at $(c_n, 0)$, $n \in \mathbb{N}_0$. It follows from the counterclockwise property of the curve Σ that $(-1)^n c_n > 0$ for all $n \in \mathbb{N}_0$. In addition, we have $|c_n| = \sqrt{1 + t_n^2}$, which implies that $c_n = (-1)^n\sqrt{1 + t_n^2}$ for $n \in \mathbb{N}_0$ and $\{|c_n|\}_{n\in\mathbb{N}_0}$ is an increasing sequence. In particular, $c_0 = 1$ and $c_1 = \sec\tau t_1 < -1$. Moreover, we claim that

$$(-1)^n v'(t_n) > 0 \qquad \text{and} \qquad (-1)^n u'(t_n) \geq 0 \qquad \text{for } n \in \mathbb{N}_0. \tag{2.33}$$

Equality in the second formula of (2.33) holds if and only if $n = 0$. In fact, we can check that $v'(t_n) \neq 0$ when $v(t_n) = 0$. This, combined with the counterclockwise property of the curve Σ, gives the first inequality in (2.33). From $u^2(t) + v^2(t) = 1 + t^2$, we have $u'(t)u(t) + v'(t)v(t) = t$ for $t \in \mathbb{R}^+$. Particularly, $u'(t_n)c_n = t_n$ for all $n \in \mathbb{N}_0$. This, combined with $(-1)^n c_n > 0$ for $n \in \mathbb{N}_0$, immediately implies the second inequality in (2.33). This proves the claim. Finally, $u^2(t)+v^2(t)=1+t^2\geq 1$ also implies that the curve is not inside the unit circle and it has only one intersection point $(1,0)$ with the unit circle.

For each $n\in\mathbb{N}_0$, let $\Sigma_n=\{(u(t),v(t)) : t\in[-t_{n+1},-t_n]\cup[t_n,t_{n+1}]\}$, which is a closed curve with $(0,0)$ inside. The curve Σ is schematically illustrated in Fig. 2.1. In the sequel, we will identify Σ with $\{u(t)+\mathrm{i}v(t) : t\in\mathbb{R}\}\subset\mathbb{C}$. The following

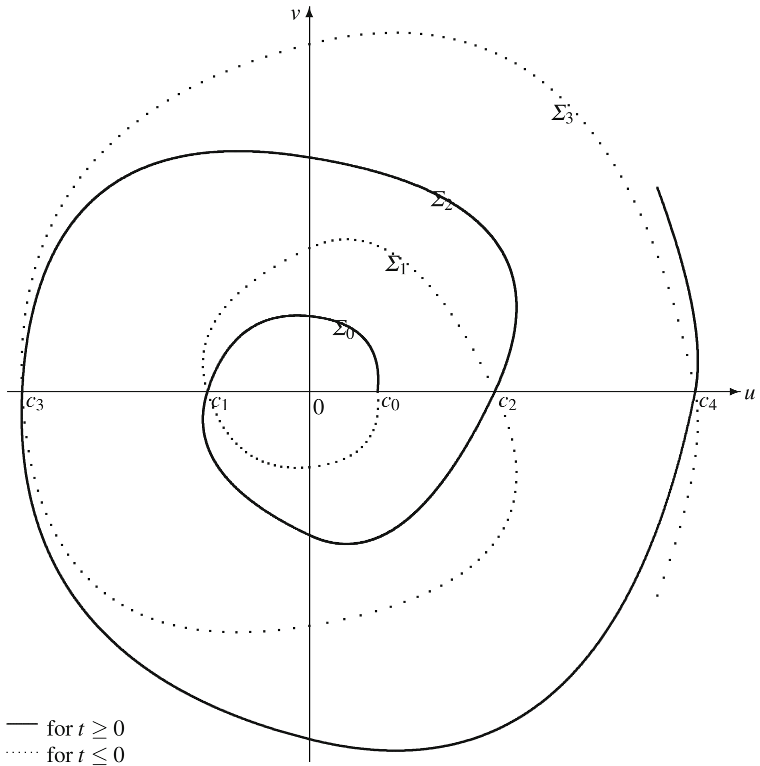

Fig. 2.1 The parametric curve Σ

lemma will play an important role in analyzing the distributions of the roots of (2.1).

Lemma 2.1. *Consider $P_z(\lambda)$ defined in (2.31) with $z\in\mathbb{C}$. Then the following statements are true:*

(i) *$P_z(\lambda)$ has a purely imaginary zero if and only if $z\in\Sigma$. Moreover, if $z=u(\theta)+\mathrm{i}v(\theta)$, then the purely imaginary zero is $\mathrm{i}\theta$ except that there is a pair of conjugate purely imaginary zeros $\pm\mathrm{i}t_n$ if $z=c_n$ for $n\in\mathbb{N}$.*

(ii) *For each fixed $z_0=u(\theta_0)+\mathrm{i}v(\theta_0)\in\Sigma$, there exist an open δ-neighborhood of z_0 in the complex plane, denoted by $B(z_0,\delta)$, and an analytic function $\lambda: B(z_0,\delta)\to\mathbb{C}$ such that $\lambda(z_0)=\mathrm{i}\theta_0$ and $\lambda(z)$ is a zero of $P_z(\lambda)$ for all $z\in$*

$B(z_0,\delta)$. Moreover, along the outward-pointing normal vector to the curve Σ at z_0, the directional derivative of $\mathrm{Re}\{\lambda(z)\}$ *at z_0 is positive.*

(iii) *$P_z(\lambda)$ has only zeros with strictly negative real parts if and only if z is inside the curve Σ_0, exactly $j \in \mathbb{N}$ zeros with positive real parts if z is between Σ_{j-1} and Σ_j. In particular, if $z \in \Sigma_0$, $P_z(\lambda)$ has either a simple real zero* 0 *(if $z = 1$) or a simple purely imaginary zero (if* $\mathrm{Im}(z) \neq 0$*), or a pair of simple purely imaginary zeros (if $z = c_1$), and all other zeros has strictly negative real parts.*

Proof. (i) $P_z(\lambda)$ has a purely imaginary zero, say $\lambda = \mathrm{i}\theta$, if and only if $\mathrm{e}^{\mathrm{i}\tau\theta}(1+\mathrm{i}\theta) = z$, which is equivalent to $z \in \Sigma$ by separating the real and imaginary parts of $\mathrm{e}^{\mathrm{i}\tau\theta}(1+\mathrm{i}\theta)$.

(ii) Note that $P_{z_0}(\mathrm{i}\theta_0) = 0$ and $\mathrm{i}\theta_0$ is a simple zero of $P_{z_0}(\lambda)$. The existence of δ and the mapping λ follow from the implicit function theorem. Moreover, $\lambda(z)$ is analytic with respect to z. Thus,

$$\lambda'(z) = \frac{\partial}{\partial a}\mathrm{Re}\{\lambda(z)\} + \mathrm{i}\frac{\partial}{\partial a}\mathrm{Im}\{\lambda(z)\} = \frac{\partial}{\partial b}\mathrm{Im}\{\lambda(z)\} - \mathrm{i}\frac{\partial}{\partial b}\mathrm{Re}\{\lambda(z)\},$$

where $a = \mathrm{Re}(z)$ and $b = \mathrm{Im}(z)$. On the other hand, differentiating $P_z(\lambda) = 0$ with respect to z and using $P_{z_0}(\mathrm{i}\theta_0) = 0$, we have

$$\lambda'(z_0) = \varepsilon_1\left[u_0\varepsilon_2 + \theta_0 v_0 + \mathrm{i}(\theta_0 u_0 - v_0\varepsilon_2)\right],$$

where $\varepsilon_1 = [(1+\tau)^2 + (\tau\theta_0)^2]^{-1}(1+\theta_0^2)^{-1}$ and $\varepsilon_2 = 1+\tau+\tau\theta_0^2$. It follows that

$$\begin{aligned}\nabla\mathrm{Re}\{\lambda(z_0)\} &= \left(\frac{\partial}{\partial a}\mathrm{Re}\{\lambda(z_0)\}, \frac{\partial}{\partial b}\mathrm{Re}\{\lambda(z_0)\}\right)^T \\ &= \varepsilon_1\left(u_0\varepsilon_2 + \theta_0 v_0, v_0\varepsilon_2 - \theta_0 u_0\right)^T.\end{aligned}$$

Let $\vartheta(\xi) = (v'(\theta_0), -u'(\theta_0))M(\xi)$, where $\xi \in (-\pi/2, \pi/2)$ and

$$M(\xi) = \begin{bmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{bmatrix}.$$

Obviously, for each fixed $\xi \in (-\pi/2, \pi/2)$, $\vartheta(\xi)$ is an outward-pointing vector to the curve Σ at z_0, Thus, the directional derivative along the vector $\vartheta(\xi)$ at z_0 is

$$\begin{aligned}\frac{d}{d\vartheta(\xi)}\mathrm{Re}\{\lambda(z_0)\} &= \varepsilon_3(v'(\theta_0), -u'(\theta_0))M(\xi)\nabla\mathrm{Re}\{\lambda(z_0)\} \\ &= \varepsilon_1\varepsilon_3(\varepsilon_2^2 + \theta_0^2)\cos\xi > 0,\end{aligned}$$

where $\varepsilon_3 = 1/\sqrt{(1+\tau)^2 + \tau^2\theta_0^2}$.

(iii) Note that $P_0(\lambda)$ has exactly one zero -1, which obviously has a negative real part. Since zeros of $P_z(\lambda)$ depend continuously on z, there exists a region Ω_0

containing $z = 0$ such that for $z \in \Omega_0$, all zeros of $P_z(\lambda)$ have negative real parts. Moreover, as z varies and passes through the boundary $\partial\Omega_0$, only one (or two if z is real) zero point of $P_z(\lambda)$ varies from a complex number with a negative real part to a purely imaginary number and then to a complex number with a positive real part. By (i), $\partial\Omega_0 = \Sigma_0$. Therefore, $P_z(\lambda)$ has only zeros with negative real parts if z in inside the curve Σ_0. If z is a real number between Σ_{j-1} and Σ_j, then one can easily show that $P_z(\lambda)$ has exactly j zeros with positive real parts (see, for example, the discussion in Chen and Wu [59]). This, combined with (i) and the continuous dependence of zeros of $P_z(\lambda)$ on z, completes the proof. □

In view of Lemma 2.1, we have the following conclusions:

(1) All zeros of $\det\Delta(\lambda)$ have negative real parts if and only if both of $\gamma_\pm$ are inside the curve Σ.
(2) If and only if $1 \neq \gamma_+ \in \Sigma$ or $1 \neq \gamma_- \in \Sigma$, $\det\Delta(\lambda)$ has a pair of simple conjugate purely imaginary zeros $\pm\mathrm{i}\omega$, where $\omega > 0$ satisfies either $u(\omega) + \mathrm{i}v(\omega) = \gamma_+$ or $u(\omega) + \mathrm{i}v(\omega) = \gamma_-$. In particular, $\omega = t_n$ if either γ_+ or γ_- is equal to c_n for some $n \in \mathbb{N}$.
(3) If and only if only one of γ_+ and γ_- is equal to 1, $\det\Delta(\lambda)$ has a simple zero $\lambda = 0$. Moreover, if $c_1 < \gamma_- < \gamma_+ = 1$, then all zeros but $\lambda = 0$ of $\det\Delta(\lambda)$ have strictly negative real parts.

If $a_{12}a_{21} > 0$ and only one of $\gamma_\pm$ lies on the curve Σ, or $a_{12}a_{21} < 0$ and $\gamma_\pm \in \Sigma$, then on the imaginary axis, the infinitesimal generator $\mathscr{A}$ has only one pair of simple purely imaginary eigenvalues $\pm\mathrm{i}\omega$. Let $\Phi = (\varphi_1, \varphi_2)$ and $\Psi = (\psi_1, \psi_2)^T$ be bases for the generalized eigenspaces $P_{\pm\mathrm{i}\omega}$ and $P^*_{\pm\mathrm{i}\omega}$ of $\mathscr{A}$ and $\mathscr{A}^*$ associated with eigenvalues $\pm\mathrm{i}\omega$, respectively. In fact, we can choose

$$\varphi_1(\theta) = \overline{\varphi}_2(\theta) = (1, d)^T \mathrm{e}^{\mathrm{i}\omega\theta}, \quad \theta \in [-\tau_1, 0],$$
$$\psi_1(\xi) = \overline{\psi}_2(\xi) = \overline{D}\left(\overline{d}, 1\right) \mathrm{e}^{\mathrm{i}\omega\xi}, \quad \xi \in [0, \tau_1],$$

and

$$d = (1 + \mathrm{i}\omega - \beta\mathrm{e}^{-\mathrm{i}\omega\tau})\mathrm{e}^{\mathrm{i}\omega\tau_1}/a_{12},$$
$$D = \{2d\,[1 + \tau(1 + \mathrm{i}\omega)]\}^{-1}.$$

Moreover, $\langle\psi_j, \varphi_k\rangle = \delta_{jk}$, $j, k = 1, 2$, where $\langle\cdot,\cdot\rangle$ is defined in (2.29) and

$$\delta_{jk} = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases}$$

Assume that $a_{12}a_{21} > 0$. If $\gamma_+ = 1$ and $\gamma_- = c_n$ or $\gamma_- = 1$ and $\gamma_+ = c_n$ for some $n \in \mathbb{N}$, then on the imaginary axis, the infinitesimal generator $\mathscr{A}$ has only simple eigenvalues 0, $\mathrm{i}t_n$, and $-\mathrm{i}t_n$. Here, we consider only the first case. Namely, assume that $a_{12}a_{21} > 0$ and $\gamma_+ = 1$ and $\gamma_- = c_n$ for some $n \in \mathbb{N}$. Let $\Phi = (q_0, q_1, \bar{q}_1)$, and $\Psi = (p_0, p_1, \bar{p}_1)^T$ be bases for the generalized eigenspaces P_Λ and P^*_Λ of $\mathscr{A}$ and $\mathscr{A}^*$ associated with $\Lambda = \{0, \mathrm{i}t_n, -\mathrm{i}t_n\}$. In fact, we can choose

$$q_0(\theta) = (1, d_0)^T, \quad q_1(\theta) = (1, d_1)^T e^{it_n\theta}, \quad \theta \in [-\tau_1, 0],$$

and

$$p_0(\xi) = D_0(d_0, 1), \quad p_1(\xi) = \overline{D}_1\left(\overline{d}_1, 1\right)e^{it_n\xi}, \quad \xi \in [0, \tau_1],$$

where $d_0 = (1-\beta)/a_{12}$, $d_1 = (1+it_n - \beta e^{-it_n\tau})e^{it_n\tau_1}/a_{12}$, $D_0 = [2d_0(1+\tau)]^{-1}$, and $D_1 = \{2d_1[1+\tau(1+it_n)]\}^{-1}$. Moreover, $\langle p_j, q_k\rangle = \delta_{jk}$ and $\langle p_j, \bar{q}_k\rangle = 0$, $j,k = 0,1$.

Assume that $a_{12}a_{21} > 0$. If $\gamma_+ = c_n$ and $\gamma_- = c_m$ for $n, m \in \mathbb{N}$ such that $c_n > c_m$, then on the imaginary axis, the infinitesimal generator $\mathscr{A}$ has only two pairs of simple purely imaginary eigenvalues $\pm i\omega_1$ and $\pm i\omega_2$, where $\omega_1 = t_n$ and $\omega_2 = t_m$. Let $\Phi = (q_1, \bar{q}_1, q_2, \bar{q}_2)$, and $\Psi = (p_1, \bar{p}_1, p_2, \bar{p}_2)^T$ be bases for the generalized eigenspaces P_Λ and P_Λ^* of $\mathscr{A}$ and $\mathscr{A}^*$ associated with $\Lambda = \{i\omega_1, -i\omega_1, i\omega_2, -i\omega_2\}$. In fact, we can choose

$$q_j(\theta) = (1, d_j)^T e^{i\omega_j\theta}, \quad \theta \in [-\tau_1, 0], \quad j = 1, 2,$$

and

$$p_j(\xi) = \overline{D}_j\left(\overline{d}_j, 1\right)e^{i\omega_j\xi}, \quad \xi \in [0, \tau_1], \quad j = 1, 2,$$

where $d_1 = (1+i\omega_j - \beta e^{-i\omega_j\tau})e^{i\omega_j\tau_1}/a_{12}$ and $D_j = \left\{2d_j[1+\tau(1+i\omega_j)]\right\}^{-1}$. Moreover, $\langle p_j, q_k\rangle = \delta_{jk}$ and $\langle p_j, \bar{q}_k\rangle = 0$, $j,k = 1,2$.

2.3 General Framework of NFDEs

Suppose that f, h: $C_{n,\tau} \to \mathbb{R}^n$ are given continuous mappings. The relation

$$\frac{d}{dt}h(x_t) = f(x_t) \tag{2.34}$$

is called a neutral functional differential equation (NFDE). The mapping h will be called the difference operator for NFDE (2.34). If $h(\varphi) = \varphi(0)$ for all φ, then (2.34) becomes (2.10). Consequently, DDEs are special cases of NFDEs.

A function x is said to be a solution of (2.34) on $[t_0, t_0 + A)$ for some $t_0 \in \mathbb{R}$ and $A > 0$ if $x \in C([t_0 - \tau, t_0 + A), \mathbb{R}^n)$, $x_t \in C_{n,\tau}$ for all $t \in [t_0, t_0 + A)$, $h(x_t)$ is continuously differentiable, and $x(t)$ satisfies (2.34) for all $t \in [t_0, t_0 + A)$.

Let D, L: $C_{n,\tau} \to \mathbb{R}^n$ be the two linearized operators of h and f at some equilibrium point, respectively. Without loss of generality, we assume that there exist two $n \times n$ matrix-valued functions $\mu, \eta : [-\tau, 0] \to \mathbb{R}^{n^2}$ whose components each have bounded variation and such that for $\varphi \in C_{n,\tau}$,

$$D\varphi = \varphi(0) - \int_{-\tau}^{0} d\mu(\theta)\varphi(\theta), \quad L\varphi = \int_{-\tau}^{0} d\eta(\theta)\varphi(\theta).$$

Moreover, we assume that D is atomic at zero, that is, $\mathrm{Var}_{[s,0]}\mu(\theta) \to 0$ as $s \to 0$ (see Hale and Verduyn Lunel [154] for more details). The linear system

$$\frac{\mathrm{d}}{\mathrm{d}t} D x_t = L x_t \tag{2.35}$$

generates a strongly continuous semigroup of linear operators with infinitesimal generator $\mathscr{A}$. The spectrum of $\mathscr{A}$, denoted by $\sigma(\mathscr{A})$, is the point spectrum. Moreover, λ is an eigenvalue of $\mathscr{A}$, i.e., $\lambda \in \sigma(\mathscr{A})$, if and only if λ satisfies $\det \Delta(\lambda) = 0$, where the characteristic matrix $\Delta(\lambda)$ is given by

$$\Delta(\lambda) = \lambda D(\mathrm{e}^{\lambda(\cdot)}\mathrm{Id}) - L(\mathrm{e}^{\lambda(\cdot)}\mathrm{Id}).$$

It is well known that $\phi \in C_{n,\tau}$ is an eigenvector of $\mathscr{A}$ associated with the eigenvalue λ if and only if $\phi(\theta) = \mathrm{e}^{\lambda\theta} b$ for $\theta \in [-\tau, 0]$ and some vector $b \in \mathbb{C}^n$ such that $\Delta(\lambda) b = 0$. Here and in the sequel, for the sake of convenience, we shall also allow functions with range in $\mathbb{C}^n$.

Let Λ be a set of some eigenvalues of $\mathscr{A}$, and denote by E_Λ the generalized eigenspace of $\mathscr{A}$ associated with Λ. It is known that $\dim E_\Lambda = m$, where m is the number of zeros of $\det \Delta(\lambda)$ in Λ, counting multiplicities. As we did earlier for DDEs, we define a bilinear form

$$\begin{aligned} \langle \psi, \varphi \rangle = \overline{\psi}(0)\varphi(0) &- \int_{-\tau}^{0} \left[\frac{\mathrm{d}}{\mathrm{d}s} \int_0^s \overline{\psi}(\xi - s) d\mu(\theta) \varphi(\xi) d\xi \right]_{s=\theta} \\ &- \int_{-\tau}^{0} \int_0^{\theta} \overline{\psi}(\xi - \theta) d\eta(\theta) \varphi(\xi) d\xi \end{aligned} \tag{2.36}$$

for $\psi \in C_{n,\tau}^*$ and $\varphi \in C_{n,\tau}$. Let Φ be a basis for E_Λ and Ψ the basis for the dual space E_Λ^* in C_n^* such that $\langle \Psi, \Phi \rangle = \mathrm{Id}_m$. The phase space $C_{n,\tau}$ is decomposed by Λ as $C_{n,\tau} = E_\Lambda \oplus Q_\Lambda$, where $Q_\Lambda = \{\phi \in C_{n,\tau} : \langle \Psi, \phi \rangle = 0\}$. Moreover, there exists an $m \times m$ constant matrix B with $\sigma(B) = \Lambda$ such that

$$\dot{\Phi} = \Phi B \quad \text{and} \quad \dot{\Psi} = -B\Psi.$$

Similarly to the previous sections for DDEs, we may enlarge the phase space $C_{n,\tau}$ such that (2.34) can be written as an abstract ODE in the Banach space BC_n. First, in BC_n, we consider an extension of the infinitesimal generator $\mathscr{A}$, still denoted by $\mathscr{A}$,

$$\mathscr{A} : BC_n \supset C_{n,\tau}^1 \ni \varphi \mapsto \dot{\varphi} + X_0[L\varphi - D\dot{\varphi}] \in BC_n,$$

where $\mathrm{Dom}(\mathscr{A}) = C_{n,\tau}^1 \overset{\mathrm{def}}{=} \{\varphi \in C_{n,\tau} : \dot{\varphi} \in C_{n,\tau}\}$. The bilinear form in $C_{n,\tau}^* \times C_{n,\tau}$ given by (2.36) is extended in a natural way to $C_{n,\tau}^* \times BC_n$ by setting $\langle \psi, X_0 \rangle = \overline{\psi}(0)$. Thus, the abstract ODE in BC_n associated with (2.34) can be rewritten in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} u = \mathscr{A} u + X_0 G(u), \tag{2.37}$$

where

$$G(u) = f(u) - Lu - \frac{\mathrm{d}}{\mathrm{d}t}\left[h(u) - Du\right]. \tag{2.38}$$

Consider the projection $\pi : BC_n \to E_\Lambda$ given by

$$\pi(\varphi + X_0\xi) = \Phi[\langle \Psi, \varphi \rangle + \overline{\Psi}(0)\xi]. \tag{2.39}$$

Obviously, π is a continuous projection onto E_Λ, which commutes with $\mathscr{A}$ in $C^1_{n,\tau}$. This allows us to decompose BC_n as a topological direct sum, $BC_n = E_\Lambda \oplus \operatorname{Ker}\pi$, where $Q_\Lambda \subset \operatorname{Ker}\pi$.

Due to the decomposition of BC_n, we can decompose u in (2.37) in the form $u = \Phi x + y$, with $x \in \mathbb{R}^m$, $y \in Q \stackrel{\mathrm{def}}{=} \operatorname{Ker}\pi \cap C^1_{n,\tau}$. Then (2.37) is equivalent to the system

$$\begin{aligned} \dot{x} &= Bx + \overline{\Psi}(0)G(\Phi x + y), \\ \tfrac{\mathrm{d}y}{\mathrm{d}t} &= \mathscr{A}_Q y + (\mathrm{I} - \pi)X_0 G(\Phi x + y), \end{aligned} \tag{2.40}$$

where $\mathscr{A}_Q$ is the restriction of $\mathscr{A}$ to Q interpreted as an operator acting in the Banach space $\operatorname{Ker}\pi$. The spectrum of $\mathscr{A}_Q$ will be very important for the construction of normal forms. Similarly, $\mathscr{A}_Q$ has only a point spectrum. Moreover, $\sigma(\mathscr{A}_Q) = \sigma(\mathscr{A}) \setminus \Lambda$.

http://www.springer.com/978-1-4614-6991-9

Bifurcation Theory of Functional Differential Equations
Guo, S.; Wu, J.
2013, IX, 289 p., Hardcover
ISBN: 978-1-4614-6991-9