Delay equations and characteristic roots: stability and more from a single curve

Dimitri Breda$^1, 2$, Giulia Menegon$^2$ and Monica Nonino$^3$

$^1$CDLab – Computational Dynamics Laboratory
$^2$Department of Mathematics, Computer Science and Physics, University of Udine, via delle scienze 206, 33100 Udine, Italy
$^3$International School for Advanced Studies – SISSA, via Bonomea 265, 34136 Trieste, Italy

Received 5 May 2017, appeared 10 October 2018
Communicated by Gergely Röst

Abstract. Delays appear always more frequently in applications, ranging, e.g., from population dynamics to automatic control, where the study of steady states is undoubtedly of major concern. As many other dynamical systems, those generated by nonlinear delay equations usually obey the celebrated principle of linearized stability. Therefore, hyperbolic equilibria inherit the stability properties of the corresponding linearizations, the study of which relies on associated characteristic equations. The transcendence of the latter, due to the presence of the delay, leads to infinitely-many roots in the complex plane. Simple algebraic manipulations show, first, that all such roots belong to the intersection of two curves. Second, only one of these curves is crucial for stability, and relevant sufficient and/or necessary criteria can be easily derived from its analysis. Other aspects can be investigated under this framework and a link to the theory of modulus semigroups and monotone semiflows is also discussed.

Keywords: delay equations, characteristic roots, equilibria, stability analysis.

2010 Mathematics Subject Classification: 34K20, 34K06.

1 Introduction

Delay equations are nowadays ubiquitous, as witnessed by the growing number of monographs or reviews dealing with either theory, methods and applications [2,4–6,13,20,21,24,27,28,31,32,34–36,39,40,45,48]. They generate dynamical systems on infinite-dimensional state spaces, usually Banach spaces of functions defined on the delay interval. Possible choices are continuous and Lebesgue-measurable functions, or spaces with a Hilbert structure, see, e.g., [18] for the former and [14,16] for the latter. Hereafter we consider the classic state space of continuous functions defined on the delay interval.

In most applications, the stability of equilibria represents a primary interest. The basic reproduction number in population dynamics is an example of scientific motivation of broad
reach. Universally known as “$R_0$”, it dictates the transition of asymptotic stability from trivial to nontrivial equilibria, thus assuming a prominent role in the study of epidemics.

In general, by resorting to linearization, the local problem is reduced to determine the position in the complex plane of the so-called characteristic roots. As largely known, these can be seen either as solutions of the associated characteristic equation or, equivalently, as eigenvalues of the infinitesimal generator of the semigroup of solution operators, see [27, Chapter 7, Lemma 2.1] or [20, Chapter IV, Corollary 3.3]. Both characterizations have pros and cons.

In the last couple of decades, for instance, the theory of strongly continuous semigroups has furnished a solid theoretical background to most numerical methods devoted to the approximation of a (dominant or rightmost) part of the spectrum, see, e.g., [10, 12, 15, 23] to name just a few specific techniques, or [13, 40] for a compendium. Nevertheless, the approach through the characteristic equation is still extensively adopted, although necessarily restricted to selected, mainly low-dimensional, yet important and interesting models. [7, 19, 30, 37] are examples of recent results along this line. One way or the other, the effort is often tuned to finding suitable stability conditions depending on varying parameters, either in terms of more or (probably) less comprehensible analytical constraints and inequalities, or by using graphical representations such as stability boundaries and charts. Perhaps, in general, the approach through the characteristic equation is more advantageous for obtaining qualitative insight, while the numerical discretization of the infinite-dimensional operators is more suitable for quantitative studies.

In the present work, resuming from part of [9], we investigate a methodology based on the characteristic equation, which seems particularly practical for the scalar prototype instances of Delay Differential Equations (DDEs) with either single (Section 2), multiple (Section 3) or distributed delays (Section 4), as well as for more general classes of delay equations that may include neutral or renewal ones (Section 5). The underlying approach relies on recognizing the characteristic roots as intersection points of two curves, only one of which is essential in view of stability. Given its vertical develop in the complex plane, in fact, it can be conveniently bounded from the right. Consequently, sufficient and/or necessary stability criteria can be easily formulated, as well as straightforwardly adopted in the practice of applications. Various computational aspects are indeed discussed in Section 6, including approximation of distributed delays. Also, other kinds of results can be obtained as directly, related to, e.g., delay-independent stability regions (Example 2.4) or bifurcation (Example 2.5).

Incidentally, the presence of a right bound and the way it is recovered fall in the lines of the theory of modulus semigroups [3, 8, 47, 49], related also to monotone semiflows [33, 43, 44], thus providing a strong theoretical support to the whole procedure. This point of view is considered in Section 7 as a concluding perspective.

Finally, although the present research is not targeted to any specific model, the hope is that future works based on these ideas can extend the understanding of the behavior of the roots with respect to varying parameters also to more general systems. In this sense the main contribution is to be found in the general methodology rather than in the specific use of the latter to address this or that application. And let us underline that this general understanding of the roots is still lacking. Indeed, despite the wide knowledge acquired by the delay community, the seemingly simple problem of, e.g., relating the characteristic roots of $x'(t) = Ax(t) + Bx(t - r)$ to the matrix coefficients or to the delay is still open, as anticipated 40 years ago by Jack Hale: “the exact region of stability as an explicit function of $A$, $B$ and $r$ is not known and probably will never be known.”, [27, p. 136]. In this respect, more comments can be
found in [9], while [43, Corollary 3.2] represents perhaps the most far-reaching result under positivity and irreducibility assumptions on the right-hand side, see also Section 7. Other works relying on positivity exist, as well as on stability in general. For a brief but nice account on stability and positivity see, e.g., [26, Section 4].

2 A single discrete delay

Consider the linear DDE with a single discrete delay

\[ x'(t) = ax(t) + bx(t - \tau) \]  

for parameters \( a, b \in \mathbb{R} \) and delay \( \tau > 0 \). The associated characteristic equation is

\[ \lambda - a - be^{-\lambda \tau} = 0, \]  

which amounts to the equality between complex numbers

\[ \lambda - a = be^{-\lambda \tau}. \]

Hereafter let us set \( \lambda = \alpha + i\omega \) for \( \alpha, \omega \in \mathbb{R} \), and restrict to \( \omega \geq 0 \) without loss of generality for conjugation-invariance.

A classic approach is used to obtain the stability chart (see, e.g., [20, Figure XI.1], [27, Figure 5.1], [48, Fig. 2.1]), but also to derive numerical routines (see, e.g., [50]). It consists in decomposing (2.3) into its real and imaginary parts:

\[ \alpha - a = be^{-\alpha \tau} \cos \omega \tau, \]

\[ -\omega = be^{-\alpha \tau} \sin \omega \tau. \]

The approach in [9], instead, relies on equating magnitude and argument, precisely the square of the first and the tangent of the second:

\[ (\alpha - a)^2 + \omega^2 = b^2 e^{-2\alpha \tau}, \]

\[ -\frac{\omega}{\alpha - a} = \tan \omega \tau. \]  

The originating motivation lies in getting rid of the trigonometric factors through the former and of the exponential factor through the latter. Equation (2.5) is solved explicitly for \( \omega \) as a function of \( \alpha \), viz.

\[ \omega = E(\alpha) := \sqrt{b^2 e^{-2\alpha \tau} - (\alpha - a)^2}, \]

while (2.6) is solved explicitly for \( \alpha \) as a function of \( \omega \), viz.

\[ \alpha = T(\omega) := a - \frac{\omega}{\tan \omega \tau}. \]

The graphs of \( E \) and \( T \) define curves in the \((\alpha, \omega)\)- and \((\omega, \alpha)\)-planes, respectively. Let us put both in the same \((\alpha, \omega)\)-plane by defining

\[ \mathcal{E} := \{(\alpha, E(\alpha)) : \alpha \in \text{dom}(E)\} \]

and

\[ \mathcal{T} := \{(T(\omega), \omega) : \omega \in \text{dom}(T)\}. \]

The following result is straightforward. Although the explicit reference to (2.1), the statement is valid for more general equations, see later on. Here, the case \( \omega = 0 \) is special because \( T \) is not defined, although the latter can be continuously prolonged by letting \( T(0) = a - 1/\tau \).
**Theorem 2.1.** If \( \lambda = \alpha + i\omega \) is a characteristic root of (2.1) with \( \omega > 0 \), then \((\alpha, \omega) \in E \cap T\). If \( \omega = 0 \), then \( \lambda \in E \cap \mathbb{R} \times \{0\} \), i.e., \( \lambda \) is a zero of \( E \).

The converse is not entirely true: “half” of the infinitely-many intersection points between \( E \) and \( T \) are spurious characteristic roots due to the squaring of the magnitude in (2.5). Deciding which half is not difficult, see [9, Proposition 14] for (2.1), but it is shown in Section 2.3 that it is not even important, when stability is the target. First it is necessary to know the geometry of both curves. A detailed analysis is left to [9], from which we resume the essential aspects in the forthcoming sections. The possible configurations for (2.1) are anticipated in Figure 2.1. Here, as well as in the following similar figures, the roots are computed with the codes provided in [13].

Figure 2.1: Curve \( E \) (thick red), curve \( T \) (thin blue) and characteristic roots (•) of (2.1) for \( b > 0 \) (left column), \( b < 0 \) (right column), \(|b| > e^{a\tau - 1}/\tau \) (top row), \(|b| = e^{a\tau - 1}/\tau \) (mid row), \(|b| < e^{a\tau - 1}/\tau \) (bottom row).
2.1 The exponential curve

Let us start with $\mathcal{E}$, graph of the function $E$ in (2.7). Simple algebra shows that

$$\text{dom}(E) = \begin{cases} (-\infty, a_0], & |b| \geq e^{a\tau - 1}/\tau, \\
(-\infty, a_2] \cup [a_1, a_0], & |b| < e^{a\tau - 1}/\tau, \end{cases}$$

(2.9)

where $a_0$ is the unique solution of

$$|b|e^{-a\tau} = a - a$$

in both cases, while $a_2 < a_1$ are the two solutions of

$$|b|e^{-a\tau} = a - a$$

(2.10)

(2.11)

in the second case. It holds

$$a_2 \leq a - 1/\tau \leq a_1 \leq a \leq a_0.$$  

The equalities $a_1 = a_2 = a - 1/\tau$ hold in the tangent case $|b| = e^{a\tau - 1}/\tau$: below, (2.11) has two distinct solutions; above, none. Also, $a_0 \to a^+, a_1 \to a^-$ and $a_2 \to -\infty$ as $|b| \to 0$, which obviously confirms that $a$ is the only eigenvalue of the ODE $x' = ax$.

In the first case of (2.9), $\mathcal{E}$ is made by a single connected component, which originates in $(a_0, 0)$ and, as clear from (2.7), it behaves exponentially in the limit $\alpha \to -\infty$, Figure 2.1 (top row). From this the name exponential curve and the use of the letters $E$ and $\mathcal{E}$. In the second case of (2.9), a similar exponential behavior is preserved in a branch arising from $(a_2, 0)$, but another connected component exists on $[a_1, a_0]$, Figure 2.1 (bottom row). While retaining the name exponential curve for the whole, let us call exponential branch and ring, respectively, the two separated components. Figure 2.1 (mid row) represents the tangent condition above recalled: the exponential branch and the ring connect to each other in $(\alpha, \omega) = (a - 1/\tau, 0)$.

The case of real characteristic roots is special in Theorem 2.1. When $\omega = 0$, it is not difficult to recover from (2.4) that $a_0$ is the rightmost characteristic root if $b > 0$, Figure 2.1 (left column), while in the case $b < 0$ there is no real characteristic root if $b < -e^{a\tau - 1}/\tau$, Figure 2.1 (top-right), one double rightmost characteristic root $a_1 = a_2$ if $b = -e^{a\tau - 1}/\tau$, Figure 2.1 (mid-right), and $a_1$ is the rightmost characteristic root if $b > -e^{a\tau - 1}/\tau$, Figure 2.1 (bottom-right).

The existence of the finite right extremum $a_0$ of dom($E$), together with Theorem 2.1, is an alternative proof of the well-known fact that the characteristic roots of (2.1) are bounded to the right of $\mathbb{C}$, see, e.g., [20, Chapter I, Theorem 4.4]. It is equally evident that $|\lambda| \to +\infty$ implies $\Re(\lambda) \to -\infty$ and that any vertical strip in $\mathbb{C}$ may contain only finitely-many roots, being the left-hand side of (2.2) analytic. Stability and other considerations are left to Section 2.3, after a glimpse of $T$.

2.2 The tangent curve

Let us give a short account of $T$, although less important as clarified in Section 2.3. Starting from the $(\omega, a)$-plane, where the graph of $T$ is meaningful, it is not difficult to grasp a prominent tangent-like behavior, namely

$$\text{dom}(T) = (0, +\infty) \setminus \{k\pi/\tau : k = 1, 2, \ldots\}$$

$$\lim_{\omega \to k\pi/\tau^\pm} T(\omega) = \mp \infty, \quad k = 1, 2, \ldots.$$
Moving to the \((a, \omega)\)-plane leads to a curve \(T\) made of infinitely-many branches, each defined on all \(\mathbb{R}\) and contained in the horizontal strip \((k\pi/\tau, (k+1)\pi/\tau)\). A further branch exists in the horizontal strip \((0, \pi/\tau)\), with domain \((a - 1/\tau, +\infty)\). The dominant behavior resembles a bundle of arctangents, Figure 2.1, though we retain the name tangent curve and, consequently, the use of the letters \(T\) and \(T\). Incidentally, notice that \(\text{dom}(T)\) above implicitly considers the continuous prolongation by \(T((2k+1)\pi/2\tau) = a\) for \(k = 1, 2, \ldots\)

### 2.3 Stability criteria

Asymptotic stability is equivalent to a rightmost characteristic root with negative real part. Any root with positive real part gives instability. Given Theorem 2.1 and the develops of \(\mathcal{E}\) and \(\mathcal{T}\), it is clear that only the positioning of \(\mathcal{E}\) with respect to the imaginary axis is essential for stability, and yet this yields a quite simple tool. As a proof of this claim, let us illustrate the following sufficient criterion. First let us rewrite (2.10) as

\[
f(a) = 0
\]

for

\[
f(a) := |b|e^{-\alpha \tau} - (a - a).
\]  

**Theorem 2.2.** Let \(\alpha_0\) be the unique zero of (2.12). If \(\alpha_0 < 0\) then (2.1) is asymptotically stable.

The proof is trivial from the arguments above. Stronger criteria follow, paying attention to the rightmost characteristic roots.

**Theorem 2.3.** Let \(\alpha_0\) be the unique zero of (2.12) and \(\alpha_1\) be the greatest solution of (2.11) for \(|b| < e^{\alpha \tau - 1}/\tau\):

(i) if \(b \geq 0\), then (2.1) is asymptotically stable iff \(\alpha_0 < 0\);

(ii) if \(-e^{\alpha \tau - 1}/\tau \leq b < 0\), then (2.1) is asymptotically stable iff \(\alpha_1 < 0\);

(iii) if \(b < -e^{\alpha \tau - 1}/\tau\), then (2.1) is asymptotically stable if \(\alpha_0 < 0\).

The proof is immediate from Figure 2.1, in particular: left column for (i), right column, respectively mid and bottom for (ii) and top for (iii).

Let us conclude this part on the single discrete delay case with a couple of examples. With very little effort we show how it is possible to recover classic results related to stability regions and Hopf bifurcations.

**Example 2.4.** If one uses \(\text{sign}(\alpha_0)\) as a stability indicator, the delay-independent stability region is obtained, Figure 2.2. Indeed, from (2.12) it is straightforward to recover

\[
\alpha_0 < 0 \quad \text{if} \quad \begin{cases} 
  b < -a & \text{for } b \geq 0, \\
  b > a & \text{for } b < 0.
\end{cases}
\]

In fact, on the one hand, the first line corresponds to Theorem 2.3 (i), a necessary and sufficient condition, thus giving the true stability boundary. On the other hand, for the second line we can rely only on Theorem 2.2, a sufficient criterion, thus furnishing a more conservative boundary. Moreover, it is not difficult to argue that both boundaries are independent of \(\tau\).
Example 2.5. The Hutchinson equation [29], or logistic DDE,
\[ y'(t) = ry(t)(1 - y(t - 1)) \]
has the nontrivial equilibrium \( y = 1 \) independently of the value of the growth parameter \( r \), assumed to be positive. The corresponding linearization
\[ x'(t) = -rx(t - 1) \]
leads to the characteristic equation
\[ \lambda + re^{-\lambda} = 0, \]
i.e., (2.3) with \( a = 0, b = -r < 0 \) and \( \tau = 1 \). It follows from (2.8) that \( \mathcal{T} \) is independent of \( r \), while \( \mathcal{E} \) moves to the right as \( r \) increases. In particular, for \( r > 1/e \), i.e., like in Figure 2.1 (top-right), \( \mathcal{E} \) has only the exponential branch, so that it is really easy to appreciate how the characteristic roots move to the right as \( r \) increases, tied to flow along the branches of \( \mathcal{T} \). Also, again from (2.8), it follows that the first crossing of the imaginary axis occurs with \( \omega = \pi/2 \) and, correspondingly from (2.5), for \( r = \pi/2 \).

Now, by taking different generalizing directions, let us expand our point of view on the use of the curve \( \mathcal{E} \) and of similar functions \( f \). As a main line to formulate stability criteria we focus on the computation of a right bound \( a_0 \) to \( \text{dom}(E) \) as a (unique or rightmost) zero of \( f \). Practical details are given in Section 6 where, however, other useful implications are also investigated.

3 Two and more discrete delays

Consider the linear DDE with two discrete delays
\[ x'(t) = ax(t) + b_1x(t - \tau_1) + b_2x(t - \tau_2) \quad (3.1) \]
for parameters \( a, b_1, b_2 \in \mathbb{R} \) and delays \( 0 \leq \tau_1 \leq \tau_2 \) with \( (\tau_1, \tau_2) \neq (0,0) \). The associated characteristic equation is
\[ \lambda - a - b_1e^{-\lambda \tau_1} - b_2e^{-\lambda \tau_2} = 0, \]
which is rewritten as
\[ \lambda - a = b_1 e^{-\lambda \tau_1} + b_2 e^{-\lambda \tau_2}. \]  

(3.2)

As previously explained, we focus principally on the magnitude equation
\[ (\alpha - a)^2 + \omega^2 = b_1^2 e^{-2\alpha \tau_1} + b_2^2 e^{-2\alpha \tau_2} + 2b_1 b_2 e^{-a(\tau_1 + \tau_2)} \cos \omega(\tau_1 - \tau_2). \]  

(3.3)

Due to the presence of the cosine term, no explicit expression of \( \omega \) as a function of \( \alpha \) can be derived. An explicit expression is in general convenient, but the following characterization as zero level curve is as useful as practical (see Section 6):
\[ E := \{ (\alpha, \omega) : S_E(\alpha, \omega) = 0 \} \]  

with
\[ S_E(\alpha, \omega) := (\alpha - a)^2 + \omega^2 - b_1^2 e^{-2\alpha \tau_1} - b_2^2 e^{-2\alpha \tau_2} - 2b_1 b_2 e^{-a(\tau_1 + \tau_2)} \cos \omega(\tau_1 - \tau_2). \]  

(3.5)

The same can be said for the tangent curve
\[ T := \{ (\alpha, \omega) : S_T(\alpha, \omega) = 0 \} \]  

with
\[ S_T(\alpha, \omega) := (\alpha - a)^2 + \omega^2 - b_1 e^{-a \tau_1} \sin \omega \tau_1 + b_2 e^{-a \tau_2} \sin \omega \tau_2 \]
\[ + \omega \left[ b_1 e^{-a \tau_1} \cos \omega \tau_1 + b_2 e^{-a \tau_2} \cos \omega \tau_2 \right]. \]  

(3.6)

The latter is used only for representing the characteristic roots as intersection points by virtue of Theorem 2.1, which holds unaltered but for (3.1) in place of (2.1) and for the fact that this formulation of \( S_T \) implies \( \mathbb{R} \times \{0\} \subset T \). Instead, in the sequel, we definitely focus on \( E \) to find useful bounds and relevant stability criteria, keeping in mind that results similar to Theorem 2.1 still hold.

**Remark 3.1.** Existence and regularity of \( E \) and \( T \) can be inferred by using the Implicit Function Theorem on \( S_E \) and \( S_T \), respectively.

By rewriting (3.3) as
\[ \omega = \sqrt{b_1^2 e^{-2a \tau_1} + b_2^2 e^{-2a \tau_2} + 2b_1 b_2 e^{-a(\tau_1 + \tau_2)} \cos \omega(\tau_1 - \tau_2) - (\alpha - a)^2}, \]  

(3.7)

it is not difficult to argue that \( E \) oscillates around the mean exponential branch
\[ E_m(\alpha) := \sqrt{b_1^2 e^{-2a \tau_1} + b_2^2 e^{-2a \tau_2} - (\alpha - a)^2}, \]

obtained by annihilating the cosine term in (3.7). The situation gets more complicated approaching the real axis, where also a ring may arise. Nevertheless, by using \( |\cos x| \leq 1 \), the upper bound
\[ \omega \leq \bar{E}(\alpha) := \sqrt{(|b_1| e^{-a \tau_1} + |b_2| e^{-a \tau_2})^2 - (\alpha - a)^2} \]  

(3.8)

holds for the \( \omega \) component of \( E \). This bound is also optimal for \( b_1, b_2 > 0 \): the graph of \( \bar{E} \) is tangent to \( E \) whenever \( \omega = 2k\pi / (\tau_1 - \tau_2) \) for \( k \in \mathbb{Z} \), with the convention \( \omega \geq 0 \). Let us study this bound.
Similarly to \((2.9)\), \(\text{dom} (E)\) is bounded to the right by \(a_0\), where \(a_0 \geq a\) is the unique zero of
\[
f(\alpha) := |b_1| e^{-\alpha \tau_1} + |b_2| e^{-\alpha \tau_2} - (\alpha - a).
\] (3.9)

Then \(\text{dom} (E) = (-\infty, a_0]\) if
\[
|b_1| e^{-\alpha \tau_1} + |b_2| e^{-\alpha \tau_2} = a - \alpha
\] (3.10)
has no solutions, in which case the graph of \(E\) is an exponential branch. Otherwise, if (3.10) has two solutions, say \(\alpha_1 > \alpha_2\), then \(\text{dom} (E) = (-\infty, a_2] \cup [a_1, a_0]\) and a ring is present, to the right of the exponential branch since \(a_2 < a_1 \leq a \leq a_0\). In the tangent case \(a_1 = a_2\) the ring is attached to the exponential branch.

By combining the right-boundedness of \(\text{dom} (E)\) with (3.8) we get the following sufficient criterion, analogous of Theorem 2.2.

**Theorem 3.2.** Let \(a_0\) be the unique zero of (3.9). If \(a_0 < 0\) then (3.1) is asymptotically stable.

Necessity demands for positive coefficients of the delayed terms, a partial analogous of Theorem 2.3.

**Theorem 3.3.** Let \(a_0\) be the unique zero of (3.9). If \(b_1, b_2 > 0\) then (3.1) is asymptotically stable iff \(a_0 < 0\).

**Proof.** Sufficiency follows from Theorem 3.2. As for necessity, by comparing (3.2) with (3.9) it is clear that \(\lambda = a_0\) is a real characteristic root. \(\Box\)

From the proof above we have that \(a_0\) is the real rightmost characteristic root if \(b_1, b_2 > 0\), Figure 3.1 (left). Otherwise, the bound \(\alpha < a_0\) for all \(\alpha = \Re (\lambda)\) and \(\lambda\) characteristic root still holds, despite being more conservative. To complete the analogous of Theorem 2.3 a deeper analysis of the exact curve \(E\) is required. Resorting to the lower bound
\[
\omega \geq \overline{E}(\alpha) := \sqrt{(|b_1| e^{-\alpha \tau_1} - |b_2| e^{-\alpha \tau_2})^2 - (\alpha - a)^2}
\] (3.11)
might help (which is optimal for \(b_1 b_2 < 0\)). An example is given in Figure 3.1 (right), for values of \(b_1\) and \(b_2\) such that
\[
|b_1| e^{-\alpha \tau_1} - |b_2| e^{-\alpha \tau_2} = a - \alpha
\] (3.12)
has two solutions \(\alpha_1 > \alpha_2\); in this case all the roots lie to the left of \(\alpha_1\), giving a less conservative bound. Once more, notice how the structure of the curve \(T\), although different from the case of a single delay, still has an horizontal develop, making it non influential as long as stability is concerned.

The above analysis can be extended straightforwardly to equations with any finite number of discrete delays. For
\[
x'(t) = ax(t) + \sum_{i=1}^{n} b_i x(t - \tau_i),
\] (3.13)
(3.5) and (3.6) become, respectively,
\[
S_E(\alpha, \omega) := (\alpha - a)^2 + \omega^2 - \sum_{i=1}^{n} b_i^2 e^{-\alpha \tau_i} - 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_i b_j e^{-\alpha (\tau_i + \tau_j)} \cos \omega (\tau_i - \tau_j)
\]
and
\[
S_T(\alpha, \omega) := (\alpha - a) \sum_{i=1}^{n} b_i e^{-\alpha \tau_i} \sin \omega \tau_i + \omega \sum_{i=1}^{n} b_i e^{-\alpha \tau_i} \cos \omega \tau_i.
\]
Figure 3.1: Curve $\mathcal{E}$ (thick red), curve $\mathcal{T}$ (thin blue) and characteristic roots (●) of (3.1): the dashed green line is the graph of $\overline{E}$ in (3.8) for $b_1, b_2 > 0$ (left), or the graph of $\underline{E}$ in (3.11) for $b_1$ and $b_2$ such that (3.12) has two solutions (right).

Their curves of level zero always give $\mathcal{E}$ and $\mathcal{T}$, but now their shape can be rather complicated and contour algorithms are definitely the only way out (see Section 6). Nevertheless, the upper bound

$$\omega \leq \overline{E}(\alpha) := \sqrt{\left( \sum_{i=1}^{n} |b_i| e^{-\alpha \tau_i} \right)^2 - (\alpha - a)^2}$$

holds, preserving the same simple structure, and it can be easily used to obtain similar stability criteria through the use of the function

$$f(\alpha) := \sum_{i=1}^{n} |b_i| e^{-\alpha \tau_i} - (\alpha - a). \quad (3.14)$$

Of course, it is not optimal anymore since, in general, not all cosine terms can be maximized simultaneously. An example is shown in Figure 3.2. More details are contained in the starting work [38].

Figure 3.2: Curve $\mathcal{E}$ (thick red), curve $\mathcal{T}$ (thin blue), graph of $\underline{E}$ (dashed green) and characteristic roots (●) of (3.13): an instance with $n = 4$. 
4 A distributed delay

Consider the linear DDE with distributed delay

$$x'(t) = ax(t) + \int_{-\tau}^{0} c(\theta)x(t + \theta)d\theta$$

(4.1)

for parameter $a \in \mathbb{R}$, kernel $c : [-\tau, 0] \to \mathbb{R}$ as smooth as necessary and delay $\tau > 0$. The associated characteristic equation is written as

$$\lambda - a = \int_{-\tau}^{0} c(\theta)e^{\lambda \theta}d\theta.$$ 

(4.2)

Let us observe that there is no natural parameterization to construct a stability chart: a general kernel cannot be identified through a finite number of parameters. Opposite, the approach of the previous sections can be applied with inessential modifications. The magnitude equation gives rise to the the curve $E$ according to (3.4) for $S_E(\alpha, \omega)$:

$$S_E(\alpha, \omega) := (\alpha - a)^2 + \omega^2 - \left(\int_{-\tau}^{0} c(\theta)e^{\alpha \theta}\cos \omega \theta d\theta\right)^2 - \left(\int_{-\tau}^{0} c(\theta)e^{\alpha \theta}\sin \omega \theta d\theta\right)^2.$$

The upper bound

$$\omega \leq \overline{E}(\alpha) := \sqrt{\left(\int_{-\tau}^{0} |c(\theta)|e^{\alpha \theta}d\theta\right)^2 - (\alpha - a)^2}$$

follows easily from (4.2) through

$$|\lambda - a| \leq \int_{-\tau}^{0} |c(\theta)|e^{\alpha \theta}d\theta.$$

The right bound $\alpha_0$ of $\text{dom}(\overline{E})$ is the unique zero of $f$ for

$$f(\alpha) := \int_{-\tau}^{0} |c(\theta)|e^{\alpha \theta}d\theta - (\alpha - a).$$

(4.3)

Again $\alpha_0 \geq a$. Let us state the following, in the spirit of Theorems 2.2 and 3.2.

**Theorem 4.1.** Let $\alpha_0$ be the unique zero of (4.3). If $\alpha_0 < 0$ then (4.1) is asymptotically stable.

A couple of examples is shown in Figure 4.1. Notice the exactness of the upper bound on the real axis. In fact,

$$\{\alpha : (\alpha, 0) \in E\} = \{\alpha : \overline{E}(\alpha) = 0\}$$

follows from the fact that

$$\left(\int_{-\tau}^{0} |c(\theta)|e^{\alpha \theta}d\theta\right)^2 - \left(\int_{-\tau}^{0} c(\theta)e^{\alpha \theta}d\theta\right)^2 = \left(\int_{-\tau}^{0} |c(\theta)| + c(\theta)\right)e^{\alpha \theta}d\theta \left(\int_{-\tau}^{0} |c(\theta)| - c(\theta)\right)e^{\alpha \theta}d\theta$$

vanishes independently of $c$. Observe, moreover, how the sign of $c$ affects the necessity counterpart of Theorem 4.1, which is seemingly ensured by a positive $c$ – Figure 4.1 (left) – thus resembling Theorem 2.3 (i). Finally, and as already seen before, more insight can be gained by analyzing the zero level curve of $S_E$, obtaining less conservative stability criteria or necessary ones.
5 Renewal, coupled and neutral delay equations: a specific case

In this section we investigate up to which extent the proposed methodology can be applied to more complicated problems.

Inspired by the recent work [19], we focus on the particular equation

$$\lambda = a + (b + c\alpha)e^{-\lambda}$$

(5.1)

for parameters $a, b, c \in \mathbb{R}$ with $c \in [-1, 1]$. In [19] it is shown how (5.1) is obtained as the characteristic equation of the nontrivial steady state of a cell population model, based on the coupling of a renewal (Volterra integral) equation with a DDE. The delay is implicitly set to be $\tau = 1$, and we refer to [19, §3] for a discussion of the model and the derivation of (5.1). Here, in the spirit of the previous sections, we continue the approach initiated in [42]. We remark again that this approach is different from the more traditional one, used, e.g., in [19], where stability charts are determined in the $(a, b)$-plane for different values of the third parameter $c$.

Clearly, $c = 0$ sends us back to Section 2, so we assume $c \neq 0$ in the sequel.

To start with, we rewrite (5.1) as the equality between complex numbers

$$\lambda - a = (b + c\alpha)e^{-\lambda}.$$ 

Separating real and imaginary parts leads to

$$\alpha - a = [(b + ca)\cos \omega + c\alpha \sin \omega]e^{-a},$$

$$\omega = [c\omega \cos \omega - (b + ca) \sin \omega]e^{-a}.$$

Squaring and summing gives the magnitude equation

$$(\alpha - a)^2 + \omega^2 = [(b + ca)^2 + c^2\omega^2]e^{-2\alpha},$$

from which $\omega$ is recovered as

$$\omega = E(\alpha) := \sqrt{\frac{(b + ca)^2e^{-2\alpha} - (\alpha - a)^2}{1 - c^2e^{-2\alpha}}}.$$ (5.2)

Again, as partial analogous of Theorem 2.1, $\lambda = \alpha + i\omega$ for $\omega \geq 0$ is a root of (5.1) only if $(\alpha, \omega) \in \mathcal{E}$ for $\mathcal{E}$ the graph of $E$. 

Figure 4.1: Curve $\mathcal{E}$ (thick red), curve $\mathcal{T}$ (thin blue), graph of $\mathcal{E}$ (dashed green) and characteristic roots (●) of (4.1) with $a = 1$ and $c(\theta) = e^\theta$ (left) and $c(\theta) = -e^{\theta/10}$ (right); inner panels: zoom of the rings.
It is evident that \( \text{dom}(E) \) is bounded to the right: \( \alpha \to +\infty \) leads to a negative argument of the square root in (5.2). On the other hand, by rewriting the latter as

\[
E(\alpha) = \sqrt{\frac{f(\alpha)}{g(\alpha)}}
\]

for

\[
f(\alpha) := e^{2a}(a - \alpha)^2 - (b + ca)^2
\]

and

\[
g(\alpha) := e^2 - e^{2a},
\]

it is also evident that \( \text{dom}(E) \) is bounded to the left, opposite to the case of DDEs: equations like (5.1) can have infinitely many solutions in a vertical strip of the complex plane, a situation characterizing many neutral dynamics (see, e.g., [19] for relevant comments). In any case, for the sake of stability, we are interested in a right bound, so we seek for

\[
a_0 := \max_{\alpha \in \mathbb{R}} \left\{ \frac{f(\alpha)}{g(\alpha)} \geq 0 \right\}. (5.5)
\]

Analyzing the sign of \( g \) yields

\[
g(\alpha) \begin{cases} > 0, & \alpha < \ln |c|, \\ < 0, & \alpha > \ln |c|. \end{cases} (5.6)
\]

The discussion on the sign of \( f \) is not as trivial. To keep it reasonably simple, observe that

\[
\lim_{\alpha \to +\infty} f(\alpha) = +\infty, \quad \lim_{\alpha \to -\infty} f(\alpha) = -\infty,
\]

so there exists at least one zero. It is not difficult to realize that \( f \) can actually change sign more than once depending on \( a, b \) and \( c \). This can lead to several “exponential branches” and “rings” as detailed in [42]. Nevertheless, the following painless geometric reasoning shows that \( f \) always has a rightmost zero \( \alpha_r \geq a \), to the right of which it is definitively positive, strictly increasing and convex.

To illustrate the argument, first define the two parabolas

\[
f_1(\alpha) := (a - \alpha)^2, \quad f_2(\alpha) := (b + ca)^2
\]

and rewrite \( f \) as

\[
f(\alpha) = f_3(\alpha) - f_2(\alpha)
\]

for

\[
f_3(\alpha) := e^{2a}f_1(\alpha).
\]

Observe that the two parabolas \( f_1 \) and \( f_2 \) are both nonnegative, tangent to the horizontal axis and with vertexes in \( a \) and \(-b/c\), respectively. The relative position of their vertical axis thus depends on whether \( b + ca \) is positive (\( f_1 \) to the right), zero (same axis) or negative (\( f_1 \) to the left). Moreover, \( f_2 \) grows more slowly when \( |c| < 1 \), while, for \( |c| = 1 \), \( f_1 \) and \( f_2 \) actually differ just by translation.

As far as \( f_3 \) is concerned, the effect of the exponential factor is to further increase the slope of the increasing branch of \( f_1 \) for positive values of \( \alpha \). More precisely, \( f_3 \) vanishes only at \( \alpha = a \), it equals \( f_1 \) again at \( \alpha = 0 \) and it increases to the right of \( a \), faster than \( f_1 \) if \( a \geq 0 \) while,
for $a < 0$, it increases first slower (for $\alpha < 0$) and then definitively faster (for $\alpha > 0$), Figure 5.1.

Now, observe that, independently of the condition on the sign of $b + ca$, there is always a rightmost intersection between the graph of $f_3$ and that of the slower parabola $f_2$, Figure 5.2. This intersection lays always to the right of $a$ (or on it if $a = -b/c = 0$), i.e., on the right increasing branch of $f_3$. Let us call $\alpha_r$ its abscissa. If $b + ca > 0$ it easily follows that $f'(\alpha)$ and $f''(\alpha)$ are both strictly positive for $\alpha > \alpha_r$, since $f_3'(\alpha) > f_2'(\alpha)$ and $f_3''(\alpha) > f_2''(\alpha)$ both hold. The configuration is almost unchanged for $b + ca = 0$, the only difference being that $\alpha_r$ can be either a simple zero or a minimum tangent to the horizontal axis, the exchange occurring when $f_3''(\alpha_r) = f_1''(\alpha_r)$. Finally, when $b + ca < 0$, the rightmost intersection lies on the right increasing branch of $f_3$ and on the decreasing branch of $f_2$, so that $f'(\alpha)$ and $f''(\alpha)$ are again both strictly positive for $a > \alpha_r$.

Concluding the above reasoning,

$$f(\alpha) \begin{cases} > 0 & \text{if } \alpha > \alpha_r, \\ < 0 & \text{if } \alpha_r - \epsilon < \alpha < \alpha_r, \end{cases}$$

for some $\epsilon > 0$: the sign of $f$ is not interesting far to the left of $\alpha_r$. Eventually, combining with (5.6), we obtain

$$\alpha_0 = \begin{cases} \alpha_r, & \alpha_r > \ln |c|, \\ \ln |c|, & \alpha_r < \ln |c|. \end{cases} (5.7)$$
Notice that in the nongeneric case \( \alpha_r = \ln |c| \), for which (5.3) has a vertical asymptote, the above determined rightmost region of positivity of \( f/g \) collapses in a single point falling out of the domain of \( E \). Consequently, \( \alpha_0 \) comes from zeros of \( f \) strictly to the left of \( \alpha_r = \ln |c| \).

Anyway, since \( \ln |c| \leq 0 \) for \( 0 < |c| \leq 1 \), this case is not harmful for stability.

As a direct consequence of the above discussion, we are again able to formulate the main result of the proposed approach in the form of the following sufficient criterion for asymptotic stability, which first concerns \( \alpha_0 \) in (5.5) and traces Theorem 2.2, Theorem 3.2 and Theorem 4.1. A corollary in terms of \( \alpha_r \) immediately follows from (5.7), recalling that \( \ln |c| \leq 0 \).

**Theorem 5.1.** Let \( \alpha_0 \) be given by (5.5). If \( \alpha_0 < 0 \) then (5.1) is asymptotically stable.

**Corollary 5.2.** Let \( \alpha_r \) be the rightmost zero of (5.4). If \( |c| < 1 \) and \( \alpha_r < 0 \) then (5.1) is asymptotically stable.

Figure 5.3 illustrates a couple of cases. In particular, the left and right panel correspond, respectively, to the first and second case of (5.7). Notice, moreover, how the roots are distributed along the vertical asymptote \( \alpha = \ln |c| \), which resembles, as anticipated, a neutral dynamics. Indeed, neutral equations can be easily obtained from renewal equations by adding terms with discrete delays.

**6 Computational aspects**

Several stability criteria have been formulated in the preceding sections based on the knowledge of the unique or rightmost zeros of certain functions. Here we comment on how to actually compute these quantities, with a digression on the approximation of distributed delays as far as Section 4 is concerned.

Let us start with the case of a single discrete delay, Section 2. Concerning Theorem 2.2, in fact, it is important to observe that \( \alpha_0 \) can be obtained easily since \( f \) in (2.12) is convex and decreasing independently of \( a, b \) and \( \tau \), so that Newton method started from \( a \) (recall that \( \alpha_0 \geq a \)) gives an automatic and quadratically convergent procedure to get an approximation of \( \alpha_0 \) with an error \( \varepsilon \) below a desired tolerance. The following pseudocode emphasizes the simplicity of this procedure:
Consider that $a_1$ and $a_2$ in (2.9) can be recovered similarly and as easily from (2.11). As an application see Example 2.4: the light-shaded region in Figure 2.2 is obtained automatically as zero level curve of the surface $a_0(a,b)$ with $a_0$ computed by Newton method as previously described. The choice of a suitable initial value for the convergence of the method is postponed to the end of the section.

Let us come now to the case of multiple discrete delays, Section 3, where the lack of an explicit expression for $\omega$ as a function of $\alpha$ is solved by characterizing the problem as zero level curve of (3.5). Computationally, this is not a true obstacle, as efficient contouring algorithms are readily available (see, e.g., [46] for Matlab contour or [11], already used in Example 2.4).

With reference to Theorem 3.2, instead, we remark that $a_0$ can be computed again by Newton method started from $a$, analogously to what discussed above.

The same situation occurs in the case of distributed delays, Section 4. In fact, $f$ in (4.3) is again convex and decreasing, and Newton method can approximate $a_0$ in Theorem 4.1 with any desired tolerance. A more interesting objective is related instead to substituting the integral in (4.1) with a quadrature sum. By using any interpolatory formula based on $n$ nodes $\theta_i \in [-\tau,0]$ and relevant weights $w_i$, $i = 1, \ldots, n$, (see, e.g., [17]), (4.1) reduces to (3.13) with $b_i = w_i c(\theta_i)$ and $\tau_i = -\theta_i$. Let us notice that quadrature is the common approach to treat distributed delay terms. This is the case, e.g., if one desires to use the Matlab package dde-biftool [1, 22], which is not tuned to treat distributed delays directly. But also other numerical methods automatically incorporate quadrature [13]. In Figure 6.1 we compare the exact curve $E$ for (4.1) with its quadrature counterpart, i.e., that of (3.13) obtained via composite Newton-Cotes (left, equidistant quadrature nodes) and Clenshaw-Curtis (right, Chebyshev extrema as quadrature nodes) formulae. In this specific case we choose $a = 1$ and $c(\theta) = e^\theta$, so that “exact curve” means explicit analytical expression. Notoriously, Clenshaw-Curtis outperforms Newton-Cotes in the presence of sufficient regularity of the integrand, as it is the case represented in Figure 6.1 given the exponential kernel. This typical behavior is reflected in the approximation of $E$. Indeed, the error with $n = 10$ nodes is still macroscopic for Newton-Cotes. Moreover, apparently, the effect of the magnitude of the roots is not much relevant: by observing, e.g., the exponential branch, the approximated curves approach the exact one “from the right” as $n$ increases, so that no root is actually intercepted. Opposite, Clenshaw-Curtis furnishes a good approximation already with few nodes, with an error increasing with the magnitude of the roots. Notice, in fact, how the curves approximating the exponential branch converge “from below” as $n$ increases. The latter fact is in general good news for stability since, typically, the rightmost root has small magnitude with respect to those lying along the exponential branch. Instead, as far as the ring is concerned (when present), Newton–Cotes converges “from inside”, giving rise to a more conservative stability test. On the other hand, the rings from Clenshaw–Curtis are already indistinguishable for low values of $n$. In any case, it emerges evident that the issue of choosing a sufficiently large number of quadrature nodes is not negligible, even if a highly (even spectrally) accurate formula is employed. As a final
consideration (or curious aspect), observe that the roots of the approximation (3.13) oscillate as rapidly and “chaotically” around a mean exponential branch as large is \( n \) (recall Section 3 and Figure 3.2) but, in the limit as \( n \to \infty \), they tidily align along the exponential branch of (4.1), since the convergence of the quadrature formula is anyway guaranteed.

![Figure 6.1: Curve \( E \) (thick red), curve \( T \) (thin blue) and characteristic roots (●) of (4.1) with \( a = 1 \) and \( c(\theta) = e^{\theta} \) and curve \( \bar{E} \) approximated by quadrature on \( n \) nodes (dashed purple tones) by using Newton-Cotes (left) and Clenshaw-Curtis (right) formulae; inner panels: zoom of the rings.](image)

Finally, concerning Section 5, the function \( f \) in (5.4) is shown to be strictly increasing and convex to the right of its rightmost zero \( \alpha_r \). Again, this is the optimal situation for Newton method to converge monotonically to \( \alpha_r \) from the right, confirming the validity of the proposed strategy. At this point it is worthy to provide a rightward estimate. If \(|c| < 1\) this immediately follows from (5.1) in the form

\[
\alpha_r < \bar{\alpha}_r := \frac{|a| + |b|}{1 - |c|},
\]

compare also with [19, Lemma 2.2]. Notice that \( \bar{\alpha}_r \) grows unbounded as \(|c| \to 1\), so that it might soon cause overflow in machine arithmetic due to the presence of the exponential factor in \( f \). A naive way to overcome this problem, given the simplicity of \( f \), is to plot its graph and choose a proper estimate by hands. Anyway, in all the cases we experimented, convergence was always obtained starting from \( \alpha^* \) satisfying \( \sqrt{\text{realmax}} = e^\alpha \alpha^* \) for \( \text{realmax} \) the greatest machine number (about \( 10^{308} \) according to IEEE754 with double precision). This is important if one wishes to implement codes that exploit Newton method with automatic choice of the starting guess, as done, e.g., to obtain Figure 2.2.

7 Relation to modulus semigroups and monotone semiflows

Most of the analysis carried out in this paper is ascribable to the use of functions like (2.12), (3.9), (3.14) and (4.3). These can be thought as obtained from the corresponding DDEs by substituting the coefficients of the delayed terms with their absolute values.

According to [3, Proposition 3.3], it is not difficult to realize that the consequent bounds represent the stability bounds of the associated modulus semigroups. In this context, the modulus semigroup is the semigroup of positive operators that minimally dominates the
relevant semigroup of solution operators. In particular, (2.1), (3.1), (3.13) and (4.1) represent the most common scalar instances of the general case considered in [3, Section 3]. Two consequences deserve attention, which further provide a solid theoretical background to the proposed methodology.

First, in the space of parameters and due to the minimal domination property, the stability domain of the original semigroup contains that of its modulus and, in practice, the latter can approximate the former very well from inside. This is in fact the situation in Figure 2.2 for Example 2.4. In particular, the (dashed and solid) stability boundaries coincide for \( b > 0 \), where indeed \( b = |b| \). This positivity leads next to the second general observation.

The dominant root of the generator of a semigroup of positive linear operators, which are equivalently monotone, is necessarily real, see [25] or [41, Chapter B-III, Theorem 1.1]. Therefore it is much easier to determine the stability boundary of the modulus semigroup than that of the original semigroup. As an evidence see [43, Corollary 3.2] and the illustrative Remark 2 few lines above, which lead us back again to the end of the introduction.

Eventually, in the case of scalar equations one can exploit this monotonicity (and possibly convexity), which explains, indeed, why Newton method is particularly appropriate to efficiently approximate the stability bound.

It is not clear whether the treatment developed in Section 5 fits in the same lines. On the one hand, the theory of modulus semigroups has been advanced to such an abstract level to include also neutral and renewal equations, see in particular [49], but also [8, 47] for completion. As a counterpart for monotone semiflows, see [33] concerning neutral problems. On the other hand, in Section 5 we start directly from a specific characteristic equation, so it might be difficult to see the relation with the modulus semigroup of an originating problem with delay (possibly the coupling of a DDE with a renewal equation, see [19]). We reserve to investigate further on this potential connection and its consequences in a future work.

Acknowledgments

The authors, in particular D.B., are indebted to Odo Diekmann (Utrecht University) for his never-ending assistance and valuable observations on the subject. Special thankfulness is expressed for bringing the theory of modulus semigroups to their attention and for fostering most of the content of Section 7. Moreover, D.B. is a member of INdAM Research group GNCS and is supported by the INdAM GNCS project “Analisi e sviluppo di metodologie numeriche per certi tipi non classici di sistemi dinamici.” (2017) and by the project PSD_2015_2017_DIMA_PRID_2017_ZANOLIN “SIDIA – Sistemi Dinamici e Applicazioni” (UNIUD).

References


