# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td><strong>Delays and Dynamics</strong></td>
<td>3</td>
</tr>
<tr>
<td>Physical Systems with Delays</td>
<td>4</td>
</tr>
<tr>
<td>Examples From Biology</td>
<td>4</td>
</tr>
<tr>
<td>Examples From Networks</td>
<td>9</td>
</tr>
<tr>
<td>Other physical examples</td>
<td>15</td>
</tr>
<tr>
<td>Mathematical Models</td>
<td>16</td>
</tr>
<tr>
<td>Integrators</td>
<td>16</td>
</tr>
<tr>
<td><strong>The Dual Nature of Delay</strong></td>
<td>17</td>
</tr>
<tr>
<td>Detrimental effects</td>
<td>18</td>
</tr>
<tr>
<td>Beneficial effects</td>
<td>20</td>
</tr>
<tr>
<td><strong>Characterizing Delayed Feedback</strong></td>
<td>22</td>
</tr>
<tr>
<td>Limitations of delayed feedback</td>
<td>22</td>
</tr>
<tr>
<td>Delay and gain bounds: explicit computations</td>
<td>25</td>
</tr>
</tbody>
</table>
Why Does Delay Feedback Work? 30

Derivative Feedback ................................................................. 31
Phase synchronization .............................................................. 34

Designing With Delays 39

Stabilizing oscillatory systems ............................................... 40
Stabilizing systems with multiple delays .................................. 42
Designing for Performance ...................................................... 43
Stabilizing Predictors ............................................................... 43
Stabilizing unstable periodic orbits ......................................... 44

Interconnections, Interference, and Complexity Issues 44

Multiple delays interference phenomena ............................... 45
Computational complexity in delay systems ............................ 47
Smith predictor interference mechanism .................................. 49
High-gains and delay sensitivity ................................................ 50

Concluding remarks 50

References 51
Introduction

A feedback control system is composed of a plant to be controlled, a set of sensors with access to information signals from the plant, a controller which implements the control algorithms, and actuators that translate the outputs of the controllers into inputs to the plant. Along the way, signals are processed, transformed, and communicated. Most control engineers however, have until recently considered the transport and processing times either too small to worry about, or too complex to handle.

Recently however, control engineers have began implementing control laws that exploit the popularity of shared communications networks such as the Internet and wireless IEEE 802.11 networks. A byproduct of the increasing usage of such networks is the interest in understanding the effect of delays on the stability and performance of feedback control systems. The delay in propagating information and control signals is in fact becoming a critical parameter in determining the performance of networked dynamical systems [101].

Delays are also present in modeling highway traffic flows, where a realistic model of the drivers’ reactions must include delayed reactions [36]. Such physiological time delays are critical in creating accurate driving models, and their estimation is a complicated task, as they must be subdivided into sensing, perception, response, selection, and finally programming delays [6]. On the other hand, material distribution and supply networks provide examples of interconnected entities subject to propagation and transport delays [84]. Such delays may be used to differentiate qualitative behaviors and the performance in the corresponding system. Finally, biology and population dynamics are application areas where delays arise naturally. Specifically, delays help in describing a chain reaction [62], a transport process (breathing process in the physiological circuit controlling the carbon-dioxide level in the blood [105]), storing nutrients or cell cycles.
in order to control the supply of nutrients to a growing population of microorganisms in a chemostat [5], latency and short intercellular phases in epidemics such as cell-to-cell spread models in the bloodstream [19], to cite only a few. This list of dynamical systems with delays is far from complete, and other examples are more thoroughly discussed throughout the article.

The aims of this article are threefold: first, to present various problems and opportunities arising because of the presence of delays in the feedback path; second, to explain when and why delayed feedback strategies are beneficial and finally, to present some limitations of delayed feedback. Note that while the modeling and estimation of delays are not explicit goals of the current article, such issues are implicitly discussed throughout the article.

Notation:

Delays and Dynamics

In various physical and biological systems, dual processes (e.g. production/consumption, demand/supply, excitatory/inhibitory, reproduction/death, synthesis/consumption, production/elimination) are taken place. The dynamics of the corresponding systems are described using distributed-parameters over time intervals, areas, or volumes. If we focus on the time-domain component of such dynamics (while neglecting their spatial distribution), standard conservation laws (mass, energy, information) may be expressed by relating the rate of change of some flow variable $x$ at time $t$, to the balance between the corresponding inflow (production, reproduction) $U_f(.)$ and outflow rates (consumption, death, elimination) $Y_f(.)$. We can then write:

$$\dot{x}(t) = U_f(x_t) - Y_f(x_t),$$

(1)
where \( x_t \) denotes the translation operator \( x_t(\theta) = x(t + \theta) \) over some non-empty but not necessarily bounded (infinite memory) time-interval \( \theta \in \mathcal{I} \subset \mathbb{R}_- \). In the case of a selective memory system, \( Y_f \) and \( U_f \) may depend on \( x \) evaluated at some discrete, past time instances, possibly including the present time, that is, \( x(t - \tau_0), x(t - \tau_1), \ldots x(t - \tau_m) \), for \( \tau_k \in \mathbb{R}_+ \), for all \( k = 0, \ldots, m \). Note that for the generic model in (1), the notion of state must be re-defined, and the piece of trajectory notion, \( x_t \), as suggested by Krasovskii [50] in the 1950s, seems more appropriate than the traditional \( x(t) \). There are several ways to represent the dynamics of the systems in (1) and throughout the article, we use the functional differential equations framework (FDE) as given in [50], [10], [40].

There are numerous examples of physical dynamical systems with delays (see, for instance, [37], [23], [34], [62], [67], [73] and the references therein) which specialize the generic model in (1). A sample of such systems are presented next.

**Physical Systems with Delays**

*Examples From Biology*

**Cyclic biochemical feedback in cell regulatory networks.** In [32], Goodwin introduced a low-order model of a dynamical system that soon became the quintessential example of biochemical oscillator. The model is at the origin of many feedback-based approaches for describing cell regulatory mechanisms (see, e.g. [33] and the references therein).

Consider the second-order system:

\[
\begin{align*}
\dot{x}_1(t) &= -\lambda_1 x_1(t) + c_1 x_2(t - \tau_1) \\
\dot{x}_2(t) &= -\lambda_2 x_2(t) + g(x_1(t - \tau_2)),
\end{align*}
\]

(2)
where $x_1$ and $x_2$ denote the concentrations of the mRNA (messenger RNA) and of the protein (end product), and the rate $dx/dt$ is defined by the balance between mRNA synthesis and the end product consumption. The variables $c_1 \in \mathbb{R}_+$ and $\lambda_i \in \mathbb{R}_+ (i = 1, 2)$ describe the translation and degradation effects, respectively. The feedback function $g$ models the propagation effects among the components of the system such that one component predicts a change (increase/decrease, inhibition/activation) in its own production rate specified by an appropriate monotone function. Note that the delay $\tau_1$ ($\tau_2$) defines the time-lag between the initiation of the translation (transcription) and the appearance of the mature protein (mRNA). See [28], [86] for the use of delays in describing chemical or biochemical kinetics [30].

Similar models are encountered in mitogen-activated protein kinase cascades [93] (reversible enzyme activation based mechanisms) or in circadian rhythm generator [89] (feedback mechanism of protein products on the transcription rate of their genes). Finally, a variation of model (2) with a single delay concentrating the translation and the transcription time-lags was discussed by [11] in the case of Hes1-mRNA self-repression. To the best of our knowledge, reference [59] was the first study to extend the classical model of Goodwin by incorporating a time-delay representing the dependency of one chemical component on another.

In order to study the local behavior of system (2), we linearize the dynamical equations around an equilibrium point $\bar{x} = (\bar{x}_1 \ \bar{x}_2)^T$. The characteristic equation of the linearized version is given by:

$$(s + \lambda_1)(s + \lambda_2) + ke^{-s\tau} = 0,$$  \hspace{1cm} (3)

where $k = -c_1g'(\bar{x}_1)$, and $\tau = \tau_1 + \tau_2$ represents the total delay (translation+transcription). Equation (3) may be interpreted as the closed-loop characteristic equation of a second-order LTI system subject to a delayed output feedback.
Using a simple geometric argument (see also Sections , ), we will prove that the overall delay $\tau$ does not affect the asymptotic stability of (3) if $|k| \leq \lambda_1 \lambda_2$. The limitations of delayed feedback controllers will be presented in Section .

There is much recent interest in Goodwin-like oscillators and their dynamics in the context of monotone systems [3] (see also [26]). In fact, the characteristic equation of the linearized version of the biochemical system in [26] is described by:

$$(s + \lambda_1) \ldots (s + \lambda_m) + ke^{-s\tau} = 0,$$

for some $m \in \mathbb{N}, m \geq 2$. Note that if, instead of the selective memory model in (2), delays were described using a $\gamma$-distribution law (as suggested and discussed by [62]), the resulting model is similar to the original delay-free Goodwin oscillatory system [32], but with an increased number of dimensions. Finally, note that combining positive and negative feedbacks, with time-delays may generate oscillations and complex behaviors in simple dynamical systems [98] as will be illustrated in Section .

**Population dynamics, epidemics and dynamical diseases.** Understanding the underlying mechanisms of biological processes and epidemics represents a serious challenge for health workers engaged in designing clinically-relevant treatment strategies. Considering epidemics and diseases as dynamical processes allows for using the powerful analysis tools of mathematical modeling, mathematical systems, and control theory as described next.

Consider first an example from hematology. As suggested by Foley and Mackey [31], circulating cell populations in one compartment may be schematically described by the following differential equation:

$$\dot{x}(t) = -\lambda x(t) + G(x(t - \tau)),$$  

where $x$ represents the circulating cell population, $\lambda$ the loss cell rate, and the monotone (typically
decreasing) function $G$ denotes the flux of cells from the previous compartment, and describes a feedback mechanism. The delay $\tau$ represents the average-time required to go through the compartment. Model (5) is standard in population dynamics, where the delay generally represents a maturing period (see also [51] for further insights on such models).

It is worth mentioning that the mechanism regulating the production of blood cells is not completely understood, and there exists a large class of mathematical models representing such a feedback mechanism. Consider next the regulation process of hematopoiesis (production of blood cells). Let $z(t, a)$ be the cell density at time $t$ and age $a$, and assume that the cells die at a rate $\lambda(t, a)$. If we consider that a cell enters (leaves) a compartment at age $a = 0$ ($a = T$), and that the aging-velocity in the compartment is $v(t)$, then the reaction-convection dynamical evolution of $z(t, a)$ is given by:

$$\frac{\partial z}{\partial t} + v(t) \frac{\partial z}{\partial a} = -\lambda(t, a)z, \quad t > 0, \quad a \in [0, T].$$

(6)

New cells enter the compartment as defined by the boundary condition $z(t, 0) = F_b(t)$, and for completeness, the initial condition $x(0, a)$ must be specified. Under appropriate assumptions (for example, constant aging velocity $v(t) = V$, and constant rate $\lambda$), the \textit{method of characteristic} eliminates the need of the age-structured population leading to a delay differential system for the \textit{total mature} population as described by:

$$\dot{x}(t) = -\lambda x(t) + V \left[ F_b(t) - F_b(t - \tau) e^{-\lambda \tau} \right]$$

(7)

where $\tau = T/V$, and

$$x(t) = \int_0^T z(t, a) da$$

denotes the total number of cells. In the case of an age-structured model for erythropoiesis (production of erythrocytes) based on precursor cells (as suggested by [63]) controlled by the
hormone EPO $E(t)$, we are able reduce the PDEs to FDEs if the aging-velocity for the precursor cells is constant (for instance, $V = 1$). Given that the precursor cells grow exponentially for a given period, then stop dividing, the corresponding set of PDEs may be written as:

$$\begin{aligned}
\frac{\partial z_1}{\partial t} + \frac{\partial z_1}{\partial a_1} &= \beta(a_1, E)z_1 \\
\frac{\partial z_2}{\partial t} + \frac{\partial z_2}{\partial a_2} &= -\lambda z_2,
\end{aligned}$$

(8)

under appropriate boundary and initial conditions. Here $\beta$ represents the birth rate for proliferating precursor cells. The concentration of EPO is governed by a standard production/degradation differential equation:

$$\frac{dE}{dt} = f\left(\int_{0}^{T_2(t)} z_2(t, a_1) da_1\right) - \lambda_E E(t),$$

where $f$ represents the rate of production (feedback action), the integral argument describes the total population of erythrocytes on the corresponding time-horizon $[0, T_2]$, and finally, $\lambda_E$ models the degradation effect. After linearization around the equilibrium, the corresponding characteristic equation has the form:

$$(s + \lambda)\left[(s + \lambda)(s + \lambda_E) + ke^{-s\tau}\right] = 0.$$  

(9)

for some appropriate $k$ that depends on the equilibrium point and on $f'$. It is important to note that equation (9) may be analyzed in a similar fashion as equation (3) since $\lambda$ and $\lambda_E$ are both positive.

Among the diseases encountered in hematology, we cite anemia and leukemia. In the case of chronic myelogenous leukemia (CML), there exist many dynamical delay models as pointed out by [22] (and the references therein), with most of them including more than one delay. Recall that Leukemia represents a cancer of the blood cells characterized by an abnormal proliferation of leucocytes. For instance, reference [22] introduces and discusses a model that
includes 4 delays of various sizes. It is of biological interest to see how large delays (1 to 8 days) interact with small ones (1 to 5 minutes) (see, for instance, [76] for such an analysis on some model of reduced-order). Finally, time-delays frequently appear in epidemic models due to incubation times, as first mentioned a century ago (see, for instance, [62], [51]).

Examples From Networks

Load balancing algorithms for parallel computation. We concentrate now on a different set of problems, namely those arising in parallel computation, and computer networking problems. A standard feature of distributed computing architectures is to utilize a networked set of computation elements to achieve performance levels that are not attainable by a single element. One common distributed architecture is a cluster of otherwise independent computers communicating through a shared network (see, e.g. [16] and the references therein). In this context, the distribution of the computational load across available resources is referred to as load balancing.

Consider a computing network consisting of \( n \) computers (nodes) all of which can communicate with each other. At start up, the computers are assigned an equal number of tasks. The change of load at each computer will lead to load imbalance, as some nodes may operate faster than others. In addition, when a node executes a particular task, it may in turn generate more tasks, so that the loads on various nodes become unequal. To balance the loads, each computer in the network sends its queue size \( q_j(t) \) to all other computers in the network. A node \( i \) receives this information from node \( j \) delayed by a finite amount of time \( \tau_{ij} \); that is, it receives \( q_j(t - \tau_{ij}) \). Each node \( i \) then uses this information to compute its local estimate...
of the average number of tasks in the queues of the $n$ computers in the network. The simple estimator
\[ \left( \sum_{j=1}^{n} q_j(t - \tau_{ij}) \right) / n, \quad (\tau_{ii} = 0) \]
which is based on the most recent observations is used. Node $i$ then compares its queue size $q_i(t)$ with its estimate of the network average as
\[ \left( q_i(t) - \left( \sum_{j=1}^{n} q_j(t - \tau_{ij}) \right) / n \right) \]
and, if this is greater than zero or some positive threshold, the node sends some of its tasks to the other nodes. If it is less than zero, no tasks are sent. Further, the tasks sent by node $i$ are received by node $j$ with a delay $h_{ij}$. The task transfer delay $h_{ij}$ depends on the number of tasks to be transferred and is much greater than the communication delay $\tau_{ij}$. The controller (load balancing algorithm) decides how often and fast to do load balancing (transfer tasks among the nodes) and how many tasks are to be sent to each node.

As just explained, each node controller (load balancing algorithm) has only delayed values of the queue lengths of the other nodes, and each transfer of data from one node to another is received only after a finite time delay. An important issue considered here is the effect of these delays on system performance. Specifically, the model developed here represents our effort to capture the effect of the delays in load balancing techniques as well as the processor constraints so that system theoretic methods could be used to analyze them.

The mathematical model of a given computing node is given by

\[
\frac{dx_i(t)}{dt} = \lambda_i - \mu_i + u_i(t) - \sum_{j \neq i} p_{ij} u_j(t - h_{ij})
\]
\[
y_i(t) = x_i(t) - \frac{\sum_{j=1}^{n} x_j(t - \tau_{ij})}{n}
\]
\[
u_i(t) = -K_i y_i
\]
where

\[ x_i(t) = \text{Expected waiting time experienced by a task inserted into the queue of the } i^{th} \text{ node.} \]

\[ q_i = \text{Number of tasks in the } i^{th} \text{ node.} \]

\[ t_{pi} = \text{Average time needed to process a task on the } i^{th} \text{ node } (x_i(t) = q_i t_{pi}). \]

\[ \lambda_i = \text{Rate of generation of new tasks on the } i^{th} \text{ node (rate of increase in } x_i). \]

\[ \mu_i = \text{Rate of service of tasks at the } i^{th} \text{ node } (\mu_i = \frac{1}{t_{pi}} = 1 \text{ for all } i). \]

\[ u_i(t) = \text{Rate of removal (transfer) of the tasks from node } i \text{ at time } t \text{ by load balancing at node } i. \]

\[ p_{ij} = \text{Fraction of } u_j(t) \text{ that node } j \text{ allocates to node } i \text{ at time } t; \sum_{i=1}^{n} p_{ij} = 1, (p_{jj} = 0) \]

\[ p_{ij} u_j(t - h_{ij}) = \text{Rate of removal (transfer) of tasks at time } t \text{ from node } j \text{ by (to) node } i. \]

\[ h_{ij} = \text{Time delay for task transfer from node } j \text{ to node } i. (h_{ii} = 0) \]

\[ \tau_{ij} = \text{Time delay for communicating the node waiting time } x_j \text{ to node } i (\tau_{ii} = 0). \]

\[ n = \text{Number of nodes} \]

All rates are in units of the rate of change of expected waiting time (time/time, which is dimensionless). In what follows, \( u_i(t) < 0 \) means tasks are being sent to other nodes while \( u_i(t) > 0 \) means the \( i^{th} \) node is receiving tasks from other nodes. A delay is experienced by transmitted tasks before they are received at the other node. Model (10) depicts the behavior of a parallel computing environment when \( u_i(t) < 0 \), but not when \( u_i(t) > 0 \). In particular, with \( u_i(t) > 0 \) the increase in tasks for the \( i^{th} \) node is seen immediately, but the decreases in the other nodes where these tasks originated are delayed due to communication times. For the purposes of this analysis, this modeling inaccuracy is considered acceptable since the actual behavior is
nonlinear.

The control law \( u_i(t) = -K_i y_i \) simply states that if the \( i^{th} \) node output \( x_i(t) \) is above the local average \( \left( \sum_{j=1}^{n} x_j(t - \tau_{ij}) \right)/n \) then send data to the other nodes, while if it is less than the local average, accept data from the other nodes. Often, the \( p_{ij} \) are functions of the state \( x_i \) so as to send a higher fraction of the data to those nodes that have less tasks. However, this is left out in this model to retain linearity of the system so that a stability analysis can be carried out. To study the stability of the model, three nodes \( (n = 3) \) are considered with \( K_1 = K_2 = K_3 = K, p_{ij} = 1/2, \) for all \( i, j, \tau_{ij} = h \) for \( i \neq j, \tau_{ii} = 0, h_{ij} = 2h \) for \( i \neq j, h_{ii} = 0 \). The transfer function from the inputs \( d_1 = \lambda_1 - \mu_1, d_2 = \lambda_2 - \mu_2, d_3 = \lambda_3 - \mu_3 \) to the output \( y_1(s) \triangleq x_1(s) - \left( x_1(s) + e^{-hs}x_2(s) + e^{-hs}x_3(s) \right)/3 \) is \( (z = e^{-hs}) \)

\[
y_1(s) = -\frac{1}{3} \left( \frac{z(s + \frac{K}{6}z(-2 + z + z^2))}{(s + \frac{K}{6}(4 + 2z + 2z^2 + z^3))(s + \frac{2K}{3}(z - 1)^2(z + 1))} \right)(d_2 + d_3) + \frac{2}{3} \left( \frac{s + \frac{K}{6}(4 - 2z - 4z^2 + z^3 + z^4)}{(s + \frac{K}{6}(4 + 2z + 2z^2 + z^3))(s + \frac{2K}{3}(z - 1)^2(z + 1))} \right) d_1
\]

\[
= -\frac{1}{3} \frac{zb_1(s, z)}{a(s, z)} (d_2 + d_3) + \frac{2}{3} \frac{b_2(s, z)}{a(s, z)} d_1
\]

where \( b_1(s, z) = s + \frac{K}{6}z(-2 + z + z^2), b_2(s, z) = s + \frac{K}{6}(4 - 2z - 4z^2 + z^3 + z^4) \). The denominator can be written as \( a(s, z) = a_1(s, z)a_2(s, z) \) where \( a_1(s, z) = (s + \frac{K}{6}(4 + 2z + 2z^2 + z^3)), a_2(s, z) = (s + \frac{2K}{3}(z - 1)^2(z + 1)) \). Not that a general description of the closed-loop polynomial of such systems is given by:

\[
(s + \lambda_1)(s + \lambda_2) + \sum_{i=1}^{N} a_i e^{-ih_s}
\]

**Delays in converging flow problems** When multiple systems communicate through a high-speed network without any explicit resource reservation, the destination node may become overloaded.
This problem is known as the *converging flows problem* [85] and may be solved by using an appropriate feedback control mechanism that increases or decreases the sending rate by the sources. One of the simplest cases of the converging flow problem is represented by a network containing one destination node and two source nodes, modeled as *queueing systems* as shown in Figure xx. As described in [85], [81], these are FIFO ("first-in first-out") $G/D/1/b_i$ queues with buffer size $b_i$, arrival rates $x_i$, queue lengths $n_i$, and service rates $d_i(t)$, for $i = 1, 2$. Let $\tau$ the propagation delay through the connection, and assume that it has the same value in both directions $\tau/2$ (symmetric connections). Thus,

$$y_i(t) = e_i(t - \tau/2), \quad d_i(t) = u_i(t - \tau/2), \quad i = 1, 2,$$

where $y_i$ represents the (output) load information stored at the destination, $e_i$ the percentage of the buffer being used, $u_i$ the service rate (amount of bandwidth allocated to each source), for $i = 1, 2$.

Consider now the control algorithm suggested by [81] (a modified version of [85]) proposed under the assumption that the sources are connected to some high-speed network without any bandwidth limitation. This algorithm consists in the following three steps: (1) computation of the explicit amount of bandwidth available $\beta(t)$ at the destination for a given time value $t$; (2) estimation of the fraction $\sigma_i(t)$ of available bandwidth to be allocated to the $i$th source such that $\sigma_1(t) + \sigma_2(t) = 1$; (3) finally, define the service rate $u_i(t)$ as $u_i(t) = \beta(t)\sigma_i(t)$, for $i = 1, 2$. If $\sigma_i$ is chosen as

$$\sigma_i(t) = \frac{e^{-k(1-y_i(t))}}{e^{-k(1-y_1(t))} + e^{-k(1-y_2(t))}}, \quad i = 1, 2,$$

for some $k > 0$, then for some asymptotic buffer occupancy rate $r$, the explicit bound available $\beta(t)$ can be defined as:

$$\beta(t) = h(r - e_0(t)) + d_0(t),$$
where \(d_0\) is the departure rate from the destination, and \(e_0\) being the percentage of the buffer being used at the destination. Now, under the assumption that the source buffer sizes are equal for both sources \(b_1(t) = b_2(t)\), then we arrive to [81]:

\[
\begin{aligned}
\dot{b}_0 \epsilon_0(t) + h e_0(t - \tau) &= hr \\
\dot{e}(t) + (h(r - e_0(t - \tau)) + d_0) \tanh(k e(t - \tau)/2) &= p,
\end{aligned}
\] (12)

where \(d_0\) and \(p\) are assumed to be constant (for fluid approximation models in converging flows problem).

**Active queue management:** Consider now the case of modeling a single random early detection (RED) controller queue interacting with several TCP sources in the idealized case (infinite buffers, sources respecting the rules of additive increasing, multiplicative decreasing) [?]. Roughly speaking, such an algorithm takes into account the probability of some packets being lost during the transmission process to actively manage the queue. To complete the discussion, we need a model of the queue dynamics as presented in [?]. Once the model is completed, the corresponding characteristic equation of its linearized version is given by:

\[
(s + \alpha_1)(s + \alpha_2) + \alpha_0 e^{-s\tau} = 0,
\]

where \(\tau\) denotes the round trip time, and constants \(\alpha_i\) \((i = 0, 1, 2)\) are obtained by linearizing the nonlinear system of FDEs with respect to the variables \(\overline{\lambda}(t), \overline{s}(t)\) and \(\overline{q}(t)\), which represent the expectations of the TCP throughput \(\lambda(t)\), the instantaneous queue length \(q(t)\) and the corresponding exponentially averaged queue length \(s(t)\). Note that once again, we end up with a second-order system with positive coefficients, and one delay value.
Other physical examples

In the following, we collect other physical examples where delays may be found.

**Heat exchanger:** The following equations represent the transition state of the heat exchange process in the case of counterflow,

\[
\frac{\partial T_1}{\partial t} + v_1 \frac{\partial T_1}{\partial t} = k_{s1}(T_s - T_1) \\
\frac{\partial T_2}{\partial t} - v_2 \frac{\partial T_2}{\partial t} = k_{s2}(T_s - T_2) \\
\frac{\partial T_s}{\partial t} = k_{1s}(T_1 - T_s) + k_{2s}(T_2 - T_s)
\]

where \( T_1, T_2 \) and \( T_s \) are the temperature of the first medium, the second medium, and the partition wall, and are functions of the time \( t \) and position \( x \). The coefficients \( k' \)'s represent heat exchange coefficients, and \( v_1 \) and \( v_2 \) are the velocities of the medium flows. Eliminating the variable \( T_s \), taking Laplace transform with respect to the variable \( t \), it can be shown that \( T_1 \) and \( T_2 \) at \( x = 0 \) and \( x = l \) can be related in terms of transfer function matrix

\[
T_{il}(s) = G_{i1}(s)T_{10}(s) + G_{i2}(s)T_{20}(s), \quad i = 1, 2
\]

where,

\[
T_{i0}(s) = T_i(s, 0), \\
T_{il}(s) = T_i(s, l)
\]

and transfer functions \( G_{ij}(s), i, j = 1, 2 \) are transcendental functions. In practice, it is often sufficient to use

\[
G_{11}(s) = \frac{be^{-st_1}}{1 + as}
\]

with similar simplification possible for other \( G_{ij}(s) \). For details, the readers are referred to Section 1.2.2 of [?]. Such a process of obtaining a time-delay equation from a partial differential equation
can be used in many different cases. Another well known example is the lossless transmission line discussed by Brayton in [107]. A more recent overview was done by Rasvan [110].

**Combustion in Liquid Propellant Motor:** A linearized model of liquid monopropellant rocket motor with a pressure feeding system was described by Crocco [115]

\[
\dot{\phi}(t) = (\gamma - 1)\phi(t) - \gamma\phi(t - \delta) + \mu(t - \delta) \\
\dot{\mu}_1(t) = \frac{1}{\xi J} \left[-\psi(t) + \frac{p_0 - p_1}{2\Delta p}\right] \\
\dot{\mu}(t) = \frac{1}{(1 - \xi) J} [-\mu(t) + \psi(t) - P\phi(t)] \\
\dot{\psi}(t) = \frac{1}{E} [\mu_1(t) - \mu(t)]
\]

This model was also used by Fiagbedzi and Pearson [116] and more recently Zheng, Cheng and Gao [117]. Internal combustion engines also provide rich examples of time-delay elements, such as engine cycle delays caused by fuel-air mixing, ignition delay. A comprehensive survey of internal combustion engine modeling can be found in the article by Cook and Powell [118]. A more recent study can be found in [119].

**Mathematical Models**

**Integrators**

The simplest example for which delays may be destabilizing is that of an integrator \( H(s) = 1/s \), subject to a proportional control law \( u(t) = -ky(t - \tau) \), with \( k > 0 \). The stability of the closed-loop system is determined by the location of the zeros of the characteristic equation: \( s + ke^{-s\tau} = 0 \), which is a quasipolynomial equation. Since a quasipolynomial is transcendental, it has an infinite number of roots. It is therefore important to know how the roots behave with respect to the parameters (e.g. whether standard continuity properties hold). Intuitively,
the continuous dependence of the roots of the characteristic equation on the delay should hold
(the easiest argument is given using Rouché’s lemma [2]). The only technical problem seems
to be that when increasing the delay \( \tau \) from its zero value to a small \( \tau = \varepsilon > 0 \) we suddenly
encounter an infinite number of roots (coming from \(-\infty\)) having started with a single root
located at \( s = -k \).

The case of multiple integrators \( H(s) = 1/s^n \) for \( n > 1 \) is also of interest since it arises
when feedback-linearizing a special class of nonlinear systems. It may in fact be possible to use
delayed feedback (as opposed to differentiating the output) to stabilize nonlinear systems. To the
best of the authors’ knowledge, such an issue was not sufficiently exploited in the literature.

Chaouki will expand further and will also review some state-space models as well as discuss
delay in inputs and delay in outputs.

The Dual Nature of Delay

The study of delays in engineering systems began in the 1930s and focused on the stability
of such systems [12], [13], [24]. One primary interest of these early works was to understand
and to explicitly study the relationship between the parameters of a given system and its stability.
There are 2 general methods that may be used to study the stability of quasipolynomials,
namely the \( D \)-decomposition [72] and the \( \tau \)-decomposition [52]. As discussed by [67], the
main difference between the methods lie on the way the delay is accounted for. The second
method (\( \tau \)-decomposition) may be adapted to define the stability domains (or intervals in the
case of a single delay) with respect to the delay parameter. We focus our presentation on the
\( \tau \)-decomposition and illustrate its salient ideas with the mathematical models of section .
Detrimental effects

Consider again the transfer function of a single integrator $H(s) = 1/s$. Assume that the integrator is feedback-controlled using a proportional control law $u(t) = -ky(t-\tau)$, with $k > 0$. As noted earlier, the rightmost root (the closest to the imaginary axis in our case) is a continuous function of the delay parameter. Further discussions may be found in [25], [20], [73], [67]. In order to determine the stability of the closed-loop system, we need to find out whether or not the roots of the closed-loop characteristic equation cross the imaginary axis when the delay value is increased. In other words, we need to find all $\tau$ solutions to the following equations:

$$ j\omega + ke^{-j\omega\tau} = 0 \Rightarrow \begin{cases} \cos(\omega\tau) = 0, \\ k\sin(\omega\tau) = \omega. \end{cases} \quad (13) $$

The first delay value for which equation (13) has a root is termed a critical delay (or delay margin) and is given by $\tau_c = \pi/(2k)$. This delay margin corresponds to the crossing frequency (imaginary axis root) $\omega_c = k$. Due to the $2\pi$ periodicity, we can actually define an infinite number of critical delay values $\tau_{c,\ell} = \pi/(2k) + (\pi\ell)/k$, for $\ell = 0, 1, 2, \ldots$, such that for each critical delay value, there exists at least one root of (13) on the imaginary axis. Using the continuity argument mentioned above, it follows that closed-loop asymptotic stability is guaranteed for all delays

$$ \tau \in \left[0, \frac{\pi}{2k}\right), $$

thus validating the terminology of delay margin for $\tau_c = \pi/(2k)$. Furthermore, for any delay value $\tau \geq \tau_c$, the system can never recover its stability, that is instability persists for large delay values.

Let us explain this delay-induced instability phenomenon (destabilizing effect of delay) by using the characteristic roots dependency on the delay parameter. We are in fact interested in the
movement of the rightmost root location as the delay is increased. Consider for example the gain $k = \sqrt{4} e^{4\pi/3}$. As seen in Fig. 3, increasing the delay value away from 0 generates fast-moving characteristic roots which originate from $-\infty$. Note that the open-loop root located at $-k$ for delay $\tau = 0$ also moves slowly to the right as the delay is increased. Finally, at some value $\tau_c$, a pair of roots (which originated at $-\infty$) cross the imaginary axis into the instability region. As expected, larger values of the gain $k$ induce smaller delay margins, since $\tau_c = 2\pi/k$. This result may also be confirmed using a Nyquist plot Note that increasing the delay value further will lead to a more unstable system as more roots (originating at $-\infty$) cross the imaginary axis into the right half plane. This may be concluded by examining the crossing direction of an imaginary axis root as a function of the delay parameter $\tau$, evaluated at the corresponding crossing frequency. In fact, note that the following quantity:

$$\text{Re}\left\{\frac{ds}{d\tau}\right\}_{s=j\omega_c} = \omega_c^2,$$

is always positive. In other words, as the delay value increases, each critical delay value corresponds to the situation when a pair of characteristic equation roots cross the imaginary axis towards instability. The behavior of the pair of roots may also be explained using a perturbation-based analysis as the delay parameter $\tau$ is perturbed by a small $\varepsilon$ to $\tau = \tau_0 + \varepsilon > 0$. For details, see [15], [67].

Remark 1: Note that the continuity argument is one of the simplest approaches to handle more complicated situations such as the case of multivariable or multi-input-multi-output (MIMO) systems (see, for instance, the matrix pencil techniques in [37], [73], [67] and the references therein). A different approach which leads to the same conclusions uses the Rekasius transformation [83] that reformulates the closed-loop characteristic polynomial as a one-parameter polynomial [79]. The new polynomial shares the same imaginary axis roots.
with the original quasipolynomial. The Rekasius transformation technique is also known as the \textit{pseudo-delay} method [62].

**Beneficial effects**

The stability of a closed-loop system may sometimes be enhanced by the presence of delays. This is best illustrated for a second-order system that starts out oscillatory (or even unstable) but becomes asymptotically stable by increasing the delay amount [1].

Consider the open-loop system with the transfer function $H(s) = 1/(s^2 + \omega_0^2)$ subject to the delayed proportional feedback control law $u(t) = ky(t - \tau)$, with $k \in (0, \omega_0^2)$. The closed-loop characteristic equation is given by the quasipolynomial $s + \omega_0^2 - ke^{-s\tau} = 0$. It is easy to see that when $\tau = 0$, there does not exist any real gain $k$ such that the closed-loop is asymptotically stable. On the other hand, the open-loop system may be stabilized using a proportional-derivative (PD) feedback, $u(t) = k_p y(t) + k_d \dot{y}(t)$. We show next, that a very small positive delay in the proportional feedback controller will induce closed-loop stability. Note that the closed-loop characteristic equation $s^2 + \omega_0^2 - ke^{-s\tau} = 0$ may also arise when closing the loop around a double-integrator described by the transfer function $H(s) = 1/s^2$ with a feedback controller $u(t) = -\omega_0^2 y(t) + ky(t - \tau)$.

Indeed, simple computations show that there exist two (simple) crossing frequencies: $\omega_{c,1} = \sqrt{\omega_0^2 - k}$, and $\omega_{c,2} = \sqrt{\omega_0^2 + k}$, leading to the critical delay values $\tau_{c,1,\ell} = (2\ell\pi)/\sqrt{\omega_0^2 - k}$, and $\tau_{c,2,\ell} = (2\ell + 1)\pi/\sqrt{\omega_0^2 + k}$, for $\ell = 0, 1, 2, \ldots$, respectively. Now, let us find the crossing direction corresponding to each crossing frequency. It is easy to see that:

$$\text{Re} \left\{ \left[ \frac{ds}{d\tau} \right]^{-1} \right\} \bigg|_{s = j\omega_c} = -\frac{2}{\omega_0^2 - \omega_c^2},$$

and thus the crossing direction is towards stability (instability) at $\omega_c = \omega_{c,1,\ell} (\omega_{c,2,\ell})$. At $\ell = 0$, May 26, 2008
$\tau_{c,1,0} = 0 < \tau_{c,2,0} = \pi/\sqrt{\omega_0^2 + k}$, it follows that for a sufficiently small $\tau = \varepsilon > 0$, the closed-loop system is asymptotically stable since the crossing direction is towards stability, no other closed-loop roots may be located in right-half plane (RHP) or on the imaginary axis ($j\mathbb{R}$). Therefore, increasing the delay value has a stabilizing effect and the first delay interval guaranteeing closed-loop asymptotic stability is given by:

$$0 < \tau < \frac{\pi}{\sqrt{\omega_0^2 + k}}.$$

More generally, the delay intervals for which stability is achieved are given by $\tau_{c,2,\ell} > \tau_{c,1,\ell}$, that is:

$$\frac{2\ell\pi}{\sqrt{\omega_0^2 - k}} < \tau < \frac{(2\ell + 1)\pi}{\sqrt{\omega_0^2 + k}}; \ell = 0, 1, 2, \ldots$$

A more detailed analysis will be presented in Section using a Nyquist argument.

Let us study the behavior of the rightmost root as the delay is increased: this particular root will wander in $\mathbb{C}_-$ for a while, then cross into $\mathbb{C}_+$ where it remains for some time before eventually turning back to $\mathbb{C}_-$. If no other characteristic roots cross the imaginary axis to $\mathbb{C}_+$ during this process, the behavior of the rightmost root completely determines the stability of the closed-loop system. As the rightmost root crosses back and forth between the left and right half planes, there will be many delay intervals that correspond to a sequence of stability/instability regions.

Recapping the results of this section, it is obvious that one can study a large class of dynamical systems with delays, as the feedback interconnection of a simple linear system (integrators) with a delay block - structure defined by a pair of parameters (gain,delay). Further remarks on controlling a chain of integrators by using delays are given in section.
Characterizing Delayed Feedback

In the sequel we consider the class of strictly proper SISO open-loop systems:

\[
\frac{P(s)}{Q(s)} = c^T(sI_n - A)^{-1}b
\]

(14)

where \((A, b, c^T)\) is a state-space representation of the open-loop system, with the controller

\[
u(t) = -ky(t - \tau).
\]

(15)

Limitations of delayed feedback

In order to determine the limitations of the delayed feedback, the following problem will be addressed and the results are described using analytical as well as geometric arguments.

**Problem 1: (Delay stabilizing effect)** Find explicit conditions on the pair \((k, \tau)\), such that the controller (15) stabilizes the system (14) if and only if the delay \(\tau\) is strictly positive.

Solving Problem 1 provides a simple controller where the delay time may be explored as a design parameter in situations where it may not be possible to use a controller without delay (see, for instance, the congestion controllers in high-speed networks [45], [46], [74]).

The characteristic polynomial of the closed-loop system obtained by interconnecting (14) and (15) is given by

\[
H(s; k, \tau) = Q(s) + kP(s)e^{-s\tau}.
\]

(16)

Two quantities will play a major role in the stabilizability study:

1) \(\text{card}(\mathcal{U}_+), \) where \(\mathcal{U}_+\) is the set of roots of \(H(s; k, 0) = Q(s) + kP(s)\), located in the closed right half plane, and \(\text{card}(\cdot)\) denotes the cardinality (number of elements).
2) \( \text{card}(\mathcal{S}_+) \), where \( \mathcal{S}_+ \) the set of strictly positive roots of the polynomial

\[
F(\omega; k) = |Q(j\omega)|^2 - k^2 |P(j\omega)|^2.
\] (17)

Both \( \text{card}(\mathcal{U}_+) \) and \( \text{card}(\mathcal{S}_+) \) depend only on the gain parameter \( k \). Note however that characterizing \( H(s; k, 0) \) corresponds to the static (non-delayed) output feedback stabilizability problem. The difficulty of this problem is well known (see, for instance, [99] and the references therein). In the SISO system case however, the static output feedback problem is reduced to a one-parameter search problem, whose solution is relatively using graphical tests (root-locus, Nyquist), or by computation of the real roots of an appropriate set of polynomials. In addition to these standard methods, we cite two interesting approaches [14], [42] based on generalized eigenvalues computation of some appropriate matrix pencils defined by the corresponding Hurwitz [14], and Hermite [42] matrices. Our approach uses such concepts and is inspired by Chen’s characterization [14] for systems without delay.

Consider the Hurwitz matrix \( H(Q) \) associated to the denominator polynomial \( Q(s) = \sum_{i=0}^{n} q_i s^{n-i} \) of the transfer function (14):

\[
H(Q) = \begin{bmatrix}
q_1 & q_3 & q_5 & \cdots & q_{2n-1} \\
q_2 & q_4 & q_6 & \cdots & q_{2n-2} \\
0 & q_1 & q_3 & \cdots & q_{2n-3} \\
0 & q_2 & q_4 & \cdots & q_{2n-4} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & q_n
\end{bmatrix} \in \mathbb{R}^{n \times n},
\] (18)

where the coefficients \( q_l = 0 \), for all \( l > n \). Similarly, let \( H(P) \) be the matrix associated with the numerator polynomial \( P(s) \) of the transfer function (14) with the understanding that \( p_l = 0 \)
for all \( l > m \). The following result is a slight modification, and generalization of Theorem 2.1 by Chen [14]:

**Lemma 1:** Let \( \lambda_1 < \lambda_2 < \ldots < \lambda_h \); \( h \leq n \), be the real eigenvalues of the matrix pencil \( \Sigma(\lambda) = \det(\lambda H(P) + H(Q)) \). Then the system (14) cannot be stabilized by the controller \( u(t) = -ky(t) \) for any \( k = \lambda_i \), \( i = 1, 2, \ldots, h \). Furthermore, if there are \( r \) unstable closed-loop roots \( 0 \leq r \leq n \) for \( k = k^* \), \( k^* \in (\lambda_i, \lambda_{i+1}) \), then, there are \( r \) unstable closed-loop roots for any gain \( k \in (\lambda_i, \lambda_{i+1}) \). In other words, \( \text{card}(U_+) \) remains constant as \( k \) varies within each interval \( (\lambda_i, \lambda_{i+1}) \). The same holds for the intervals \( (-\infty, \lambda_1) \) and \( (\lambda_h, \infty) \).

Thus by computing the generalized eigenvalues of the matrix pencil \( \Sigma(\lambda) \), yielding the critical gain values, and computing \( U_+ \) for intermediate gain values, the function \( k \to \text{card}(U_+)(k) \) is completely determined.

**Calculating** \( \text{card}(S_+) \): Without loss of generality assume that \( P(0) \neq 0 \). Then:

**Proposition 1:** Assume that \( \text{card}(S_+) \) changes at a gain value \( k^* \). Then there exists a frequency \( \omega^* \geq 0 \), such that for \( \omega = \omega^* \):

\[
|Q(j\omega)|^2 - k^*^2 |P(j\omega)|^2 = 0 \tag{19}
\]

and

\[
|P(j\omega)|^2 \frac{d}{d\omega} |Q(j\omega)|^2 - |Q(j\omega)|^2 \frac{d}{d\omega} |P(j\omega)|^2 = 0. \tag{20}
\]

Proposition 1 allows us to systematically compute the behavior of \( \text{card}(S^+) \) as a function of the gain \( k \) as follows. First, we find the real roots of the polynomial (20). Then, the critical values of the gain \( k \) are found from (19). The characterization of \( \text{card}(S^+) \) is completed by computing \( S_+ \) for intermediate gain values, which again corresponds to finding the roots of a polynomial.
Remark 2: Using the symmetry of the function $F(w; k)$ in (17), we conclude that the pairs $(\omega, k) = (0, \pm \frac{Q(0)}{P(0)})$ always satisfy (19)-(20) and furthermore, at these $k$-values, $\text{card}(S_+)$ always changes by $\pm 1$. It also follows from (17) that at other $k$-values, $\text{card}(S_+)$ can only change by $\pm 2$.

Delay and gain bounds: explicit computations

By exploiting the dependence of $\text{card}(S_+)$, and $\text{card}(U_+)$ on gain $k$, we obtain next an explicit solution to problem 1 under the following assumptions:

Assumption 1: Let the gain $k \in \mathbb{R}$ be such that

1) all the roots of $F$ are simple,

2) $0 \not\in U_+$,

3) $\text{card}(U_+) \neq 0$.

Note that the first condition is satisfied for almost all $k$. The second condition is necessary for stabilization because it excludes a characteristic root at zero, the latter being invariant w.r.t. delay changes. The third assumption excludes the trivial case where the system is asymptotically stable when the delay $\tau = 0$. We then recall the following result from [17]:

Theorem 1: The characteristic equation has a root $j\omega$ for some delay value $\tau_0$ if and only if

$$\omega \in S_+.$$  \hfill (21)

Furthermore, for any $\omega$ satisfying (21), the set of corresponding delay values is given by

$$T_\omega = \left\{ \frac{1}{\omega} \left[ -j \log \left( -\frac{kP(j\omega)}{Q(j\omega)} \right) + 2\pi l \right] \geq 0, \ l \in \mathbb{Z} \right\}. \hfill (22)$$

Note that “Log” denotes the principal value of the logarithm. Consequently for $|z| = 1$, $\log(z) = j \arg(z)$ with $\arg(z) \in (-\pi, \pi]$. Finally, a the delay increases, the corresponding crossing
direction of the characteristic roots is towards instability (stability) when $F'(\omega) > 0 \ (< 0)$.

Remark 3: Theorem 1 is due to the fact that $H(j\omega; k, \tau) = 0$ is an equation in two independent unknown variables, $j\omega \in j\mathbb{R}$ and $z := \exp(-j\omega\tau)$. Taking the modulus of $H(.)$ allows us to eliminate the second variable $z$. As seen in [73], the first variable may also be eliminated via matrix pencils techniques, leading to an alternative characterization of the stability regions.

Theorem 1, along with the continuous dependence of the characteristic roots on the delay, allows to completely characterize the stability/instability regions in the delay parameter. Indeed, the set

$$\mathcal{T} = \bigcup_{\omega \in \mathcal{S}_+} \mathcal{T}_\omega$$

partitions the delay space ($\mathbb{R}^+$) into intervals within each, the number of roots in the open right half plane is constant. Moreover, we have the following result:

Remark 4: (crossing direction) Assume that $\mathcal{S}_+$ is not empty and denote its elements in descending order with $\omega_1 > \omega_2 > \ldots$. Since $\lim_{\omega \to \infty} F(\omega) = +\infty$ and the roots $\{\omega_1, \omega_2, \ldots\}$ of $F$ are simple, the sign of $F'$ at these roots alternates, with $F'(\omega_1) > 0$. As a consequence, as the delay is monotonically increased, all root crossings for delay values $\mathcal{T}_{\omega_1}$ are towards instability, while all root crossings for delay values $\mathcal{T}_{\omega_2}$ towards stability etc.

Taking into account the number of unstable roots for $\tau = 0$ and the crossing directions in $\mathcal{T}$, the number of unstable roots for each delay value may be found.

Based on Theorem 1, we can examine the delay stabilization problem. First, we analyze cases for which the delay stabilization problem has no solution, then present the simplest case (in terms of $\text{card}(\mathcal{S}_+)$, and $\text{card}(\mathcal{U}_+)$) for which the delay stabilization problem has a positive
answer. Finally, as a consequence of all the cases treated, we present the necessary, and sufficient conditions such that a pair \((k, \tau)\) stabilizes the SISO system (14).

*Proposition 2:* Assume that \(\text{card}(U_+)\) is an odd number. Then the delay stabilization problem has no solution.

The proof of this result is relatively simple. In fact, a strictly positive root of the characteristic equation exists for all delay values if, the delay-free system has such a root, and if the number of roots in the open right half plane is odd. Similar results using a variety of proofs have already been given in the literature (see, for instance, [95], [34], [64] to cite only a few).

*Proposition 3:* If either \(\text{card}(S_+^c) = 0\) or \(\text{card}(S_+) = 1\), then the delay stabilizing problem has no solution.

Note that the condition \(\text{card}(S_+) = 0\), corresponds to the delay-independent hyperbolicity property (fixed number of unstable roots for all positive delay values), as defined in [39] (see also [40]). For the remaining cases, one needs to count the roots crossing the imaginary axis towards stability/instability, and to define the corresponding delay intervals (see also [73], chapters 4, and 7).

It follows from the previous two propositions that the first case for which the delay parameter has a stabilizing effect takes place when \(\text{card}(U_+) = 2, \text{card}(S_+) \geq 2\), and the first crossing is towards stability. Furthermore, as proved in the sequel, such a condition becomes also necessary if \(\text{card}(S_+) \leq 3\).

*Proposition 4:* Assume that \(\text{card}(S_+) = 2\) or \(\text{card}(S_+) = 3\). Then the delay stabilizing problem has a solution if and only if
1) $\text{card}(U_+) = 2$

2) $\tau_- < \tau_+$, where

$$
\tau_- = \min \bigcup_{\omega \in S_+} \{ \omega : F'(\omega) < 0 \} T_\omega,
$$

$$
\tau_+ = \min \bigcup_{\omega \in S_+} \{ \omega : F'(\omega) > 0 \} T_\omega \setminus \{0\}.
$$

If the open-loop system is stabilizable, all delay values $\tau \in (\tau_-, \tau_+)$ are stabilizing.

\[\square\]

Remark 5: In order to check stabilizability, one only needs to find the first imaginary axis root crossing as the delay is increased from zero. This is particularly useful when numerically investigating stabilizability by computing the rightmost characteristic roots as a function of delay.

In the case when $\text{card}(S_+) = 2$, the set of all stabilizing delay values may be expressed analytically as shown next:

**Corollary 2:** Assume that the conditions of Proposition 4 are satisfied, and in addition $\text{card}(S_+) = 2$. Then all stabilizing delay values are given by $\tau \in (\tau_l, \tau_l')$, $l = 0, 1, 2, \ldots, l_m$, where

$$
\tau_l = \tau_- + \frac{2\pi l}{\omega_-}, \quad \tau_l' = \tau_+ + \frac{2\pi l}{\omega_+},
$$

$S_+ = \{\omega_+, \omega_-\}$ with $\omega_+ > \omega_-$, and $l_m$ is the largest integer for which $\tau_l < \tau_l'$, which can be expressed as

$$
l_m = \max_{l \in \mathbb{Z}} \left\{ l < \frac{\omega_+ - \omega_-}{\omega_+ - \omega_-} \cdot \frac{\tau_+ - \tau_-}{\omega_+ - \omega_-} \right\}.
$$

(23)

\[\square\]

**Proposition 5:** Assume that $\text{card}(S_+) = 2n$ or $\text{card}(S_+) = 2n + 1$, with $n \geq 1$. Assume further that $\text{card}(U_+) > 2n$. Then the delay stabilizing problem has no solution.

\[\square\]
Define now the following quantities:

\[ n_+ (\tau) = \sum_{\omega \in S^+, F'(\omega) > 0} \text{card} \{ T_\omega \cap (0, \tau) \}, \quad (24) \]
\[ n_- (\tau) = \sum_{\omega \in S^+, F'(\omega) < 0} \text{card} \{ T_\omega \cap [0, \tau] \}, \quad (25) \]

for some positive \( \tau > 0 \). Furthermore, introduce the sets \( T^+ \) and \( T^- \), which represent a partition of \( T \) as a function of the sign of the derivative \( F' \) evaluated at the corresponding crossing frequency, that is:

\[ T^+ = \bigcup_{\omega \in S^+, F'(\omega) > 0} T_\omega \setminus \{ 0 \}, \]
\[ T^- = \bigcup_{\omega \in S^+, F'(\omega) < 0} T_\omega. \]

Based on the conditions and the notations above, we conclude with the following result:

**Proposition 6:** For a given gain \( k \), the stabilizing control problem has a solution of the form \( u(t) = -ky(t - \tau) \) if and only if the following conditions hold simultaneously:

(i) \( \text{card}(U_+(k)) \) is a strictly positive even integer, which satisfies the inequality \( \text{card}(U_+(k)) \leq \text{card}(S_+(k)) \), and

(ii) there exists at least one delay value \( \hat{\tau} \in T \), such that the following equality is verified:

\[ 2n_- (\hat{\tau}) = 2n_+ (\hat{\tau}) + \text{card}(U_+(k)). \quad (26) \]

Then all delay values \( \tau \in (\hat{\tau}, \hat{\tau}_+) \), with

\[ \hat{\tau}_+ = \min \{ T^+ \cap (\hat{\tau}, +\infty) \} \quad (27) \]

guarantee closed-loop asymptotic stability.

\[ \square \]

The main results of this section are displayed in Table I. Recall that \( \text{card}(U_+) \) and \( \text{card}(S_+) \) only depend on the gain \( k \). Also recall that \( \text{card}(U_+) \) may be obtained by computing the generalized
eigenvalues of a matrix pencil (Proposition 1), while \( \text{card}(S_+)^+ \) is determined by computing the roots of a polynomial (Proposition 1).

The procedure of finding a stabilizing pair \((k, \tau)\) (if any) is summarized as follows:

- Compute \( \text{card}(S_+) \) and \( \text{card}(U_+) \) as functions of the gain parameter \( k \), then select possible gain intervals \((k, \bar{k})\) such that the condition (i) of Proposition 6 holds;
- Next, for a given \( k \), search for stabilizing delay values. In the special case where \( \text{card}(S_+) = 2 \) or \( \text{card}(S_+) = 3 \) and \( \text{card}(U_+) = 2 \), Proposition 4 may be applied, i.e. it is sufficient to investigate only whether the first root crossing of the imaginary axis is toward stability as the delay is increased from zero. In general, a more complete characterization of stability/instability regions becomes necessary, as we illustrate in the next section. Indeed, one has to compute the set \( T \), and next the partition \( T^+ \) and \( T^- \), which gives the roots crossings (function of the delay values) towards instability and stability, respectively. Condition (26) together with (27) completely characterizes the corresponding stabilizing delay intervals.

**Why Does Delay Feedback Work?**

It is generally assumed that delays in the feedback law may lead to instability. Examples of such behavior have already been mentioned in the previous section (first- and second-order systems with large delays in the feedback law, or high gains in the case of proper systems). There are many cases however, where a delay in a feedback law leads to stable closed-loop schemes, as shown for the simple oscillatory system with positive delayed feedback [1] presented in the previous section. To the best of the authors’ knowledge however, a systematic study of the stabilizing effect of delays is yet to be provided.
TABLE I

OUTPUT FEEDBACK STABILIZABILITY CONDITIONS WHEN USING THE DELAY AS CONTROLLER PARAMETER. NECESSARY AND SUFFICIENT CONDITIONS ARE GIVEN BY PROPOSITION 6. IN CASE OF \( \text{card}(\mathcal{U}_+) = 2 \) AND \( \text{card}(\mathcal{S}_+) \in \{2, 3\} \), PROPOSITION 4 CAN BE APPLIED.

We believe that the stabilizing effect of delayed feedback may be properly understood using one of two interpretations: either as an approximate derivative feedback, or as an instrument of phase synchronization as will be shown in the sequel.

**Derivative Feedback**

Delays may arise in practice when using a finite difference approximation of signal derivative. For example a finite difference of angular velocities was used in [108] instead of the angular acceleration in implementing an adaptive control of robot dynamics. In order to explain
this approach consider the simple linear system:

\[ \ddot{x}(t) - 0.1\dot{x}(t) + x(t) = u(t). \]  

(28)

The open-loop system is unstable due to the negative damping term \(-0.1\dot{x}(t)\). A derivative feedback

\[ u(t) = -k\dot{x}(t) \]  

(29)

with \( k > 0.1 \) moves the open-loop poles originally located \( 0.05 \pm j0.9987 \) into the left half plane. Alternatively, we may use the delayed-feedback control law

\[ u(t) = x(t - r) - x(t). \]  

(30)

The control law (30) may be interpreted as a finite difference control law with a gain \( r \) such that

\[ u(t) = -r \frac{x(t) - x(t - r)}{r}. \]

For small values of the delay \( r \), equation (refdc-fd) approximates the derivative control (29) with \( k = r \) reasonably well. Indeed, it may be shown that system (28) is stabilized by moving the two right-half-plane poles to the left half plane for \( r \in (0.1002, 1.7178) \). This example was used in Chapter 4 of [37] and [?] to assess the effectiveness of some stability tests.

In [75], Niculescu and Michiels explored the possibility of using a combination of \( m \) distinct delays

\[ u(t) = -\sum_{i=1}^{m} k_i x(t - \tau_i) \]  

(31)

to stabilize the chain of integrators:

\[ x^{(n)}(t) = u(t). \]  

(32)
It was found in [75] that a judicious use of \( n \) distinct delays, one of which may be zero, is sufficient to achieve stability. For the chain of integrators system, the delays are not even required to be small. Indeed, if (31) stabilizes (32), then so will

\[
u(t) = - \sum_{i=1}^{n} \frac{k_i}{\rho} x(t - \rho \tau_i).
\]

This may be easily seen by scaling the time variable

\[t = \hat{t}/\rho.\]

One of the controllers proposed in [75] uses a direct approximation of derivatives combined with a scaling of time, leading to the controller

\[
u(t) = - \left( e^n q_0 \frac{e^{n-1}q_1}{(-1)} \frac{2^n e^{n-2}q_2}{(-1)^2} \cdots \frac{(n-1)! e q_{n-1}}{(-1)^{n-1}} \cdot T^{-1}(\tau) \right) \begin{pmatrix} x(t - \tau_1) \\ x(t - \tau_2) \\ \vdots \\ x(t - \tau_n) \end{pmatrix}
\]

where \( \tau_i, i = 1, 2, \ldots, n \) satisfy \( 0 \leq \tau_1 < \tau_2 < \cdots < \tau_n \); \( q_i, i = 0, 1, \ldots, n - 1 \) are chosen such that the closed-loop system with the derivative feedback control

\[
u(t) = - \sum_{i=0}^{n-1} q_i x^{(i)}(t)
\]

is stable. Note that \( T(\tau) \) is the Vandermonde matrix

\[
T(\tau) = \begin{pmatrix} 1 & \tau_1 & \tau_1^2 & \cdots & \tau_1^{n-1} \\ 1 & \tau_2 & \tau_2^2 & \cdots & \tau_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \tau_n & \tau_n^2 & \cdots & \tau_n^{n-1} \end{pmatrix}
\]

and \( \varepsilon > 0 \) is sufficiently small. Note that Kharitonov et al showed in [47], that the system can never be stabilized by (31) if \( m < n \), which parallels the requirement of stabilizability using
derivative feedback. Using finite differences to approximate derivatives may not however be valid when the derivative feedback is not well-posed [65]. Consider for example the system:

\[ \dot{x}(t) = x(t) + u(t). \]

The derivative feedback

\[ u(t) = 2\dot{x}(t) \]

stabilizes the system but feedback system is not well-posed. Indeed, it can be shown that no controllers of the form

\[ u(t) = H(x(t) - x(t - T)) \]

can stabilize the system [109].

**Phase synchronization**

In the following we study a more general version of equation (16). We study the behavior of the zeros of the function

\[ H(s; \tau, \epsilon) := f(s) + \epsilon H(s)e^{-st}, \quad (33) \]

where \( f : \mathbb{C} \to \mathbb{C} \) and \( g : \mathbb{C} \to \mathbb{C} \) are entire functions of the gain parameter \( \epsilon \in \mathbb{R} \) and the delay parameter \( \tau \geq 0 \). The following result characterizes the behavior of the zeros of (33) as a function of the delay parameter:

**Proposition 7:** Let \( \hat{s} \) be a zero of \( f \) with multiplicity \( m \geq 1 \) that is not a zero of \( g \). Let \( \Omega \subset \mathbb{C} \) be a compact set, which contains \( \hat{s} \) but no other zeros of \( f \), and such that \( \partial \Omega \) is a closed, simple contour not containing \( \hat{s} \). Then for all \( \hat{\tau} > 0 \) there exists a number \( \hat{\epsilon} > 0 \) such that the following holds:

1) \( H(s; \tau, \epsilon) \) has exactly \( m \) zeros in \( \Omega \), for all \( (\tau, \epsilon) \in [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}] \);
2) there are $m$ functions $r_i: [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}] \rightarrow \Omega,$

$$(\tau, \epsilon) \mapsto r_i(\tau, \epsilon), \quad i = 1, \ldots, m,$$

which satisfy $r_i(0, 0) = \hat{s},$ $H(r_i(\tau, \epsilon); \tau, \epsilon) = 0,$ $\forall (\tau, \epsilon) \in [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}],$ and which can be decomposed as:

$$r_i(\tau, \epsilon) = \hat{s} + |\epsilon|^{\frac{1}{m}} \mu_i(\tau, \epsilon),$$

where

$$\lim_{|\epsilon| \rightarrow 0^+} \max_{\tau \in [0, \hat{\tau}]} \left| \mu_i(\tau, \epsilon) - \left( -\text{sign}(\epsilon) \frac{m g(\hat{s})}{m g(\hat{s})} \right) \frac{1}{m} e^{2 \pi (i-1) \frac{\epsilon}{m}} e^{-\frac{i}{m} \tau} \right| = 0, \quad i = 1, \ldots, m. \quad (34)$$

□

For the case of a simple zero ($m = 1$) Proposition 7 simplifies to:

**Corollary 3:** Let $\hat{s}$ be a zero of $f$ with multiplicity one that is not a zero of $g.$ Let $\Omega \subset \mathbb{C}$ be a compact set, which contains $\hat{s}$ but no other zeros of $f,$ and such that $\partial\Omega$ is a closed, simple contour not containing $\hat{s}.$ Then for all $\hat{\tau} > 0$ there exists a number $\hat{\epsilon} > 0$ such that $H(s; \tau, \epsilon)$ has exactly one zero in $\Omega,$ for all $(\tau, \epsilon) \in [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}].$ Furthermore, there is a (unique) function $r: [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}] \rightarrow \Omega,$ $(\tau, \epsilon) \mapsto r(\tau, \epsilon),$ which satisfies $r(0, 0) = \hat{s},$

$$H(r(\tau, \epsilon); \tau, \epsilon) = 0, \quad \forall (\tau, \epsilon) \in [0, \hat{\tau}] \times [-\hat{\epsilon}, \hat{\epsilon}],$$

and which can be decomposed as:

$$r(\tau, \epsilon) = \hat{s} + \epsilon \mu(\tau, \epsilon),$$

where

$$\lim_{|\epsilon| \rightarrow 0^+} \max_{\tau \in [0, \hat{\tau}]} \left| \mu(\tau, \epsilon) + \frac{g(\hat{s})}{f'(\hat{s})} e^{-\hat{s} \tau} \right| = 0. \quad (35)$$

□
We briefly discuss the above results, starting with the special case treated in the
Corollary 3. Expression (35) implies that for small values of the gain parameter \( \epsilon \), an isolated
zero behaves as the function

\[
\tau \mapsto \hat{s} - \epsilon \frac{g(\hat{s})}{f'(\hat{s})} e^{-\hat{s}\tau}. \tag{36}
\]

Consequently, if \( \Im(\hat{s}) > 0 \), it has an oscillatory behavior as a function of the delay parameter.
If, furthermore, \( \hat{s} \in \mathbb{C}_- (\mathbb{C}_+) \), then the zero behaves like an unstable (stable) spiral. This is
illustrated in Figure 4 for the quasi-polynomial

\[
H(s; \tau, \epsilon) = \prod_{i=1}^{6} (s - s_i) + \epsilon e^{-s\tau}, \tag{37}
\]

where

\[
s_{1,2} = -\frac{1}{20} \pm j, \quad s_{3,4} = \pm \frac{3}{2} j, \quad s_{5,6} = \frac{1}{40} \pm 2 j,
\]

\( \tau \in [0, 36] \) and \( \epsilon = 1/8 \) (left), respectively \( \epsilon = 1/16 \) (right). The zeros of (37) were computed
using the package DDE-BIFTOOL [27]. Notice from (35) that \( \mu(\tau, \epsilon) \) is the offset of the zero, i.e.
r(\tau, \epsilon) - \hat{s}, scaled by the gain parameter \( \epsilon \). This explains why the 'amplitudes' of the oscillations
in Figure 4 are smaller for \( \epsilon = 1/16 \) than for \( \epsilon = 1/8 \).

In the case where \( f \) has a zero with multiplicity \( m > 1 \), covered by Proposition 7, the
\( m \) corresponding zeros of \( H \) behave for fixed small values of the gain parameter as the equally
'shifted' spiral curves,

\[
\tau \mapsto \hat{s} + \left( -\epsilon \frac{m!g(\hat{s})}{d^m f(\hat{s})} \right) \frac{1}{m} e^{-\frac{\hat{s}}{m} \tau} e^{\frac{2\pi i (i-1)}{m}}, \quad i = 1, \ldots, m.
\]

To illustrate this we show in Figure 5 the real parts of the rightmost zeros of

\[
H(s; \tau, \epsilon) = (s^2 + 1)^3 + \epsilon e^{-s\tau} \tag{38}
\]
as a function of $\tau$ for $\epsilon = 1/640$. The (initial) delay-shift of $2\pi$ between the functions (corresponding to a phase shift of $2\pi/3$) is a consequence of the presence of a zero of $f(s) = h(s; \tau, 0)$ with multiplicity three.

In this section we assume that the rightmost zeros of $f$ are on the imaginary axis. We further assume that

$$f(\bar{s}) = \overline{f(s)}, \ g(\bar{s}) = \overline{H(s)}, \ \forall s \in \mathbb{C}.$$  

As a consequence, the zeros of $H$ appear in complex conjugate pairs. Simple zeros: If all zeros of $f$ on the imaginary axis are simple, then the corresponding functions of the form (36) have a sinusoidal real part, exhibiting different frequencies. As a consequence, the asymptotic stability of $H$ for small values of $\epsilon$ is related to having an appropriate phase of these sinusoidal functions, which depends of the delay parameter only. This relation between the stability of $H$ and a phase 'synchronization' problem is clarified in the following proposition:

**Proposition 8**: Assume that there is a constant $\gamma > 0$ such that

$$\lim_{R \to \infty} \sup \left\{ \left| \frac{H(s)}{f(s)} \right| : \Re(s) \geq -\gamma, \ |s| \geq R \right\} = 0. \quad (39)$$

Assume further that all zeros of $f$ are in the closed left half plane. Denote by $j\omega_i$, $i = 1, \ldots, \nu$, the zeros of $f$ on the positive imaginary axis, which all have multiplicity one.

If the delay parameter $\tau$ is such that for all $i = 1, \ldots, \nu$:

$$\Re \left( \frac{g(j\omega_i)}{f'(j\omega_i)} e^{-j\omega_i \tau} \right) > 0 \ (\leq 0), \quad (40)$$

then all zeros of $H(s; \tau, \epsilon)$ are in the open left half plane for sufficiently small $\epsilon > 0 \ (\epsilon < 0)$.

**Remark 6**: The assumption (39) is technical and serves to exclude the situation where increasing $|\epsilon|$ from zero leads to the introduction of additional zero’s in the right half plane that
come from infinity. It is fulfilled in most applications of interest. For instance, it is satisfied if $f$ and $g$ are polynomials satisfying $\deg(g) < \deg(f)$.

Remark 7: Proposition 8 generalizes Proposition 5.3 of [?], where an alternative proof is presented, based on the sensitivity of the zeros of $H$ with respect to the parameter $\epsilon$ for a fixed value of $\tau$. More precisely, define a function $s_i(\tau, \epsilon)$ satisfying $s_i(0, 0) = j\omega_i$ and $H(s_i(\tau, \epsilon); \tau, \epsilon) = 0$. Then we get:

$$\frac{\partial s_i(\tau, \epsilon)}{\partial \epsilon} \bigg|_{\epsilon=0} = -\frac{g(j\omega_i)e^{-j\omega_i\tau}}{f'(j\omega_i)}, \quad i = 1, \ldots, \nu.$$ (41)

Expression (40) is equivalent to requiring

$$\frac{\partial \Re(s_i(\tau, \epsilon))}{\partial \epsilon} \bigg|_{\epsilon=0} < 0, \quad i = 1, \ldots, \nu.$$  

The next result addresses a case where the condition (40) can always be met by an appropriate choice of the delay parameter:

**Corollary 4:** Assume that all zeros of $f$ are in the closed left half plane. Denote by $j\omega_i, \ i = 1, \ldots, \nu$, the zeros of $f$ on the positive imaginary axis, which all have multiplicity one. Assume further that (39) holds. If the nonzero frequencies $\omega_i$ are rationally independent, then there always exist values of $\epsilon$ and $\tau$ such that the zeros of $H(s; \tau, \epsilon)$ are in $\mathbb{C}_-$. Recall that The real numbers $(r_1, r_2, \ldots, r_m)$ are rational independent if and only if $\sum_{i=1}^m n_ir_i = 0, \ n_i \in \mathbb{Z}$, implies $n_i = 0, \ i = 1, \ldots, m$. For example, two numbers are rationally independent if and only if their ratio is an irrational number.
If the frequencies $\omega_i, \ i = 1, \ldots, \nu$, are commensurate (multiples of the same number), then the condition (40) can be easily checked because the sinusoidal functions

$$\tau \mapsto \Re \left( \frac{g(j\omega_i)}{f'(j\omega_i)} e^{-j\omega_i \tau} \right), \ i = 1, \ldots, \nu,$$

need to be evaluated only on a finite interval of length $2\pi/(\min_i \omega_i)$. With the following example we illustrate that situations might occur where the condition (40) is violated, whatever the value of the delay.

**Designing With Delays**

In the following, we describe the use of delay as a design parameter for various feedback controllers in linear systems. Consider the general model (4):

$$\Pi_{i=1}^m (s + \lambda_i)^m + ke^{-s\tau} = 0 \quad (42)$$

briefly presented in section , and let us concentrate on the delay-block $(k, \tau)$, and more precisely on dependence between the gain $k$ and the delay $\tau$ in order to guarantee the closed-loop asymptotic stability.

Geometrically speaking, it is easy to see that if $k \leq \Pi_{i=1}^m \lambda_i$, the characteristic (42) has no roots on the imaginary axis, since the graph of the stable open-loop transfer $H(s) = -1/(s + \lambda_1) \ldots s(+\lambda_m)$ evaluated on at $s = j\omega$, with $\omega > 0$ is completely within the unit circle of the complex plane. As mentioned in Section (Tsypkin-type argument), we can the conclude the closed-loop system is stable independent of the delay. In what follows, we describe strategies to stabilize various linear systems,= by a judicious choice of the delay parameter.
Stabilizing oscillatory systems

Consider the following second-order linear, time-invariant plant:

\[ H(s) = \frac{Y(s)}{U(s)} = \frac{1}{s^2 + \omega_n^2} \]  \hspace{1cm} (43)

where \( \omega_n \) is the natural frequency of the system. Let a static, output-feedback delay compensator be given by:

\[ U(s) = C(s)U(s); \quad C(s) = ke^{-st} \] \hspace{1cm} (44)

where \( k \) and \( \tau \) are both design parameters. The closed-loop transfer function is given by:

\[ \frac{Y(s)}{R(s)} = \frac{ke^{-st}}{s^2 + \omega_n^2 - ke^{-st}} \] \hspace{1cm} (45)

Let us consider the Nyquist plot of the open-loop system \( H(s)C(s) \) given by:

\[ H(j\omega)C(j\omega) = \frac{Ke^{-j\omega\tau}}{\omega_n^2 - \omega^2} \] \hspace{1cm} (46)

We divide the Nyquist graph of (46) into three regions: the first, when the frequency is less than the natural frequency, or \( \omega < \omega_n \); the second, when they are equal, i.e. \( \omega = \omega_n \); and the third when the frequency is greater than the natural frequency, so that \( \omega > \omega_n \). Note that the magnitude \( ||H(j\omega)C(j\omega)||_{\omega=\omega_n} \) is infinite, so our analysis focuses on the cases were \( \omega \neq \omega_n \).

For those regions, the magnitude of the open-loop gain is:

\[ |G(j\omega)C(j\omega)| = \frac{K}{\omega_n^2 - \omega^2}; \quad 0 \leq \omega < \omega_n \] \hspace{1cm} (47)

\[ |G(j\omega)C(j\omega)| = \frac{K}{\omega^2 - \omega_n^2}; \quad \omega > \omega_n \]

and its phase is given by

\[ \theta(\omega) = -\pi - \omega \tau; \quad \text{for} \quad 0 \leq \omega < \omega_n \] \hspace{1cm} (48)

\[ \theta(\omega) = -\omega \tau; \quad \text{for} \quad \omega > \omega_n \]
The intersections of the polar plot with the negative real axis take place at the frequencies \( w_c \) where

\[
\begin{align*}
w_c &= \frac{2n\pi}{\tau}, \quad 0 \leq w_c < w_n \\
w_c &= \frac{(2n+1)\pi}{\tau}, \quad w_c > w_n
\end{align*}
\]  

(49)

In order to guarantee asymptotic stability of closed-loop system, the magnitude \(|G(jw)C(jw)|\) evaluated at \( w_c \) must be less than 1 so that the -1 point is not encircled. Therefore:

\[
\begin{align*}
k \frac{w_n^2 - (2n\pi)^2/\tau^2}{w_n^2 - k} &< 1, \text{ for } 0 \leq 2n\pi/\tau < w_n \\
k \frac{(2n+1)^2/\tau^2 - w_n^2}{(2n+1)^2/\tau^2 - k} &< 1, \text{ for } (2n+1)\pi/\tau > w_n
\end{align*}
\]

(50)

Combining the last two conditions we find the lower and upper bounds of the stability region for positive gain \( k > 0 \)

\[
\frac{2n\pi}{\sqrt{w_n^2 - k}} < \tau < \frac{(2n+1)\pi}{\sqrt{w_n^2 + k}}
\]

\[
0 < k \leq \frac{1 + 4n}{1 + 4n + 8n^2} w_n^2
\]

(51)

Following a similar analysis, we find the complete stability region for negative gains \( k < 0 \) as follows

\[
\frac{2n\pi}{\sqrt{w_n^2 - k}} > \tau > \frac{(2n-1)\pi}{\sqrt{w_n^2 + k}}
\]

\[
0 > k \geq \frac{1 + 4n}{1 + 4n + 8n^2} w_n^2
\]

(52)

Combining the stability regions for positive and negative gains, we obtain the graph shown in Figure 7 for \( w_n^2 = 1 \). As expected, the stabilizing gain regions decrease as the delay increases along the vertical axis. Using this simple graphical tool, we may choose the delay and gain values to guarantee that the closed-loop system is stable. In order to illustrate the process, assume that the open-loop plant is described by the transfer function:

\[
H(s) = \frac{1}{s^2 + 1}
\]

(53)
The output signal is subject to a delay that is randomly distributed between 0 and 3 seconds according to a uniform distribution. Let the initial conditions be $y(0) = y(0) = 0.1$. We show via a Matlab simulation, how the closed-loop system is stable with a choice of a small gain $k = 0.1$ despite the fact that the delay is randomly varying (see Figure 8). Note that this simply illustrates that the delayed feedback controller is somewhat robust to changes in the delay as may be encountered across a shared communication network. Our analytical results however do not guarantee closed-loop stability in general, since the rate of change in the delay needs to be accounted for. In order to decrease the magnitude of the gain in the controller, we choose a tenth of the estimated gain, being closer of the delay axis and not so at the edge of stability.

**Stabilizing systems with multiple delays**

Consider the two-delay feedback block used by Izmailov [45] to model a congestion control algorithm in a communications network [74]. Consider a deterministic model of a single connection between a source controlled by an access regulator and a distant node with a constant transmission capacity $\mu$ as given in [45] by:

$$
\begin{align*}
\dot{x}_1(t) &= x_2(t - \tau_1) - \mu \\
\dot{x}_2(t) &= -a(x_1(t - \tau_2) - \bar{X}) - b(x_1(t - \tau_2 - \tau) - \bar{X}) ,
\end{align*}
$$

where $x_1$ represents the buffer contents, $x_2$ the current input rate, and $\bar{X}$ the buffer target value. Using the error variable $y(t) = x_1(t) - \bar{X}$, (54) leads to the following second-order delay equations with two discrete and independent delays $\tau = \tau_1 + \tau_2$ and $r$:

$$
\ddot{y}(t) + ay(t - \tau) + by(t - \tau - r) = 0 ,
$$
where $\tau$ defines the round-trip time, and $r$ denotes the corresponding control-time interval. The corresponding characteristic equation is given by

$$s^2 + ae^{-s\tau} + be^{-s(\tau+r)} = 0.$$ 

The closed-loop system is composed of the feedback interconnection of a double integrator and a two-delay block, as described in the flow-based dynamics models of Section.

As suggested by [74], the particular form of the stability regions in the delay parameter space $\tau r$ proves useful in defining a wait-and-act strategy that provides stability robustness with respect to the round-trip time $\tau$ as depicted in Figure 6.

**Designing for Performance**

**Chaouki will include Tillbury’s work**

**Stabilizing Predictors**

Consider two systems $A$, and $B$, where $A$ acts as a predictor of system $B$. To obtain a feedback of the prediction error, the delayed output of $A$ is compared with the output of $B$ and used for feedback construction. In fact, given:

- System $B$

$$\dot{x} = Ax$$

$$y = Cx$$

- Predictor $A$

$$\dot{z}(t) = Az(t) + K(Cz(t - \tau) - y(t))$$

We then obtain the error dynamics:

$$\dot{e}(t) = Ae(t) + KCe(t - \tau)$$
where $e = z(t - \tau) - x$ is the prediction error. Because the predictor is stable if and only if the zero solution of (58) is asymptotically stable, the problem is reduced to a standard stabilization problem in the presence of delay.

**Stabilizing unstable periodic orbits**

**Chaouki & Silviu**

- Stabilizing unstable period (Pyragas stuff [82]): the controller does not change the periodic orbit under consideration, the delay is not introduced in the first place to stabilize the system but to keep the nominal solution unchanged.

  More precisely, the control law is given by $u(t) = k(x(t) - x(t - \tau))$, for some appropriate gain $k$, where $\tau$ represents the period of the system.

- Takens Theorem, OGY et al. strategies.

**Interconnections, Interference, and Complexity Issues**

Networks are a prevalent metaphor to describe interconnected systems that transfer materials, energy, and information. Delays are the simplest mathematical elements that model transport and propagation across a network. One interesting byproduct of the presence of delays in networks is the sensitivity of the overall system to changes in delay. In addition to their large-scale character, one of the distinguishing features of flow models in networks is the presence of multiple delays [57]. In biological networks for example, the large variations in the time-scales of delays may dramatically affect the stability of the underlying systems [22], [76]. Recall that one of the simplest stability types of time-delay systems is delay-independent stability, where a system remains stable for all positive delay values. This property was discussed in the literature.
since the 1950s (see, e.g. [25], [73], [37], [67] and the references therein). In the linear case, analysis methods ranging from semi-definite-programming (SDP) to matrix pencils and symbolic computation are available. In the sequel, we present specific results to illustrate the effects of multiple delays in a delay-independent stability situation.

**Multiple delays interference phenomena**

In the 1980s, reference [60] proposed an interesting biological model where the mutual coupling between the delays of individual dynamical systems was essential for stability of the overall system. As stated in the title of [60], *two delays may not destabilize the system although either delay can*. Such a stability phenomenon is known as *delay interference* [61], [67] and describes the *fragility* of the *delay-independent* stability property along a particular direction (ray) in the space of delays. The delay-independent stability property is shown not to be robust with respect to arbitrary small perturbations in the direction of a ray (see also [21], [55]). In order to illustrate the interference phenomenon, consider the simple example of a first-order system described by:

\[
\dot{x}(t) = -x(t) - x(t - \tau_1) - \frac{1}{2} x(t - \tau_2).
\] (59)

The rays for which delay-independent stability holds are represented by the axes of the delay-parameter space \( \tau_1 = \tau_2 = 0 \), and by the particular ray \( \tau_2 = 2\tau_1 \) as described next. Let us consider the case \( \tau_2 = 0 \), but \( \tau_1 = \tau \neq 0 \). It is easy to see that the characteristic equation of system (59) is given by \( s + 3/2 + e^{-s\tau} = 0 \). The delay-free system is stable (with characteristic root at \( -5/2 \)), and the characteristic equation has no roots on the imaginary axis regardless of the delay value \( \tau \). Indeed, the graph of \( H(j\omega) = -1/(j\omega + 3/2) \) is located inside unit circle, and thus \( |H(j\omega)| \neq 1 \), for all \( \omega > 0 \), and \( 1 - H(j\omega)e^{-j\omega\tau} \neq 0 \), for all \( \omega \in \mathbb{R} \), all \( \tau \in \mathbb{R}_+ \).
conclusion, by the continuity property mentioned in Section , the stability of the corresponding system is *delay-independent*. A similar property holds if \( \tau_1 = 0 \), and \( \tau_2 \neq 0 \).

**Remark 8:** The delay-independent stability property is a result of the *Tsypkin (frequency-sweeping) criterion* which guarantees the *stability robustness* of the interconnection of a stable open-loop SISO transfer connected via a unitary delayed feedback [104] (see also [73]). □

Consider now the case \( \tau_2 = 2\tau_1 = 2\tau \). The corresponding characteristic equation becomes

\[
s + 1 + e^{-s\tau} + 1/2e^{-2s\tau} = 0.
\]

Similarly to the previous case, we need to find the characteristic equation imaginary roots, that is \( j\omega + 1 + e^{-j\omega\tau} + 1/2e^{-2j\omega\tau} = 0 \). Equivalently, we search for the solutions \( z \in [-1, 1] : (z = \cos(\omega\tau)) \) to the equation \( 1/2z^2 + z + 1 = 0 \) which correspond to the real part of the characteristic equation on \( j\mathbb{R} \). Simple computations show that such an equation has no solutions, and in conclusion \( j\omega + 1 + e^{-j\omega\tau} + 1/2e^{-2j\omega\tau} \neq 0 \) for all \( \omega \in \mathbb{R} \), \( \tau \in \mathbb{R}_+ \). This shows once again the *delay-independent* stability of the closed-loop system, but for the ray \( \tau_2 = 2\tau_1 \) in the delay-parameter space.

Consider now a perturbation of the ray \( \tau_2 = 2\tau_1 \) into: \( \tau_2 = (2 + \varepsilon)\tau_1 \), for some small \( \varepsilon > 0 \). Since the system (59) is not asymptotically stable for *all positive delays* \( \tau_1 \) and \( \tau_2 \) (it actually has imaginary roots at \( j/2 \) for \( \tau_1 = 2\pi \), and \( \tau_2 = 3\pi \), a natural question to ask is whether such a ray is still stable, or whether it intersects a stability crossing curve. It can be shown that there indeed exists a sequence \( \{\varepsilon_n\}_{n \geq 1} \to 0 \), where \( \varepsilon_n = 1/(2(2n + 1)) \) such that the ray with \( \varepsilon = \varepsilon_n \) is not stable. More precisely, for some delay values \( \tau_1 > 2(2n + 1)\pi \), the system becomes unstable on the ray corresponding \( \varepsilon = \varepsilon_n \). Indeed, we obtain an imaginary root \( j/2 \) at \( \tau_1 = 2(2n + 1)\pi \) (see, e.g. [21], [67]).
Consider now the generalization [67] of (59) given by:
\begin{equation}
\dot{x}(t) = -ax(t) - x(t - \tau_1) - \frac{1}{2}x(t - \tau_2),
\end{equation}
where $a \in \mathbb{R}_+$ is a parameter. It is easy to see that this system is *delay-independent* stable for all $a \geq 3/2$. Indeed, it is easy to see that for all $a \geq 3/2$, \( |H_1(j\omega)| + |H_2(j\omega)| < 1 \), for all $\omega > 0$, where $H_1 = 1/(a + j\omega)$ and $H_2(j\omega) = 1/2(a + j\omega)$, and thus $1 + H_1(j\omega)e^{-j\omega\tau_1} + H_2(j\omega)e^{-j\omega\tau_2} \neq 0$, for all $\omega \in \mathbb{R}$, $\tau_1, \tau_2 \in \mathbb{R}_+$. Since the delay-free system is stable, we conclude that the system remains stable independent of the size of delays for all $\tau_1, \tau_2 \in \mathbb{R}_+$ (by extension of the Tsypkin’s criterion [104] to the multiple delay case, see also [25], [39], [37]).

Next consider the case of $a = 1$ and note that there are *only three stable rays*: the axis $O\tau_1 (\tau_2 = 0)$, $O\tau_2 (\tau_1 = 0)$, and the ray $\tau_2 = 2\tau_1$. In Figure 9, the complete stability/instability regions of (60) in the delay-parameters space are shown for $a = 1$ and $a = 1.3$. The solid lines correspond to delay values for which there are characteristic roots on the imaginary axis. The dashed lines indicate the stable rays. Note that small perturbations in the slope of the dashed lines lead to their intersecting the solid curves, as predicted by the delay-interference phenomenon. As $a \to 3/2$, we move towards delay-independent stability, as characterized by the increase of the number of stable rays. That number becomes arbitrarily large when we approach the *bifurcation* value $a = 3/2$, for which we completely recover the delay-independent stability property.

**Computational complexity in delay systems**

As shown above, the characterization of stability regions in the delay-parameter space becomes quite complicated due to the presence of stability rays even for small a number of delays. The *computational complexity* of determining the stability of time-delay systems becomes a concern as the size (dimension, number of delays) of the system grows. Let us first consider the
small delay case. Consider a square MIMO system with an open-loop transfer function matrix $P(s)$,

$$y(s) = P(s)u(s) \quad (61)$$

with the feedback control

$$u(s) = -y(s)$$

Assume that the closed-loop feedback control system is stable, and let us examine whether the system remains stable when the feedback is subject to a small delay as modeled by,

$$u(s) = -e^{-s\Lambda}y(s) \quad (62)$$

where

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$$

and $\lambda_i, i = 1, 2, \ldots, m$ are independent small delays of each input-output pair. For the SISO case, Willems showed in [106] that the system is robustly stable under small delay if the system is stable in the usual sense $|P(\infty)| < 1$. Meinsma, Fu and Iwasaki [65] showed that the necessary and sufficient condition for closed-loop stability in the MIMO case is that $(I + P)^{-1}$ stable, and $\mu(P(\infty)) < 1$, where $\mu(\cdot)$ is the structured singular value given by $\mu(M) =

\inf \left\{ \beta > 0 \left| \begin{array}{c} |I_n + M\text{diag}(\delta_1, \delta_2, \ldots, \delta_n)| \neq 0 \\ \text{for all } \delta_k \in \mathbb{C}, |\delta_k| \leq 1/\beta, \\
\quad k = 1, 2, \ldots, n \end{array} \right. \right\}$

As was shown by Toker [102] however, (see also [103]), finding the structured singular value is NP-hard.

For the not necessarily small time-delay case and a large number of delays, the computational complexity may be even larger. Consider again the feedback system consisting
of (61) and (62). We would like to find out whether the system is stable for arbitrary delays \( \lambda_k \geq 0, \ k = 1, 2, \ldots n \). A sufficient condition given in Section 3.4 of [37] is that

\[ \mu(P(j\omega)) < 1 \quad \text{for all } \omega \geq 0 \]  

(63)

The computation of (63) requires sweeping through all positive frequencies, and solving a \( \mu \) problem for each given frequency. Note that less conservative stability conditions are even more computationally expensive. On the other hand, most practical problems have only a small number of delays, greatly reducing the complexity of the problem.

**Smith predictor interference mechanism**

In light of the results presented so far, consider the standard Smith predictor control scheme [97], [44], [80] for a given SISO transfer function \( H(s) = H_0(s)e^{-s\tau} \), where \( H_0(s) \) is a *strictly proper* and stable. Assume that the delay \( \tau \) is not exactly known such that \( |\tau - \tau_0| \leq \delta \) where \( \tau_0 \) is a nominal delay value and \( \delta > 0 \) is a given bound. Let \( K_0(s) \) be a stabilizing controller for \( H_0(s) \). The Smith controller computed for the nominal delay case \( \tau = \tau_0 \) (no modeling errors or uncertainties) has the form:

\[ K(s) = \frac{K_0(s)}{K_0(s)H_0(s)(1 - e^{-s\tau_0})}. \]

Let \( H_{cl,0}(s) = K_0(s)H_0(s)/(1 + K_0(s)H_0(s)) \) the transfer of the delay-free closed-loop system.

For the uncertain delay case, the closed-loop transfer function becomes:

\[ H_{cl}(s) = \frac{H_{cl,0}(s)e^{-s\tau}}{1 - H_{cl,0}(s)e^{-s\tau_0}(1 - e^{-s(\tau - \tau_0)})}, \]  

(64)

The stability of (64) is determined by the zero locations of the meromorphic function: \( 1 - H_{cl,0}(s)e^{-s\tau_1} + H_{cl,0}(s)e^{-s\tau_2} \), where \( \tau_2 = \tau, \tau_1 = \tau_0 \). It is easy to see that if the closed-loop system is not *practically stable* (that is there exists some frequency \( \omega_0 > 0 \) such that \( |H_{cl,0}(j\omega_0)| > 1/2 \)).
then the ray $\tau_2 = \tau_1$ ($\tau = \tau_n$) is subject to *interference* as described in the previous paragraph. A more detailed characterization of the geometry of the stability regions (in the spirit of [38]) for the Smith predictor in the delay-parameter space may be found in [71].

**High-gains and delay sensitivity**

Wim and Silviu Need to discuss the case of proper systems subject to high-gain are extremely sensitive to infinitesimal delays in the loop (as seen in some examples of [54], see also [66]).

**Concluding remarks**

what we did not cover:

1) MIMO
2) random delays
3) discrete-time systems
4) LMI analysis
5) randomized algorithms

**Acknowledgements**

ANTONIS PAPACHRISTODOULO (Oxford, UK), DORON LEVY (Maryland, USA) (other names to be added)
References


[21] Datko, R.: Time delay perturbations and robust stability. in *Differential Equations, Dynam-


IEEE Infocom'95 (1995).


Silviu-Ilian Niculescu is
Chaouki Abdallah obtained his Ph.D. in electrical engineering from the Georgia Institute of Technology in 1988. He joined the Electrical and Computer Engineering Department at the University of New Mexico, where he is currently professor and department chair. He conducts research and teaches courses in the general area of systems theory with focus on control, communications, and computing systems. He is a co-founder of the ISTEC consortium, which currently includes more than 150 universities in the US, Spain, and Latin America. (corresponding author; Electrical and Computer Engineering, MSC01 1100, The University of New Mexico, Albuquerque NM 87131; chaouki@ece.unm.edu)

Wim Michiels is

Keqin Gu is
Figure 1. System with delays.
Figure 2. Controlling across a shared communications network. Control signals, measurements of the plant state, and external inputs travel from their source to their destination through the links of a communication network. The signals experience random delays.

Figure 3. Rightmost characteristic roots location of the closed-loop quasipolynomial $s + ke^{-s\tau}$ for various values of $\tau \in [0, 0.4)$ with $k = \sqrt{2}e^{3\pi/4}$.
Figure 4. The real parts of the rightmost zeros of (37) as a function of the delay parameter $\tau$, for small values of the gain parameter $\epsilon$.

Figure 5. The real parts of the rightmost zeros of (38) as a function of the delay parameter $\tau$, for $\epsilon = 1/640$. 
Figure 6. Stability regions in the delay parameter space \((\tau, r)\) for the Izmailov’s model (54)

Figure 7. Stability Regions for \(w_n^2 = 1\)
Figure 8. Simulation of Second Order System with Uniform Variable Delay, Positive Feedback
Figure 9. Stability/instability regions of the zero solution of (60) in the \((\tau_1, \tau_2)\)-space, for \(a = 1\) (top) and \(a = 1.3\) (bottom). For \(a = 1\), three stable rays (dashed lines) including the axis; for \(a = 1.3\), seven stable rays including the axis.