Well-posedness of Problems involving Time-varying Delays.

Erik I. Verriest

***Abstract*— Problems in defining state, state space, trajectory and stability for systems with time varying delays are expounded. First, it is shown that the information structure, providing the necessary side information, usually understood, but never made explicit, plays a significant role. It is then shown that when the delay derivative satisfies $\dot{\tau} < 1$, state space and hence stability are well defined. In the absence of this constraint, causality of the model may fail depending on the information structure. It is shown that embedding the delay in a larger fixed delay interval is also not always satisfactory, creating problems with consistency, minimality and well-posedness. Two causalizations are proposed which are useful within different information structures: lossless causalization for complete (omniscient) information and forgetful causalization if the delay is not predictable. Next, we consider the practically relevant problem (for networked control) when the delay is piecewise constant, and relate this to the recently introduced multi-mode multi-dimensional (M^3D) systems. Finally, we observe that in the discrete case, this M^3D model is viable.**

I. INTRODUCTION

The past decade has seen an explosive growth of papers involving LMI methods applied towards the stability analysis of time delay systems. Starting from a modest conference paper [1] (see also [2]), where the LMI condition was stated as a Riccati condition, ever increasing generality was sought to obtain less restrictive stability conditions. One of the first extensions was for time variant linear delay systems with time-varying delays. In [3] we investigated the stability of

$$\dot{x}(t) = A(t)x(t) + B(t)x(t - \tau(t))$$

A key issue was the condition $\frac{\mathrm{d}}{\mathrm{d}t}\tau < 1$ appearing in the proof of stability via the Lyapunov-Krasovskii theory. Choosing a functional

$$V(t, \phi) = \phi'(0)P(t)\phi(0) + \int_{t-\tau(t)}^{0} \phi'(s)Q(t, t + s)\phi(s)\,\mathrm{d}s,$$

we found that indeed $\dot{V}$, along trajectories of the system was equal to

$$x'(\dot{P} + A'P + PA + Q + PB\tilde{Q}^{-1}B'P)x - (1 - \dot{\tau})[\cdots]'[\cdots].$$

The second term is negative semidefinite if the condition $\dot{\tau} \leq 1$ holds, leaving the Riccati equality condition: $\exists P(\cdot) > 0, Q(\cdot) > 0, R(\cdot) > 0$ such that $\frac{\partial Q(t,s)}{\partial t} \leq 0$

$$\dot{P} + A'P + PA + Q + PB\tilde{Q}^{-1}B'P + R = 0$$

as a sufficient condition for stability.

E.I. Verriest is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA erik.verriest@ece.gatech.edu

Many authors have interpreted the condition $\dot{\tau} < 1$ as a major restriction and have "proved" stability conditions for the case were this condition is violated. The purpose of this paper is to point out that this is *not* merely a *technical* condition to make a proof work, but an essential and existential condition to give sense to the delay system itself! Indeed, we will show that casuality of the model is at stake when it is not imposed. Moreover, related problems of non-minimality of the realization occur when the condition is absent. Part of the confusion seems to originate in what exactly should be understood as *state space* for such a system as well as what the *information structure* should be. Moreover, not just stability, but even the question of *existence* of solutions to equations with time-varying point delays with L_p initial conditions does not have an adequate answer [4].

A. Information Structure

In order to clarify what we could mean by the information structure, let us consider the two equations

$$\dot{x}(t) = a(t)x(t) \quad (1)$$
$$\dot{x}(t) = u(t)x(t). \quad (2)$$

To the mathematician these two are of course the same, but the engineer may *interpret* these differently based on an ingrown conditioning. Indeed the first may conjure up a time variant but otherwise non-controlled system, the second one a bilinear system, controlled by the input $u(t)$. What is their difference? The engineer might argue that $a(\cdot)$ is some known function, predetermined by say a periodic or other phenomenon. In the second equation, $u(t)$ is at the whim of the controller, and therefore not predeterminable. Hence the difference in these two cases lies in the information structure. Whereas $a(\cdot)$ is entirely known for all t, $u(\tau)$ might only be known for $\tau \leq t$, or even just at $\tau = t$ if there is no memory implied of past inputs. In fact, the information structure is also involved in distinguishing *parameter varying systems* $\dot{x}(t) = p(t)x(t)$ (PV) from *time varying systems*[1]. In PV systems, the information at time t may consist only of bounds on present an future values of a, although in some cases $p(t)$ is known at time t. The information structure for a PV system should thus be something intermediate between the previous ones.

In fact, side information may even be needed at a more fundamental level. For a (classical) physicist the differential equations of mechanics are bi-directional. Neither the laws of mechanics nor mechanical observables give us a

[1]Many authors also seem unclear about this issue.

ISBN 978-963-311-370-7

direction of time, unless such a direction has been defined previously e.g., by reference to some irreversible process [5].

We conclude then that (1) or (2) cannot be complete descriptions of the system without the necessary side information (the information structure). Besides the notion of a *state* (see below), structural side information is needed as well. Of course, the notion of a *filtration* from stochastic processes guides us to the concept of this information structure. We recall some notions from measure theory first.

Knowledge of a function $f : \mathbb{R} \to \mathbb{R}$ entails that for every Borel set B of $\mathbb{R}$ the set $f^{-1}(B) = \{s | f(s) \in B\}$ is measurable in some sense. This means that one should have defined a measurable space $(\mathbb{R}, \mathcal{G})$ such that $f^{-1}(B) \in \mathcal{G}$. This is the usual concept of a measurable function. The need for the algebra axioms stems from the (Aristotelean) logic structure in dealing with more fundamental yes/no questions such as "Is $f(t)$ in B?" The countable union axiom for the σ-algebra construction is thrown in for mathematical reasons (limits), but is not postulated by the logic [6].

So far there is no dynamic, no flow of time implied. In order to match the information with a temporal structure, we first need a time set $\mathbb{T}$. In continuous time, we let this be $\mathbb{T} = \mathbb{R}$ or $\mathbb{R}_+$, but for discrete time systems, this would be $\mathbb{Z}$ or $\mathbb{N}$. We further need to define a linear order: Only two can be chosen: the forward or the backward one. Choosing the forward order, we associate a family of sigma algebras $\mathcal{G}_{(-\infty,t]}$ with the property that for $s < t$, $\mathcal{G}_{(-\infty,s]} \subset \mathcal{G}_{(-\infty,t]}$. For instance, we may choose the open intervals in $(-\infty, t]$ as the generating family for the Borel sets (denoted by $\mathcal{B}$) in $(-\infty, t]$. Complete knowledge (w.r.t. $\mathcal{G}$) of $u(\cdot)$ up to time t is then interpreted as $u : ((-\infty, t], \mathcal{G}_{(-\infty,t]}) \to (\mathbb{R}, \mathcal{B})$ being a measurable function. It is customary to denote this by $u(t) \in \mathcal{G}_{(-\infty,t]}$. If, on the other hand, the function $a(\cdot)$ is prespecified, in the sense that the entire map $a : \mathbb{R} \to \mathbb{R}$ is known at all times t, then we denote $a(\cdot) \in \mathcal{G}_{(-\infty,t]}$ for all t. i.e., for any Borel set B of he real line $a^{-1}(B) = \{s | a(s) \in B\} \in \mathcal{G}_{(-\infty,t]}$. This is formalized by first extending the σ-algebra by one generator

$$\overline{\mathcal{G}}^t = \sigma\{\mathcal{G}_{(-\infty,t]} \cup (t, \infty]\} = \mathcal{G}_{(-\infty,t]} \cup ([t, \infty) \cup \mathbb{R},$$

take the product σ-algebra $\overline{\mathcal{G}}^t \otimes \mathcal{B} = \mathcal{G}^t$, and define

$$\overline{a} : (\mathbb{T} \times \mathbb{R}, \overline{\mathcal{G}}^t \otimes \mathcal{B}) \to (\mathbb{R}, \mathcal{B}) : \overline{a}(t, s) = a(s), \ \forall t$$

since then for any Borel set B

$$\begin{aligned} \overline{a}^{-1}(B) &= \{(t, s) | \overline{a}(t, s) \in B\} \\ &= \{(t, s) | a(s) \in B\} \\ &= \mathbb{T} \times a^{-1}(B) \in \mathcal{G}^t \end{aligned}$$

since $a^{-1}(B) \in \mathcal{B}$, and $\mathbb{T} \in \overline{\mathcal{G}}^t$.

It is more intricate to model PV systems. For instance if $|p(s)| \leq 1$, for $s > t$, the filtration in the image space must be limited to the σ-algebra generated by the interval $[-1, 1]$.

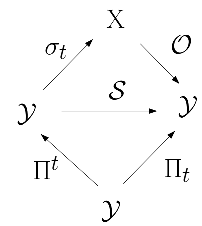

Fig. 1. State Space X by splitting

B. State Space

Here we shall be interested in linear evolution systems, meaning that they are not influenced by external inputs. In finite dimensions. such a system is modeled by $\dot{x}(t) = A(t)x(t)$. We shall assume that the signal $y(t) = c(t)x(t)$ is available, and A and c are known at all times. It is well known that the solution $x(t)$ exists, and is unique if $x(\cdot)$ is specified at some time. Let $\mathcal{Y}$ be the space of all possible outputs generated by such a system. Define the projection operators Π^t and Π_t such that for any arbitrary function ϕ

$$(\Pi^t \phi)(s) = \begin{cases} \phi(s) & s \leq t \\ 0 & s > t \end{cases}$$

$$(\Pi_t \phi)(s) = \begin{cases} \phi(s) & s \geq t \\ 0 & s < t \end{cases}$$

By choosing the forward direction of time flow, we understand that we can associate a map from past to future

$$\mathcal{S} : \Pi^t \mathcal{Y} \to \Pi_t \mathcal{Y} : \Pi^t y \to \Pi_t y$$

by splitting y at time t. This is done as follows: Since the sets of orthogonal projectors form resolutions of the identity, $\mathcal{P}^* = \{\Pi^t | t \in \mathbb{T}\}$ and $\mathcal{P}_* = \{\Pi_t | t \in \mathbb{T}\}$, the pairs $(\mathcal{Y}, \mathcal{P}^*)$ and $(\mathcal{Y}, \mathcal{P}_*)$ are Hilbert resolution spaces [7]. An operator $\mathbf{A}$ on a resolution space $(\mathcal{Y}, \mathcal{P}^*)$ is called causal if for all $\Pi^t \in \mathcal{P}^*$ and any $x, y \in \mathcal{Y}$, the equality $\Pi^t x = \Pi^t y$ implies that $\Pi^t \mathbf{A} x = \Pi^t \mathbf{A} y$. It is easily shown that this definition is equivalent with the more natural operator characterization: $\Pi^t \mathbf{A} \Pi_t = 0$, where $\mathbf{A}$ is the system operator. This may be paraphrased as "No past effect caused by a future event."

We note that the past information may be summarized into the state $x(t) \in \mathbb{R}^n$ by the *observer* map at t: $\sigma_t(y) = x(t)$. The future is given by the output map $\mathcal{O}_t : X = \mathbb{R}^n \to C^n([t, \infty), \mathbb{R}^n) : x \to \{c e^{A(s-t)} x | s \geq t\}$ (in the time invariant case). Then clearly, the commutative diagram in (1) holds for $X = \mathbb{R}^n$. A state space is the structure which provides the interface, X_t (at t), in the factorization of the map $\Pi^t f \to \Pi_t f$. For more detail, see [7], [8], [9].

Now we merge the concepts of information structure and state space. If at time t we have $x \in X_t$ and $(A(\cdot), c(\cdot))$ is known at t, then $\dot{x}(t)$ is known (as forward and as backward derivative), and that is all. We need $(A(t), c(t)) \in \mathcal{G}^t$, for all t in order to integrate forward, and determine $x(s)$ for $s \geq t$, or $(A(t), c(t)) \in \mathcal{G}_t$ for all t, in order to integrate backward (retrodict). We say that $A(t)$ and $c(t)$ need to be adapted to the information flow.

There is one caveat. The construction is local: a state space is defined at *each* instant of time. Unlike the finite dimensional LTI case where $X = \mathbb{R}^n$ at each t, no clear relationship may exist among state spaces defined at different times. This implies that state trajectories are not functions since at each time they take values in different spaces. If one wants to speak of trajectories evolving in the state space, this structure should necessarily be static. State trajectories can be defined as Hilbert space valued functions for which continuity and bounded variation can be shown [10].

II. Time delay systems

Consider a functional differential equation in the form

$$\dot{x}(t) = \int_{-\tau}^{0} \mathrm{d}\eta(s)x(t+s) \tag{3}$$

where τ is fixed (known), and η is a measure. Let $\mathcal{L}(\phi) = \int_{-\tau}^{0} d\eta(s)\phi(s)$, and note that it is a bounded operator on $C^n([-\tau,0],\mathbb{R})$. It follows then that for each $(\phi, x_0) \in X = C^n([-\tau,0],\mathbb{R}) \times \mathbb{R}^n$ there corresponds a unique solution to (3), defined for $t > 0$ [11]. Introduce the evolution operator

$$S(t) : X \to X : (\phi, x_0) \to (x_t, x(t)),$$

then the set $\{S_t\}_{t\geq 0}$ forms a C_0-semigroup under composition. Its infinitesimal generator $\mathcal{A}$:

i) has domain $\mathcal{D}(\mathcal{A}) = \{(\phi, x_0)\,|\,\dot{\phi} \in L_2^n(-\tau,0), \phi(0) = x_0, \phi \text{ absolutely continuous}\}$

ii) If $(\phi, x_0) \in \mathcal{DA}$, then $\mathcal{A}(\phi, x_0) = (\dot{\phi}, \mathcal{L}(\phi))$.

iii) There exist $c \in \mathbb{R}$ s. t. Spec(A) lies to the left of c, and for each $\epsilon > 0$, there exists M_ϵ such that

$$\|S(t)(\phi, x_0)\| \leq M_\epsilon \mathrm{e}^{(c+\epsilon)t}\|(\phi, x_0)\|$$

iv) The spectrum of $\mathcal{A}$ is a purely point spectrum, and $\lambda \in \mathrm{Spec}\,(\mathcal{A})$ iff $\det(sI - \int_{\tau}^{0} \mathrm{d}\eta(\theta)\mathrm{e}^{s\theta}) = 0$.

Thus it is clear that here the space X is well defined and its structure is independent of time, just as it was in the finite dimensional case. An equation of the form $\dot{x}(t) = Ax(t) + Bx(t-\tau)$ may be represented in the above form.

A. Initial Value Problem

Whereas the equation $\dot{x}(t) = Ax(t)$ is integrable forward as well as back ward in time, it may not be possible to integrate the delay equation backward in time without encountering singularities. Indeed with $x(t)$ and $x(t-\tau)$ in $\mathcal{G}_{(0,t]}$, it follows that the differential $\mathrm{d}_+x(t) \in \mathcal{G}_{(0,t]}$. With knowledge of x in the interval $(t, t+\tau)$, the relation $x(t) = b^{-1}(\dot{x}_-(t+\tau) - ax(t+\tau))$ determines the backwards solution. However, this is not possible if a and b are only known at time $t+\tau$ instead of being adapted to $\mathcal{G}_{([t,\infty)}$.

B. Well-posed Problem

Here we focus on the forward time direction, and hence only need a semigroup property. Hadamard, defined a well-posed initial value problem as a problem for which

i) a solution exists

ii) The solution is unique

iii) The solution depends continuously on the data in a reasonable topology.

Clearly the cases discussed above satisfy these properties.

III. Time-varying delay

Consider now $\dot{x} = f(x(t-\tau(t))$. If one interprets $\tau(t)$ to mean that the segment of $x(\cdot)$ over the interval $(t-\tau(t), t)$ is to be considered as the state in a general delay system, what precisely is then the *state space* for such a system? Obviously, the first requirement is that $\tau(t)$ belongs to the information structure for all t. But as discussed for ODE's that only guarantees that the forward derivative $\dot{x}(t)$ can be determined at t. Hence let us first assume for simplicity that $\tau(\cdot)$ is known at all t. Defining the state space to be $C((-\tau(t),0],\mathbb{R}^n)$, or any other function space over $(-\tau(\cdot),0]$ is not be a good choice, as the state space itself should be a fixed time-independent structure. One way around this is to find a time transformation $\lambda : \mathbb{T} \to \mathbb{T} : t \to \lambda(t)$, which deforms the requisite delay interval to make it constant: i.e., require $\lambda(t) - \lambda(\tau(t))$ = constant. Of course, one wants such a transformation to be invertible. In a now standard notation we shall write the above equation in a functional form $\dot{x} = f(x_t)$, where $x_t = \{x(t+\theta)| -\tau(t) < \theta < 0\}$.

A. Construction of the State Space

Assuming that $\tau(\cdot) \in \mathcal{G}^t$, a global time transformation $\lambda : \mathbb{T} \to \mathbb{T} : t \to \lambda(t)$, which deforms the state interval to make it constant, will take us out of the impasse: i.e., require $\lambda(t) - \lambda(\tau(t))$ = constant. Of course, we want such a transformation to be invertible (order preserving).

Example 1: Consider for $t > 0$, the delay system

$$\dot{x}(t) = A(t)x(t) + B(t)x(t-\tau(t)),$$

where $\tau(t) = 1 + \gamma t$, with $0 < \gamma < 1$. Let the initial data be $x(\theta) = \phi(\theta)$, with $-1 < \theta \leq 0$. Letting $\lambda(t) = k\log(t + \frac{1}{\gamma})$, for $k = -1/\log(1-\gamma) > 0$, yields

$$\lambda(t) - \lambda(t - \tau(t)) = 1.$$

With the new independent variable, λ, and letting $\overline{x}(\lambda) = \overline{x}(\lambda(t)) = x(t)$, the system is represented by

$$\frac{\mathrm{d}\overline{x}(\lambda)}{\mathrm{d}\lambda} = \overline{A}(\lambda)\overline{x}(\lambda) + \overline{B}(\lambda)\overline{x}(\lambda - 1).$$

The parameters are $\overline{A}(\lambda) = \frac{\mathrm{d}\mathcal{L}(\lambda)}{\mathrm{d}\lambda}A(\mathcal{L}(\lambda))$, and $\overline{B}(\lambda) = \frac{\mathrm{d}\mathcal{L}(\lambda)}{\mathrm{d}\lambda}B(\mathcal{L}(\lambda))$, the function $\mathcal{L}$ being the inverse of λ, i.e., $\lambda(t) = s \Leftrightarrow \mathcal{L}(s) = t$. In the above example, $\mathcal{L}(\lambda) = \exp(\lambda/k) - \frac{1}{\gamma}$. If A and B were originally constant, the delay equation for $\overline{x}$ is time varying, but with *constant* delay. In Figure 2, the horizontal axis is the time t, while the vertical axis gives θ in the interval $(-\tau(t), 0]$. The hue encodes the value $x_t(\theta)$ as function of t and θ in the original system. The delay character is easily detected by the uniform shifts of the

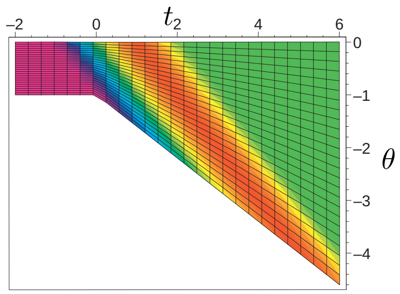

Fig. 2. Time varying delay

shaded bands (which follow the characteristics). Indeed, the infinitesimal update of the state is given by

$$x_{t+\mathrm{d}t}(0) = x_t(0) + [Ax_t(0) + Bx_t(-\tau(t)]\,\mathrm{d}t \quad (4)$$
$$x_{t+\mathrm{d}t}(\theta) = x_t(\theta + \mathrm{d}t), \quad \text{for } \theta \neq 0. \quad (5)$$

It is then readily seen that for $x_t(\theta)$ in the open domain $t_0 < t < T$ and bounded by $\theta = 0$ and $\theta = -\tau(t)$, the following PDE holds:

$$\frac{\partial x_t(\theta)}{\partial t} - \frac{\partial x_t(\theta)}{\partial \theta} = 0,$$

with boundary condion: $\frac{\partial x_t(0)}{\partial t} = Ax_t(0)+Bx_t(-\tau(t))$. The global time transformation $t \longrightarrow \lambda(t)$, reduces the problem to one with a fixed 'delay', shown in figure 3, with independent variable λ. The 'state' at λ is the function segment in the interval $-1 < \overline{\theta} < 0$. Its solution is easily obtained with the method of steps [11]. We have again the PDE

$$\frac{\partial \overline{x}_\lambda(\overline{\theta})}{\partial \lambda} - \frac{\partial \overline{x}_\lambda(\overline{\theta})}{\partial \overline{\theta}} = 0$$

in the domain $(\lambda_0, \lambda_f) \times (-1, 0)$, with boundary condition $\frac{\partial x_\lambda(0)}{\partial t} = \overline{A}(\lambda)x_\lambda(0) + \overline{B}(\lambda)x_\lambda(-1)$. Note that in this example $\overline{x}(\lambda, \overline{\theta})$ corresponds to $x(t, \theta)$ with $t = \left(\frac{\mathrm{e}^\lambda}{k} - \frac{1}{\gamma}\right)$, and $\theta = \left((t + \frac{1}{\gamma})[\mathrm{e}^{\overline{\theta}/k} - 1]\right)$.
More generally, the state space of a time delay system with time varying delay satisfying $\dot{\tau} < 1$ is homeomorphic to $C((-1, 0], \mathbb{R}^n)$. We remark that the transformation is itself obtained as the solution of a functional equation. Even with the delay function not specified, (but $\tau(t)$ adapted), its computability is guaranteed. Obviously, when $t - \tau(t)$ decreases (which happens when $\dot{\tau} > 1$), $\mathcal{L}$ does not exist and *this construction fails*.

There is another possibility way to produce a system with $C((-1, 0], \mathbb{R})$ as state space. In order to reduce the problem to one with a fixed delay, the 'state' x_t is squeezed (by uniform stretching or compression) to an interval of length 1. Thus $-1 < \overline{\theta} < 0$. This is displayed in Figure 4. The delay character of the equation is now not preserved as evidenced by the warped color bands. In fact, since here

$$\overline{x}_t(\overline{\theta}) = x(t + \tau(t)\overline{\theta}),$$

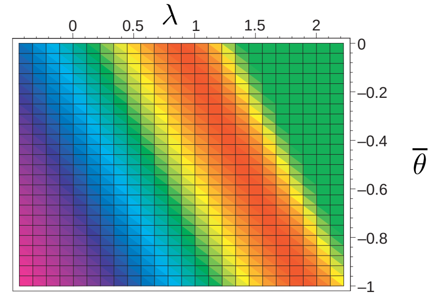

Fig. 3. Transformation to fixed delay

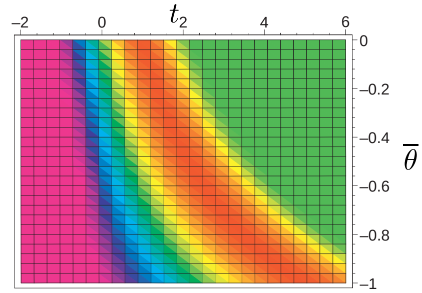

Fig. 4. Time transformation

we get

$$\begin{aligned} x_{t+\mathrm{d}t}(\overline{\theta}) &= x(t + \mathrm{d}t + \tau(t + \mathrm{d}t)\overline{\theta}) \\ &= x(t + \tau(t)\overline{\theta} + [1 + \dot{\tau}(t)\overline{\theta}]\mathrm{d}t) \\ &= \overline{x}_t\left(\overline{\theta} + \frac{1 + \dot{\tau}(t)\overline{\theta}}{\tau(t)}\,\mathrm{d}t\right). \end{aligned}$$

This leads to the PDE

$$\frac{\partial \overline{x}_t}{\partial t} - \left(\frac{1 + \dot{\tau}(t)\overline{\theta}}{\tau(t)}\right)\frac{\partial \overline{x}_t}{\partial \overline{\theta}} = 0,$$

and this advection equation no longer corresponds to a delay equation.

The above becomes problematic if $\dot{\tau}$ increases beyond 1. Then for some $-1 \leq \overline{\theta} < 0$, the second term in the advection equation vanishes and the characteristic is parallel to the t-axis. Beyond this, the slope of the characteristic becomes positive. However, boundary information at $\overline{\theta}$ is not available, and the solution to the PDE is undetermined. The problem is no longer well-posed in the sense of Hadamard. In fact, [4] gives the example $\dot{x}(t) = x(t - \tau(t))$ with $\dot{\tau} \leq 1$, for which the solution does not depend continuously on the initial time t_0.

To illustrate the problem, let's look at the scalar linear delay system $\dot{x}(t) = Ax(t - \tau(t))$, where for $t > t_0$, $\dot{\tau} > 1$. It follows that the left endpoint of the delay interval, $t_b = t - \tau(t)$ travels backwards as t increases. If $\eta(t)$ is *any arbitrary* $\mathbb{R}^n$-valued function, then the system has a forward component, $\dot{x}(t) = A\eta(t)$ and a backward component $x(t) \in \eta(\sigma(t)) + \mathcal{N}(A)$ for $t < t_0 - \tau(t_0)$. Here $t = \sigma(t_b)$ is the inverse function of $t_b = t - \tau(t)$, and $\mathcal{N}$ denotes the

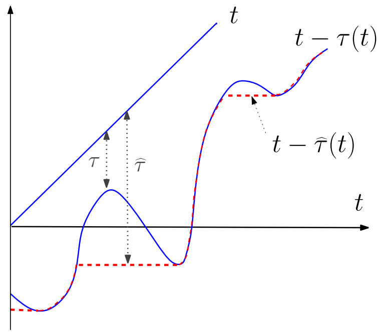

Fig. 5. Lossless Causalization

nullspace. The problem with this is that the future predicts the past, violating causality. Conversely, at time t, only x in the interval $(t - \tau(t), t)$ needs to be known to compute how $x(\cdot)$ will evolve immediately after that time. But one cannot extend this much further, let alone indefinitely, since an infinitesimal instant δ later, this past knowledge no longer suffices to determine its future behavior. The problem is that x in the time interval $(t - \tau(t), t)$ no longer is a sufficient statistic to determine future behavior uniquely, and therefore fails to be a state in the sense conceived by Nerode [12].

We conclude that conditions for stability in the time-varying case appearing in the literature should be carefully scrutinized. Similarly, when applied to systems with delays that jump between constant values (e.g. hybrid delay systems), the transition problem (from smaller to larger delay) is not well posed, as the delay derivative gets unbounded. The remaining part of this paper shows how such problems should be recast in a causal way.

B. Lossless causalization

One obvious fix is to embed the delay time $\tau(t)$ in a bigger delay, i.e., $\hat{\tau}(t)$, such that the graph of $t - \hat{\tau}(t)$ is nondecreasing, but in some minimal sense. This can be done by retaining $\hat{\tau}(t) = \tau(t)$ if $\dot{\tau}(t) < 1$, and truncating $t - \tau(t)$ otherwise. See Figure 5. We refer to this embedding as *lossless causalization*.

This requires taking

$$t - \hat{\tau}(t) = \inf_{s > t}(s - \tau(s)). \tag{6}$$

However, this induces another problem. At each time t the *future* behavior of the delay, $\tau(s), s \geq t$ must be known. This only *moves the noncausality* to another part of the problem! Surely if τ is only adapted, as when the time variation stems from state dependence, $\tau(x(t))$, such information is not available. Of course one could be super cautious, and sweep all problems away by letting $\tau = \infty$ from the beginning and cast all systems in a nonparsimonious way as *Volterra systems with infinite aftereffect.*

$$\dot{x}(t) = f\left(t, \int_{-\infty}^{t} K(t, \theta, x(\theta))\,\mathrm{d}\theta\right).$$

C. Unbounded delay

Also, this "fix" is bound to incur some problems. In [13] the stability of a nonlinear functional differential equation with time varying delay is considered. The notion of state space is not mentioned, but the authors assumed that the initial data was given for a segment of length $\sup_{u>0} \tau(u)$, thus in principle allowing also an infinite delay. However, the theorems assume that either $\tau(t)$ remains bounded or $\tau(t) \to \tau_0 < \infty$, hence the effective state space may be limited to a suitable function space defined over the interval $(-\tau_0 - \epsilon, 0]$, for arbitrary $\epsilon > 0$. Stability is interpreted in the sense that $|x(t)| \leq M\mathrm{e}^{-}\delta t$ for $t \geq T$. Note that this implies that the norm in the function space converges to 0. In [14], the authors "extend" the results to an infinite distributed delay, however they limit the discussion to kernels with rational (Laplace) transforms so that effectively the delay equation may be transformed to a finite set of ODE's. The equivalent requisite initial data is then obtained for the original initial data $\{\phi(s)|s < 0\}$ by a finite number of linear functionals of ϕ. A simple version of the problem is illustrated by the autonomous system

$$\dot{x}(t) = \int_0^\infty k(s)x(t-s)\,\mathrm{d}s, \quad x_0 = \{\phi(s)|s < 0\},$$

where $\dot{k} = -\alpha k$. Letting $u(t)$ be the right hand side, the equivalent dynamics is given by

$$\begin{aligned} \dot{u} &= k(0)x - \alpha u \\ \dot{x} &= u. \end{aligned}$$

The initial conditions are $x(0) = \phi(0)$ and $u(0) = \int_0^\infty k(s)\phi(-s)\,\mathrm{d}s$. Hence the original problem is only infinite dimensional in disguise, and it surely lacks the observability property, as ϕ cannot uniquely be determined from the pair (x, u).

Various criteria for the existence and/or absence of oscillatory solutions to time varying delay systems have been obtained, even allowing unbounded delay. However, the notion of state does not seem to be necessary top tackle the problem [15], [16]. What one should choose as state space when the delay is unbounded is indeed far from trivial [17], [18].

Example 2: The system $\dot{x}(t) = -ax(t) + x(t - \tau(t))$ has unbounded delay for $\tau(t) = 2t$. With initial data $\phi(t) = \mathrm{e}^{\sigma}t$ for $t < 0$, the 'solution' to the IVP is $x(t) = (\mathrm{e}^t + 1)\mathrm{e}^{-2t}$, and hence converges pointwise to zero. If however we take $L_2((-\infty, 0])$ as state space, then $\|x_t\|^2 \to \frac{23}{12}$, so that the solution does not converge to the null solution. On the other hand, the weighted L_2 norm with exponential weight $\|x_t\|^2 = \int_{-\infty}^{t} \mathrm{e}^{s-t}x(s)^2\,\mathrm{d}s$, does converge to zero.

Example 3: Consider the system $\dot{x}(t) = Ax(t - \tau(t))$, where $\tau(t) = 0$ for $t < 0$ and $\tau(t) = 2t$ for $t > 0$ (thus $\tau(\cdot) \in \mathcal{G}^t$ for all t.) At time $t = 0$, let it have the state x_0. This allows us to retrodict its state for $t < 0$: $x(t) = \exp(At)x_0$. Integrating forward for $t > 0$ the equation $\dot{x}(t) = Ax(-t)$

yields then $x(t) = (2I - \exp(-At))x_0$.

Example 4: Consider the system as in example 3, assume $A \neq 0$, and $\tau(t) = 1$ for $t < 0$, and $\tau(t) = 1 + 2t$ for $t > 0$. At time $t = 0$ the "state" is the function segment ϕ defined over $(-1, 0)$. Let's even assume that ϕ is infinitely many times differentiable there, and let also $x_0 = \phi(0)$. It is possible to extend the initial data backwards, But a jump may occur at $t = -1$ and successively higher singularities at the decreasing integers (< -1) [11, p.51]. For $-k - 1 < t < k$, $x = \mathbf{B}^k\phi$, where $\mathbf{B}$ is the operator $\frac{1}{A}\mathbf{D}\mathbf{T}_1$, where $\mathbf{D}$ is differentiation, and $\mathbf{T}_a$ is the shift operator defined by $(\mathbf{T}_a x)(t) = x(t + a)$. If $a > 0$ this is a left shift. Hence, in principle retrodiction is possible, and it is easily seen that this will hold also in the more general case if $\dot{\tau} < 1$ for $t < 0$. Consequently, under this last condition, ϕ over $(-\tau(0), 0]$ suffices to determine the entire solution to the FDE, but not in a causal way.

Interpreting the problem as a *Cauchy initial value problem* (IVP), one assumes no model for the dynamics for $t < 0$. Hence the system in Example 3,

$$\dot{x}(t) = Ax(-t); \quad x(0) = x_0,$$

exhibits a continuum of continuous solutions. Indeed, if $\psi \in C((-\infty, 0), \mathbb{R}^n)$, with $\psi(0) = x_0$, then $x(t) = \psi(t)$ for $t < 0$ and $x(t) = x_0 + A\int_0^t \psi(-s)\,\mathrm{d}s$ for $t > 0$ is a solution to the IVP passing through $(0, x_0)$.
The Cauchy problem associated with Example 4 is: $\dot{x}(t) = Ax(-1 - t)$ with $x(\theta) = \phi(\theta)$ for $-1 < \theta \leq 0$. All its continuous solutions of are $x = \phi$ in $(-1, 0)$, $x = \psi$ in $(-\infty, -1)$ and $x(t) = \phi(0) + A\int_0^t \psi(-1 - s)\,\mathrm{d}s$ for $t > 0$, where $\psi \in C^\infty((-\infty, -1], \mathbb{R}^n)$ with $\psi(-1) = \phi(1)$, Note that in this case only the boundary $\phi(0)$ and $\phi(-1)$ of the data is required, but the solution is otherwise also degenerate.
With lossless causalization, the state space for these examples is the set of generalized functions defined over the negative reals. If we're only interest in a solution up to some time $T > 0$, then the state space intervals $(-T, 0)$ and $(-T-1, 0)$ respectively suffice for the examples 3 and 4.

Even when one finds this system causalization acceptable, another problem crops up: The *nonminimality* of such a representation.
The details are expounded in [19]. Here we just reconsider example 3. Suppose we're interested in this system over the interval $(-T, T)$. For $0 < t < T$, we get $\widehat{\tau}(t) = t - \inf_{t<s<T}(s - \tau(s)) = t - \inf_{t<s<T}(-s) = t + T$. This means that one has to remember all of $\{x(s), -T < s < 0\}$ when $t > 0$. But this *infinite dimensional* system obtained for $t > 0$ by causalization is highly redundant. Indeed, we found the solution via the system of equations $\dot{x}(t) = Ax(t)$, for $t < 0$ and $\dot{x}(t) = Ax(-t)$, for $t > 0$. But we can decouple the past from the future in these two. Let $y(t)$ be defined, for *positive* t only, as $y(t) = x(-t)$. Then we obtain a *second order* system:

$$\begin{aligned} \dot{y}(t) &= -Ay(t) \\ \dot{x}(t) &= Ay(t) \end{aligned} \tag{7}$$

defined for $t > 0$, with initial condition $y(0) = x(0)$. This system captures the *same* behavior as the infinite dimensional delay system. Consequently, the delay system cannot be a minimal realization for the behavior it exhibits. If a linear system is not minimal, then it cannot be jointly reachable and observable. Without an input, reachable is a mote point, but it is shown in [19] that observability fails.

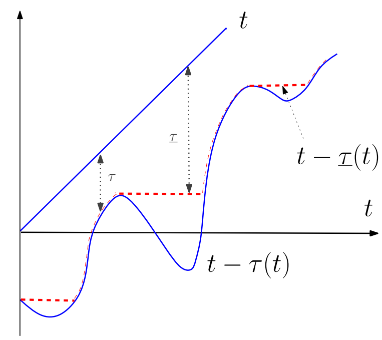

Fig. 6. Forgetful Causalization

D. Forgetful Causalization

Another approach is possible, especially if $\tau(t)$ is only adapted to $\mathcal{G}^t$. By strictly adhering to the fact that at time t, only data in the interval $(t - \tau(t), t]$ is remembered, all other history being irretrievably lost, one must *choose* values for $x(\cdot)$ that are missing. If the system operates near its equilibrium, as expected if it is stable, it seems then reasonable to substitute this equilibrium value for this missing information. This means that at time $t > 0$, one only effectively remembers the data in $(t - \underline{\tau}(t), t)$, where now

$$\underline{\tau}(t) = \sup_{s \leq t}(s - \tau(s)). \tag{8}$$

We'll call this the *forgetful causalization*. See Figure 6. Obviously forgetful causalization does not yield the same solution as lossless causalization, but it has the advantage of not requiring future information. Hence a truly causal structure results. Or to state our principle more bluntly: "One cannot remember what one has already forgotten." (This point seems to be missed by many authors).

IV. Delay Systems as M^3D

In [20] we considered a system with control switched between an open loop steering and an autonomous delayed feedback. This results in a hybrid system with two modes:

$$\begin{aligned} \text{non}-\text{delay}(ND): \quad & \dot{x}(t) = A_0 x(t) + bu(t), \\ \text{delay}(D): \quad & \dot{x}(t) = Ax(t) + Bx(t - \tau). \end{aligned}$$

In fact, the above models two instances: If $A_0 = A$, and $B = bk$, a delayed feedback is modeled, whereas the non-delay mode is an open loop control. The other instance, for $A_0 = A + B$, gives a sustained feedback operation but with intermittent delay (an external input term may then also be included in this delay mode). This can be cast as a particular instance of a system with a *time-varying* delay,

$$\dot{x}(t) = Ax(t) + Bx(t - \tau(t)) + bu(t) \tag{9}$$

for the special case where the delay $\tau(t)$ only assumes values τ and 0. More generally, assume $\tau : \mathbb{T} \to \{t_i\}_{i \in \mathcal{I} \subset \mathbb{N}}$. Assuming the adapted model, explained in [20], a transition with $\tau_{i+1} < \tau_i$ is allowed, but a transition with $\tau_{i+1} > \tau_i$, requires a reset of the initial state to retain causality This corresponds to the *forgetful causalization* explained in the previous section.

Thus if at time t the delay switches from the non-delay (ND) to the delay (D) mode, one needs to *choose* an initialization for $x(\cdot)$ in the interval $(t - \tau, t)$. If the system operates near its equilibrium, as we expect if it is stable, we shall load a zero (equilibrium) value for this missing information. For the transition from the delay (D) to the non-delay (ND) mode, we simply load the last state value in the new mode. With the notation x_t denoting the state of the delay mode, i.e., $x_t = \{x(t + \theta) | -\tau < \theta < 0\}$, we thus have at the switching instant t:

$$\begin{array}{ll} ND \to D: & x_{t_+}(\theta) = 0,\ \theta < 0,\ \text{and } x_{t_+}(0) = x(t_-) \\ D \to ND: & x(t_+) = x_{t_-}(0), \end{array}$$

Instantaneous switching with forgetful causalization does not loose the delay character of the system. The delay is built up in the following interval of length τ, (unless the mode is switched again before that), after which the system operates as a full delay system. This is precisely a multi-mode multi-dimensional (M^3D) system as described in [20], [22].

A. Multi-mode multi-dimensional system

A state space for M^3D systems was motivated as a bundle over the mode set, Ξ, where each fiber is a finite dimensional vector space $\mathbb{R}^{n_a}$ at $a \in \Xi$. The base space contains the mode information, and the fiber the partial state x of dimension n_a in mode $a \in \Xi$. In addition there is a switching structure

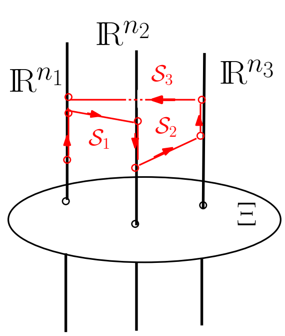

Fig. 7. The state bundle for an M^3D system

at the mode transitions. For the mode switching (say at time τ_i) from mode i to mode $i+1$ we have

$$x_{i+1}(k) = S_{i+1,i}x_i(k) + v_i.$$

with $S_{i+1,i} \in \mathbb{R}^{n_{i+1} \times n_i}$, and $v_i \in \mathbb{R}^{n_{i+1}}$ and the *pseudo-continuity* constraints. The latter express that the state should not change if the system takes a complete cycle through all its modes in an infinitesimally short time, starting with the lowest dimension. This implies (assuming that $n_1 = \min_{1 \le i \le N}\{\dim n_i\}$) from

$$x_{N+1} = \underbrace{S_N \cdots S_1}_{\mathcal{P}_1^N} x_1 + \underbrace{[S_N \cdots S_2, S_N \cdots S_3, \cdots, S_N, I]}_{\mathcal{R}_1} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}}_{\overline{v}}$$

that $\mathcal{P}_1^N = I \quad , \quad \mathcal{R}_1\overline{v} = 0.$
We note that $v_i \equiv 0$ and

$$S_i = [\, I \,,\, 0 \,] \quad \text{if} \quad n_{i+1} < n_i,$$
$$S_i = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad \text{if} \quad n_{i+1} > n_i.$$

is consistent with forgetful causalization.

V. Discrete Delay Systems

Here the time set is $\mathbb{T} = \mathbb{Z}$, and $\mathbb{T} = \mathbb{N}$ for the corresponding initial value problem. The typical linear system is

$$x_{k+1} = A_k x_k + B_{k-h(k)},$$

where $h(\cdot)$ is the integer valued delay function.

With a causal information structure, $h(k)$ is adapted with respect to some filtration $\mathcal{G}^k$. Causality is maintained as long as $h(k+1) \le h(k) + 1$. Unlike the continuous time case, a time transformation to render the system as one with constant delay is nonexistent. Consequently, finding a useful notion of a state space with a time invariant structure is more difficult. On the other hand, a discrete time delay system is inherently finite dimensional, albeit with a varying dimension. State augmentation leads to a one step transition

$$\begin{bmatrix} x_{k+1} \\ x_k \\ \vdots \\ x_{k-h(k)+1} \end{bmatrix} = \begin{bmatrix} A_k & \cdots & \cdots & B_k \\ I & & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-h(k)} \end{bmatrix}$$

But since at step $k+1$ the delay is $h(k+1) \le h(k) + 1$ in the causal case, the left hand side, $\mathcal{A}_k\chi_k$, is not generally χ_{k+1}. There is an additional *mode transition* needed, i.e.,

$$\chi_{k+1} = S_{k+1}\mathcal{A}_k\chi_k$$

in order to adapt to the new delay. We defined where $\chi'_k = [x'_k, \cdots, x'_{k-h(k)}]$ with $n_k = \dim \chi_k = n(h(k)+1)$, and $S_{i+1} \in \mathbb{R}^{n_{i+1} \times n_i}$. We note that, since

$$S_{i+1,i} = [\, I \,,\, 0 \,] \quad \text{if} \quad n_{i+1} < n_i,$$

$$S_{i+1,i} = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad \text{if} \quad n_{i+1} > n_i.$$

pseudo continuity holds in this discrete M^3D model, although a precise meaning as in the continuous case is lost.

In case the delay satisfies $h(k+1) > h(k)+1$, the model is no longer causal. Just as in the continuous time case, problems with uniqueness of the solution and minimality occur.

If the system is "omniscient" (full knowledge) all of $h(\cdot)$ is known at any step, k. In principle, a nonminimal lifted description can be given with the maximum delay if it exists. This gives a nonparsimonious model of fixed dimension $n(\max_k h(k) + 1)$. Problems are expected if the delay is unbounded.

Alternatively, lossless causalization will lead to a multi-mode multi-dimensional discrete time model, whose augmented state at step k has the (forward looking) dimension $n_k = n(\max_{\ell \geq k} h(\ell) + 1)$. Expressed differently, lossless causalization requires that from step k on the partial state $x_{k-h(k)}$ still needs to be remembered up to step ℓ where $\min_{\ell > k}(\ell - h(\ell)) = k - h(k)$, but will be used for the last time at that k-th step if $\min_{\ell > k}(\ell - h(\ell)) = k - h(k)$. The transitions for the augmented partial state are of the form

$$\chi_{k+1} = M_{k+1,k}\chi_k$$

with dimension at most increasing by $n = \dim x_k$ at each step.

The map $\widehat{h}$ is the discrete equivalent of the $\widehat{h}(t)$ for continuous time lossless causalization. The matrix $M_{k+1,k}$ is in a generalized block top-companion form: a block companion with deleted rightmost block columns if $h(k+1) < h(k)$, or with an appended block row $[0, \cdots, 0, I]$ at the bottom if $\widehat{h}(k+1) = \widehat{h}(k) + 1$.

Forgetful causalization leads to a similar model, and works in the case the delay map h is not known into the future.

VI. Conclusions

The problem of defining the proper notion of state and state space for systems with time-varying delay has been considered. The main issue is that the notion of state space should entail a constant structure, as only then it makes sense to speak about a trajectory for the system. Likewise the notion of stability requires the notion of a norm on this fixed state space. It was shown that when the delay derivative satisfies $\dot{\tau} < 1$, a state space and hence stability can be well defined. Many problems are associated with a violation of this constraint. We also clarified that part of the problems are associated with the interpretation of the system in terms of the side information: the essential question being: "Are future delays known at the present time or not?" To answer this we have formalized the information structure of a systems. Lossless causalization is then associated with an omniscient information structure, whereas for adapted information, forgetful causalization is the sole answer. An example illustrated that if the delay can grow unboundedly, none of these solutions may be satisfactory. In the discrete time case, the only recourse was to model the system as an M^3D system, for which a state space is modeled as a fibre bundle. Forgetfull and lossless causalization were defined in this context.

References

[1] E.I. Verriest and A.F. Ivanov, (1991) "Robust Stabilization of Systems with Delayed Feedback," *Proceedings of the Second International Symposium on Implicit and Robust Systems*, Warszawa, Poland, pp. 190-193, July 1991.

[2] E.I. Verriest and A.F. Ivanov, (1994) "Robust Stability of Systems with Delayed Feedback," *Circuits, Systems, and Signal Processing*, vol. 13, no. 2/3, pp. 213-222.

[3] E.I. Verriest, (1994) "Stabilization of Deterministic and Stochastic Systems with Uncertain Time Delays," *Proceedings of the 33rd IEEE Conference on Decision and Control*, Orlando, FL, pp. 3829-3834.

[4] F. Colonius, A. Manitius and D. Salamon, "'Structure theory and duality for time-varying retarded functional differential equations. *Journal of Differential Equations*, 78, pp. 320-353, 1989.

[5] H. Reichenbach, *The Direction of Time*. Dover, 1999.

[6] J.M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, 1968.

[7] R. Saeks, Causality in Hilbert Space. *SIAM Review*, Vol. 12, No. 3, pp. 357-383, 1970.

[8] R. Saeks. State in Hilbert space. *SIAM Review*, Vol. 15, No. 2, pp. 283-308, 1973.

[9] A. Feintuch and R. Saeks, *System Theory: A Hilbert Space Approach*, Academic Press, 1982.

[10] J.L. Aravena, State trajectories in Hilbert resolution spaces. *Math Systems Theory* 19, 95-101 (1986).

[11] Hale, J.K. and S. M. Verduyn Lunel (1993). *Introduction to functional differential equations*, Applied Math. Sciences Series, **99**, Springer-Verlag, New York.

[12] A. Nerode, "Linear Automaton Transformations". *Proceedings of the American Mathematical Society*, Vol. 9, No. 4 (Aug., 1958), pp. 541-544.

[13] B.G. Zhang and K. Gopalsamy, Global attactivity in the delay logistic equation with variable parameters, *Math. Proc. Camb. Phil. Soc.* **107** 579-590. (1990).

[14] K. Gopalsamy and X.-Z. He, Dynamics of an almost periodic logistic integrodifferential equation. *Methods and Applications of Analysis* **2** (1) pp. 38-66, 1995.

[15] C. Yuanji, "Oscillation in nonautoinomous scalar differential equations with deviating arguments." *Proceedings of the American Mathematical Society*, Vol. 110, No. 3, pp. 711-719, 1990.

[16] G. Ladas, C. Qian, and J. Yan, "A comparison result for the oscillation of delay differential equations," *Proceedings of the American Mathematical Society*, Vol. 114, No. 4, pp. 939-947, 1992.

[17] F. Kappel. On degeneracy of functional-differential equations. *J. Differential Eqs.*, 22:250267, 1976.

[18] F. Kappel and W. Schappacher. Some considerations to the fundamental theory of infinite delay equations. *J. Differential Eqs.*, 37:141183, 1980.

[19] E.I. Verriest, "Inconsistencies in systems with time-varying delays and their resolution." (2010) submitted.

[20] E.I. Verriest, (2009) Causal Behavior of Switched Delay Systems as Multi-Mode Multi-Dimensional Systems, *Proceedings of the 8-th IFAC International Symposium on Time-Delay Systems*, Sinaia, Romania, September 2009.

[21] E.I. Verriest, (2009) Multi-Mode Multi-Dimensional Systems with Poissonian Sequencing, The Brockett Legacy Issue of *Communications in Information and Systems*, Vol. 9, No. 1, 2009.

[22] E.I. Verriest, "Multi-Mode Multi-Dimensional Systems with Application to Switched Systems with Delay," *Proceedings of the 48th Conference on Decision and Control and 28th Chinese Control Conference*, pp. 3958-3963, Shangai, PRC, December 2009.