A nonlinear time-delay realization for gastroparesis in patients with diabetes

Claudia Califano\textsuperscript{a}, Emeric Scharbarg\textsuperscript{b}, Nicolas Magdelaine\textsuperscript{b} and Claude H. Moog\textsuperscript{b}

\textsuperscript{a}Claudia Califano is with DIAG, Università di Roma, via Ariosto 25, 00184 Roma, Italy. claudia.califano@uniroma1.it
\textsuperscript{b}Emeric Scharbarg, Nicolas Magdelaine and Claude H. Moog are with LS2N, UMR CNRS 6004, BP 92101, 44321 Nantes Cedex 3, France. emeric.scharbarg@ls2n.fr

Abstract

The realization problem is fundamental in control systems theory and is well understood for linear systems. The situation is quite different for nonlinear systems with or without time delays. New results are given herein either as an abstract characterization of the class of delay–free nonlinear systems which are known to admit a realization, or by characterizing the existence of a realization for a wider class of retarded type or neutral type time delay nonlinear systems. The problem is motivated by the very practical modelling and realization problem of diabetes with or without gastroparesis. Up to 4% to 12% patients with diabetes are affected by the so-called gastroparesis which delays the digestion process. Several concurrent state space mathematical models do exist for diabetes and require a re-evaluation based on physiologic facts. These are reviewed herein to derive a (minimal) realization.

Keywords: Biomedical systems, Systems biology, time-delay systems, realization, glucose dynamics

1. Introduction

The realization problem is one of those very fundamental problems tracing back to the early history of control systems theory, just like the controllability problem and the observability one. Its solution is well established for linear systems but surprisingly remains open for the class of nonlinear systems despite the availability of partial solutions. The contribution in this paper is threefold. First, a state of the art is given on the realization problem for a broad class of delay–free nonlinear control systems. An abstract definition is given for the class of systems for which the realization problem is well posed. Secondly, the existence of a realization is characterized for a wide class of nonlinear time-delay systems. Finally, a nonlinear and minimal realization is derived for diabetes of type 1 with gastroparesis, which involves a delay in the digestion.

Though realization is a fundamental problem in control theory, major contributions were obtained only recently for nonlinear systems. Some milestones are as follows. A necessary condition for the existence of a realization was given in Crouch et al. (1992) which states that the higher order input-output differential equation has to be affine in
the highest order derivative of the input. In practice, there is no state realization for
the input-output system described by the equation $\ddot{y} = \dot{u}^2$, although the latter is strictly
proper. Stronger conditions are available for Hamiltonian systems in Crouch et al.
The necessary and sufficient polynomial structure which is required for the input-
output differential equation to admit a realization was described in Liu et al. (1994)
thanks to a so-called realization algorithm. Unfortunately, no concise abstract interpre-
tation of those conditions was given.
The first abstract necessary and sufficient condition for the existence of a realization for
continuous time nonlinear systems is found in Moog (1999) in terms of integrability
conditions.
The discrete time case is nicely treated in Kotta et al. (2001) within a suitable algebraic
framework. In this context also, the main result for the existence of a realization relies
on concise integrability necessary and sufficient conditions.
The differential algebraic framework is updated and completed in Zhang et al. (2010)
for continuous time nonlinear systems.
Minimality of realizations is worked out in Kotta et al. (2018) to characterize the
largest class of delay free nonlinear systems for which the realization problem is well
posed.
A necessary and sufficient condition for the existence of a strongly observable realiza-
tion is given in Kaldmäe et al. (2018) for input-output systems with delays of retarded
type.
Major open problems about the realization for nonlinear systems with or without
delays are as follows.
1. Characterization of the largest class of delay free nonlinear systems which admit
   a realization (through a concise and abstract definition).
2. Characterization of the largest class of delay free nonlinear systems which admit
   a minimal realization (through a concise and abstract definition).
3. Characterization of the existence of a general (weakly observable or not) realiza-
   tion for retarded type input-output systems.
4. Characterization of the existence of a realization for neutral type input-output
   systems.
Some of those principles are illustrated by deriving a real-life state space realization of
the plasma glucose and insulin dynamics. There are basically two antagonistic prac-
tical or theoretic constraints to cope with. The first one is imposed by the practical
Flexible Insulin Therapy (FIT) and yields an affine dynamics, i.e. a linear dynamics up
to some nonzero basal equilibrium. The second constraint is imposed by the positivity
of concentrations and inputs. These positivity properties display the limits of the affine
dynamics and necessarily impose a nonlinear functions instead of constant model pa-
rameters. These points are argued in Section 3 and a resulting Diabetes Concordance
Model is derived.
Basically, the system which is considered is a cascade composition of a second
order insulin subsystem driven by the exogenous insulin infusion for patients with di-
abetes, a second order digestion subsystem driven by the ingested carbohydrates and a
first order plasma glucose subsystem. The latter is the focus of our argument in Section 3. The state space model is widely used to develop the so-called artificial pancreas which closes the loop for an automatic regulation of the blood glucose thanks to an automatic infusion of exogenous insulin. One mainstream for this regulation of blood glucose relies on MPC (Model Predictive Control) which is computed from the state space model.

In addition, patients with diabetes, the evolution of this chronic disease goes together with the risk of gastroparesis, that is, with delayed gastric emptying (GE). Gastroparesis is characterized by a constellation of gastrointestinal symptoms Shin et al. (2013). It introduces a time delay in the mathematical model.

This paper aims to contribute to some of the above mentioned open theoretical problems as well as to derive a suitable mathematical state space model for type 1 diabetes with gastroparesis. Preliminary results are found in Califano et al. (2018) without any argument on realizability, but a focus on observability and an observer design for the special diabetes model.

Section 2 is fully devoted to the realization theory including recalls and theoretical contributions as announced above. After recalling some basic facts from the so-called Bergman model in Section 3, the diabetes state space model is detailed step by step. The process ends up with the so-called Diabetes Concordance Model as it is consistent with the constraints of positive systems, and with the medical practice that is flexible insulin therapy. Clinical data are worked out in Section 4 to illustrate the efficiency of the Diabetes Concordance Model developed in Section 3, with and without delay. In the Appendix of the paper, some notions of observability for time-delay systems are shortly argued for completeness.

2. Realizations and minimal realizations

2.1. Delay free nonlinear systems

Let us recall some facts from Liu et al. (1994), Kotta et al. (2001), Zhang et al. (2010). Consider a higher order nonlinear retarded type input-output equation as

\[ y^{(n)} = F(y, \dot{y}, ..., y^{(n-1)}, u, \dot{u}, ..., u^{(s)}). \]  

(1)

The nice feature of (1) is that the differential ideal spanned by \( y^{(n)} - F(\cdot) \) is prime, since \( y^{(n)} - F(\cdot) \) cannot be factored out. Implicit systems as \( \dot{y}^2 - 3u\dot{y} + 2u^2 = 0 \) are not considered since \( \dot{y}^2 - 3u\dot{y} + 2u^2 = (\dot{y} - u)(\dot{y} - 2u) \).

In opposition to linear systems, with or without delays, it is not true that any nonlinear system (1), with \( n > s \), admits a state realization in the form

\[
\begin{align*}
\dot{x} &= f(x,u) \\
y &= h(x).
\end{align*}
\]

An easy counter-example is given by \( \ddot{y} = u^2 \). Recall now a necessary and sufficient condition for the existence of such a (possibly non-minimal) realization with dimension \( n \). Given (1), associate the following extended system which is defined from the
notation $z_1 = y, z_2 = \dot{y}, \ldots, z_n = y^{(n-1)}, z_{n+1} = u, \ldots, z_{n+s+1} = u^{(s)}$ so that

\[
\dot{z} = \begin{pmatrix}
    z_2 \\
    \vdots \\
    z_n \\
    z_{(n+2)} \\
    \vdots \\
    z_{n+s+1} \\
    0
\end{pmatrix} + \begin{pmatrix}
    0 \\
    \vdots \\
    0 \\
    0 \\
    \vdots \\
    0 \\
    1
\end{pmatrix} u^{(s+1)}
\]

(2)

\[
y = z_1.
\]

The so-called accessibility filtration $\mathcal{H}_k$ associated to (2) is defined by $\mathcal{H}_1 = \text{span}\{dz\}$ and $\mathcal{H}_{k+1} = \{\omega \in \mathcal{H}_k | \dot{\omega} \in \mathcal{H}_k\}$. Then, the following result was proven in (Conte et al. (2014), Theorem 2.16).

**Theorem 1.** A (possibly non-minimal) realization exists for (1) if and only if

(i) $n > s$, and

(ii) $\mathcal{H}_{s+2}$ is integrable.

This condition is not fulfilled for the 2nd-order system $\ddot{y} = \dot{u}^2$; the dimension of the extended system is 4:

\[
\frac{d}{dt} \begin{pmatrix}
    y \\
    \dot{y} \\
    u \\
    \dot{u}
\end{pmatrix} = \begin{pmatrix}
    \dot{y} \\
    \dot{\dot{y}} \\
    \dot{u}^2 \\
    \dot{u}
\end{pmatrix} + \begin{pmatrix}
    0 \\
    0 \\
    0 \\
    1
\end{pmatrix} \dot{u},
\]

and $\mathcal{H}_3 = \text{span}\{dy, 2\dot{u}u - dy\}$ is not integrable.

**Minimality of the state realization for nonlinear systems**

For linear systems, minimality in the sense of a minimal number of state variables is equivalent to minimality in the sense of simultaneous observability and controllability. This feature is much more arguable for nonlinear systems.

Consider the second order input/output equation $\ddot{y}u - \dot{y}\dot{u} + \dot{y}u = 0$ proposed in Kotta et al. (2013). A state space realization is easily derived when writing down the associated extended system (2) and applying Theorem 1. Since $s = 1, (2)$ becomes

\[
\frac{d}{dt} \begin{pmatrix}
    y \\
    \dot{y} \\
    u \\
    \dot{u}
\end{pmatrix} = \begin{pmatrix}
    \dot{y} \\
    \ddot{y} \\
    \dot{u}^2 - \dot{y} \\
    \dot{u}
\end{pmatrix} + \begin{pmatrix}
    0 \\
    0 \\
    0 \\
    1
\end{pmatrix} \dot{u}.
\]

Compute $\mathcal{H}_3 = \text{span}\{dy, d(\dot{x}/u)\}$; setting $x_1 = y$ and $x_2 = \dot{y}/u$ for $u \neq 0$ yields the second order non-minimal realization

\[
\begin{cases}
    \dot{x}_1 = x_2u \\
    \dot{x}_2 = 0 \\
    y = x_1.
\end{cases}
\]
Obviously, this realization is not accessible. This is due to the fact that the differential ideal spanned by $\dot{y}u - \dot{y}u + \dot{y}u$ is not equal to its radical. The latter is in fact spanned by $\frac{\partial}{\partial u}$. As a matter of fact, in Kotta et al. (2013) it was shown that an accessible and observable first order realization may be derived for $\dot{y}u - \dot{y}u + \dot{y}u = 0$, but it would necessarily be time-varying.

Recall that any ideal $I$ has a primary decomposition, i.e. $I$ can be written as the intersection of a finite number of prime ideals: $I = \bigcap_j I_j$. Define the closure $\bar{I}$ of $I$ as the ideal $\bar{I} = \bigcup_j \sqrt{I_j}$, where $\sqrt{I_j}$ denotes the radical of $I_j$.

**Examples**

(i) Let $I = \langle \frac{\partial}{\partial t} (\dot{y} - u)(\dot{y} - 2u) \rangle$ which is prime. We get $\bar{I} = \langle \dot{y} - u, (\dot{y} - 2u) \rangle$ which has two irredundant generators.

(ii) Let $I = \langle \frac{\partial}{\partial t} (\dot{y} - u)^2 + y^{(3)} - \ddot{u} \rangle$ which is prime. Its closure $\bar{I} = \langle \dot{y} - u \rangle$ has one single generator ; consequently $\bar{I}$ is an irreducible ideal.

The following result can now be stated which gives sufficient conditions for the existence of a state space realization for a given input-output equation.

**Theorem 2.** Given the input-output equation (1), denote by 

$I = \langle y^{(n)} - F(y,\dot{y},...,y^{(n-1)},u,\dot{u},...,u^{(s)}) \rangle$

the differential ideal spanned by $y^{(n)} - F(y,\dot{y},...,y^{(n-1)},u,\dot{u},...,u^{(s)})$. There exists a minimal realization for (1) if

1. the closure $\bar{I}$ of $I$ is an irreducible ideal, i.e. it is generated by one single element $\varphi(y,\dot{y},...,y^{(n')},u,\dot{u},...,u^{(s')})$, such that
2. $\varphi(0,...,0) = 0$, and
3. $\varphi(\cdot)$ fulfills the conditions of Theorem 1.

Consider for instance the following counter example.

\[ \ddot{y} = \frac{3\ddot{y}u - 4\dot{u}\ddot{y}}{2\ddot{y} - 3u}, \]

which has the form (1) and which can be rewritten as $\frac{\partial}{\partial t} (\dot{y} - u)(\dot{y} - 2u) = 0$. The differential ideal $I$ is generated by $\frac{\partial}{\partial t} (\dot{y} - u)(\dot{y} - 2u)$ and its closure $\bar{I} = \langle \dot{y} - u, (\dot{y} - 2u) \rangle$ is generated by both $(\dot{y} - u)$ and $(\dot{y} - 2u)$. The first condition in Theorem 2 is not fulfilled.

The second condition in Theorem 2 is designed to avoid equivalence of the systems $\ddot{y} = \dot{u}, \ddot{y} = u + 1$ and $\ddot{y} = u$. The two last systems should not have the same first order realization.

**Proof.** From item 1. in Theorem 2, there exists one single function $\varphi(y,\dot{y},...,y^{(n')},u,\dot{u},...,u^{(s')})$, such that the given input-output equation can be written in terms $\varphi$ and its derivatives. Set $\varphi(y,\dot{y},...,y^{(n')},u,\dot{u},...,u^{(s')}) = 0$, then the implicit function theorem yields (locally) $y^{(n')} = \psi(y,\dot{y},...,y^{(n'-1)},u,\dot{u},...,u^{(s')})$ for some $\psi$. The latter is in the form (1). Item 3. in Theorem 2 and Theorem 1 yield the result.

\[ \square \]
2.2. Nonlinear time-delay systems

In Kaldmäe et al. (2018), it was shown that a retarded type input/output equation of the form

\[ \Phi(y^{(n)}(t), y^{(n-1)}(t), y^{(n-1)}(t - 1), \ldots, u^{(n-1)}(t), u^{(n-1)}(t - 1), \ldots) = 0 \]

admits a strongly observable realization if and only if some \( \mathcal{H}_{n+1} \) is integrable, where \( \mathcal{H}_{n+1} \) is defined in analogy with the delay free case above but on the non-commutative ring \( \mathcal{H}(\delta) \)

\[ \mathcal{H}_1 = \text{span}_{\mathcal{H}(\delta)} \{ dy, \ldots, dy^{(n-1)}, du, \ldots, du^{(r)} \} \]

\[ \vdots \]

\[ \mathcal{H}_k = \text{span}_{\mathcal{H}(\delta)} \{ \omega(x, \delta) \in \mathcal{H}_{k-1}, \omega(x, \delta) \in \mathcal{H}_{k-1} \}. \]

Example

Consider the 1st order retarded type system borrowed from Garcia-Ramirez et al. (2016) (Example 10):

\[ \dot{x}(t) = x(t - 1)u(t) \]
\[ y(t) = x(t) + x(t - 1). \]

Straightforward, but tedious, computations to eliminate the state yield the 1st order neutral type input/output differential equation

\[ y(t)[u(t - 2) - u(t - 1)] + y(t - 1)[u(t - 1) - u(t)] + \\
+ [u(t - 1) - u(t)][y(t - 1)(u(t - 2) - u(t - 1)) - y(t - 2)u(t - 2)] + \\
- u(t - 1)y(t - 1)[u(t - 2) - u(t - 1)] = 0. \]

According to Halas et al. (2013) (Theorem 13), there is an equivalent higher order retarded type input/output differential equation which represents the system. In the special case of our example, we get

\[ \dot{y}(t) = \frac{\dot{x}(t) - y(t - 1)}{u(t - 1)}[u(t) - u(t - 1)]u(t - 1) - u(t - 1) + u(t - 2)u(t - 1)] + \\
+ y(t)[u(t)u(t - 1) + u(t - 1) - u(t - 2)u(t - 1)] + \\
+ y(t - 2)u(t - 2)u(t - 1). \]

---

1 Denote by \( \mathcal{H} \) the field of causal meromorphic functions \( f(x(t), \ldots, x(t - s), u(t), \ldots, u(t - j)) \), with \( s, j \in \mathbb{N} \). Accordingly, \( f(-1) = f(x(t - 1), \ldots, x(t - s - 1), u(t - 1), \ldots, u(t - j - 1)) \). Let \( \delta \) be the backward time-shift operator: for \( a(\cdot), f(\cdot) \in \mathcal{H} : \delta[adf] = a(-1)\delta df = a(-1)df(-1) \). Denoting by \( e \) the vector space spanned by the differentials \( dx(t - i) ; i \in \mathbb{N} \) over the field \( \mathcal{H} \), the elements of \( e \) are called one-forms. The shift operator is applied to the vector space \( e \) in this way: if \( \omega \) is the one-form \( \omega = \sum_{j=0}^{s} \sum_{i=0}^{j} a_{j-i} dx(i-t-j) \) then \( \delta \omega \) is given by \( \delta \omega = \sum_{j=0}^{s} \sum_{i=0}^{j} \delta(a_{j-i}) dx(i-t-j) \). In this framework \( \mathcal{H}(\delta) \) is the left ring of polynomials in \( \delta \) with coefficients in \( \mathcal{H} \). Every element of \( \mathcal{H}(\delta) \) may be rewritten as \( \alpha(\delta) = \delta^{r_{\alpha}} \sum_{j=0}^{s} a_{j-i} \delta^{j} \) with \( a_{j-i} \in \mathbb{N} \) and \( r_{\alpha} = \text{deg}(\alpha(\delta)) \) the polynomial degree in \( \delta \).
The existence of a 2nd order retarded type, strongly observable (and non-minimal) realization may be checked for the latter using the techniques in Kaldmæe et al. (2018). Nevertheless, a retarded type minimal realization is obtained setting

\[ x(t) = y(t) - \frac{1}{u(t) - u(t-1)}[\dot{y}(t) - u(t-1)y(t-1)] \]

which yields the first order dynamics (4).

3. Mathematical modeling of the glucose-insulin dynamics

The so-called Bergman model (Bergman et al. 1981) is certainly one of the most popular mathematical models for the glucose-insulin dynamics, and it gave rise to numerous modifications and improvements.

The two first equations in (Bergman et al. 1981, page 1458) read as

\[ \begin{align*}
\dot{G} &= -(p_1 + X)G + p_4 \\
\dot{X} &= -p_2X + p_3I
\end{align*} \] (5)

for some constant parameters \( p_1, p_2, p_3, p_4 \) and where \( G \) denotes the glucose concentration in plasma, \( X \) is the active insulin in some "remote compartment" and \( I \) represents the plasma insulin concentration. Assume that (5) is driven by some constant \( I_{eq} \); then the equilibrium points \((G_{eq},X_{eq})\) of (5) are easily computed as

\[ \begin{align*}
G_{eq} &= \frac{p_2p_4}{p_1p_2 + p_3I_{eq}} \\
X_{eq} &= \frac{p_3I_{eq}}{p_2}
\end{align*} \]

Note that any constant (basal candidate) \( I_{eq} \) is able to stabilize any plasma glucose concentration \( G_{eq} \). In fact, Bergman model displays a one to one relationship between constant plasma insulin values and constant plasma glucose values. Thus, it yields a nice disputation argument as two contradictory theoretical and practical features stand in apparent opposition, as detailed right now.

3.1. Two antagonistic facts

1. On one hand, the physiology imposes that concentrations as the plasma glucose \( G \) and insulin \( I \) remain positive under any condition although an almost zero plasma glucose concentration does not make sense in practice. Plasma glucose concentration decreases under the action of insulin, at a rate \( k_{is} \) which is known as the insulin sensitivity factor. Thus the decay of plasma glucose concentration incorporates a term \( -k_{is}(\cdot)I \). To avoid plasma glucose concentration to become negative at very low values of \( G \) whatever other conditions are, it is mandatory that the insulin sensitivity factor vanishes when \( G \) vanishes

2. On the other hand, the mathematical model is assumed to be consistent with the well accepted flexible insulin therapy. The latter implies linearity of the glucose dynamics, and insulin sensitivity factor should just be constant.
Next, we start with the physiologic description which leads to a (positive) nonlinear dynamics. Then, the model underlying the flexible insulin therapy is just an approximation valid in some domain of interest.

3.2. The diabetes concordance model

The full model of the glucose-insulin dynamics includes the behaviour of several quantities, and in particular of $G$ - the blood plasma glucose concentration, $I_p$ - the blood plasma insulin concentration, $I_{sc}$ - the subcutaneous insulin concentration, $X_s$ - the amount of assimilated carbohydrates that are transferred from the stomach into the duodenum, $X_d$ - the amount of assimilated carbohydrates that is transferred from the duodenum to the plasma. Consequently the raise of blood glucose concentration when a meal is digested can be taken into account by adding a linear term of the form $\theta_4 X_d$ with $\theta_4 > 0$ in the plasma glucose concentration dynamics.

$u$ - the injected insulin rate
$r$ - the carbohydrates absorbed through the meal.

Glucose dynamics

Basically, plasma glucose concentration is subject to glucose consumption $f_m$ by the muscle cells (under the action of insulin), to glucose consumption by the brain $f_b$ which is not insulin dependent, to glucose storage or delivery by the liver $F_l$ and to glucose release from the digestion of carbohydrates.

The glucose dynamics is represented as a first order subsystem subject to two inputs, the blood insulin concentration $I_p$ which has an hypoglycemic effect, and to some exogenous input $X_d$ which results from the digestion of carbohydrates and which has an hyperglycemic effect.

At this stage, the general input-output equation is as follows for the plasma glucose subsystem.

$$\dot{G} = -f_m(G, I_p) - f_b(G) + F_l(G, I_p) + \theta_4 X_d \quad (6)$$

At the cellular level the glucose dynamics is governed directly by GLUT transporters Lodish H et al. (2013). These are membrane proteins that allow the transport of glucose inside the cells. The main membrane proteins are GLUT1 (red blood cells), GLUT2 (hepatocytes), GLUT3 (neurons) and GLUT4 (muscle and fat cells). These transporters are involved in a specific way in the glucose metabolism depending on the tissue where they are expressed and physiological conditions. They each have distinct kinetics according to their functional roles.

Glucose consumption by the muscle cells

The GLUT4 is responsible for glucose storage regulated by insulin which will allow assimilation of glucose by the muscle and fat cells Bryant et al. (2002).

The consumption of glucose by the muscles partly results from osmosis, but is mainly governed by the GLUT4 transporter as described above. The term $f_{mi}(G, I_p)$ decomposes into the insulin independent term $\epsilon G$ due to osmosis and the term $f_{mi}(G, I_p)$
due to the GLUT4 transporter. The insulin sensitivity factor is defined as $\frac{\partial f_m}{\partial I_p}$. It is supposed to depend of $G$ only Tolic et al. (2000) and obviously has to vanish when $G$ becomes zero:

$$f_m(G, I_p) = \varepsilon G + \theta_2 (1 - \exp(-\frac{G}{G_c})) I_p$$  \hspace{1cm} (7)$$

and, considering fasting, i.e. $X_d = 0$

$$\dot{G} = -\varepsilon G - \theta_2 (1 - \exp(-\frac{G}{G_c})) I_p - f_b(G) + F_l(G, I_p)$$ \hspace{1cm} (8)$$

where $\varepsilon$, $\theta_2$ and $G_c$ are some constant parameters. $G_c^-\text{ is a lower critical value for plasma glucose concentration below which the insulin sensitivity factor decreases to zero. Above this critical value, the insulin sensitivity factor increases asymptotically towards } \theta_2 \text{ as } G \text{ increases.}$

**Glucose consumption by the brain**

The GLUT3 transporters are mostly expressed in neurons, this glucose transporters are not mediated by insulin. They have a very high affinity for glucose even in case of hypoglycemia. The glucose consumption by the brain is assumed to be almost constant, except for some very low values of plasma glucose concentration as it has to be zero when plasma glucose concentration is zero. The glucose consumption $f_b(G)$ by the brain is then taken as Tolic et al. (2000), eq. (8).

$$f_b(G) = \theta_b (1 - \exp(-\frac{G}{G_c})).$$ \hspace{1cm} (9)$$

For simplicity, the same critical value $G_c^-$ is taken as in (8), so that (8) can be rewritten as:

$$\dot{G} = -\varepsilon G - \theta_2 (1 - \exp(-\frac{G}{G_c})) I_p - \theta_b (1 - \exp(-\frac{G}{G_c})) + F_l(G, I_p).$$ \hspace{1cm} (10)$$

Obviously, this glucose consumption by the brain vanishes when plasma glucose concentration approaches zero and will not hinder the positivity of the system Farina et al. (2000).

**Glucose storage and glucose delivery by the liver**

The liver has a fundamental role in the regulation of plasma glucose because it is able to store and release glucose, so its transporter must be active all the time. After a meal, dietary carbohydrates are digested in monosaccharides, and they are transported into various tissues as a primary fuel for ATP generation (adenosine triphosphate). The postprandial hyperglycemia increases insulin secretion from pancreas (beta cells). In the liver excess glucose is stored mainly in the form of glycogen via glycogenogenesis. If glucose is too abundant, it can be stored into fatty acids via lipogenesis.

This action induces a decrease in plasma glucose concentration and contributes as $-f^l_i(G)$ in $F_l(G, I_p)$, with

$$f^l_i(G) = C_0 \left[ \frac{\exp(G/G_c^+)}{\exp(G/G_c^+)} + 1 \right].$$
$G_c^+$ denotes a higher critical value of plasma glucose concentration whose level decides about the storage function of the liver; $f_s^l(G)$ is approximately zero for low plasma glucose concentration values, and approximately constant (equal to $C_0$) for high plasma glucose concentration values (in a post prandial phase).

Between two meals, the liver is the only organ capable of releasing glucose into the blood by glycogenolysis or gluconeogenesis. Under fasting periods, the liver has a major role in releasing glucose for the brain, red blood cells, muscles and others tissues. In that case, the alpha cells from the pancreas secrete glucagon in response to hypoglycemia which initiates the cascade of molecular reaction that releases glucose from the stored glycogen via glycogenolysis.

This action induces an increase in plasma glucose concentration represented by $+f_d^l(I)$ which contributes to $F_l$ so that

$$F_l = f_d^l(I_p) - f_s^l(G).$$

$f_d^l(I)$ is similar to the function $f_5$ in Tolic et al. (2000), equation (11):

$$f_d^l(I_p) = \frac{2\theta_1}{1 + \exp I_p / I_c},$$

where $I_c$ is a critical value above which $f_d^l$ almost vanishes (after some meal).

At this stage, we did not take into account the glucose increase due to the digestion of carbohydrates and the glucose dynamics in equation (10) becomes

$$\dot{G} = -\varepsilon G - \theta_2 (1 - \exp (-\frac{G}{G_c})) I_p - \theta_b (1 - \exp (-\frac{G}{G_c})) + \frac{2\theta_1}{1 + \exp I_p / I_c} - C_0 \left[ \frac{\exp (G/G_c^+) - 1}{\exp (G/G_c^+) + 1} \right].$$

Now, taking into account a post prandial phase, and denoting $X_d$ the amount of glucose due to the digestion of carbohydrates, the overall glucose dynamics now takes the form

$$\dot{G} = -\varepsilon G - \theta_2 (1 - \exp (-\frac{G}{G_c})) I_p - \theta_b (1 - \exp (-\frac{G}{G_c})) + \frac{2\theta_1}{1 + \exp I_p / I_c} - C_0 \left[ \frac{\exp (G/G_c^+) - 1}{\exp (G/G_c^+) + 1} \right] + \theta_4 X_d(t).$$

**Positivity of the glucose dynamics**

The dynamics (12) is subject to the two positive inputs $I_p$ and $X_d$ and has the form $\dot{G} = F(G, I_p, X_d)$. First order linear dynamics are always positive because the origin is always an equilibrium, whereas nonlinear dynamics are not always positive.

The dynamics $\dot{G} = F(G, I_p, X_d)$ is positive if and only if $F(0, I_p, X_d) \geq 0$, for any $I_p \geq 0$ and any $X_d \geq 0$.

The right hand side of (12) is evaluated at $G = 0$ and reduces to

$$\frac{2\theta_1}{1 + \exp I_p / I_c} + \theta_4 X_d(t).$$

The latter is positive for any positive input.
Insulin dynamics

The subcutaneous insulin dynamics is described by the differential equation

\[
\dot{I}_{sc} = -\frac{1}{\theta_3} I_{sc} + \frac{\theta_u}{V_4 \theta_3} u,
\]

where $\theta_3 > 0$ is the response time of the insulin subsystem. The dynamics of the plasma insulin $I_p$ instead reads

\[
\dot{I}_p = \frac{1}{\theta_3} I_{sc} - \frac{1}{\theta_3} I_p.
\]

Digestion dynamics

At this point, considering the input $X_d$ to be zero, one gets the minimal model during a fasting period. When carbohydrates (CHO) are ingested during a meal, then a two compartment digestion subsystem is modeled as follows. The stomach compartment is fed by the input $r(t)$, that is the amount of CHO contained in the meal:

\[
\dot{X}_s(t) = -\frac{1}{\theta_5} X_s(t) + \frac{1}{\theta_5} r(t).
\]

The duodenum compartment dynamics is instead described by the differential equation

\[
\dot{X}_d(t) = -\frac{1}{\theta_5} X_d(t) + \frac{1}{\theta_5} X_s(t - \tau),
\]

where the gastric emptying time $\tau$ characterizing gastroparesis is displayed. For the majority of patients with diabetes, i.e. without gastroparesis, the response time $\theta_5$ of the digestion subsystem is smaller than the response time $\theta_3$ of the insulin subsystem. Besides the delay $\tau$, the patients with diabetes affected by gastroparesis are also characterized by the relationship $\theta_5 > \theta_3$ which will be assumed in the rest of this paper.

Consistency with flexible insulin therapy

On one hand, for a minority of patients, one has $G^- < G^+$ but these two levels of plasma glucose concentration remain in the same range. Consequently, $f_b$ and $f_m$ do not simultaneously stabilize at $\theta_b$ and $\theta_2$ respectively as in equations (7) and (9). The dynamics (12) remains nonlinear for any value of $G$. Those patients have a plasma glucose concentration which is difficult to stabilize. Standard FIT should be likely revised for these patients.

On the other hand, for a majority of patients with diabetes, the critical values of plasma glucose concentration are away from each other so that $G^- << G^+$. This means that there is a significantly large window from moderate hypoglycemia to moderate hyperglycemia in which the following approximations hold:

\[
\theta_2 (1 - \exp(-\frac{G}{G^+_c})) = \theta_2, \quad \theta_b (1 - \exp(-\frac{G}{G^-_c})) = \theta_b
\]

and

\[
\frac{2 \theta}{1 + \exp(G/\theta)} - C_0 \left[ \frac{\exp(G/\theta^+_c + 1)}{\exp(G/\theta^+_c + 1)} - 1 \right] = \theta_l.
\]
Within this window, the model simplifies dramatically and the principles of Flexible Insulin Therapy fully apply.

Setting $\theta_1 = \theta_I - \theta_b$, equation (12) is approximated by the linear dynamics which is consistent with standard FIT:

$$G(t) = -\varepsilon G(t) - \theta_2 I_p + \theta_1 + \theta_3 X_d(t).$$  \hspace{1cm} (17)

As in Magdelaine et al. (2015), the parameters of FIT can be computed from the mathematical model. For instance, the insulin basal rate $I_b$ is computed for $X_d \equiv 0$ and equals $\theta_1/\theta_2$ whenever $\varepsilon << 0$. $I_b = \frac{\theta_1 - \varepsilon G}{2}$ whenever $\varepsilon \neq 0$ in which case the basal insulin rate slightly depends on the value of the plasma glucose concentration.

Considering a fasting phase, $X_d \equiv 0$, considering that $\varepsilon$ can be neglected, and setting $G = x_1$, $I_c = x_2$, $I_p = x_3$, $X_s = x_4$ and $X_d = x_5$ the complete simplified model (consistent with standard FIT) is as follows.

\[
\begin{align*}
\dot{x}_1(t) &= -\theta_2 x_3(t) + \theta_1 + \theta_4 x_5(t) \\
\dot{x}_2(t) &= -\frac{1}{\theta_3} x_2(t) + \frac{\theta_1}{\theta_3} u(t) \\
\dot{x}_3(t) &= \frac{1}{\theta_3} x_2(t) - \frac{1}{\theta_5} x_3(t) \\
\dot{x}_4(t) &= -\frac{1}{\theta_5} x_4(t) + \frac{1}{\theta_5} r(t) \\
\dot{x}_5(t) &= -\frac{1}{\theta_5} x_5(t) + \frac{1}{\theta_5} x_4(t - \tau) \\
y(t) &= x_1(t).
\end{align*}
\]  \hspace{1cm} (18)

The latter is now linear within the range from moderate hypoglycemia to moderate hyperglycemia, and consistent with the basics of flexible insulin therapy.

4. Parameter identification

4.1. Methodology

Figure 1 displays identification results with and without delay.

The data set was divided into a training set used to estimate the parameters, and a test set in order to test the model’s ability to predict new data. The training set is 12.5 hours long and the test set is 6.5 hours long. The ratio between the number of sample of the training set (fit) and the number of samples of the data set if 66% which is quite common in the holdout cross-validation (cv) method:

- The model-fit was made from 17:30 to 06:00 using least square error on the output i.e. plasma glucose concentration. The identification of constant parameters is thus achieved using clinical data over a range of about 12 hours.

- This full model with specific constant parameters is further simulated over 6 additional hours, to evaluate the prediction ability of the model: cross-validation was made from 06:00 to 12:30.

The Root Mean Square Error (RMSE) criterion $J = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\varepsilon_i^2)}$, where $\varepsilon_i$ represents the error for sample $i$, is computed for the training fit and the cross-validation
with and without delay:
without delay: \( J_{fu} = 18 \text{mg/dl} ; J_{cv} = 10 \text{mg/dl} \)
with delay: \( J_{fu} = 8 \text{mg/dl} ; J_{cv} = 24 \text{mg/dl} \)

Although the cross-validation criterion of the estimated model with delay is worse than without delay, its value is still very good for a 6 hours and 30 min long prediction. This has to be compared with:
- 35 mg/dl RMSE for 1 hour prediction in Zhao et al. (2014)
- no cross-validation for the UVA/Padova simulator in Vinsentin et al. (2014).

### 4.2. Identification result

The identification results with and without a non zero delay \( \tau \) are as follows and are recalled from Califano et al. (2018).

<table>
<thead>
<tr>
<th>parameters</th>
<th>without delay</th>
<th>with delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 ) (mg/U/min)</td>
<td>0.14</td>
<td>0.32</td>
</tr>
<tr>
<td>( \theta_2 ) (mg/dl/min)</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>( \theta_3 ) (min)</td>
<td>33</td>
<td>70</td>
</tr>
<tr>
<td>( \theta_4 ) (dl(^{-1}))</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>( \theta_5 ) (min)</td>
<td>49</td>
<td>22</td>
</tr>
<tr>
<td>( \tau ) (min)</td>
<td>0</td>
<td>95</td>
</tr>
<tr>
<td>( \theta_u/V_f ) (min/dl)</td>
<td>750</td>
<td>750</td>
</tr>
</tbody>
</table>
It illustrates that taking into account the delay, and setting $\tau \neq 0$, allows to emulate better the behaviour of the plasma glucose concentration. It becomes obvious in Figure 1 around time 22:30, as the simulated trajectory tracks much better the data.

The cross-validation simulates the model with the parameters obtained from the fit. Cross-validation shows that the Carbo-to-Insulin Ratio could be re-estimated for breakfast which is a typical meal with fast carbs, as at 12:30 simulated plasma glucose concentration drifts from continuous glucose monitor (CGM) data.

Further improvements go through a LPV identification of the insulin sensitivity factor ($\text{ISF} = \theta_2 * \theta_u / V_i$) and carbo-to-insulin ratio ($\text{ISF} / \theta_3$). This appears to be suitable as breakfast, lunch and dinner are quite different type of meals, and ISF is known to have a circadian variability. Training data could be two days long and cross-validation would then be performed on the third day.

For alternative models, the reader is referred for instance to Cobelli et al. (2014) in which delays refer to a second phase insulin secretion and to the CGM and insulin pump technologies, to Palumbo et al. (2013) in which the tissue glucose uptake is delayed with respect to the insulin action and the insulin pancreatic secretion is delayed with respect to the plasma glucose action (whenever the pancreas insulin secretion is non zero), or to Reiterer et al. (2015) where no delay is considered but the essential features of the glucose-insulin dynamics are modeled by linear terms as in (18).

5. Conclusion

The realization problem is challenging for nonlinear systems and some of the main open problems have been sketched herein as well as hints for their solutions. As delays occur in many areas, amongst in biology, we considered the motivating example of diabetes. Independently from the theoretic open questions related to realizations and minimal realizations, modelling of diabetes deserves inspection of basic physiology.

The contributions is this research announcement are as follows:

1. A new time delay model is introduced for the subpopulation of patients with diabetes affected by gastroparesis. From the current literature Borri et al. (2017), time-delay models have shown to be suitable to take into account external delays due to measurement devices or information processing. In the present paper, the internal delay is inherent to the organism and may be measured through clinical tests or estimated by means of identification of a glycemic holter.

2. The realization problem was discussed with reference to the class of nonlinear time delay systems and a result concerning minimality was presented.

3. Starting from the realization of time delay system through neutral type input–output equation, a new definition of observability is given which may play an important role with respect to minimal realizations.

Future work will focus on necessary and sufficient conditions for the existence of a realization of minimal order.
Acknowledgement

Emeric Scharbarg acknowledges the financial support by "Fondation pour la Recherche Médicale" under grant FRM FDM20180606247. Claude H. Moog acknowledges the financial support by University of Rome "La Sapienza".

References


APPENDIX: Observability of time-delay systems

Weak observability

In plain words, weak observability requires that at some time $t$, the (eventually delayed) value of the full state can be computed from the knowledge the external signals, i.e. from the output $y(t)$, the input $u(t)$ and their past values. This is the case for the following linear example

$$
\begin{align*}
\dot{x}(t) &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t-1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t-2) \\
y(t) &= x_1(t-1)
\end{align*}
$$

since $x_1(t-1) = y(t)$ and $x_2(t-2) = \dot{y}(t)$.

More generally, using the shift operator $\delta$, weak observability of a linear time delay system

$$
\begin{align*}
\dot{x}(t) &= A(\delta)x(t) + B(\delta)u(t) \\
y(t) &= C(\delta)x(t)
\end{align*}
$$

is easily characterized the full rank condition of the observability matrix

$$
\begin{pmatrix}
C(\delta) \\
C(\delta)A(\delta) \\
\vdots \\
C(\delta)A^{n-1}(\delta)
\end{pmatrix}
$$

where $n = \text{dim} x(t)$.

For a nonlinear time-delay system

$$
\begin{align*}
\dot{x}(t) &= f(x(t), x(t-1),..., u(t), u(t-1),...) \\
y(t) &= h(x(t), x(t-1),...),
\end{align*}
$$
where \( f \) and \( h \) are meromorphic functions, the \( \delta \) notation still makes sense in the linearization

\[
dy(t) = \left[ \frac{\partial h}{\partial x(t)} + \frac{\partial h}{\partial x(t-1)} \delta + \ldots \right] dx(t).
\]

Abusing the notation, the latter will simply be denoted \( \left[ \frac{\partial y}{\partial x}(\delta) \right] \) and is a polynomial in \( \delta \) whose coefficients are meromorphic functions in their arguments. The characterization of weak observability for delayed nonlinear systems is essentially unchanged, but the results are in general only locally valid. This feature follows from nonlinearity and not from delays. Consider for instance

\[
\begin{align*}
\dot{x}(t) &= 0 \\
y(t) &= x^2(t-1).
\end{align*}
\]

The associated "observability matrix" is then \( \left[ \frac{\partial y}{\partial x}(\delta) \right] = [2x(t-1)\delta] \) which has full rank equal to 1 everywhere except at the singularity \( x(\cdot) = 0 \). Thus, away from the singularity, the following "local" observation equations hold: \( x(t-1) = \sqrt{y}/2 \) when \( x(\cdot) > 0 \) and \( x(t-1) = -\sqrt{y}/2 \) when \( x(\cdot) < 0 \).

"Global" strong observability is a matter of case study as illustrated by the system

\[
\begin{align*}
\dot{x}(t) &= 1 \\
y(t) &= x^2(t-1).
\end{align*}
\]

It is still true that \( \frac{\partial y(t)}{\partial x(t-1)} = 2x(t-1) \), but the further information is available: \( \frac{\partial y(t)}{\partial x(t-1)} = \frac{\partial y(t)}{\partial x(t-1)} \frac{\partial y(t)}{\partial x(t-1)} = 2 \). The "extended" observability matrix \( \left[ \begin{array}{c} \frac{\partial y(t)}{\partial x(t-1)} \\ \frac{\partial y(t)}{\partial x(t-1)} \end{array} \right] = \left[ \begin{array}{c} 2x(t-1) \\ 2 \end{array} \right] \) has full rank everywhere and no singularity anymore. The new observability equation \( x(t-1) = \frac{y(t)}{2} \) holds globally everywhere.

The conclusion is that (local) weak observability of a nonlinear time-delay system is characterized by

\[
\text{rank} \left[ \begin{array}{c} (\partial y/\partial x)(\delta) \\ (\partial y/\partial x)(\delta) \\ \vdots \\ (\partial y^{(N)}/\partial x)(\delta) \end{array} \right] = n
\]

for some integer \( N \geq n - 1 \). The above polynomial matrix will display the minimal number singularities.

**Strong observability**

Strong observability states the stronger requirement that at some time \( t \), the full state \( x(t) \) can be computed from the knowledge the external signals, i.e. from the output \( y(t) \), the input \( u(t) \) and their past values.
A linear time delay systems is strongly observable if and only if its observability matrix
\[
\begin{pmatrix}
C(\delta) \\
C(\delta)A(\delta) \\
\vdots \\
C(\delta)A^{n-1}(\delta)
\end{pmatrix}
\]
is unimodular, i.e. it admits a polynomial inverse.
In the nonlinear case, two notions of observability are derived.

**Definition 1.** Garcia-Ramirez et al. (2016) A nonlinear time delay system is strongly observable in a strict sense if
\[
\begin{pmatrix}
(\partial y/\partial x)(\delta) \\
(\partial \dot{y}/\partial x)(\delta) \\
\vdots \\
(\partial y^{(n-1)}/\partial x)(\delta)
\end{pmatrix}
\]
is unimodular.

This condition is weakened as follows.

**Definition 2.** A nonlinear time delay system is strongly observable in a general sense if there exist \(N \geq n-1\) and \(\kappa \geq 0\) such that
\[
\begin{pmatrix}
(\partial y(t)/\partial x)(\delta) \\
\vdots \\
(\partial y(t-\kappa)/\partial x)(\delta) \\
\vdots \\
(\partial y^{(N)}(t)/\partial x)(\delta)
\end{pmatrix}
\]
adsits a polynomial left inverse.

Consider for instance, the nonlinear retarded dynamics
\[
\begin{align*}
\dot{x}(t) &= x(t-1)u(t) \\
y(t) &= x(t) + x(t-1).
\end{align*}
\]

This system does not fulfill the condition in Definition 1 as \([1 + \delta]\) has no polynomial inverse. However, when computing in addition \(y(t-1)\) and \(\dot{y}(t)\), the jacobian matrix in Definition 2 becomes
\[
\begin{pmatrix}
1 + \delta \\
\delta + \delta^2 \\
u(t)\delta + u(t-1)\delta^2
\end{pmatrix}
\]
whose left inverse
\[
\begin{pmatrix}
\frac{1}{u(t-1)} \quad -\frac{1}{u(t-1)}
\end{pmatrix}
\]
is polynomial.