Contents

1 Hidden Input Observability

We consider a mapping $L^2([0,T])^{\otimes m} \to L^2([0,T])^{\otimes p}$ defined by

$$y(t) := C \int_0^t \mathrm{e}^{A(t-\tau)} D w(\tau) \,\mathrm{d}\tau \tag{1}$$

where $w : [0,T] \to \mathbb{R}^m$, $y : [0,T] \to \mathbb{R}^p$, $A \in \mathbb{R}^{n\times n}$, $D \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$.

We assume that in an biological system the functions w and y are defined on a natural interval of time that covers $[0,T]$. This means that for a small ϵ the model could be extended to an interval $(0-\epsilon, T+\epsilon)$ and thus differentiation of w and y at $t=0$ and $t=T$ makes sense. To denote this idea without introducing ϵ we use the notation $[0^-, T^+]$.

It is clear, that if D is not injective, we can find nonzero $w \in \operatorname{kernel} D$ that produce zero output. For this, we will only consider injective D. Furthermore if C is not surjective, we can reduce the number of observables since $(p - \operatorname{rank} C)$ observables are redundant. Thus we will only consider surjective C. Then also $p \leq n$ and $p \leq n$.

We firstly rewrite the problem by expanding (1),

$$y(t) = \sum_{k=0}^{\infty} C A^k D \int_0^t \frac{(t-\tau)^k}{k!} w(\tau) \,\mathrm{d}\tau \quad . \tag{2}$$

Writing $(CA^kD)_{\cdot\mu}$ means the μ-th column and w_μ the μ-th component and if we only consider the ν-th component of y we get

$$y_\nu(t) = \sum_{k=0}^{\infty} \sum_{\mu=1}^{m} \int_0^t \frac{(t-\tau)^k}{k!} w_\mu(\tau) \,\mathrm{d}\tau (CA^kD)_{\nu\mu} \quad . \tag{3}$$

We introduce the coefficients $c_k^{\mu\nu} := (CA^kD)_{\nu\mu}$ to get

$$y_\nu(t) = \sum_{\mu=1}^{m} \sum_{k=0}^{\infty} c_k^{\mu\nu} \int_0^t \frac{(t-\tau)^k}{k!} w_\mu(\tau) \,\mathrm{d}\tau \quad \forall \nu = 1,2,\ldots,p \quad . \tag{4}$$

Here, it is legit to commute the sums, since the infinite sum over k is absolutely convergent.

1.1 Basic Properties

It is now convenient to introduce some operators.

Definition 1

Let $v \in L^2$ be a function and $(c_k)_{k\in\mathbb{N}_0}$ a real-valued sequence. For the sake of readability we write $c = (c_k)_{k\in\mathbb{N}_0}$ for sequences.

1. $\phi_k[v](t) := \int_0^t \frac{(t-\tau)^k}{k!} v(\tau)\,\mathrm{d}\tau$
2. $\Phi_c[v] := \sum_{k=0}^{\infty} c_k \phi_k[v]$

We always exclude the sequence $c_k = 0 \forall k$, since this obviously produces the operator that maps everything to the zero function and thus cannot be injective.

We rewrite (4) as

$$y_\nu(t) = \sum_{\mu=1}^{m} \Phi_{c^{\mu\nu}}[w_\mu](t) \quad \forall \nu = 1,2,\ldots,p \tag{5}$$

and discuss some properties.

Lemma 1: Properties of ϕ_k

1. $\phi_k[v](0) = 0$
2. $\frac{\mathrm{d}}{\mathrm{d}t}\phi_0[v](t) = v(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\phi_k[v](t) = \phi_{k-1}[v](t)$

Proof. The first property is trivial.
Computing $\phi_k[v](t+\Delta t)$ yields

$$\phi_k[v](t+\Delta t) = \int_0^t \frac{(t-\tau+\Delta t)^k}{k!} v(\tau)\,\mathrm{d}\tau + \int_t^{t+\Delta t} \frac{(t-\tau+\Delta t)^k}{k!} v(\tau)\,\mathrm{d}\tau \tag{6}$$

and using $(t+\Delta t)^k = t^k + k t^{k-1}\Delta t + \mathcal{O}(\Delta t^2)$ for $k > 0$ and $(t+\Delta t)^0 = 1$ and a small Δt leads to

$$\phi_k[v](t+\Delta t) \simeq \int_0^t \frac{(t-\tau)^k}{k!} v(t)\,\mathrm{d}\tau + \Delta t \int_0^t \frac{(t-\tau)^{k-1}}{(k-1)!} v(t)\,\mathrm{d}\tau \tag{7}$$

if $k > 0$ and to

$$\int_0^t \frac{(t-\tau)^0}{k!} v(t)\,\mathrm{d}\tau + \Delta t\, v(t) \tag{8}$$

if $k=0$. Comparing this with the definitions of the operators we get $\phi_k[v](t+\Delta t) \simeq \phi_k(t)+\Delta t\phi_{k-1}[v](t)$ and $\phi_0[v](t+\Delta t) \simeq \phi_0[v](t)+\Delta t v(t)$. Taking the limit $\Delta t \to 0$ ends the proof of the derivation rules.

□

1.2 SISO-System

We start with a *Single Input Single Output (SISO)*-System. That means (5) simplifies to

$$y(t) = \Phi_c[v](t) \tag{9}$$

with a real-valued y and v and a scalar $c_k = (CA^k D)$.

Proposition 1

Let v be an integrable function on $[0^-, T^+]$ and let $\Phi_c[v] \equiv 0$. Then

$$\sum_{l=0}^{q} c_l v^{(q-l)}(0) = 0 \quad \forall q \in \mathbb{N}_0 \tag{$\star$}$$

where $v^{(q)}(0)$ denotes the q-th derivative of v at $t=0$.

Proof. Consider

$$\Phi_c[v] = c_0\phi_0[v] + c_1\phi_1[v] + \ldots + c_{q-1}\phi_{q-1}[v] + c_q\phi_q[v] + c_{q+1}\phi_{q+1}[v] + \ldots \tag{10}$$

and the $(q+1)$-th derivative

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{q+1} \Phi_c[v] = c_0 v^{(q)} + c_1 v^{(q-1)} + \ldots + c_{q-1} v^{(1)} + c_q v^{(0)} + c_{q+1}\phi_0[v] + \ldots \tag{11}$$

for short

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{q+1} \Phi_c[v] = \sum_{l=0}^{q} c_l v^{(q-l)} + \sum_{l=0}^{\infty} c_{q+1+l}\phi_l[v] \quad . \tag{12}$$

Evaluating the latter expression at $t=0$ completes the proof. □

To illustrate the idea of the following lemma and theorem we write down some instances of $(\star)$:

$$\begin{array}{lccccccccc}
q=0 & c_0 v^{(0)} & & & & & & & & \\
q=1 & c_0 v^{(1)} & + & c_1 v^{(0)} & & & & & & \\
q=2 & c_0 v^{(2)} & + & c_1 v^{(1)} & + & c_2 v^{(0)} & & & & \\
q=3 & c_0 v^{(3)} & + & c_1 v^{(2)} & + & c_2 v^{(1)} & + & c_3 v^{(0)} & & \\
q=4 & c_0 v^{(4)} & + & c_1 v^{(3)} & + & c_2 v^{(2)} & + & c_3 v^{(1)} & + & c_4 v^{(0)} \\
\vdots & & \ddots & & \ddots & & \ddots & & \ddots & & \ddots
\end{array}$$

For a better readability each $v^{(q)}$ is understood to be evaluated at $t=0$. The structure of this triangle remains the same if $c_0=0$, i.e. the first column vanishes, and if $v^{(0)}=0$, i.e. the diagonal at the top vanishes.

Lemma 2: Induction Step

Assume proposition 1 holds and let c_K be the first nonzero coefficient. If there is a $r \in \mathbb{N}_0$ with $v^{(0)}(0) = v^{(1)}(0) = \ldots = v^{(r-1)}(0) = 0$ then $v^{(r)}(0) = 0$.

Proof. Using $(\star)$ with $q = r+K$ yields

$$\sum_{l=0}^{r+K} c_l v^{(r+K-l)}(0) = c_K v^{(r)}(0) = 0 \tag{13}$$

since all other terms of the sum vanish due to $c_l = 0$ or $v^{(l)} = 0$. This shows that also $v^{(r)}(0) = 0$. □

Theorem 1

Let $v \in L^2([0,T]) \cap C^\infty(0^-, 0^+)$ and c a sequence. Then

$$\Phi_c[v] \equiv 0 \quad \Rightarrow \quad v^{(q)}(0) = 0 \quad \forall q \in \mathbb{N}_0 \quad . \tag{14}$$

Proof. Let c_K be the first nonzero coefficient. Using $(\star)$ with $q = K$ yields

$$c_K v^{(0)} = 0 \tag{15}$$

which shows that $v^{(0)} = 0$. Using lemma 2 completes the inductive proof. □

Corollary 1

If v can be represented by its Taylor-expansion, then $\Phi_c[v] \equiv 0$ implies $v \equiv 0$.

If there is a disjoint union $[0^-, T^+] = I_1 \dot{\cap} I_2 \dot{\cap} \ldots$ such that v has a valid Taylor-expansion on each interval $I_j = (t_{j-1}^-, t_j^+)$, we can argue that $v \equiv 0$ on I_1. This leads to $\phi_k[v](t_1) = 0$ which allows us to get a modification of $(\star)$, written out

$$\sum_{l=0}^{q} c_l v^{(q-l)}(t_1) = 0 \quad \forall q \in \mathbb{N}_0 \quad . \tag{16}$$

Formally this can again be handled as an inductive proof to show, that v must vanish on each interval.

A SISO-system is HIO if and only if there is an $k \in \mathbb{N}_0$ $CA^kD \neq 0$.

1.3 MIMO-Systems

As equation (5) shows, we usually do not have a simple $y(t) = \Phi_c[v](t)$ relation but a sum with different sequences $c^{\mu\nu}$ and functions w_μ. Since summation and differentiation are linear operations, we directly get the following extension of proposition 1.

Proposition 2

Let w_μ integrable on $[0^-, T^+]$ for $\mu = 1, 2, \ldots, m$ and let $c^{\mu\nu}$ be sequences. For each $\nu = 1, 2, \ldots, p$

$$\sum_{\mu=1}^{m} \Phi_{c^{\mu\nu}}[w_\mu] \equiv 0 \quad \Rightarrow \quad \sum_{\mu=1}^{m} \sum_{l=0}^{q} c_l^{\mu\nu} w_\mu^{(q-l)}(0) = 0 \quad \forall q \in \mathbb{N}_0 \tag{$\star\star$}$$

Proof. The proof works analogous to that of proposition 1. □

Whereas theorem 1 holds for any sequence c and function v, proposition 2 allows cancellation of different functions w_μ. We shortly demonstrate, that a fully observed system will always be hidden input observable. Considering a fully observed system with possible hidden inputs on each state, i.e. $C = D = \mathbb{1}$ and $p = n = m$, we directly get $c_0^{\mu\nu} = \delta_{\mu\nu}$. Inserting this into $(\star\star)$ with $q = 0$ yields

$$w_\nu^{(0)} = 0 \quad \text{for} \quad \nu = 1, 2, \ldots, n \tag{17}$$

Following the idea of lemma 2 we proceed with an induction step to get

$$w_\nu^{(q)} = 0 \quad \forall q \in \mathbb{N}_0 \quad \text{for} \quad \nu = 1, 2, \ldots, n \quad . \tag{18}$$

As argued in corollary 1, if we assume that each w_μ can be represented by a Taylor series, we know, that this system is hidden input observable.

1.3.1 Directly Observed Hidden Inputs

To generalize the idea of an fully observed system we find the following lemma in analogy to lemma 2.

Lemma 3: Induction Step

Assume proposition 2 holds and assume for an integer K that $CD = CAD = \ldots = CA^{K-1}D = 0$ and $CA^K D$ has rank m. If there is an $r \in \mathbb{N}$ such that $w_\mu^{(0)}(0) = \ldots = w_\mu^{(r-1)}(0) = 0$ for all $\mu = 1, 2, \ldots, m$, then $w_\mu^{(r)}(0) = 0$ for all μ.

Proof. Using ($\star\star$) with $q = r + K$ yields

$$\sum_{\mu=1}^{m} c_K^{\mu\nu} w_\mu^{(r)}(0) = 0 \tag{19}$$

which is by definition of $c_K^{\mu\nu}$ equivalent to the linear equation

$$CA^K D w^{(r)}(0) = 0 \tag{20}$$

where $w^{(r)}(0)$ is the vector $(w_1^{(r)}(0), \ldots, w_m^{(r)}(0))$. If and only if $\operatorname{rank} CA^K D = m$ this implies $w^{(r)}(0) = 0 \in \mathbb{R}^m$. $\square$

Theorem 2: Directly Observed Hidden Inputs

Let $w \in L^2([0,T])^{\otimes m} \cap C^\infty(0^-, 0^+)^{\otimes m}$ and (A, C, D) the matrices of (1) and assume that $CA^K D$ is the first nonvanishing coefficient matrix.
If $\operatorname{rank} CA^K D = m$, then

$$y \equiv 0 \quad \Rightarrow \quad w_\mu^{(q)}(0) = 0 \quad \forall q \in \mathbb{N}_0 \quad . \tag{21}$$

Proof. Due to the assumptions we can use proposition 2. Equation ($\star\star$) with $q = K$ yields

$$\sum_{\mu=1}^{m} c_K^{\mu\nu} w_\mu^{(0)}(0) = \sum_{\mu=1}^{m} \left(CA^K D\right)_{\nu\mu} w_\mu^{(0)}(0) = 0 \quad \forall \nu \in \{1,2,\dots,p\} \tag{22}$$

which implies $w_\mu^{(0)}(0) = 0 \forall \mu$, since $(CA^K D) : \mathbb{R}^m \to \mathbb{R}^p$ is injective. This allows us to use lemma 3 as an induction step to complete the proof. □

1.3.2 Indirectly Observed Hidden Inputs

In the following example you can see a composition of two HIO systems. Though it should again be HIO, we cannot prove this by using theorem 2.

Example 1

Consider the obviously HIO system

$$A = 0 \quad , \quad D = 1 \quad , \quad C = 1 \quad . \tag{23}$$

Furthermore consider the system

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 1 \end{pmatrix} \tag{24}$$

which yields $CD = 0$ and $CAD = 1$ and therefore is also HIO. If we simply write these two systems into one system

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{25}$$

we get $\operatorname{rank} CD = 1$ which means we cannot apply theorem 2 to prove HIO.

To get an extension of theorem 2 we need a new construct.

Definition 2

Consider CA^kD and define an index set $\mathscr{I}^k$ for each $k \in \mathbb{N}_0$ by

$$\mu \in \{1,2,\dots,m\} \quad \text{is element of} \quad \mathscr{I}^k \quad \Leftrightarrow \quad (CA^kD)_{.\mu} \neq 0 \quad . \tag{26}$$

Here $(CA^kD)_{.i}$ denotes the i-th column.

For a better readability $\mathscr{R}^k := \mathscr{I}^k \backslash (\mathscr{I}^{k-1} \cup \ldots \cup \mathscr{I}^0)$.

The *reduced matrices* $(CA^kD)^*$ are defined by

$$(CD)^* = \left[CD_{.\mu}\right]_{\mu\in\mathscr{I}^0} \tag{27}$$

and

$$(CA^kD)^* = \left[(CA^{k-1}D)^* \left[CA^kD_{.\mu}\right]_{\mu\in\mathscr{R}^k}\right] \quad . \tag{28}$$

The square brackets mean the composition of columns to a new matrix.

Example 2

Let us again consider the latter example.

$$CD = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad CAD = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{29}$$

We find $\mathscr{I}^0 = \{1\}$ and $\mathscr{I}^1 = \{2\}$. The reduced matrices are

$$(CD)^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad (CAD)^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{30}$$

The first matrix contains only the nonzero columns of CD. The second matrix is the composition of $(CD)^*$ and the columns of (CAD) with index in $\mathscr{I}^1 \backslash \mathscr{I}^0 = \{2\}$. In this example it is the composition of the first column of CD and the second column of CAD.
We can see that both reduced matrices have full column rank and that the reduced matrix of the highest power, $(CAD)^*$, has rank m.

Using this construct we can generalize theorem 2.

Theorem 3: Indirectly Observed Hidden Inputs

Let $w:[0^-,T^+]\to\mathbb{R}^m$ be element of $L^2([0,T])^{\otimes m}\cap C^\infty(0^-,0^+)^{\otimes m}$ and let (A,C,D) be the matrices of a linear dynamic system (1).
If each reduced matrix $(CA^kD)^*$ is injective, i.e. has full column rank, and if there is an integer K such that $(CA^KD)^*$ has rank m, then

$$y\equiv 0 \quad\Rightarrow\quad w_\mu^{(q)}(0)=0 \quad \forall q\in\mathbb{N}_0 \quad . \tag{31}$$

Proof. Once more we prove by induction. Writing $w_\mu^{(q)}$ always means evaluated at $t=0$. Let us set $y\equiv 0$.

1. Proposition 2 with $q=0$ yields

$$\sum_{\mu=1}^m c_0^{\mu\nu}w_\mu^{(0)}=0\;\forall\nu \quad\text{which implies}\quad \sum_{\mu=1}^m w_\mu^{(0)}CD_{.\mu}=0 \quad . \tag{32}$$

By definition of $\mathscr{I}^0$ we get

$$\sum_{\mu\in\mathscr{I}^0} w_\mu^{(0)}(CD)_{.\mu}=0 \quad . \tag{33}$$

Since this is a linear combination of the columns of an injective matrix $(CD)^*$ we get

$$w_\mu^{(0)}=0 \quad \forall\mu\in\mathscr{I}^0 \quad . \tag{34}$$

2. Before we do the induction step, let us use proposition 2 with $q=1$ to get

$$\sum_{\mu\in\mathscr{I}^0} c_0^{\mu\nu}w_\mu^{(1)}+\sum_{\mu\in\mathscr{R}^1} c_1^{\mu\nu}w_\mu^{(0)}=0 \quad \forall\nu \quad . \tag{35}$$

Again written as a vector equation this is equivalent to

$$\sum_{\mu\in\mathscr{I}^0} w_\mu^{(1)}(CD)_{.\mu}+\sum_{\mu\in\mathscr{R}^1} w_\mu^{(0)}(CAD)_{.\mu}=0 \quad . \tag{36}$$

This again is a linear combination of columns of $(CAD)^*$. Since $(CAD)^*$ is injective we get

$$w_\mu^{(0)}=0 \quad \forall\mu\in\mathscr{I}^0\cup\mathscr{R}^1 \quad\text{and}\quad w_\mu^{(1)}=0 \quad \forall\mu\in\mathscr{I}^0 \quad . \tag{37}$$

We note that $\mathscr{R}^i\cup\mathscr{I}^{i-1}\cup\ldots\cup\mathscr{I}^0$ is equal to $\mathscr{I}^i\cup\mathscr{I}^{i-1}\cup\ldots\cup\mathscr{I}^0$.

3. Following the idea of step 2 we do the induction step:
   Assume we have found an integer r such that

$$\begin{aligned} w_\mu^{(0)} &= 0 \ \forall \mu \in \bigcup_{\rho=0}^{r} \mathscr{I}^\rho \\ w_\mu^{(1)} &= 0 \ \forall \mu \in \bigcup_{\rho=0}^{r-1} \mathscr{I}^\rho \\ &\vdots \\ w_\mu^{(r)} &= 0 \ \forall \mu \in \mathscr{I}^0 \end{aligned}$$

   then use proposition 2 with $q = r+1$ to get for all ν

$$\sum_{\mu=1}^{m} c_0^{\mu\nu} w_\mu^{(r+1)} + \sum_{\mu\notin\mathscr{I}^0} c_1^{\mu\nu} w_\mu^{(r)} + \ldots + \sum_{\mu\notin\mathscr{I}^0\cup\ldots\cup\mathscr{I}^r} c_{r+1}^{\mu\nu} w_\mu^{(0)} = 0 \quad . \tag{38}$$

   As a vector equation we get

$$\sum_{\mu=1}^{m} w_\mu^{(r+1)} (CD)_{.\mu} + \sum_{\mu\notin\mathscr{I}^0} w_\mu^{(r)} (CAD)_{.\mu} + \ldots + \sum_{\mu\notin\mathscr{I}^0\cup\ldots\cup\mathscr{I}^r} w_\mu^{(0)} (CA^{r+1}D)_{.\mu} = 0 \quad . \tag{39}$$

   We easily see that the combination $\mu \in \mathscr{I}^{\rho+1}$ and $\mu \notin \mathscr{I}^\rho \cup \ldots \cup \mathscr{I}^0$ is the same as $\mu \in \mathscr{R}^{q+1}$. Using the knowledge of zero columns we get

$$\sum_{\mu\in\mathscr{I}^0}^{m} w_\mu^{(r+1)} (CD)_{.\mu} + \sum_{\mu\in\mathscr{R}^1} w_\mu^{(r)} (CAD)_{.\mu} + \ldots + \sum_{\mu\in\mathscr{R}^{r+1}} w_\mu^{(0)} (CA^{r+1}D)_{.\mu} = 0 \quad . \tag{40}$$

   This is a linear combination of columns of $(CA^{r+1}D)^*$ which is injective. Therefore

$$\begin{aligned} w_\mu^{(0)} &= 0 \ \forall \mu \in \bigcup_{\rho=0}^{r+1} \mathscr{I}^\rho \\ w_\mu^{(1)} &= 0 \ \forall \mu \in \bigcup_{\rho=0}^{r} \mathscr{I}^\rho \\ &\vdots \\ w_\mu^{(r)} &= 0 \ \forall \mu \in \bigcup_{\rho=0}^{1} \mathscr{I}^\rho \\ w_\mu^{(r+1)} &= 0 \ \forall \mu \in \mathscr{I}^0 \end{aligned}$$

   This completes the induction.

4. Finally we show that the maximum $\bigcup_\rho \mathscr{I}^\rho = \{1,2,\dots,m\}$ is always reached. Assume there is an $m' < m$ such that $|\bigcup_\rho \mathscr{I}^\rho| = m'$ for all $\rho \geq \rho_0$ but there is no ρ such that $|\bigcup_\rho \mathscr{I}^\rho| > m'$. Here $|.|$ denotes the number of elements.
   (a) By construction $(CD)^*$ has $|\mathscr{I}^0|$ columns.
   (b) Then $(CAD)^*$ has exactly $|\mathscr{I}^0| + |\mathscr{R}^1|$ columns.
   (c) $(CA^kD)^*$ has exactly $|\mathscr{I}^0| + |\mathscr{R}^1| + \ldots + |\mathscr{R}^k|$ columns.
   (d) There is an integer K such that $(CA^KD)^*$ has column rank m. This means $|\mathscr{R}^{K+1}| = 0$.

□

We complete the MIMO-systems with the last corollary in analogy to corollary 1.

Corollary 2

If there are a disjoint unions $[0^-, T^+] = I_1^\mu \dot{\cap} I_2^\mu \dot{\cap} \ldots$ such that each w_μ can be represented by its Taylor-expansion on each interval $I_j^\mu = ((t_{j-1}^\mu)^-, (t_j^\mu)^+)$ and if theorem 2 or theorem 3 holds, then the MIMO-system is HIO.

To have any CA^kD of rank m, we need $p \geq m$.

1.3.3 Examples

Example 3

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{41}$$

A SISO-system and $CAD = 1$ thus it is HIO.

Example 4

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{42}$$

A SISO-system and $CA^kD = 0$ for any integer k thus it is not HIO.

Example 5

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{43}$$

A MIMO-system and CD is the 2×2 identity matrix. Thus it is HIO.

2 Zero Dynamics

We follow the ideas of [1]. The *zero dynamics (ZD)* of a linear system (A, D, C) is a set of triplets $(x, w, y) : [0, T] \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ that solve the system equations and produce $y \equiv 0$.

Assume we have found such an triplet $(x^*, w^*, 0)$. Due to linearity we can perturb any other triplet (x, w, y) that solves the system equations by a ZD triplet, since $(x + x^*, w + w^*, y)$ again solves the system equations and produces the same output y as the unperturbed system.

Definition 3

A subset V of $\mathbb{R}^n$ is called (A, D) invariant if

$$AV \subseteq V + \operatorname{Im} D \quad . \tag{44}$$

For the largest (A, D) invariant V we write

$$\mathfrak{V} := V \cap \operatorname{kernel} C \quad . \tag{45}$$

We say a system has *trivial zero dynamics*, if the only triplet (x, w, y) with zero output is $(0, 0, 0)$.

Corollary 3

If the matrix D is not injective, then the ZD cannot be trivial. This is clear because we can always choose $x = 0 \in \mathfrak{V}$ and $w \in \operatorname{kernel} D$ which produces zero output.

If $\operatorname{kernel} C \cap \operatorname{kernel} A \neq \{0\}$, then the ZD cannot be trivial. To see this, choose $V = \operatorname{kernel} C \cap \operatorname{kernel} A$. This is clearly (A, D)-invariant and subset of $\operatorname{kernel} C$.

If D is not injective, the system is not HIO and has nontrivial ZD.

Proposition 3

A trajectory $x : [0, T] \to \mathfrak{V}$ that solves the system equations is part of the zero dynamics of the corresponding system, i.e. if we have an initial value $x_0 \in \mathfrak{V}$, it is possible to find an input u such that $x(t) \in \mathfrak{V}$ at all times.

Proof. We do not give a rigorous proof but only make the proposition plausible. A more comprehensive treatment can be found in [1].
Assume $x(0) \in \mathfrak{V}$. We define recursively

$$\begin{aligned} Ax(0) &= \gamma_0 + \lambda_0 \\ A\gamma_i &= \gamma_{i+1} + \lambda_{i+1} \end{aligned}$$

where each $\gamma_i \in \mathfrak{V}$ and each $\lambda_i \in \operatorname{Im} B$. By the definition of $\mathfrak{V}$ we know that this is possible. We do a short inductive proof to show

$$A^k(\gamma_0 + \lambda_0) = \gamma_k + \sum_{l=0}^{k} A^l \lambda_{k-l} \quad . \tag{46}$$

1. $k = 0$ is trivial.

2. Assume the statement is correct.

$$\begin{aligned} A^{k+1}(\gamma_0 + \lambda_0) = A A^k(\gamma_0 + \lambda_0) &= A\left(\gamma_k + \sum_{l=0}^{k} A^l \lambda_{k-l}\right) \\ &= \gamma_{k+1} + \lambda_{k+1} + \sum_{l=0}^{k} A^{l+1} \lambda_{k-l} = \gamma_{k+1} + \lambda_{k+1} + \sum_{l=1}^{k+1} A^l \lambda_{k-(l-1)} \\ &= \gamma_{k+1} + \sum_{l=0}^{k+1} A^l \lambda_{k+1-l} \quad \text{q.e.d.} \end{aligned}$$

Furthermore we know that

$$x(t) = \mathrm{e}^{At} x(0) + \int_0^t \mathrm{e}^{A(t-\tau)} B w(\tau)\, \mathrm{d}\tau \quad . \tag{47}$$

Expanding the exponentials and shifting the index in the $x(0)$-term yields

$$x(t) = x(0) + \sum_{k=0}^{\infty} A^k \left(\frac{t^{k+1}}{(k+1)!}(\gamma_0 + \lambda_0) + \int_0^t \frac{(t-\tau)^k}{k!} B w(\tau)\, \mathrm{d}\tau \right) \tag{48}$$

and using the letter statement

$$x(t) = x(0) + \sum_{k=0}^{\infty} \left\{ \frac{t^{k+1}}{(k+1)!} \left(\gamma_k + \sum_{l=0}^{k} A^l \lambda_{k-l} \right) + A^k \int_0^t \frac{(t-\tau)^k}{k!} B w(\tau)\, \mathrm{d}\tau \right\} \quad . \tag{49}$$

Applying C leads to

$$y(t) = C \sum_{k=0}^{\infty} \left\{ \frac{t^{k+1}}{(k+1)!} \sum_{l=0}^{k} A^l \lambda_{k-l} + A^k \int_0^t \frac{(t-\tau)^k}{k!} B w(\tau) \, \mathrm{d}\tau \right\} \tag{50}$$

At order $CA^{k'}$ we have

$$\sum_{r=0}^{\infty} \frac{t^{k'+1+r}}{(k'+1+r)!} \lambda_r + \int_0^t \frac{(t-\tau)^{k'}}{k'!} B w(\tau) \, \mathrm{d}\tau \quad . \tag{51}$$

Since each λ_r is element of $\operatorname{Im} B$ we may find a w that suppresses each term in the summation separately. Without proof we assume that this is always possible.
As a special case we consider $\operatorname{Im} B \subseteq \operatorname{kernel} C$, which makes the problem trivial. $\square$

2.1 Zero Dynamics vs. HIO

In the following example we see, that HIO and ZD are closely related but not equivalent.

Example 6

Consider a dynamic system with the matrices

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad . \tag{52}$$

For $\mathfrak{V} := \text{span}\{(0,0,1)^T\}$ we find $A\mathfrak{V} = \text{span}\{(1,0,0)^T\}$ is a subset of $\text{Im}\, B$ and thus $\mathfrak{V}$ is (A,B)-invariant.
For instance the set

$$x^*(0) = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} \quad \text{and} \quad w^*(t) = \begin{pmatrix} -a \\ 0 \end{pmatrix} \tag{53}$$

leads to a constant

$$x^*(t) = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} \tag{54}$$

which produces zero output $y = Cx^*$ with nonvanishing hidden inputs w^*. This shows that the zero dynamics of the system is not trivial.
At the same time we find that CB is injective which means the system is HIO and thus zero output should imply zero hidden intput.

As illustrated by the example, HIO does not imply trivial ZD but it is easy to see that trivial ZD imply HIO.

Proposition 4: HIO is necessary for Trivial ZD

For a linear system (A,D,C)

$$\text{trivial ZD} \quad \Rightarrow \quad \text{HIO} \quad . \tag{55}$$

Proof. Since we assume trivial ZD, by corollary 3 we know that D is injective.
If $\text{kernel}\, A \cap \text{kernel}\, C$ is nontrivial, the ZD would again be nontrivial.
By the definition of $\mathfrak{V}$ we know that

$$\forall \quad x \in \mathfrak{V} \quad \exists \quad (\tilde{x}, w) \in \mathfrak{V} \times \mathbb{R}^m \quad | \quad Ax = \tilde{x} + Dw \quad . \tag{56}$$

We make the assumption that CD is not injective. Then $\text{kernel}\, C \cap \text{Im}\, D \supset \{0\}$. That means

$$\exists \quad w \neq 0 \in \mathbb{R}^m \quad | \quad Dw \neq 0 \quad \text{and} \quad CDw = 0 \quad . \tag{57}$$

If we now chose such an w and set $\xi := Dw$ and $V = \operatorname{span} \xi$ we know

$$A\xi = Bw \tag{58}$$

$\square$

References

[1] Achim Ilchmann Thomas Berger. “Zero dynamics of time-varying linear systems”. In: (2010).