Available online at www.sciencedirect.com
ScienceDirect
Journal of the Franklin Institute 347 (2010) 533–544
Journal of The Franklin Institute
www.elsevier.com/locate/jfranklin

# Embedding singular linear system with initial condition into distribution space ☆

Zhibin Yan

*Natural Science Research Center, Science Park, P. O. Box 3041, Harbin Institute of Technology, No. 2, Yikuang Street, Harbin 150080, China*

Received 20 May 2008; received in revised form 28 October 2009; accepted 16 November 2009

**Abstract**

For inconsistent initial value problem of singular linear system, passing from classical solution to distributional solution is formulated as a process of embedding the system with initial condition into distribution space. Three technical notions, function space on $\mathbb{R}_+$ with given initial value, consistent extension of the pair of the function space and input, and order exchange of the two operations of embedding and differentiating, result in a rigorous definition of distributional solution.


*Keywords:* Singular system; Descriptor system; Differential algebraic equations

## 1. Introduction

In singular linear system, bounded and sufficiently smooth input can excite impulse when the system is given inconsistent initial condition [1,2]. Such impulse phenomenon distinguishes the singular linear system from the standard one [3–8].

Under a regularity assumption, singular linear system can be transformed into two subsystems through Weierstrass decomposition [9,2]. One has the form of the standard linear system, called finite (or slow) subsystem, which does not exhibit impulse behavior; another is of the form (1.1a) called infinite (or fast) subsystem [1, p. 11]. For an arbitrary

☆This work was supported by the Harbin Institute of Technology Science Foundation under Grant HITC200712.
*E-mail address:* zbyan@hit.edu.cn

0016-0032/$32.00 
doi:10.1016/j.jfranklin.2009.11.008

initial condition (1.1b), in general, the initial value problem of singular linear differential equation

$$N\dot{x}(t) = x(t) + Bu(t), \quad t \geq 0, \tag{1.1a}$$

$$x(0) = x_0 \tag{1.1b}$$

has no solution in the sense of classical differentiable function, where $N \in \mathbb{R}^{n\times n}$ is nilpotent, $B \in \mathbb{R}^{n\times r}$, $u$ is a sufficiently smooth function mapping $\mathbb{R}_+$ to $\mathbb{R}^r$, and $x_0 \in \mathbb{R}^n$. On the other hand, a mathematical description of impulse uses Dirac delta generalized function or distribution [10]. So a theory interpreting the impulse phenomenon can be established by introducing some generalized or distributional solution to Eq. (1.1). See [2] for details, or [11] for a brief. We denote the index of nilpotency of $N$ by $h$, that is

$$h = \min\{k : k \geq 1, N^k = 0\}. \tag{1.2}$$

Based on the Jordan canonical form of the nilpotent matrix $N$, the system (1.1a) can be decomposed into some subsystems further, among which those corresponding to the Jordan blocks with index of nilpotency one, if exist, are called to be nondynamic [1, p. 12].

To problem (1.1), an extensively used distributional solution formula in the literature is [1, Eq. (3.25), p. 11], [2, Eqs. (1–4.9), p. 18]

$$x(t) = -\sum_{i=0}^{h-1} N^i B u^{(i)}(t) - \sum_{i=1}^{h-1} \delta^{(i-1)}(t) N^i x_0, \tag{1.3}$$

with $u^{(i)}$ understood as distributional derivative of $i$-th order. In [12, p. 224], [11,13], another distributional solution formula

$$x(t) = -\sum_{i=0}^{h-1} N^i B u^{(i)}(t) - \sum_{k=1}^{h-1} \delta^{(k-1)}(t) N^k \left\{ x_0 + \sum_{i=0}^{h-1} N^i B u^{(i)}(0^+) \right\} \tag{1.4}$$

is obtained by Laplace transform. The formula (1.4) is used to clarify many problems about impulse analysis of singular systems [11,15]. A similar formula for time varying case, in which Laplace transform is not available, is also obtained in [14].

To see the motivation of this paper, we need to observe further the following two existing approaches to define "distributional solution" of (1.1) in the literature.

(1) *Laplace transform approach to distributional solution*: It says the solution should be the Laplace inverse transform of $X(s)$, where $X(s)$ satisfies [16]

$$(sN - I)X(s) = B\mathcal{L}[u](s) + Nx_0. \tag{1.5}$$

In following, we use $U$ to denote the Laplace transform $\mathcal{L}[u]$. By the nilpotency of $N$, it is easy to see that

$$X(s) = (sN - I)^{-1}(BU(s) + Nx_0) = -\sum_{i=1}^{h-1} s^{i-1} N^i x_0 - \sum_{i=0}^{h-1} s^i N^i BU(s), \tag{1.6}$$

which does not have Laplace inverse transform generally in the sense of classical function (i.e., function mapping $\mathbb{R}_+$ to $\mathbb{R}^n$). Besides using the familiar formula $\mathcal{L}[\delta^{(i)}](s) = s^i$, regarding (1.3) as the Laplace inverse transform of (1.6) stems from the using of the formula

$$\mathcal{L}[u^{(i)}](s) = s^i U(s), \quad i = 1, 2, \ldots, \tag{1.7}$$

in some questionable sense of distribution. On the other hand, passing from time domain Eq. (1.1) to the frequency domain one (1.5) uses the different formula

$$\mathcal{L}[\dot{x}](s) = sX(s) - x_0. \tag{1.8}$$

So the distributional solution (1.3) cannot justify itself coherently according to the Laplace transform approach. In [12, p. 224], [13,11], the formula

$$\mathcal{L}[u^{(i)}](s) = s^i U(s) - \sum_{k=0}^{i-1} s^k u^{(i-1-k)}(0^+), \tag{1.9}$$

instead of (1.7), is used to arrive at the distributional solution (1.4) through some computation. Here are also some questionable points. For example, in (1.9), $u$ is a classical sufficiently differentiable function, and the Laplace transform is understood as a usual Lebesgue integral. In this sense, (1.9) is obviously correct, but it cannot interpret (1.8) for a *distribution* $x$. Sometimes the initial condition (1.1b) is interpreted as [16]

$$x(0^-) = x_0, \tag{1.10}$$

the Laplace transform is understood as $\mathcal{L}_-$ to operate on functions which may contain distribution, and the formula for derivative is written as [17, p. 717]

$$\mathcal{L}_-[u^{(i)}](s) = s^i U(s) - \sum_{k=0}^{i-1} s^k u^{(i-1-k)}(0^-).$$

But why should these $u^{(k)}(0^-)$, $k = 0, 1, \ldots, h$ be zero? Note that the original problem (1.1) is independent of whether $u$ has definition on $t<0$ or not and how it is defined if it has. We see that the Laplace transform seems to be used randomly in this approach to distributional solution.

(2) *Embedding approach to distributional solution*: It suggests that we should be considering (see [2, p. 18], [18,19])

$$N\mathcal{D}_{\mathrm{d}}x = x + Bu + \delta N x_0 \tag{1.11}$$

in some distribution space, where $\mathcal{D}_{\mathrm{d}}x$ represents the derivative of $x$ in the sense of distribution (see (2.9) or [10] for detail). Passing from the Eq. (1.1) in classical function space to the Eq. (1.11) in distribution space is some embedding process, so we name this the embedding approach to distributional solution. But why should the initial condition be included in this direct manner? How does the Eq. (1.11) export classical solution as special case if Eq. (1.1) has classical solution? What a specific distribution space is used for (1.11)? In [19], the test function space of the infinitely differentiable functions with upper-bounded support and the corresponding distribution space are used. Are the more commonly used test function space and the corresponding distribution space, i.e., the space of the infinitely differentiable functions with compact support and its dual space, respectively, not suitable for the embedding? The mathematical meaning of (1.11) and its relationship to the original Eq. (1.1) need to explore further.

To clarify various kinds of questionable points like above mentioned, a rigorous distributional solution theory for the initial value problem (1.1) is still expected. In the present paper, we formulate the embedding process rigorously and give such a theory based on embedding. The key ideas helping to dissolve the difficulty are as follows.

(1) Write the classical initial value problem (1.1) into a special differential operator form, where the initial value constraint is automatically involved in beforehand and does not need additional treatment. If there is no initial value constraint (1.1b), it will be trivial to solve the Eq. (1.1a).
(2) Realize that in order to embed into distribution space, the known function $u$ and the unknown function $x$ need to be extended consistently with each other to have definition for $t<0$. Zero extension is only one kind of such extensions.
(3) In the process of embedding, exchanging the operation order of embedding and differentiating will produce an extra term. This explains the occurrence of the mysterious term in (1.11).

## 2. Formalization of embedding

For an interval $I \subseteq \mathbb{R}$, Symbol $\mathcal{C}^k(I, \mathbb{R}^m)$ denotes the set of all $k$ times differentiable functions mapping $I$ to $\mathbb{R}^m$. From now on we consider

$$N\dot{x}(t) = x(t) + f(t), \quad t \geq 0, \tag{2.1a}$$

$$x(0) = x_0, \tag{2.1b}$$

where $x_0 \in \mathbb{R}^n$ and $f \in \mathcal{C}^h(\mathbb{R}_+, \mathbb{R}^n)$ are fixed.

### *2.1. Reformulation of initial value problem*

**Definition 2.1.** (1). The set

$$\mathcal{X}_{x_0} = \{x : x \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}^n); x(0) = x_0\} \tag{2.2}$$

is called the $\mathcal{C}^1$ function space on $\mathbb{R}_+$ with initial value $x_0$. 2). The classical differential operator $\mathcal{D} : \mathcal{X}_{x_0} \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$ is defined according to pointwise differentiation as

$$\mathcal{D}[x](t) = \begin{cases} \dot{x}(t), & t > 0, \\ \dot{x}(0^+), & t = 0. \end{cases}$$

Note that according to the definition of $\mathcal{X}_{x_0}$, we have

$$\lim_{t \to 0, t > 0} \frac{1}{t}(x(t) - x(0)) = \lim_{t \to 0, t > 0} \dot{x}(t)$$

for any $x \in \mathcal{X}_{x_0}$. The right derivative notation $\dot{x}(0^+)$ represents any one of the two limits. Similarly, the $i$-th right derivative notation $f^{(i)}(0^+)$ for $f \in \mathcal{C}^h(\mathbb{R}_+, \mathbb{R}^n)$ is also well-defined, $i = 1, 2, \ldots, h$.

Now finding the classical solution of the initial value problem (2.1a) can be reformulated as the operator form: find $x \in \mathcal{X}_{x_0}$ such that

$$N\mathcal{D}[x] = x + f.$$

Let $\mathcal{L}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ denote the set of all locally Lebesgue integrable functions from $\mathbb{R}$ to $\mathbb{R}^n$. We want to extend the definition domain of the related functions from $\mathbb{R}_+$ to $\mathbb{R}$, and further to extend the operator $\mathcal{D}$. One maybe thinks that this process should take the following form

$$\tilde{w}(t) = \begin{cases} w(t), & t \geq 0, \\ 0, & t < 0 \end{cases}$$

for $w \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$ of course. However, the following study shows that there exist other more convenient ways.

In following, for a function $g \in \mathcal{L}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ and an interval $I \subset \mathbb{R}$, we use the notation $g|_I$ to represent the restriction of $g$ on $I$, i.e.,

$$g|_I(t) = g(t) \quad \forall t \in I.$$

**Definition 2.2.** A pair $(\tilde{\mathcal{X}_{x_0}}, \tilde{f})$ of $\tilde{\mathcal{X}_{x_0}} \subset \mathcal{L}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ and $\tilde{f} \in \mathcal{L}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ is called an extension of the pair $(\mathcal{X}_{x_0}, f)$, if the following conditions hold:

1. $\tilde{f}|_{\mathbb{R}_+} = f$ and $\tilde{\mathcal{X}_{x_0}}|_{\mathbb{R}_+} = t\{\tilde{x}|_{\mathbb{R}_+} : \tilde{x} \in \tilde{\mathcal{X}_{x_0}}\} = \mathcal{X}_{x_0}$;
2. $\tilde{f}$ and $\tilde{\mathcal{X}_{x_0}}$ are consistent on $t < 0$, i.e., $\tilde{f}|_{(-\infty,0)} \in \mathcal{C}^h((-\infty, 0), \mathbb{R}^n)$, $\lim_{t \to 0, t<0} \tilde{f}^{(i)}(t) = \tilde{f}^{(i)}(0^-)$, $i = 0, 1, \ldots, h$ exist and for any $\tilde{x} \in \tilde{\mathcal{X}_{x_0}}$,
$$N\tilde{x}'(t) = \tilde{x}(t) + \tilde{f}(t), t < 0. \tag{2.3}$$

**Remark 2.1.** From (2.3) it follows that arbitrary $\tilde{x} \in \tilde{\mathcal{X}_{x_0}}$ satisfies

$$\tilde{x}(t) = -\sum_{i=0}^{h-1} N^i \tilde{f}^{(i)}(t) \quad \forall t < 0. \tag{2.4}$$

So an extension $(\tilde{\mathcal{X}_{x_0}}, \tilde{f})$ is uniquely determined by $\tilde{f}|_{(-\infty,0)}$.

**Example 2.1** (*Zero Extension*). Take

$$\tilde{f}(t) = \begin{cases} f(t), & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{2.5}$$

Then

$$\tilde{\mathcal{X}_{x_0}} = \{\tilde{x} : \tilde{x}(t) = 0, \forall t<0; \tilde{x}|_{\mathbb{R}_+} \in \mathcal{X}_{x_0}\}.$$

**Example 2.2** (*f Smooth Extension*). Arbitrarily take $\tilde{f} \in \mathcal{C}^h(\mathbb{R}, \mathbb{R}^n)$ with $\tilde{f}|_{\mathbb{R}_+} = f$, e.g.,

$$\tilde{f}(t) = \begin{cases} f(t), & t \geq 0, \\ \sum_{i=0}^{h} \dfrac{f^{(i)}(0)}{i!} t^i, & t<0. \end{cases}$$

Then

$$\tilde{\mathcal{X}_{x_0}} = \left\{ \tilde{x} : \tilde{x}(t) = -\sum_{i=0}^{h-1} N^i \tilde{f}^{(i)}(t), \forall t<0; \tilde{x}|_{\mathbb{R}_+} \in \mathcal{X}_{x_0} \right\}. \tag{2.6}$$

According to pointwise differentiation, $\mathcal{D}$ can operate on $\tilde{\mathcal{X}_{x_0}}$ as

$$\mathcal{D}[\tilde{x}](t) = \begin{cases} \tilde{x}'(t), & t \neq 0, \\ \tilde{x}'(0^+), & t = 0 \end{cases}$$

for any $\tilde{x} \in \tilde{\mathcal{X}_{x_0}}$. Further, the higher differentiations

$$\mathcal{D}^{(i)}[\tilde{f}](t) = \begin{cases} \tilde{f}^{(i)}(t), & t \neq 0, \\ \tilde{f}^{(i)}(0^+), & t = 0, \end{cases} \tag{2.7}$$

$i = 0, 1, \ldots, h$ are also well-defined. Note that $\tilde{f}^{(i)}(0^+) = f^{(i)}(0^+)$.

**Remark 2.2.** Let $(\tilde{\mathcal{X}_{x_0}}, \tilde{f})$ be an extension of $(\mathcal{X}_{x_0}, f)$. Finding the classical solution of the initial value problem (2.1a) can be reformulated further as: find $\tilde{x} \in \tilde{\mathcal{X}_{x_0}}$ such that

$$N\mathcal{D}[\tilde{x}] = \tilde{x} + \tilde{f}. \tag{2.8}$$

The advantage of introducing the set $\mathcal{X}_{x_0}$ and $\tilde{\mathcal{X}_{x_0}}$ is that the initial value is involved in automatically and then need not additional treatment.

### 2.2. Embedding in distribution space

Let $\mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R}^n)$ be the space of infinitely differentiable functions from $\mathbb{R}$ to $\mathbb{R}^n$ with compact support, and $\mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R}^n)'$ be its dual space, i.e., the distribution space [10]. A distribution $w \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R}^n)'$ then is a linear continuous (relative to a topology [10]) functional on $\mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R}^n)$. The value, a real number, of $w$ on $\lambda \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R}^n)$ will be denoted by $\langle w, \lambda \rangle$. The Dirac delta distribution $\delta \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R})'$ is defined as

$$\langle \delta, \lambda \rangle = \lambda(0)$$

for $\forall \lambda \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R})$. For any distribution $w \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R})'$, its $k$-th distributional derivative $\mathcal{D}_\mathrm{d}^{(k)} w \in \mathcal{C}_C^\infty(\mathbb{R}, \mathbb{R})'$ is defined as

$$\langle \mathcal{D}_\mathrm{d}^{(k)} w, \lambda \rangle = (-1)^k \langle w, \lambda^{(k)} \rangle \tag{2.9}$$

for $\forall \lambda \in \mathcal{C}_C^\infty(\mathbb{R},\mathbb{R})$, where $\lambda^{(k)}$ denotes the $k$-th usual derivative. The embedding map $\mathcal{E}: \mathcal{L}_{\text{loc}}(\mathbb{R},\mathbb{R}^n) \to \mathcal{C}_C^\infty(\mathbb{R},\mathbb{R}^n)'$ is defined as

$$\langle \mathcal{E}z, \lambda \rangle = \int_{-\infty}^{+\infty} z(t)^{\mathrm{T}} \lambda(t)\, dt \tag{2.10}$$

for $\forall z \in \mathcal{L}_{\text{loc}}(\mathbb{R},\mathbb{R}^n)$ and $\forall \lambda \in \mathcal{C}_C^\infty(\mathbb{R},\mathbb{R}^n)$. Here $z(t)^{\mathrm{T}}$ represents the transpose of $z(t)$, and the integral is in the sense of Lebesgue.

Let $(\tilde{\mathcal{X}}_{x_0},\tilde{f})$ be an extension of $(\mathcal{X}_{x_0},f)$ and $\tilde{x} \in \tilde{\mathcal{X}}_{x_0}$ be a solution of (2.8). Then the two sides of (2.8) belong to $\mathcal{L}_{\text{loc}}(\mathbb{R},\mathbb{R}^n)$ and can be embedded into distribution space as follows

$$\mathcal{E}t(N\mathcal{D}[\tilde{x}]) = \mathcal{E}(\tilde{x} + \tilde{f}).$$

It is obvious that

$$\mathcal{E}(N\mathcal{D}[\tilde{x}]) = N\mathcal{E}\mathcal{D}[\tilde{x}],$$

$$\mathcal{E}(\tilde{x} + \tilde{f}) = \mathcal{E}(\tilde{x}) + \mathcal{E}(\tilde{f}).$$

About the composition $\mathcal{E}\mathcal{D}$ of the two operators $\mathcal{E}$ and $\mathcal{D}$, we have

**Proposition 2.1.** *Let $(\tilde{\mathcal{X}}_{x_0},\tilde{f})$ be an extension of $(\mathcal{X}_{x_0},f)$ and $\tilde{x} \in \tilde{\mathcal{X}}_{x_0}$. Then*

$$\mathcal{E}\mathcal{D}[\tilde{x}] = \mathcal{D}_{\mathrm{d}}\mathcal{E}[\tilde{x}] - \delta \cdot \left\{ x_0 + \sum_{i=0}^{h-1} N^i \tilde{f}^{(i)}(0^-) \right\}.$$

**Proof.** First, (2.4) implies

$$\tilde{x}(0^-) = -\sum_{i=0}^{h-1} N^i \tilde{f}^{(i)}(0^-).$$

For every $\lambda \in \mathcal{C}_C^\infty(\mathbb{R},\mathbb{R}^n)$, we have

$$\begin{aligned}
\langle \mathcal{E}\mathcal{D}[\tilde{x}], \lambda \rangle &= \int_{-\infty}^{+\infty} \mathcal{D}[\tilde{x}](t)^{\mathrm{T}} \lambda(t)\, dt \\
&= \int_{-\infty}^{0} \tilde{x}'(t)^{\mathrm{T}} \lambda(t)\, dt \\
&\quad + \int_{0}^{+\infty} \tilde{x}'(t)^{\mathrm{T}} \lambda(t)\, dt \\
&= -\int_{-\infty}^{0} \tilde{x}(t)^{\mathrm{T}} \lambda'(t)\, dt \\
&\quad + \tilde{x}(0^-)^{\mathrm{T}} \lambda(0) - \int_{0}^{+\infty} \tilde{x}(t)^{\mathrm{T}} \lambda'(t)\, dt - \tilde{x}(0^+)^{\mathrm{T}} \lambda(0) \\
&= -\int_{-\infty}^{+\infty} \tilde{x}(t)^{\mathrm{T}} \lambda'(t)\, dt - (x_0 - \tilde{x}(0^-))^{\mathrm{T}} \lambda(0),
\end{aligned}$$

and

$$\langle \mathcal{D}_{\mathrm{d}}\mathcal{E}[\tilde{x}]-\delta\cdot(x_0-\tilde{x}(0^-)),\lambda\rangle = \langle \mathcal{D}_{\mathrm{d}}\mathcal{E}[\tilde{x}],\lambda\rangle-\langle\delta\cdot(x_0-\tilde{x}(0^-)),\lambda\rangle$$
$$= -\int_{-\infty}^{+\infty}\tilde{x}(t)^{\mathrm{T}}\lambda'(t)dt-(x_0-\tilde{x}(0^-))^{\mathrm{T}}\lambda(0).$$

This completes the proof. $\square$

Now the Eq. (2.8) in the classical function space obtains its distributional version

$$N\mathcal{D}_{\mathrm{d}}\mathcal{E}(\tilde{x}) = \mathcal{E}(\tilde{x}) + \mathcal{E}(\tilde{f}) + \delta\cdot N\left\{x_0 + \sum_{i=0}^{h-1}N^i\tilde{f}^{(i)}(0^-)\right\}. \tag{2.11}$$

This inspires the following definition.

**Definition 2.3.** Let $\tilde{f}\in\mathcal{L}_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^n)$ satisfy $\tilde{f}|_{\mathbb{R}_+}=f$, $\tilde{f}|_{(-\infty,0)}\in\mathcal{C}^h((-\infty,0),\mathbb{R}^n)$, and $\lim_{t\to 0,t<0}\tilde{f}^{(i)}(t)=\tilde{f}^{(i)}(0^-)$, $i=0,1,\ldots,h$ exist. Then solution of the following differential equation

$$N\mathcal{D}_{\mathrm{d}}z = z + \mathcal{E}(\tilde{f}) + \delta\cdot N\left\{x_0 + \sum_{i=0}^{h-1}N^i\tilde{f}^{(i)}(0^-)\right\} \tag{2.12}$$

in distribution space is called *the distributional solution of the initial value problem* (2.1a) *according to the extension* $\tilde{f}$, where $z$ is the unknown distribution to be found.

**Remark 2.3.**

(1) Let $(\tilde{\mathcal{X}}_{x_0},\tilde{f})$ be an extension of $(\mathcal{X}_{x_0},f)$ and $\tilde{x}\in\tilde{\mathcal{X}}_{x_0}$ be a solution of (2.8). Then it follows from (2.11) that the embedding image $\mathcal{E}(\tilde{x})$ of the classical solution $\tilde{x}$ is a distributional solution. In this sense, our distributional solution returns to classical solution as the latter exists.

(2) An $\tilde{f}$ satisfying the requirement in Definition 2.3 will be called an extension of $f$ for short in following.

(3). For given extension $\tilde{f}$, solution of (2.12), if exists, is unique. This follows from the following simple computation. Let $z_1$ and $z_2$ be two solutions, and Let $w=z_1-z_2$. Then $N\mathcal{D}_{\mathrm{d}}w=w$ and then $w=N\mathcal{D}_{\mathrm{d}}(N\mathcal{D}_{\mathrm{d}}w)=N^2(\mathcal{D}_{\mathrm{d}}^{(2)}w)=\cdots=N^h(\mathcal{D}_{\mathrm{d}}^{(h)}w)=0$. Further, for different extensions, although the corresponding solutions of (2.12) are different, they will have the same restriction on $t\geq 0$ (see Remark 3.1 for precise meaning).

## 3. Distributional solution formula and uniqueness

**Lemma 3.1.** *For* (2.12) *we have*

$$z = -\sum_{i=0}^{h-1}N^i\mathcal{D}_{\mathrm{d}}^{(i)}\mathcal{E}(\tilde{f})-\sum_{k=1}^{h-1}\delta^{(k-1)}N^k\left\{x_0+\sum_{i=0}^{h-1}N^i\tilde{f}^{(i)}(0^-)\right\}. \tag{3.1}$$

**Proof.** Differentiating in the sense of distribution and multiplying two sides of (2.12) by $N$ from left $k$ times, we get

$$N^{k+1}\mathcal{D}_{\mathrm{d}}^{(k+1)}z = N^k\mathcal{D}_{\mathrm{d}}^{(k)}z + N^k\mathcal{D}_{\mathrm{d}}^{(k)}\mathcal{E}(\tilde{f}) + \delta^{(k)} \cdot N^{k+1}\left\{x_0 + \sum_{i=0}^{h-1} N^i\tilde{f}^{(i)}(0^-)\right\}, \quad k = 0,1,2,\ldots,h-1.$$

noting that $N^h = 0$, the sum of the above all equations gives the result immediately. □

**Lemma 3.2.** $\mathcal{D}_{\mathrm{d}}^{(i)}\mathcal{E}(\tilde{f}) = \mathcal{E}(\mathcal{D}^{(i)}\tilde{f}) + \sum_{k=0}^{i-1}\delta^{(k)}(f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-))$.

**Proof.** For any test function $\lambda \in \mathcal{C}_C^\infty(\mathbb{R},\mathbb{R}^n)$, we have

$$\begin{aligned}\langle \mathcal{D}_{\mathrm{d}}^{(i)}\mathcal{E}(\tilde{f}),\lambda\rangle &= (-1)^i\langle \mathcal{E}(\tilde{f}),\lambda^{(i)}\rangle \\ &= (-1)^i\int_{-\infty}^{+\infty}\tilde{f}(t)^{\mathrm{T}}\lambda^{(i)}(t)\,dt \\ &= (-1)^i\int_0^{+\infty} f(t)^{\mathrm{T}}\lambda^{(i)}(t)\,dt + (-1)^i\int_{-\infty}^0\tilde{f}(t)^{\mathrm{T}}\lambda^{(i)}(t)\,dt,\end{aligned}$$

$$\begin{aligned}\langle \mathcal{E}(\mathcal{D}^{(i)}\tilde{f}),\lambda\rangle &= \int_{-\infty}^{+\infty}(\mathcal{D}^{(i)}\tilde{f})^{\mathrm{T}}\lambda(t)\,dt \\ &= \int_0^{+\infty} f^{(i)}(t)^{\mathrm{T}}\lambda(t)\,dt + \int_{-\infty}^0 \tilde{f}^{(i)}(t)^{\mathrm{T}}\lambda(t)\,dt\end{aligned}$$

and

$$\begin{aligned}\langle \delta^{(k)},\lambda\rangle &= (-1)^k\langle \delta,\lambda^{(k)}\rangle \\ &= (-1)^k\lambda^{(k)}(0).\end{aligned}$$

Then the result follows from the $i$ times successive applications of integration method by parts to $\int_0^{+\infty} f(t)^T\lambda^{(i)}(t)\,dt$ and $\int_{-\infty}^0 \tilde{f}(t)^T\lambda^{(i)}(t)\,dt$. □

**Theorem 3.1.** *Let $\tilde{f}$ be an extension of $f$. Then the distributional solution of the initial value problem* (2.1a) *according to $\tilde{f}$ is*

$$z = -\mathcal{E}\left(\sum_{i=0}^{h-1} N^i\mathcal{D}^{(i)}(\tilde{f})\right) - \sum_{k=1}^{h-1}\delta^{(k-1)}N^k\left\{x_0 + \sum_{i=0}^{h-1} N^i f^{(i)}(0^+)\right\}. \tag{3.2}$$

**Proof.** By Lemma 3.2, we have

$$\begin{aligned}-\sum_{i=0}^{h-1} N^i\mathcal{D}_{\mathrm{d}}^{(i)}\mathcal{E}(\tilde{f}) &= -\sum_{i=0}^{h-1} N^i\left\{\mathcal{E}(\mathcal{D}^{(i)}\tilde{f}) + \sum_{k=0}^{i-1}\delta^{(k)}(f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-))\right\} \\ &= -\mathcal{E}\left(\sum_{i=0}^{h-1} N^i\mathcal{D}^{(i)}\tilde{f}\right) - \sum_{i=0}^{h-1} N^i\left\{\sum_{k=0}^{i-1}\delta^{(k)}(f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-))\right\}.\end{aligned} \tag{3.3}$$

However,

$$
\begin{aligned}
&\sum_{i=0}^{h-1} N^i \left\{ \sum_{k=0}^{i-1} \delta^{(k)} (f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-)) \right\} \\
&= \sum_{i=0}^{h-1} \sum_{k=0}^{i-1} (\delta^{(k)} N^{k+1}) N^{i-(k+1)} (f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-)) \\
&= \sum_{k=0}^{h-2} (\delta^{(k)} N^{k+1}) \left\{ \sum_{i=k+1}^{h-1} N^{i-(k+1)} (f^{(i-(k+1))}(0^+) - \tilde{f}^{(i-(k+1))}(0^-)) \right\} \\
&= \sum_{l=1}^{h-1} (\delta^{(l-1)} N^l) \left\{ \sum_{i=l}^{h-1} N^{i-l} (f^{(i-l)}(0^+) - \tilde{f}^{(i-l)}(0^-)) \right\} \\
&= \sum_{l=1}^{h-1} (\delta^{(l-1)} N^l) \left\{ \sum_{j=0}^{h-1-l} N^j (f^{(j)}(0^+) - \tilde{f}^{(j)}(0^-)) \right\} \\
&= \sum_{l=1}^{h-1} (\delta^{(l-1)} N^l) \left\{ \sum_{j=0}^{h-1} N^j (f^{(j)}(0^+) - \tilde{f}^{(j)}(0^-)) \right\},
\end{aligned}
\tag{3.4}
$$

where the last equality follows from $N^h = 0$. Substituting (3.3) and then (3.4) in (3.1), we obtain the result immediately. $\square$

**Remark 3.1.** Comparing with (3.1), in the formula (3.2), only the first term has relation to the extension manner, which is the embedding image of the classical function $\sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} \tilde{f} \in \mathcal{L}_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$. Because the function $\sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} \tilde{f}$ has the same restriction $\sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} f$ on $\mathbb{R}_+$ for all extensions obviously, in this sense, we say that **the distributional solution of (2.1a) is unique**.

If we denote the embedding image $\mathcal{E}(\sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} \tilde{f})$ simply by $\sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} \tilde{f}$ itself, then the formula (3.2) can be written simply as

$$z = -\left( \sum_{i=0}^{h-1} N^i \mathcal{D}^{(i)} \tilde{f} \right) - \sum_{k=1}^{h-1} \delta^{(k-1)} N^k \left\{ x_0 + \sum_{i=0}^{h-1} N^i f^{(i)}(0^+) \right\},$$

and further

$$z(t) = -\sum_{i=0}^{h-1} N^i f^{(i)}(t) - \sum_{k=1}^{h-1} \delta^{(k-1)}(t) N^k \left\{ x_0 + \sum_{i=0}^{h-1} N^i f^{(i)}(0^+) \right\}, \quad t \geq 0, \tag{3.5}$$

where $\mathcal{D}^{(i)}$ is in the sense of classical pointwise differentiation (see (2.7)).

## 4. Discussion of the two examples

We have proved different extensions result in the same distributional solution to (2.1a). But it is still very interesting to see the two examples of extension.

### 4.1. Zero extension

In this case, $\tilde{f}$ is defined by (2.5). So we have

$$\tilde{f}^{(i)}(0^-) = 0$$

and (2.12) becomes

$$N\mathcal{D}_{\mathrm{d}}z = z + \mathcal{E}(\tilde{f}) + \delta \cdot Nx_0.$$

This should be the precise meaning of (1.11). Although this special approach seems to be the most natural way to define distributional solution, the solving process will involve somewhat complicated computing (like Lemma 3.2 and Eq. (3.4)) which may make to arrive at the final formula (3.2) some difficult.

### 4.2. $f$ Smooth extension

For this extension, since $\tilde{f} \in \mathcal{C}^h(\mathbb{R}, \mathbb{R}^n)$, we see

$$\tilde{f}^{(i)}(0^-) = f^{(i)}(0^+).$$

So (2.12) becomes

$$N\mathcal{D}_{\mathrm{d}}z = z + \mathcal{E}(\tilde{f}) + \delta \cdot N\left\{x_0 + \sum_{i=0}^{h-1} N^i f^{(i)}(0^+)\right\}. \tag{4.1}$$

For smooth function $\tilde{f} \in \mathcal{C}^h(\mathbb{R}, \mathbb{R}^n)$, it is very easy to see

$$\mathcal{D}_{\mathrm{d}}^{(i)}\mathcal{E}(\tilde{f}) = \mathcal{E}(\mathcal{D}^{(i)}\tilde{f}).$$

So arriving at the final formula (3.2) from (4.1) will have not any suspense.

**Remark 4.1.** With the Zero Extension preoccupying in mind, it has great appeal to think that the impulse term $\sum_{k=1}^{h-1} \delta^{(k-1)} N^k \{\sum_{i=0}^{h-1} N^i f^{(i)}(0^+)\}$ in the distributional solution (3.2) is caused by the jumps $f^{(i)}(0^+) - \tilde{f}^{(i)}(0^-) = f^{(i)}(0^+)$. But we have seen even if $f$ has smooth extension on $t<0$, the impulse term occurs still in the solution. From our theory, the complete impulse term $\sum_{k=1}^{h-1} \delta^{(k-1)} N^k \{x_0 + \sum_{i=0}^{h-1} N^i f^{(i)}(0^+)\}$ is caused neither by initial value $x_0$, nor by input $f$, but by the inconsistency between them, i.e., the inequality

$$x_0 \neq -\sum_{i=0}^{h-1} N^i f^{(i)}(0^+),$$

which is a property independent of the behavior before initial time. On the other hand, as a pure mathematical problem, (2.1a) does not involve anything on $t<0$. The initial condition written into the form (1.10) should only be considered to emphasize the physical background: the initial value is the result of the physical system running up to zero time instant [20]. From mathematical viewpoint, now that what we need to find is a function on $t \geq 0$, any requirement on $t<0$ is unnecessary.

## 5. Conclusions

A rigorous mathematical theory to interpret the phenomenon that inconsistency between initial condition and input causes impulse is established for singular linear system. The theory is based on an embedding approach to the distributional solution of initial value problem of such system. Explicit distributional solution formula is derived according to this approach. Among future related research directions, maybe the closest one is a generalization of the work to the case of time varying singular linear system.

## References

[1] F.L. Lewis, A survey of linear singular systems, Circuits, Systems and Signal Processing 5 (1) (1986) 3–36.
[2] L. Dai, Singular Control Systems, Springer, Berlin, Heidelberg, 1989.
[3] L. Dai, Impulsive modes and causality in singular systems, International Journal of Control 50 (4) (1989) 1267–1281.
[4] R. Byers, T. Geerts, V. Mehrmann, Descriptor systems without controllability at infinity, SIAM Journal on Control and Optimization 35 (2) (1997) 462–479.
[5] J.Y. Ishihara, M.H. Terra, Impulse controllability and observability of rectangular descriptor systems, IEEE Transactions on Automatic Control 46 (6) (2001) 991–994.
[6] W. Wang, Y. Zou, Analysis of impulsive modes and Luenberger observers for descriptor systems, Systems and Control Letters 44 (2001) 347–353.
[7] M. Hou, Controllability and elimination of impulsive modes in descriptor systems, IEEE Transactions on Automatic Control 49 (10) (2004) 1723–1727.
[8] P. Losse, V. Mehrmann, Controllability and Observability of Second Order Descriptor Systems, SIAM Journal on Control and Optimization 47 (3) (2008) 1351–1379.
[9] F.R. Gantmacher, The Theory of Matrices, vol. 2, Higher Education Press, Beijing, 1955 (Chinese translation from Russian).
[10] I.M. Gelfand, G.E. Shilov, Generalized Functions, vol. I, Science Press, Beijing, 1965 (Chinese translation from Russian).
[11] Z. Yan, G. Duan, Time domain solution to descriptor variable systems, IEEE Transactions on Automatic Control 50 (11) (2005) 1796–1799.
[12] A.I.G. Vardulakis, Linear Multivariable Control—Algebraic Analysis and Synthesis Methods, Wiley, New York, 1991.
[13] P.C. Müller, Remark on the solution of linear time-invariant descriptor systems, Proceedings in Applied Mathematics and Mechanics PAMM 5 (2005) 175–176.
[14] Z. Yan, G. Duan, Impulse analysis of linear time-varying singular systems, IEEE Transactions on Automatic Control 51 (12) (2006) 1975–1979.
[15] Z. Yan, Geometric analysis of impulse controllability for descriptor system, Systems and Control Letters 56 (1) (2007) 1–6.
[16] G.C. Verghese, B.C. Lévy, T. Kailath, A generalized state-space for singular systems, IEEE Transactions on Automatic Control AC-26 (4) (1981) 811–831.
[17] A.V. Oppenheim, A.S. Willsky, S.H. Nawab, Signals and Systems, Prentice-Hall, Englewood Cliffs, NJ, 1997.
[18] G. Doetsch, Introduction to the Theory and Application of the Laplace Transformation, Springer, Berlin, New York, 1974.
[19] T. Geerts, Solvability conditions consistency, and weak consistency for linear differential-algebra equations and time-invariant systems: the general case, Linear Algebra and its Applications 181 (1993) 111–130.
[20] G. Reißig, H. Boche, P.I. Barton, On inconsistent initial conditions for linear time-invariant differential-algebraic equations, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 49 (11) (2002) 1646–1648.