Approximating distributional behaviour of LTI differential systems using Gaussian function and its derivatives

Athanasios A. Pantelous[1*], **Nicos Karcanias**[2] and **George Halikias**[2]

[1]Department of Mathematical Sciences, University of Liverpool, UK

[2]School of Engineering and Mathematical Sciences, City University, UK

Abstract

The paper is concerned with the approximation of the destributional behaviour of linear, time-invariant (LTI) systems. First, we review the different types of approximations of distributions by smooth functions and explain their significance in characterizing system properties. Secondly, for controllable LTI differential systems, we establish an interesting relation between the time and volatility parameters of the Gaussian function and its derivatives in the approximation of distributional solutions. An algorithm is then proposed for calculating the distributional input and its smooth approximation which minimizes the distance to an arbitrary target state. The optimal choice of the volatility parameter for the state transition is also derived. Finally, some complementary distance problems are also considered. The main results of the paper are illustrated by numerous examples.

AMS (Classification): 93C05, 34K45, 34A37

Keywords: Linear Systems; Approximating Distributional Behaviour; Gaussian Function and its Derivatives

***Corresponding Author**: Dr Athanasios Pantelous, Department of Mathematical Sciences, University of Liverpool, Peach Street, L69 7ZL, UK: E-mail: A.Pantelous@liverpool.ac.uk.

1. Introduction

The use of distributions in the study of LTI differential system problems is a well-established subject going back to [3, 4-6, 8, 10, 12, 14-21] and references therein. The work so far has dealt with the characterisation of basic system properties such as infinite poles and zeros [17, 18] for regular and singular (implicit) systems, as well as the study of fundamental control problems where the solution is expressed in terms of distributions. Typical problems are those dealing with the notions of almost (A,B)-invariance and almost controllability subspaces [12], [20].

Indeed, the study of distributional solutions plays a key role in many areas in systems and control such as:

(i) Controllability, Observability.

(ii) Infinite zero characteristic behaviour.

(iii) Almost invariant subspaces, almost controllability spaces.

(iv) Dynamics of singular systems etc.

The distributional characterization is also linked to the solution of a number of control problems. The solution of these problems has only theoretical significance, given that distributions cannot be constructed and only smooth functions can be implemented. The idea of approximating distributional inputs with smooth functions that achieve a similar control objective was first introduced by Gupta and Hasdorff [10] (see also [11]).

In the present paper, which actually extends and provides a rigorous reformulation of the early ideas presented in [10], we consider the problem of approximating Dirac distributions with smooth functions of infinite support, and more specifically using the Gaussian distribution and its derivatives. Analytically, in Section 2 we present the problem formulation for a LTI differential system. In Section 3 we provide a brief review of the different types of approximations of distributions by smooth functions and explain their significance in characterizing system properties. In Section 4 we assume that the system is controllable, and under this assumption we establish an interesting connection between a time-parameter t and a volatility parameter σ of the Gaussian density function used in the approximation. It turns out that the fraction t/σ can be fixed (to a sufficiently large value) and in this case parameter t (or σ) controls the state-transition time and the accuracy of the approximation (which can be interpreted probabilistically). A new algorithm is proposed for calculating the smooth input signal that approximates the distributional input which transfers the origin of the state-space to an arbitrary target point (subject to a controllability assumption) and the distance (Euclidean norm) between the actual terminal state and the target state; this distance is subsequently minimized subject to magnitude constraints imposed on the coefficients of the control signal. Finally, in Section 5 we define the distance from the origin using the Euclidean norm. Moreover, we consider the problem of maximising the distance from the origin with constrained input. Section 6 concludes the paper.

2. Problem Definition

We consider the linear time invariant (LTI) system

$$\underline{x}'(t) = A\underline{x}(t) + \underline{b}u_o(t), \tag{2.1}$$

where $\underline{x}(t) \in \mathcal{C}^{\infty}\left(\mathbb{F}, \mathcal{M}(n\times 1;\mathbb{F})\right)$ (smooth function over the field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, whose elements belong to the algebra $\mathcal{M}(n\times 1;\mathbb{F})$), and $u_o(t) \in \mathcal{D}'_{n-1}$ (where $\mathcal{D}'_{n-1}$ is the space of Dirac distribution having derivatives up to an order $n-1$) are the state vector, and the impulsive input, respectively and $A \in \mathcal{M}(n\times n;R)$, $\underline{b} \in \mathcal{M}(n\times 1;R)$. Following also [10], we assume that A is simple and expressed as

$$A = diag\{\lambda_1, \lambda_2, ..., \lambda_n\}, \tag{2.2}$$

where $\lambda_i \neq \lambda_j \neq 0$ for every $i \in \underset{\sim}{n}$ ($\underset{\sim}{n} := \{1, 2, \ldots, n\}$). This assumption can be relaxed, see for more details Remark 4.1.

It is assumed that the input to the LTI is a linear combination of Dirac δ -function and its first $n-1$ derivatives, i.e.

$$u_o(t) = \sum_{k=0}^{n-1} a_k \delta^{(k)}(t), \tag{2.3}$$

where $\delta^{(k)}$ or $\frac{d^k\delta}{dt^k}$ is the k^{th} derivative of the Dirac δ -function, and a_k for $i \in \underset{\sim}{n}_o$ ($\underset{\sim}{n}_o := \{0,1,\dots n-1\}$) are the magnitudes of the delta function and its derivatives. We shall denote the initial state of the system at time $t = 0^-$ as $\underline{x}(0^-)$ and the final state defined at time $t = 0^+$ as $\underline{x}(0^+)$. It is assumed that $\underline{x}(0^-) = [0 \quad 0 \quad \dots \quad 0]^T$ and $\underline{x}(0^+) =$ $[x_1 \quad x_2 \quad \dots \quad x_n]^T$. Furthermore, it is assumed that the system is *controllable* and thus any $\underline{x}(0^+) \in \mathbb{R}^n$ can be achieved. In general, the existence of an input that transfers the state of the system (2.1) from $\underline{x}(0^-) = \underline{0}$ to $\underline{x}(0^+)$ requires that the vector $\underline{x}(0^+)$ belongs to the controllable subspace of the pair $(A, \underline{b})$, i.e. $\underline{x}(0^+) \in \{A | \underline{b}\}$. In this case, the necessary and sufficient condition for transferring the state of the system (1) from $\underline{x}(0^-) = \underline{0}$ to $\underline{x}(0^+)$ by the action of the control input defined in (2.3) is that $\underline{x}(t) = \sum_{k=0}^{n-1} \beta_k \delta^{(k)}(t)$ where the coefficients β_k, $k \in \underset{\sim}{n}_o$, are the components of $\underline{x}(0^+)$ along the directions $\{\underline{b}, A\underline{b}, A^2\underline{b}, \dots, A^{n-1}\underline{b}\}$, respectively, defined according to some projections law.

In the next section, some background results as a brief review on the approximation of Dirac delta function are presented.

3. Approximations of Dirac Delta Function

The approximation of distributions by smooth functions is a problem which has been considered in the literature. In this section, we review the main results in this area and suggest a systematic and rigorous procedure for approximating distributions and their derivatives. If the standard approximating technique of the Dirac δ-function is followed, (see e.g. [7-8, 11-13, 22]), the change of the state in some minimum practical time depends mainly on the *accuracy* of the approximations that have been generated. The relation between the type of approximation used and the duration of the resulting state-transition is one of the important issues considered in this section.

The Dirac δ-function can be viewed as the limit of a sequence of functions

$$\delta(t) = \lim_{a\to 0} \delta_a(t), \tag{3.1}$$

where $\delta_a(t)$ is referred to as a *nascent* delta function. The limit is in the sense that

$$\lim_{a\to 0} \int_{-\infty}^{+\infty} \delta_a(t) f(t)\, dt = f(0). \tag{3.2}$$

These properties can be approximately enforced by using a smooth, finite approximation of the Dirac distribution. Such approximations have additional advantages. Approximating the Dirac distribution by a smooth function may actually be a better representation of the solution sought in a particular problem, especially if the effective width of the approximation function can be coupled to the physics of the problem. Following Cohen and Kirschner [7], a suitable approximating function, which is convenient for computations, should satisfy the following properties everywhere on the domain under consideration:

1. Its limit with some defining parameter is the Dirac distribution (see eq. (3.1)).

2. It is positive, decreases monotonically from a finite maximum at a source point (for instance 0), and tends to zero at the domain's extremes.

3. Its derivative exists and is a continuous function.

4. It is symmetric about the source point, for instance 0 (see eq. (3.1) and (3.2)).

5. It can be represented by a reasonably simple Fourier integral (for infinite domains) or Fourier series (for finite domains).

Next, we discuss the approximation of the Dirac delta function by functions having infinite support.

3.1. Infinite Time-support approximations

The choice of the "best" nascent delta function depends on the particular application. Choices which have proved useful in applications are listed below and include the Gaussian and Cauchy distributions, the rectangular function, the derivative of the sigmoid (or Fermi-Dirac) function, the Airy function, etc, see for instance [8, 11, 13, 22]; for approximations based on finite difference methods see [2]:

● The Cauchy distribution:

$$\delta_a(t) = \frac{1}{\pi}\frac{a}{a^2+t^2} = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{ikt-|ak|}dk\ ,$$

● The rectangular function:

$$\delta_a(t) = \frac{rect(t/a)}{a} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\sin c\left(\frac{ak}{2\pi}\right)e^{ikt}dk\ ,$$

where

$$rect(t) = \begin{cases} 1, & -1 \le t \le 1 \\ 0, & |t| > 1 \end{cases}.$$

● The partial derivative of the sigmoid (or Fermi-Dirac) function,

$$\delta_a(t) = \partial_t \frac{1}{1+e^{-t/a}} = -\partial_t \frac{1}{1+e^{t/a}}\ ,$$

● The Airy function

$$\delta_a(t) = \frac{1}{a}A_i\left(\frac{t}{a}\right).$$

The finite difference formulae may be easily converted into sequences that approximate the derivatives of the Dirac delta function in one dimension [2]. Recall the definition of the rectangular function

$$\delta_a(t) = \begin{cases} \frac{1}{a}, & -\frac{a}{2} < t < \frac{a}{2} \\ 0, & |t| > \frac{a}{2} \end{cases}, \tag{3.3}$$

which approaches $\delta(t)$ as $a \to 0$. An expression for the derivatives of $\delta(t)$ is given by

$$\frac{d^k}{dx^k}\delta(x) = \lim_{\substack{a\to 0 \\ h\to 0}}\left[\left(\frac{1}{h}\right)^k \sum_{j=0}^{k} a_j \delta_a\left(x + b_j h\right)\right], \tag{3.4}$$

where $x = t_o - t$, the a_j are appropriate constants defining the finite differences [2], and

$$\frac{d^k}{du^k}\delta(u)|_{t-t_o} = (-1)^k \frac{d^k}{du^k}\delta(u)|_x .$$

Expression (3.4) is exactly what we would obtain by making the substitution $f(t) \to \delta_a(t)$ in the following finite difference approximation for the k^{th} derivative of a smooth test function f evaluated at t_o :

$$\frac{d^k}{dt^k} f(t)|_{t=t_o} \approx \left(\frac{1}{h}\right)^k \sum_{j=0}^{k} a_j f\left(t_o + b_j h\right). \tag{3.5}$$

Note that a_j and b_j are suitably chosen constants and (3.5) becomes exact in the limit $h \to 0$. Note also that

$$\frac{d^k}{dt^k} f(t)|_{t=t_o} = \lim_{h\to 0}\left\{\left(\frac{1}{h}\right)^k \sum_{j=0}^{k} a_j \int_{-\infty}^{+\infty} \delta\left(t - t_o - b_j h\right) f(t)\, dt\right\}. \tag{3.6}$$

since f is sampled at discrete points.

In our case the Gaussian distribution may be considered as a good approximation of the Dirac distribution on an infinite domain.

3.2 Input signal structure

Since our time-domain is *infinite*, the Gaussian distribution, i.e.

$$\delta(t) = \lim_{\sigma\to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2} = \lim_{\sigma\to 0} \frac{1}{\sigma}\Phi\left(\frac{t}{\sigma}\right) \text{ where } \Phi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (3.8)$$

is considered. Hence, the approximate expression for the input signal (2.3) is given by

$$u_\sigma(t) = \sum_{k=0}^{n-1} a_k \frac{1}{\sigma^{k+1}} \Phi^{(k)}\left(\frac{t}{\sigma}\right), \quad (3.9)$$

where

$$\Phi^{(i)}\left(\frac{t}{\sigma}\right) = \left(\frac{d^i}{d(t/\sigma)^i}\right)\Phi\left(\frac{t}{\sigma}\right).$$

The impulsive response of the system is recovered in the limit:

$$u_o(t) = \lim_{\sigma\to 0} u_\sigma(t). \quad (3.10)$$

A natural question arising in the context of the zero-time state-transition problem considered in this work is why attention is restricted to impulsive control signals expressed as a sum of Dirac delta functions (and its derivatives). The answer to this question involves the idea of *single-layer distributions* [8, 13, 22] which is briefly introduced in the next few paragraphs:

Lemma 3.1 [22] *If* $\mathcal{U}$ *is a bounded closed set in* $\mathbb{F}$ *and* $\mathcal{Y}$ *is a neighbourhood of* $\mathcal{U}$, *then there exists a function such that* $n=1$ *on* $\mathcal{U}$, $n=0$ *outside* $\mathcal{Y}$, *and* $0 \le n \le 1$ *over* $\mathbb{F}$. □

Definition 3.1 *Let $\mathcal{S}$ be a piecewise regular curve in $\mathbb{F}$ and σ is a locally integrable function defined on $\mathcal{S}$. The linear continuous functional $\sigma\delta_S$ on the space $\mathcal{D}$ of infinitely differentiable complex-valued functions on $\mathbb{F}$ with compact support is defined as*

$$\langle \sigma\delta_S, \varphi \rangle = \int_S \varphi(\xi)\sigma(\xi)\delta S,$$

$\forall\ \varphi \in \mathcal{D}$ *and is called single (or simple) layer on $\mathcal{S}$ with density σ.* □

Note that $\sigma\delta_S(x) = \int_S \delta(x-\xi)\sigma(\xi)\delta S$.

Definition 3.2 *Let $\mathcal{S}$ be a piecewise regular curve in $\mathbb{F}$ and $\mu\delta_S$. The linear continuous functional $-d/dt(\mu\delta_S)$ on the space $\mathcal{D}$ of infinitely differentiable complex-valued functions on $\mathbb{F}$ with bounded support is defined as*

$$\langle -d/dt(\sigma\delta_S), \varphi \rangle = \int_S \sigma(\xi)\frac{d\varphi(x-\xi)}{dt}\delta S \ \ \forall\ \varphi \in \mathcal{D}. \quad \square$$

It can be shown that every distribution $\sigma\delta_S(x)$ that has compact support is of finite order, see [8, 22]. Thus, every distribution $\sigma\delta_S(x)$ whose support is the point $x = \tau$ has the form $\sum_{k=0}^{n-1} c_k \delta^{(k)}(t-\tau)$, i.e. it can be expressed as a linear combination of the Dirac δ-function and its first $n-1$ derivatives.

Thus, the zero-time state transfer problem considered on this work, involving the state transfer of system (2.1) from $\underline{x}(0^-)$ to $\underline{x}(0^+)$, corresponds to a single support point $\tau = 0$ and hence (2.3) is appropriate.

4. Design of Approximate Input Signal

In this section, we attempt to answer the following question: "*What are the coefficients a_k $, k \in \underset{\sim}{n}$, and what is the optimal volatility parameter σ such that the input* *signal defined in equation (2.3) transfers the state from* $\underline{x}\left(0^{-}\right)$ *to* $\underline{x}\left(0^{+}\right)$ *?"* In attempting to answer this question the following standard result is significant.

Lemma 4.1 [11] *The solution of system (2.1) is given by*

$$\underline{x}(t)=e^{At}\int_{-\infty}^{t}e^{-A\tau}\underline{b}u_{o}(\tau)\,d\tau, \tag{4.1}$$

where $u_{o}(\tau)$ *is defined in equations (3.9) and (3.10).* □

Remark 4.1 Recall that for simplicity it is assumed that matrix A is diagonal with distinct eigenvalues. This reduces the complexity of various mathematical expressions and the number of technicalities involved, without introducing any real loss of generality. The general case can be tackled by defining a $n\times n$ non-singular similarity transformation $Q=\left[v_{1},v_{2},\ldots,v_{n}\right]\in\mathcal{M}(n\times n;\mathbb{F})$ that takes A into the Jordan canonical form.

The solution of system (2.1) to the input defined in equations (3.9) and (3.10) is

$$\underline{x}(t)=\lim_{\sigma\to 0}\left\{e^{At}\int_{-\infty}^{t}e^{-A\tau}\underline{b}u_{\sigma}(\tau)\,d\tau\right\},$$

or equivalently

$$\underline{x}(t)=e^{At}\lim_{\sigma\to 0}\left[\int_{-\infty}^{t}e^{-A\tau}\underline{b}\sum_{k=0}^{n-1}a_{k}\frac{1}{\sigma^{k+1}}\Phi^{(k)}\left(\frac{\tau}{\sigma}\right)d\tau\right]. \tag{4.2}$$

As $\sigma\to 0$, the energy of the input signal "concentrates" around $\tau=0$. Hence the zero-time state-transition problem involves setting $t=0^{+}$ and selecting the coefficients a_{k} so that (an arbitrary) $\underline{x}(0^{+})\in\mathbb{R}^{n}$ is reached (recall that controllability of the pair $(A,\underline{b})$ is assumed).

Remark 4.2 To reduce the complexity of the solution (due to the large number of terms involved) we exploit the fact that

$$\Phi\left(t/\sigma\right)=\frac{1}{\sqrt{2\pi}}\;e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2},$$

and its derivatives tend to zero very *strongly* with $t/\sigma\rightarrow\infty$. Define $t/\sigma\triangleq K\left(t,\sigma\right)$ and assume that *t* is fixed to a positive value, so that $K\left(t,\sigma\right)\rightarrow\infty$ as $\sigma\rightarrow 0$. Then,

$$\Phi\left(t/\sigma\right)\triangleq\Phi\left(K\left(t,\sigma\right)\right)\overset{K(t,\sigma)\rightarrow\infty}{\rightarrow}0,$$

and its derivatives

$$\Phi^{(k)}\left(t/\sigma\right)\triangleq\Phi^{(k)}\left(K\left(t,\sigma\right)\right)\overset{K(t,\sigma)\rightarrow\infty}{\rightarrow}0,\quad k\in\underset{\sim}{n}_o.$$

where $\Phi^{(0)}\left(t/\sigma\right)\triangleq\Phi\left(t/\sigma\right)$.

A suitable choice of $K\left(t,\sigma\right)$ depends on the choice of the transition time-variable t and the volatility-parameter σ. In practice, t can be fixed, since we can pre-define the duration of the (almost zero) transition between the initial and final (target) state of the system when solving the (almost) zero-time state transition problem (e.g., we can select *t* to be of the order of $t\propto 10^{-6}$seconds, say). This is the approximate version of the exact problem and can be formulated as follows:

For a fixed value of the time parameter $t=t^*$ *and a fixed* $\varepsilon>0$ *determine*

$$\sigma^*=\sup\left\{\sigma\in R_+:\left\|\underline{x}\left(t^*\right)-\hat{\underline{x}}\left(t^*\right)\right\|<\varepsilon\right\},\tag{4.3}$$

where $\underline{x}\left(t^*\right)$ *is the target state and* $\hat{\underline{x}}\left(t^*\right)$ *is the actual terminal state resulting from the approximation of the input signal, see equation (4.1).*

This is in the form of a distance-approximation problem. Roughly, for a fixed state-transition time-duration, we seek the “smoothest” input signal for which the error tolerance of the distance between the target and actual terminal state is kept within a pre-defined level ε. Note, that since this distance tends to zero as $\sigma\rightarrow 0$ and the only source of error arises from the approximation of the Dirac delta function and its derivatives, an alternative equivalent formulation of the problem is to determine (for a fixed value $t = t^*$),

$$\sigma^* = \sup\left\{\sigma \in R_+ : \left|\Phi^{(k)}\left(K\left(t^*,\sigma\right)\right)\right| < \varepsilon_k, k \in \underset{\sim}{n_o}\right\},$$

where the ε_k are suitable positive constants.

The following lemma is required for subsequent developments. The objective is to develop approximation bounds for the terminal state when the impulsive inputs in equation (4.1) are substituted by their smooth approximations.

Lemma 4.2 *Consider* $u_\sigma(t)$ *defined in equations (3.9). Then*

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} u_\sigma\left(\tau\right) d\tau = \sum_{k=0}^{n-1} a_k \left\{ e^{-\lambda_i t} \sum_{m=1}^{k} \left(\frac{\lambda_i^{m-k+1}}{\sigma^{k-m+1}} \Phi^{(k-m)}\left(\frac{t}{\sigma}\right)\right) + \lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2} \Phi^{-1}\left(\frac{t}{\sigma} + \lambda_i \sigma\right)\right\}, \quad (4.4)$$

where $\Phi^{(0)}\left(x\right) \triangleq \Phi\left(x\right)$, $\Phi^{(-1)}\left(x\right) \triangleq \int_{-\infty}^{x} \Phi\left(y\right) dy = \sqrt{2}\ \mathrm{erf}^{-1}\left(2x-1\right)$, $x \in \left(0,1\right)$.

Proof Substituting expression (3.9) into the integral $\int_{-\infty}^{t} e^{-\lambda_i \tau} u_\sigma\left(\tau\right) d\tau$, we obtain

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \sum_{k=0}^{n-1} \frac{a_k}{\sigma^{k+1}} \Phi^{(k)}\left(\tau/\sigma\right) d\tau = \sum_{k=0}^{n-1} a_k \int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\Phi^{(k)}\left(\tau/\sigma\right)}{\sigma^{k+1}} d\tau.$$

Consider first the term corresponding to $k = 0$,

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\Phi\left(\tau/\sigma\right)}{\sigma} d\tau = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\lambda_i^2\sigma^2} \int_{-\infty}^{t} e^{-\frac{1}{2}\left(\frac{\tau}{\sigma}+\lambda_i\sigma\right)^2} d\tau = e^{\frac{1}{2}\lambda_i^2\sigma^2} \Phi^{-1}\left(\frac{t}{\sigma} + \lambda_i \sigma\right).$$

Consider next the term corresponding to $k = 1$. Integration by parts and using the equation above gives

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\Phi'\left(\tau/\sigma\right)}{\sigma^2} d\tau = e^{-\lambda_i \tau} \frac{1}{\sigma} \Phi\left(\tau/\sigma\right)\Big|_{-\infty}^{t} + \lambda_i \int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\Phi\left(\tau/\sigma\right)}{\sigma} d\tau$$

$$= e^{-\lambda_i t}\frac{1}{\sigma}\Phi\left(t/\sigma\right)+\lambda_i e^{\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right).$$

Similarly,

$$\int_{-\infty}^{t} e^{-\lambda_i\tau}\frac{\Phi''\left(\tau/\sigma\right)}{\sigma^3}d\tau = e^{-\lambda_i\tau}\frac{1}{\sigma^2}\Phi'\left(\tau/\sigma\right)\Big|_{-\infty}^{t}+\lambda_i\int_{-\infty}^{t} e^{-\lambda_i\tau}\frac{\Phi'\left(\tau/\sigma\right)}{\sigma^2}d\tau$$

$$= e^{-\lambda_i t}\left[\lambda_i\frac{1}{\sigma}\Phi\left(t/\sigma\right)+\frac{1}{\sigma^2}\Phi'\left(t/\sigma\right)\right]+\lambda_i^2 e^{\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right).$$

A recursive application of this procedure gives

$$\int_{-\infty}^{t} e^{-\lambda_i\tau}\frac{\Phi^{(k)}\left(\tau/\sigma\right)}{\sigma^{k+1}}d\tau = e^{-\lambda_i t}\sum_{m=1}^{k}\lambda_i^{m-k+1}\frac{1}{\sigma^{k-m+1}}\Phi^{(k-m)}\left(\frac{t}{\sigma}\right)+\lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right),$$

from which the result follows. □

Choose $0^+/\sigma \triangleq K\left(0^+,\sigma\right)$ sufficiently large so that $\Phi^{(k)}\left(0^+/\sigma\right)\triangleq\Phi^{(k)}\left(K\left(0^+,\sigma\right)\right)$ ≈ 0, $k\in \underset{\sim}{n}_o$. Then the following approximation is valid

$$\int_{-\infty}^{0^+} e^{-\lambda_i\tau}\frac{\Phi^{(k)}\left(\tau/\sigma\right)}{\sigma^{k+1}}d\tau \approx \lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right).$$

Combining expressions (4.2) and (4.4) then gives

$$\underline{x}_i\left(K\left(0^+,\sigma\right)\sigma\right)\approx b_i e^{\lambda_i K\left(0^+,\sigma\right)\sigma+\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\sum_{k=0}^{n-1}a_k\lambda_i^k, \tag{4.5}$$

for $i=1,2,\ldots,n$. The *approximate almost zero-time state-transfer problem* can now be defined as follows: Suppose that parameters $(0^+,\sigma)$ have been chosen so that $\Phi^{(k)}\left(0^+/\sigma\right)\triangleq\Phi^{(k)}\left(K\left(0^+,\sigma\right)\right)\approx 0$, $k\in \underset{\sim}{n}_o$. Then, given $\underline{x}\left(0^+\right)\in\mathbb{R}^n$ determine real scalars a_k, $k\in \underset{\sim}{n}_o$ such that (4.5) are satisfied with equality for all $i\in\{1,2,\ldots,n\}$. Note that the impulsive response is recovered as $\sigma\to 0$ in which case the approximation in the above equation becomes exact; in this case we also have that $\underline{x}_i(0^+) \rightarrow \hat{\underline{x}}_i(0^+)$, $\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right) \rightarrow 1$, and

$$\hat{\underline{x}}_i\left(0^+\right) = b_i e^{\lambda_i 0^+} \sum_{k=0}^{n-1} a_k \lambda_i^k, \quad i = 1, 2, \ldots, n$$

so that

$$\underline{x}_i\left(0^+\right) \approx e^{\frac{1}{2}\lambda_i^2\sigma^2} \Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right) \hat{\underline{x}}_i\left(0^+\right), \quad i = 1, 2, \ldots, n.$$

The following Theorem now follows.

Theorem 4.1 *Let* $A = diag\{\lambda_1, \lambda_2, \ldots \lambda_n\}$ *with* $\lambda_i \neq \lambda_j$ *for* $i \neq j$, $\underline{b} = (b_1 \quad b_2 \quad \ldots \quad b_n)^T$ *and assume that the pair* $(A, \underline{b})$ *is controllable. Let also* $\hat{B} = diag\{1/b_1, 1/b_2, \ldots, 1/b_n\}$ *and denote by* $V \triangleq V\left(\lambda_1, \lambda_2, ..., \lambda_n\right)$ *the Vandermonde matrix*

$$V \triangleq V\left(\lambda_1, \lambda_2, ..., \lambda_n\right) = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ & & \vdots & & \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix}.$$

Then the coefficient vector $\underline{a} = \left[a_o \quad a_1 \quad \cdots \quad a_{n-1}\right]^T$ *of the input signal defined in (3.9) which solves the almost zero-time state-transfer problem is given by*

$$\underline{a} = V^{-1} e^{-A \cdot 0^+} \hat{B}^{-1} \hat{\underline{x}}\left(0^+\right), \tag{4.6}$$

where

$$\hat{\underline{x}}_i\left(0^+\right) \triangleq \frac{\underline{x}_i\left(K\left(0^+,\sigma\right)\sigma\right)}{e^{\frac{1}{2}\lambda_i^2\sigma^2} \Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)}, \ i \in \underset{\sim}{n}. \tag{4.7}$$

Proof Expression (4.4) can be re-written as

$$\hat{x}_i(0^+) = b_i e^{\lambda_i 0^+} \sum_{k=0}^{n-1} a_k \lambda^k \cong \frac{x_i(0^+)}{e^{\lambda_i^2\sigma^2/2}\Phi^{-1}(0^+ + \lambda_i\sigma)} \text{ for } i \in \underset{\sim}{n}.$$

Thus we can write $\underline{\hat{x}}\left(0^+\right) = \hat{B}e^{A\cdot 0^+}V\underline{a}$ or equivalently (4.6). Note that the indicated inverses V^{-1} and $\hat{B}^{-1}$ exist due the assumption that the eigenvalues of A are distinct, and the assumed controllability of $(A,\underline{b})$, respectively. □

Ideally the parameters $t^* = 0^+$ and σ should be chosen so that the distance

$$\left\|\underline{x}\left(t^*\right) - \underline{\hat{x}}\left(t^*\right)\right\|_2 = \sqrt{\sum_{i=1}^n \left[\underline{x}_i\left(K\left(t,\sigma\right)\sigma\right) - \underline{\hat{x}}_i\left(K\left(t,\sigma\right)\sigma\right)\right]^2}$$

is “small”. Clearly the distance is zero provided that $K\left(t,\sigma\right)$ is selected so that

$$\Phi^{-1}\left(K\left(t,\sigma\right) + \lambda_i\sigma\right) - e^{-\frac{1}{2}\lambda_i^2\sigma^2} = 0 \tag{4.8}$$

for all *i* which requires $\sigma \to 0$, in which case (4.8) implies that

$$\lim_{\sigma\to 0}\Phi^{-1}\left(K\left(t,\sigma\right)\right) = 1 \Leftrightarrow \lim_{\sigma\to 0}\int_{-\infty}^{K(t,\sigma)}\Phi\left(x\right)dx = 1 \Leftrightarrow K(t,\sigma)\to\infty. \tag{4.9}$$

In probability theory and statistics, the normal or Gaussian distribution $\Phi\left(x\right)$ is widely used. The graph of $\Phi\left(x\right)$ is bell-shaped and is known as the Gaussian function or bell curve. Actually, in this case we are interested in

$$\int_{-\infty}^{K(t,\sigma)}\Phi\left(x\right)dx,$$

which is the *cumulative distribution function* (cdf) of a random variable $X \sim N(0,1)$ evaluated at the upper limit of the integral $K\left(t,\sigma\right)$, denoting the probability that $X \le K\left(t,\sigma\right)$. In practice, if $|\lambda_i\sigma| << 1$ for all *i*, we can assume that equation (4.8) is approximately satisfied if $K_0 \triangleq K\left(t,\sigma\right) \ge 3.9$ (in which case $\Phi^{-1}\left(K_0\right) > 1 - 10^{-4}$). Thus, a reasonable choice for the volatility parameter is $\sigma^* = K_0^{-1}t^* \approx 0.256t^*$.

The results of the section are summarised in the following algorithm.

Algorithm TIAZT (**T**ransfer **I**n **A**lmost **Z**ero **T**ime)

1st Step: Define the terminal (target) state of the transition $\underline{x}\left(0^{+}\right)$.

2nd Step: Using the required transition time $t^{*}\left(\equiv 0^{+}\right)$ define the optimal volatility parameter $\sigma^{*}=0.256t^{*}$.

3rd Step: Finally, the coefficients if the input signal $\underline{a}=\begin{bmatrix}a_o & a_1 & \cdots & a_{n-1}\end{bmatrix}^T$ defined in equation (3.9) are obtained by (4.6), i.e. $\underline{a}=V^{-1}e^{-A\cdot 0^{+}}\hat{B}^{-1}\underline{\hat{x}}\left(0^{+}\right)$ where all variables are defined in Theorem 4.1.

Remark 4.4 From the control viewpoint it is important to choose an appropriate time duration for the state transition. This ultimately depends on the type of application, e.g. due to control signal magnitude or "slew-rate" limitations. It is clear from the imposed proportionality $\sigma^{*}=K_0^{-1}t^{*}$ that increasing the duration of the state-transition results is "smoother" input signals, which is often desirable. For example, if the system operates in a feedback loop (in which case the input signal is generated by a feedback controller), highly discontinuous signals typically correspond to system overdesign (e.g. excessive closed-loop bandwidth) and may have detrimental effects on the stability and performance characteristics, e.g. in terms of reduced robust stability margins and sensor noise amplification.

Example 4.1 (See Gupta, 1966) Consider the system

$$\begin{bmatrix}x_1'(t)\\x_2'(t)\end{bmatrix}=\begin{bmatrix}-2 & 0\\0 & -3\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}u_o(t)$$

where $\underline{x}(t)$ and $u_o(t)$ are the state and the input signals, respectively. Suppose we wish to transfer the state of the system from $\underline{x}(0)=\begin{pmatrix}0 & 0\end{pmatrix}^T$ to $\underline{x}\left(0^{+}\right)=\begin{pmatrix}3 & 4\end{pmatrix}^T$ at time $0^{+}=1\mu s$ (1 microsecond). Application of the TIAZT algorithm gives

1st Step: Here the desired state is $\underline{x}\left(0^+\right)=\left(3 \quad 4\right)^T$.

2nd Step: The transition duration has been pre-determined as $0^+ = 10^{-6}$ s, so the optimal volatility parameter is $\sigma^* = 2.56 \cdot 10^{-7}$ (taking $K_0 = 3.9$).

3rd Step: Here, $\hat{\underline{x}}_1\left(10^{-6}\right) \approx \underline{x}_1\left(10^{-6}\right) = 3$ and $\hat{\underline{x}}_2\left(10^{-6}\right) \approx \underline{x}_2\left(10^{-6}\right) = 4$. The inverse of the Vandermonde matrix is:

$$V^{-1} = V^{-1}\left(-2,-3\right) = \begin{bmatrix} 1 & -2 \\ 1 & -3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix}.$$

Thus, the coefficient vector $\underline{a} = \left[a_0 \quad a_1\right]^T$ is calculated as:

$$\underline{a} = \begin{bmatrix} a_o \\ a_1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \exp(2\times 10^{-6}) & 0 \\ 0 & \exp(3\times 10^{-6}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \approx \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

5. Distance Problems

5.1. Distance from the origin in state-space

In this section, we define the distance from the origin corresponding to a state transition of the system (2.1) from the zero (or ground) state, $\underline{x}\left(0^-\right) = \left[0 \quad 0 \quad \cdots \quad 0\right]^T$. Using the Euclidean norm this is defined as

$$r^2 \triangleq \left\|\underline{x}\left(0^+\right) - \underline{x}\left(0^-\right)\right\|^2 = \underline{x}^T\left(0^+\right)\underline{x}\left(0^+\right) = \sum_{i=1}^{n} x_i^2\left(0^+\right) , \qquad (5.1)$$

(see Fig 1). The time interval of the transition has been defined in previous sections as 0^+ (t^*) and the target state is $\hat{\underline{x}}\left(0^+\right)$.

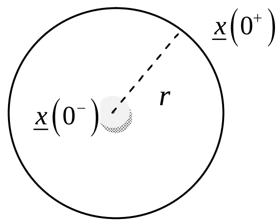

Fig. 1: 2-ball with centre $\underline{x}(0^-)$ and radius $\hat{r}$

However, if the Dirac delta function and its derivatives are replaced by smooth signals (Gaussian distribution function and its derivatives), this target state will not be reached exactly, in general. The distance in terms of the target state $\hat{\underline{x}}(0^+)$ is defined as

$$\hat{r}^2 \triangleq \sum_{i=1}^{n} \hat{x}_i^2(0^+) = \sum_{i=1}^{n} \frac{x_i^2\left(K(0^+,\sigma)\sigma\right)}{e^{\lambda_i^2\sigma^2}\left[\Phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)\right]^2},$$

where (4.7) has been used. Note that fixing $K(t,\sigma)$ and taking $\sigma \to 0$, we get $\hat{r} \to r$.

Example 5.1 Consider the system:

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_o(t),$$

where $\underline{x}(t) \in \mathcal{C}^\infty\left(R, \mathcal{M}(2\times 1; R)\right)$ and $u_o(t)$ are the state vector and the input, respectively. Let $\underline{x}(0^-) = 0$ and $\underline{x}(0^+) = [3 \quad 4]^T$. Then

$$\hat{r} = \left\|\hat{\underline{x}}(0^+) - \underline{x}(0^-)\right\| = \sqrt{\sum_{i=1}^{2} \frac{x_i^2\left(K(0^+,\sigma)\sigma\right)}{\left[\underbrace{e^{\lambda_i K(0^+,\sigma)\sigma + \frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)}_{\beta_i}\right]^2}} = \sqrt{\frac{9}{\beta_1^2} + \frac{16}{\beta_2^2}}.$$

As $\beta_1, \beta_2 \to 1$, $\hat{r} \to r = 5$.

5.2 Maximum distance from the origin with constrained input

Here we assume that the system (2.1) starts from the zero state at time $t=0^-$ and consider the problem of maximising the distance to the terminal state in an (almost) zero-time state transition. This problem of course makes sense if the input signal is constrained in some sense. Here we impose constraints on the coefficient vector of the input signal $\underline{a}=\begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1}\end{bmatrix}^T$ in terms of the Euclidian and the infinity norms.

Lemma 5.1 *Let* $\lambda_i \neq 0$, $i=1,2,\ldots,n$. *Then* $\sum_{i=1}^n |\lambda_i|^{p-1} \leq \max\left\{n, \sum_{i=1}^n |\lambda_i|^{n-1}\right\}$ *for all* $p=1,2,\ldots,n$.

Proof Define function $f(x)=\sum_{i=1}^n |\lambda_i|^{x-1}$ which can be written as $f(x)=\sum_{i=1}^n e^{m_i(x-1)}$ by setting $m_i=\ln|\lambda_i|$. Since $f''(x)=\sum_{i=1}^n m_i^2 e^{m_i(x-1)}>0$ for all $x\in\mathbb{R}$, function is convex for all $x\in\mathbb{R}$ and specifically in the interval $1\leq x\leq n$. Thus $f(x)$ attains its maximum at an edge of the interval $1\leq x\leq n$, i.e.

$$\sum_{i=1}^n |\lambda_i|^{p-1} \leq \max_{1\leq x\leq n} f(x)=\max\{f(1), f(n)\}=\max\left\{n, \sum_{i=1}^n |\lambda_i|^{n-1}\right\},$$

for every $p=1,2,\ldots,n$ as required. □

Theorem 5.1 *Let* $A=diag\{\lambda_1,\lambda_2,\ldots\lambda_n\}$, $\underline{b}=\begin{pmatrix} b_1 & b_2 & \ldots & b_n\end{pmatrix}^T$ *and assume that the pair* $(A,\underline{b})$ *is controllable. Define* $\hat{B}=diag\{1/b_1,1/b_2,\ldots,1/b_n\}$ *and denote by* $V \triangleq V(\lambda_1,\lambda_2,...,\lambda_n)$ *the Vandermonde matrix*

$$V=V(\lambda_1,\lambda_2,...,\lambda_n)=\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ & & \vdots & & \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix}.$$

Let $\underline{a} = \begin{bmatrix} a_o & a_1 & \cdots & a_{n-1} \end{bmatrix}^T$ *be the coefficient vector of the input signal* $u_o(t) = \sum_{i=0}^{n-1} a_i \delta^{(i)}(t)$ *defined in (3.9). Then, if* $\hat{\underline{x}}(0^+)$ *denotes the terminal state of the zero-time state-transition problem with* $\hat{\underline{x}}(0^-) = \underline{0}$,

$$\max_{\|a\|=1} \left\| \underline{x}(t^+) \right\| = \left\| \hat{B} e^{A \cdot 0^+} V \right\| \le \frac{t^+ \rho(A) \sqrt{n}}{\min_{i \in n} |b_i|} \max \left\{ n, \sum_{i=1}^{n} |\lambda_i|^{n-1} \right\}, \tag{5.2}$$

where the indicated matrix norm denotes the largest singular value (spectral norm) and $\rho(A)$ *denotes the spectral radius of A.*

Proof In the notation of Theorem 4.1 the terminal state of the transition is $\hat{\underline{x}}(0^+) = \hat{B} e^{A \cdot 0^+} V \underline{a}$. Thus $\max_{\|a\|=1} \left\| \hat{\underline{x}}(0^+) \right\| = \left\| \hat{B} e^{A \cdot 0^+} V \right\|$, while the maximizing coefficient vector $\underline{a}$ is the (normalised) singular vector of $\hat{B} e^{A \cdot 0^+} V$ corresponding to the largest singular value. (If the largest singular value is repeated we can choose any linear combination of unit length of the singular vectors corresponding to the repeated largest singular value). Note also that

$$\left\| \hat{B} e^{A \cdot 0^+} V \right\| \le \left\| \hat{B} \right\| \left\| e^{A \cdot 0^+} \right\| \|V\| = \frac{t^* \max_{i \in n} |\lambda_i(A)|}{\min_{i \in n} |b_i|} \|V\| = \frac{t^* \rho(A)}{\min_{i \in n} |b_i|} \|V\|. \tag{5.3}$$

Now,

$$\|V\| \le \sqrt{n} \|V\|_1 = \sqrt{n} \left\| V^T \right\|_\infty = \sqrt{n} \max_{p=1,2,\ldots,n} \sum_{i=1}^{n} |\lambda_i|^{p-1} = \sqrt{n} \max \left\{ n, \sum_{i=1}^{n} |\lambda_i|^{n-1} \right\}, \tag{5.4}$$

see Lemma 5.1 and [9], where $\| \cdot \|_1$ and $\| \cdot \|_\infty$ denote the induced 1 and ∞-matrix norms, respectively. Equation (5.2) follows by combining (5.3) and (5.4). □

Remark 5.1 Consider the almost zero-time state transition problem in which $K(t^+, \sigma) = t^+ / \sigma$ has been fixed and σ has been chosen sufficiently small so that $|\lambda_i \sigma| << 1$ for all *i* and approximation [9] is valid. Then we have $\underline{x}(0^+) = \Gamma B e^{A 0^+} V \underline{a}$, where $\Gamma = diag \left\{ \lambda_i^2 \sigma^2 \Phi^{-1} \left(K(0^+, \sigma) + \lambda_i \sigma \right) / 2 \right\}$.

It follows that in this case

$$\max_{\|\underline{a}\|=1}\left\|\underline{x}(0^+)\right\| = \left\|\Gamma B e^{A0^+} V\right\| \le \psi(n)\max_{i\in\underline{n}}\left\{|b_i|\,\lambda_i^2\, \mathrm{e}^{\lambda_i 0^+}\ \Phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)\right\},$$

where

$$\psi(n) = \frac{\sqrt{n}\sigma^2}{2}\max\left\{n, \sum_{i=1}^{n}|\lambda_i|^{n-1}\right\},$$

while the maximizing coefficient vector $\underline{a}$ is the (normalised) singular vector of $\Gamma B e^{A0^+} V$ corresponding to the largest singular value.

Next, we impose magnitude constraints on the coefficients defining the distributional input signal. Again we assume that $\hat{\underline{x}}\left(0^-\right) = \underline{x}(0^-) = \underline{0}$ and seek to maximize $\left\|\hat{\underline{x}}\left(0^+\right)\right\|$ using the impulsive input $u_0(t)$ in equation (3.10) (or $\left\|\underline{x}\left(0^+\right)\right\|$ using its smooth approximation $u_\sigma(t)$ in (3.9)) subject to the constraint:

$$|a_i| \le c_i\,,\ c_i > 0\,,\ \text{for } i \in \underset{\sim}{n} \qquad (5.5)$$

(see also [9]). Geometrically, we seek constants a_i for $i \in \underset{\sim}{n}$ in the ranges defined by (5.5) such as the radius $\hat{r}$ depicted in fig. 2 is maximised, (starting from $\hat{\underline{x}}\left(0^-\right) = \underline{0}$) where

$$\hat{r}^2 = \left\|\hat{\underline{x}}\left(0^+\right)\right\|^2 = \sum_{i=1}^{n}\hat{x}_i^2\left(0^+\right) = \sum_{i=1}^{n} b_i^2 e^{2\lambda_i 0^+}\sum_{j=1}^{n}\sum_{s=1}^{n}\lambda_i^{j+s-2} a_{j-1} a_{s-1} \qquad (5.6)$$

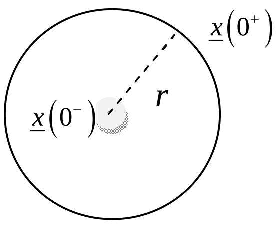

Fig. 2: *n*-ball with centre $\underline{x}\left(0^-\right)$ and radius r

Again, if the smooth approximation signal $u_\sigma(t)$ is applied, equation (4.6) should be used; substitution into equation (5.6) shows that in this case we seek to maximise:

$$r^2 = \sum_{i=1}^{n} x_i^2\left(0^+\right) = \sum_{i=1}^{n} \hat{x}_i^2\left(0^+\right)\left[e^{\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2.$$

Next note that equation (4.5) gives:

$$\hat{x}_i\left(0^+\right) = b_i e^{\lambda_i 0+}\sum_{j=1}^{n}\lambda_i^{j-1}a_{j-1},$$

and hence

$$\hat{x}_i^2\left(0^+\right) = b_i^2 e^{2\lambda_i 0+}\sum_{j=1}^{n}\sum_{s=1}^{n}\lambda_i^{j+s-2}a_{j-1}a_{s-1}, \quad i \in \underset{\sim}{n} \tag{5.7}$$

Substituting, (5.7) into (5.6), gives

$$r^2 = \left\|\underline{x}\left(0^+\right)\right\|^2 = \sum_{i=1}^{n}\left[b_i e^{\lambda_i 0^+ + \frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2\sum_{j=1}^{n}\sum_{s=1}^{n}\lambda_i^{j+s-2}a_{j-1}a_{s-1}. \tag{5.8}$$

Define the symmetric matrix

$$Q(\sigma) = V^T D^2(\sigma)V,\; D = diag\left(b_i e^{\lambda_i 0^+ + \frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right).$$

Note that due to the assumed controllability of $(A,\underline{b})$ (which implies that $b_i \neq 0$, $i \in \underset{\sim}{n}$) and the assumption that the eigenvalues of A are distinct (which implies that $\det(V) \neq 0$), we have that $Q(\sigma) = Q^T(\sigma) > 0$. The two distance maximisation problems now have the form

$$\max\ r^2 = \left\|\underline{x}\left(0^+\right)\right\|^2 = a^T Q(\sigma)a \quad s.t.\ -c_i \le a_i \le c_i,\ \ i \in \underset{\sim}{n}$$

and

$$\max\ \hat{r}^2 = \left\|\underline{\hat{x}}\left(0^+\right)\right\|^2 = a^T Q(0)a \quad s.t.\ -c_i \le a_i \le c_i,\ \ i \in \underset{\sim}{n},$$

which are *Quadratic Programming* optimization problems with "box" constraints. Since the cost function ($f(a)=a^TQ(\sigma)a$) which is maximized is convex, the constrained maximum is achieved in a vertex of a hyper-cube $|a_i|=c_i,\ i\in\underline{n}$.

Thus, we obtain

$$\left\{(-\lambda_i)^{j+s-2}\operatorname{sgn}a_{j-1}\operatorname{sgn}a_{s-1}\right\}>0 \text{ for all } j \text{ and } k.$$

This can be easily derived if we assume that

$$\operatorname{sgn}a_{j-1}=(-1)^{j-1} \text{ and } \operatorname{sgn}a_{s-1}=(-1)^{s-1},$$

so we obtain

$$\operatorname{sgn}a_{j-1}\operatorname{sgn}a_{s-1}=(-1)^{j-1}(-1)^{s-1}=(-1)^{j+s-2}.$$

So, the maximum distance is given by

$$r\triangleq\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{s=1}^{n}|\lambda_i|^{j+s-2}c_{j-1}c_{s-1}\left[e^{\lambda_iK(0^+,\sigma)\sigma+\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2. \quad (5.9)$$

Finally, again if we assume that $t^*=K\left(t^*,\sigma^*\right)\sigma^*\to 0$, and $K\left(t^*,\sigma^*\right)$ to be equal or greater to 3.90, we obtain

$$r\triangleq\left\|\underline{x}\left(t^*\right)-\underline{x}\left(0^-\right)\right\|^2=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{s=1}^{n}|\lambda_i|^{j+s-2}c_{j-1}c_{s-1}. \quad (5.10)$$

The following numerical example illustrates some of the results of this section.

Example 5.2 Consider the (almost) zero state transition problem for the system defined in example 5.1 with $\underline{x}\left(0^-\right)=\underline{0}$. Suppose that the following constraints are imposed on the coefficients of the input signal

$$|a_0|\le c_0=1, \text{ and } |a_1|\le c_1=2.$$

Subject to these constraints, the maximum distance from the zero state is:

$$\left\|\underline{x}\left(0^+\right)-\underline{x}\left(0^-\right)\right\|^2=\sum_{i=1}^{2}x_i^2\left(0^+\right)$$
$$=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{s=1}^{2}\left|\lambda_i\right|^{j+s-2}c_{j-1}c_{s-1}\left[e^{\lambda_i K\left(0^+,\sigma\right)\sigma+\frac{1}{2}\lambda_i^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2.$$
$$=\sum_{j=1}^{2}\sum_{s=1}^{2}\left|\lambda_1\right|^{j+s-2}c_{j-1}c_{s-1}\left[\underbrace{e^{\lambda_1 K\left(0^+,\sigma\right)\sigma+\frac{1}{2}\lambda_1^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_1\sigma\right)}_{\beta_1}\right]^2$$
$$+\sum_{j=1}^{2}\sum_{s=1}^{2}\left|\lambda_2\right|^{j+s-2}c_{j-1}c_{s-1}\left[\underbrace{e^{\lambda_2 K\left(0^+,\sigma\right)\sigma+\frac{1}{2}\lambda_2^2\sigma^2}\Phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_2\sigma\right)}_{\beta_2}\right]^2.$$
$$=\sum_{s=1}^{2}\left|\lambda_1\right|^{s-1}c_0c_{s-1}\beta_1^2+\sum_{s=1}^{2}\left|\lambda_1\right|^{s}c_1c_{s-1}\beta_1^2+\sum_{s=1}^{2}\left|\lambda_2\right|^{s-1}c_0c_{s-1}\beta_2^2+\sum_{s=1}^{2}\left|\lambda_2\right|^{s}c_1c_{s-1}\beta_2^2$$
$$=2\left(\beta_1^2+\beta_2^2\right)c_0c_0+2\left(\left|\lambda_1\right|\beta_1^2+\left|\lambda_2\right|\beta_2^2\right)c_0c_1+\left(\left|\lambda_1\right|^2\beta_1^2+\left|\lambda_2\right|^2\beta_1^2\right)c_1c_1$$

In this example, $c_0=1$, $c_1=2$ and $\left|\lambda_1\right|=2$, $\left|\lambda_2\right|=3$.

So, the maximum radius is given by

$$r\triangleq\left\|\underline{x}\left(0^+\right)-\underline{x}\left(0^-\right)\right\|=\sqrt{4\left(\beta_1^2+\beta_2^2\right)+4\left(2\beta_1^2+3\beta_2^2\right)+2\left(4\beta_1^2+9\beta_1^2\right)}=\sqrt{20\beta_1^2+34\beta_2^2}\ .$$

Now, for the case that $t^*=K\left(t^*,\sigma^*\right)\sigma^*\to 0$, we have $\beta_1^2,\beta_2^2\to 1$ and

$$r\triangleq\left\|\underline{x}\left(t^*\right)-\underline{x}\left(0^-\right)\right\|=\sqrt{4+4\left(2+3\right)+2\left(4+9\right)}=\sqrt{54}\approx 7.35.$$

6. Conclusions

In this paper, a novel methodology has been proposed for approximating the distributional trajectory that transfers the state of a LTI differential system in (almost) zero time by using an impulsive input. It has been shown that no loss of generality is introduced if the impulsive input signal is chosen as a linear combination of the Dirac δ-function and its first $n-1$ derivatives, where n is the order of the system. Approximations of the impulsive input signal were considered using the Gaussian (Normal) function, and the resulting response of the system was analysed. The work has addressed the following three distinct problems:

(i) We have determined the (unique) impulsive input signal (and its smooth approximation) which transfers the state of the system from the origin to an arbitrary point in state space in zero (almost-zero) time, subject to appropriate controllability assumptions. To simplify our presentation, the simplest set of assumptions has been selected (full system controllability, single control input, distinct set of eigenvalues in the system matrix); however, extension to the general case is straightforward at the expense of possible loss of uniqueness and considerable additional complexity in the resulting mathematical expressions.

(ii) A Euclidean metric has been defined to quantify the approximation error in the state-trajectories of the system resulting from substituting impulsive input signals by smooth signals. The optimal choice of two parameters (time and volatility) characterising the family of all smooth approximating functions has been obtained, along with an interesting probabilistic interpretation.

(iii) The solution of two state-space maximum-distance problems in the context of (almost) zero-time state-transition has been presented. These correspond to two different types of constraints on the coefficients of the impulsive input signal and its smooth approximation, involving the Euclidian and infinity norms of the vector of coefficients. Both problems are tractable and can be solved via an SVD and the solution of a quadratic programming problem with box constraints, respectively.

Future work will attempt to: (i) extend the results of this paper to more general classes of systems (e.g. descriptor, singular), (ii) investigate the numerical properties of simulating impulsive trajectories and their smooth approximation, and (iii) develop alternative energy-based approximation techniques of impulsive behaviour especially in the context of large-scale systems and model reduction.

Acknowledgment: Part of the present version of the paper has been presented in the 4th *IFAC Symposium on System, Structure and Control conference,* 15-17th of September

2010, Ancona, Italy and in the Workshop "*Advances in Mathematical Control Theory and Applications*", 19-20th of May 2011, Liverpool, UK. We are very grateful to Prof. Michael Malabre, Prof. Malcolm Smith, Prof. Jean-Jacques Loiseau, and Dr Efstathios Antoniou for their helpful comments and remarks.

References

[1]. Bowen, J.M. (1994) Delta function terms arising from classical point source fields, *American Journal of Physics* Vol. **62**, pp. 511-515.

[2]. Boykin, T.B. (2003) Derivatives of the Dirac delta function by explicit construction of sequences, *American Journal of Physics* Vol. **72**, pp. 462-468.

[3]. Campbell, S.L. (1980) *Singular systems of differential equations*. I. Research Notes in Mathematics 40. Pitman, Boston.

[4]. Campbell, S.L. (1982) *Singular systems of differential equations*. II. Research Notes in Mathematics 61. Pitman, Boston.

[5]. Cobb, J.D. (1982) On the Solutions of Linear Differential Equations with Singular Coefficients, *Journal of Differential Equations* Vol. **46**, No. 3, pp. 310-323.

[6]. Cobb, J.D. (1983) Descriptor Variable Systems and Optimal State Regulation, *IEEE Transactions on Automatic Control* Vol. **28**, No. 5, pp. 601-611.

[7]. Cohen and Kirchner (1991), Approximating the Dirac distribution for Fourier analysis, *Journal of Computational Physic* Vol. **93**, pp. 312-324.

[8]. Estrada, R. and Kanwal, R.P. (2000) *Singular integral equations*, Birkhauser, Boston.

[9]. Gautshi, W. (1975), Optimally conditioned Vandermode matrices, *Numerische Mathematik* Vol. 24, pp 1-12.

[10]. Gupta, S.C. and Hasdorff, L. (1963), Changing the state of a linear system by use of normal function and its derivatives, *International Journal of Electronics* Vol. **14**, pp. 351-359.

[11]. Gupta, S.C. (1966) *Transform and state variable methods in linear systems*, Wiley New York, U.S.A.

[12]. Jaffe, S. and Karcanias, N. (1981), Matrix pencil characterization of almost (A, B)- invariant subspaces: a classification of geometric concepts. *International Journal of Control* Vol. **33**, pp. 51-93.

[13]. Kanwal, R.P. (2004) *Generalized Functions: Theory and applications*, Birkäuser, 3rd edition, USA..

[14]. Karcanias, N. and Kouvaritakis, B. (1979), Zero time adjustment of initial conditions and its relationship to controllability subspaces, *International Journal of Control* Vol. **29** (5), pp. 749-765.

[15]. Karcanias, N. and Hayton, G.E. (1982), Generalised autonomous dynamical systems, algebraic duality and geometric theory. *Proceedings of 8th IFAC World Congress*, Kyoto, Japan, Pergamon Press, pp. 289-294.

[16]. Karcanias, N. and Kalogeropoulos, G.I. (1989), Geometric Theory and Feedback Invariants of Generalized Linear Systems: A Matrix Pencil Approach, *Circuits, Systems and Signal Processing* (Special issue on Singular Systems) Vol. **8** (3), pp. 395-397.

[17]. Verghese, G.C. (1979), *Infinite Frequency Behaviour of Generalised Dynamical Systems*, PhD Thesis, Stanford University, 1979.

[18]. Verghese, G.C. and Kailath, T. (1979), Impulsive Behaviour and Dynamical Systems, *Proceedings of 4th Symposium on Mathematical Theory of Networks and Systems*, Delft, The Netherlands, pp. 162-168.

[19]. Willems, J.C. (1981). Almost Invariant spaces: An Approach to High Gain Feedback-Part I: Almost controlled invariant subspaces, *IEEE Transactions Automatic Control* Vol. AC-**26**, pp. 235-252.

[20]. Willems, J.C. (1991), Paradigms and Puzzles in the Theory of Dynamical Systems, *IEEE Transactions on Automatic Control* Vol. **36**, pp 259-294.

[21]. Zadeh, Z.A. and Desoer, C.A. (1963), *Linear system theory, the state space approach*, McGraw-Hill, New York, USA.

[22]. Zemanian, A.H. (1987), *Distribution theory and transform analysis: An introduction to generalized functions with applications*, Dover Publications, Inc, New York, A.