General Delta Operator and Transform

C. TORRES
Dept. Matemática Aplicada
Univ. Politécnica de Madrid
Campus Montegancedo
28660, Madrid, Spain
ctorres@fi.upm.es

F. MONASTERIO
Dept. Tecnologías Aplicadas
Univ. Politécnica de Madrid
Ciudad Universitaria
28040, Madrid, Spain
felix@robolab.unex.es

V. GIMÉNEZ
Dept. Matemática Aplicada
Univ. Politécnica de Madrid
Campus Montegancedo
28660, Madrid, Spain
vgimenez@fi.upm.es

Abstract: - Some significant problems of numerical sensitivity appear when a continuous time system is discretized with techniques based on the $\mathcal{Z}$-transform. In this article, a new operator and its corresponding complex transform are defined, giving rise to a new method of discretization. This new method is a generalization of the delta transform studied by Middleton and Goddwin. The discrete models that represent the continuous models are necessary in order to be able to work in such fields as digital control, filter design, signal processing or robotics. The model that we present is an alternative for those models known in the shift operator and the $\mathcal{Z}$-transform, improving them noticeable, since they solve the sensitivity problems in digital systems that carry those other models. By means of an illustrative example we show the difference in numerical sensitivity between a discrete transfer function in the zeta form and its corresponding in the delta form.

Key-Words: - discrete-time systems, sampled-data systems, sampling speed, step response.

1 Introduction

The problem of choosing which discrete model best represents a given continuous linear system has been very relevant in the last two decades, due to the great development of digital technology. Its resolution evolved since the appearance of the theory of $\mathcal{Z}$-transform, by Tsypkin in 1958. This resolution has not yet been completed, since the method of $\mathcal{Z}$-transform is only valid in infinite word length, where there are no quantification errors of the continuous system coefficients.

The most common methods of discretization used by control and signal processing engineers are based in the $\mathcal{Z}$-transform. In all of them, a rational discrete-time transfer function, in the complex variable z of the $\mathcal{Z}$-transform, is obtained from a rational continuous time transfer function, in the complex variable s of the Laplace transform.

It has been proved that in all of these methods, at high sampling speeds, the poles and zeros (which are not of excess) of the discrete-time transfer function gather at point (1,0) of the complex plane; point where the system is very sensitive to numerical errors, so that on doing an implementation in finite word length, the system can become unstable.

Also, non-minimum phase transfer functions can occur [2] which are not suitable, for example, to model reference adaptive control. Besides, the discrete models in the shift operator sometimes hide properties of its corresponding continuous models. This is the case of the frequency folding and aliasing, when a signal is sampled, or the case of the hidden oscillations in control systems.

As we see, the discrete models in the shift operator do not converge to its counterpart continuous time models when the sampling period tends to zero, which would be a highly desirable property. Therefore, we can conclude that the shift operator and its corresponding $\mathcal{Z}$-transform, are not appropriate for high sampling speeds.

Some authors in the field of deterministic digital control and in the field of stochastic digital control [6], [9], saw the need to define a new and more appropriate operator in high sampling speeds. So, in the last decade, an alternative operator to the shift operator and its corresponding transform has emerged, giving birth to a new method of discretization and transform called the Unified Transform, obtaining a unified theory of continuous and discrete time systems [7]. We can find a summary of such theory in [5].

In this article we rely on Middleton and Goodwin's ideas to develop a more general way of discretization, introducing an additional degree of freedom, through the variable n_2, at the time of choosing the discrete model that will represent the continuous model. This new construction involves the method proposed by Middleton and Goodwin as a particular case.

In section 2 we give the definition of a new operator, that we call *general delta operator*. We study how to obtain a discrete-time system with this new operator. We also give the definition of the *General Delta transform* that corresponds to the new operator. We incorporate an important result about the convergence of a sampled system in the general delta form to its counterpart continuous system at sufficiently high sampling speeds.

In section 3 we show the advantages of the General Delta transform opposite to the $\mathcal{Z}$-transform from the point of view of the sensitivity of the poles due to variations in the coefficients, that is, we show some of the numerical advantages of the models in the general delta form opposite to the models in the zeta form.

2 General Delta Operator and Transform

In this article we define a new operator and its respective complex transform, more general than the operator and transform given by Middleton and Goodwin in 1986, to which we will call, respectively, general delta operator and transform (we will call the delta operator defined in [6] delta operator of Middleton). The necessary theory of systems for the new operator and transform is developed, showing the connections of this theory with the shift operator, the $\mathcal{Z}$-transform and with the formulations of continuous time.

The new concept of transform is not a function in a variable, but a set of functions, according to the value that we may give to a new parameter called n_2. This parameter n_2 allow us to be more accurate in regards to the discrete-time model that we wish to choose as the representative of the continuous time model from which we begin.

2.1 Time domain

It is necessary to define a temporal operator from which we can obtain a linear discrete time approximation of a linear differential equation, that must verify the following important condition: when the discretization parameter tends to zero, the discrete equation tends to the continuous equation. In order to do so, the linear general delta operator (δ) is defined depending on the discretization parameter (T) in the following way.

Given a differentiable and continuous function $y(t)$, it is defined

$$\delta y(t) = \frac{y(t+T) - y(t)}{T} + f(t,T) \qquad (1)$$

where $f(t,T)$ is any function, so that $f(t,0) = 0$ for all t.

So we insure that when $T \to 0$ it is verified that

$$\lim_{T\to 0} \delta y(t) = \frac{dy(t)}{dt} \qquad (2)$$

The following particular case is studied

$$f(t,T) = n_2(\delta y(t+T) - \delta y(t)) \quad \text{with } n_2 \in \mathbb{R} \quad (3)$$

With this definition for the function $f(t,T)$, the *general delta operator* is defined by

$$n_1 \delta y(t) - n_2 \delta y(t+T) = \frac{y(t+T) - y(t)}{T} \qquad (4)$$

with $n_1 - n_2 = 1$

Using the shift operator ($q(y(t) = y(t+T))$ and assuming that the identity operator is represented by the constant 1, the following expression is obtained

$$\delta(n_1 - n_2 q) y(t) = \frac{(q-1)}{T} y(t)$$

with $n_1, n_2 \in \mathbb{R}$ and $n_1 - n_2 = 1$. Due to the fact that the product of operators $(n_1 - n_2 q)^{-1}(q-1)$ commutes, the following definition of *general delta operator* is obtained accordingly to the shift operator

$$\delta = \frac{q-1}{T(n_1 - n_2 q)} \quad \text{with } n_1 - n_2 = 1 \qquad (5)$$

In the case $n_2 = 0$ it coincides with the delta operator of Middleton [6].

In the theorem 2.5 it will be shown that the model in a general delta operator converges to the continuous model, while the model in the shift operator does not converge to the continuous model and places the poles at the point $z = 1$ of the complex plane, to which the numerical errors will seriously affect.

The models in the general delta operator from a continuous model in the differential operator are obtained in the following way. Consider a given continuous time linear state space model of the form

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \qquad (6)$$

$$y(t) = Cx(t) \qquad (7)$$

where $x(t)$ is the state vector, $u(t)$ the input variable, and $y(t)$ the output variable. The general solution to the state equation at t and $t+T$ has the form

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)} Bu(\tau) d\tau \qquad (8)$$

$$x(t+T) = e^{AT}x(t) + e^{AT} \int_t^{t+T} e^{A(t-\tau)} Bu(\tau) d\tau \qquad (9)$$

Therefore, since δ is lineal it is verified that

$$\delta x(t+T) = e^{AT}\delta x(t)+e^{AT}\delta\left[\int_t^{t+T} e^{A(t-\tau)}Bu(\tau)d\tau\right] \quad (10)$$

Applying the expression (4) and (9) in (10) the following result is obtained

$$\left[(n_1 I - n_2 e^{AT})\delta + \frac{1}{T}(I - e^{AT})\right] x(t) = (1 + n_2 T\delta)\left[\frac{1}{T}\int_0^T e^{A\tau}Bu(t+T-\tau)d\tau\right] \quad (11)$$

Remark 2.1 The expression $\frac{1}{T}\int_0^T e^{A\tau}Bu(t+T-\tau)d\tau$ converges to $Bu(t)$ when T tends to zero, assuming the continuity of $u(t)$ in the interval $[t, t+T]$.

If, for example, we assume that $u(t)$ is constant at intervals of constant width T, the expression (11) take the form

$$A(\delta)x(t) = B(\delta)u(t) \quad \text{with}$$

$$A(\delta) = \left[(n_1 I - n_2 e^{AT})\delta + \frac{1}{T}(I - e^{AT})\right] \quad (12)$$

$$B(\delta) = (1 + n_2 T\delta)\left[\frac{1}{T}\int_0^T e^{A\tau}d\tau\right] B \quad (13)$$

For other functions of $u(t)$ other expressions of $B(\delta)$ will be obtained.

Since the purpose of this article is to study the problem of discretization of continuous time systems, we will express the temporal parameter in the form t_k, t_{k+1}, etc... From now on, we will consider that the interval of discretization is constant and of width the parameter of discretization, which we will call the sampling period T. So $t_k = kT$.

In this way the following discrete-time system in the operator δ is obtained

$$\delta x(kT) = (n_1 I - n_2 e^{AT})^{-1}\left[\frac{1}{T}(e^{AT} - I)x(kT) + (1 + n_2 T\delta)\left[\frac{1}{T}\int_0^T e^{A\tau}Bu(kT+T-\tau)d\tau\right]\right] \quad (14)$$

and if we assume that $u(t)$ is constant in intervals of constant width T we obtain the following state equation

$$\delta x(kT) = (n_1 I - n_2 e^{AT})^{-1}\tfrac{1}{T}(e^{AT} - I)x(kT) + (n_1 I - n_2 e^{AT})^{-1}(1 + n_2 T\delta)\left[\tfrac{1}{T}\int_0^T e^{A\tau}d\tau\right]Bu(kT) \quad (15)$$

This method of discretization must not be confused with those based on the discrete approximation of a derivative or integral. These latter numerical methods, such as the method of Tustin, carry on the discretization substituting in the differential equation the differential operator by the corresponding approximation.

Remark 2.2 The matrix expressions in the δ operator obtained in (15) are analytical expressions and their numerical computation must be done without going through the exponential matrix, which is the matrix that appears in the discrete-time systems obtained with the shift operator.

Taking into account that the matrices $(e^{AT} - I)$ and $(n_1 I - n_2 e^{AT})^{-1}$ commute and defining the matrix

$$\Omega = \frac{1}{T}\int_0^T e^{A\tau}d\tau = I + \frac{AT}{2!} + \frac{A^2T^2}{3!} + \ldots \quad (16)$$

which converges to the identity matrix when the sampling period T tends to zero, we obtain the following discrete-time system in the general delta operator.

$$\begin{cases} \delta x(kT) = (I - n_2 T\Omega A)^{-1}\Omega A x(kT) + \\ \qquad (I - n_2 T\Omega A)^{-1}\Omega B(1 + n_2 T\delta)u(kT) \\ y(kT) = Cx(kT) \end{cases} \quad (17)$$

which converges to the continuous time system when T tends to zero, since

$$\frac{1}{T}(e^{AT} - I) = \Omega A = A\Omega \overset{T\to 0}{\longrightarrow} A \quad (18)$$

$$(n_1 I - n_2 e^{AT})^{-1} = (I - n_2 T\Omega A)^{-1} \overset{T\to 0}{\longrightarrow} I \quad (19)$$

$$\left[\frac{1}{T}\int_0^T e^{A\tau}d\tau\right] B = \Omega B \overset{T\to 0}{\longrightarrow} B \quad (20)$$

The equations (17) are the ones we must use in order to discretize a continuous time system given in state variables, in the general delta domain. If we use the equations (15) the possible numerical errors made in the shift operator domain (that is, in the exponential) will be transmitted to the delta domain which is by no means desirable.

2.2 Complex Domain

In the analysis of lineal systems, and in particular for certain applications such as the design of filters and controllers, it will be very convenient to know the frequency components of the solution of a linear differential equation. In order to do so we normally use the Laplace transform defined in the following form.

$$\mathcal{L}(y(t)) = \int_0^\infty e^{-st} y(t) dt$$

Applying the Laplace transform to the expression (4) the following is obtained

$$\begin{aligned}\mathcal{L}(n_1 \delta y(t) - n_2 \delta y(t+T)) \\ &= \frac{1}{T}\left[e^{sT} Y(s) - e^{sT}\int_0^T y(\tau)e^{-s\tau}d\tau - Y(s)\right] \\ &= \frac{e^{sT}-1}{T} Y(s) - e^{sT}\frac{\int_0^T y(\tau)e^{-s\tau}d\tau}{T} \qquad (21)\end{aligned}$$

Applying limit when $T \to 0$ in both sides of the previous equation we obtain

$$\begin{aligned}\lim_{T\to 0} \mathcal{L}(n_1 \delta y(t) - n_2 \delta y(t+T)) &= \\ = \mathcal{L}((n_1 - n_2)\delta y(t)) = \mathcal{L}(\delta y(t)) \\ \lim_{T\to 0}\left[\frac{e^{sT}-1}{T} Y(s) - e^{sT}\frac{\int_0^T y(\tau)e^{-s\tau}d\tau}{T}\right] &= \\ = sY(s) - \lim_{T\to 0} y(T)e^{-sT} = sY(s) - y(0)\end{aligned}$$

Therefore, we obtain the following relation between the δ operator and the Laplace transform

$$\mathcal{L}(\delta y(t)) = sY(s) - y(0) \qquad (22)$$

that coincides with the Laplace transform the derivative of a function, that is

$$\mathcal{L}(\delta y(t)) = \mathcal{L}\left(\frac{dy(t)}{dt}\right) \qquad (23)$$

On the other hand, if we apply again the Laplace transform to the expression (4), taking into account the relation (22) we have

$$\begin{aligned}\mathcal{L}(n_1 \delta y(t) - n_2 \delta y(t+T)) = s(n_1 - n_2 e^{sT})Y(s) + \\ s n_2 e^{sT}\int_0^T y(\tau)e^{-s\tau}d\tau - n_1 y(0) + n_2 y(T)\end{aligned} \qquad (24)$$

then from expression (21) it follows

$$\begin{aligned}s(n_1 - n_2 e^{sT})Y(s) = \tfrac{e^{sT}-1}{T} Y(s) - \\ (1 + n_2 T s)e^{sT}\left[\frac{\int_0^T y(\tau)e^{-s\tau}d\tau}{T}\right] + n_1 y(0) - n_2 y(T)\end{aligned} \qquad (25)$$

which verifies that

$$\begin{aligned}\lim_{T\to 0}(1 + n_2 T s)e^{sT}\left[\frac{\int_0^T y(\tau)e^{-s\tau}d\tau}{T}\right] &= y(0) \\ \lim_{T\to 0}(n_1 y(0) - n_2 y(T)) &= y(0)\end{aligned}$$

The expression (25) suggest the introduction of a complex variable γ defined in the following way

$$\gamma = \frac{e^{sT}-1}{T(n_1 - n_2 e^{sT})} \quad \text{with} \quad n_1 - n_2 = 1 \qquad (26)$$

or inversely

$$e^{sT} = \frac{1 + n_1 \gamma T}{1 + n_2 \gamma T} \qquad (27)$$

It is easy to verify that if $T \to 0$ then $\gamma \to s$. This does not occur with the complex variable z of the $\mathcal{Z}$-transform, since $z = e^{sT}$ converges to 1 when T tends to zero.

It is well known that the output $y^*(t)$ of an ideal sampler defined as a pulse train modulation of a continuous time signal $y(t)$, has Laplace transform

$$Y^*(s) \equiv \mathcal{L}(y^*(t)) = \sum_{k=-\infty}^{\infty} y(kT)e^{-kTs}$$

This expression suggests the following definition of the *General Delta transform*, $\mathcal{D}$. We define a complex transformation, in the γ variable of the sampled signal $y^*(t)$ in the following way

$$\mathcal{D}(y(t)) = T\mathcal{L}(y^*(t)) \qquad (28)$$

The reason why we add the factor T is because upon sampling a signal we increase the gain in a factor of $1/T$, and thus we compensate it. This way the discrete-time frequency response, will have the same amplitude as the continuous time frequency response.

Taking into account the relation between e^{sT} and γ given by (27), the *General Delta transform* will stay in the form

$$\mathcal{D}(y(t)) \equiv T\sum_{k=-\infty}^{\infty} y(kT)\left(\frac{1 + n_1\gamma T}{1 + n_2 \gamma T}\right)^{-k} \qquad (29)$$

Using the $\mathcal{Z}$-transform defined in the following way

$$\mathcal{Z}(y(t)) = \sum_{i=-\infty}^{\infty} y(i)z^{-i}$$

we obtain

$$\mathcal{D}(y(t)) = T\mathcal{Z}(y(t))\big|_{z=\frac{1+n_1\gamma T}{1+n_2\gamma T}} \qquad (30)$$

The relations between the complex variables z, s and γ are the following

$$z = e^{sT} = \frac{1 + n_1\gamma T}{1 + n_2 \gamma T} \qquad (31)$$

$$\gamma = \frac{z-1}{T(n_1 - n_2 z)} \qquad (32)$$

From this we can conclude that the *General Delta transform*, $\mathcal{D}$, can be easily realized through the Z-transform by simply doing a change in the z (taking into account that this will be done in the analytical studies, but not in numerical implementations, as we have aforementioned). This way we also define the gamma transfer function of a lineal system, which allow us to generalize all the well known techniques of linear systems analysis in the complex and frequency domains.

The following theorem whose demonstration can be found in [4], shows how the General Delta transform operates on the general delta operator.

Theorem 2.3

$$\mathcal{D}(\delta y(t)) = \gamma Y(\gamma) - (1 + n_1 T\gamma)(y(0) + n_2 T\delta y(0)) \quad (33)$$

Next, we offer a lemma that relates a model in the δ operator of a sampled system with its corresponding Delta transform.

Lemma 2.4 Let us consider the following discrete-time input-output model in the general delta operator

$$A(\delta)y(kT) = (1 + n_2\delta T)B(\delta)u(kT)$$

with $A(\delta) = \delta^n + a_{n-1}\delta^{n-1} + \cdots + a_0$, $B(\delta) = b_m\delta^m + \cdots + b_0$, $m < n$, T the sampling period, $u(kT)$ the input variable, $y(kT)$ the output variable and $A(\delta)$, $B(\delta)$ two coprime polynomials.

Taking General Delta transform on both sides of the previous equation, taking into account how the Delta transform acts on the general delta operator (see theorem 2.3) and assuming zero initial conditions, we obtain

$$Y(\gamma) = G(\gamma)U(\gamma) \quad \text{with} \quad G(\gamma) = \frac{(1 + n_2\gamma T)B(\gamma)}{A(\gamma)} \quad (34)$$

where $U(\gamma)$ and $Y(\gamma)$ are the General Delta transform of the input and output variables. $G(\gamma)$ is defined as the system transfer function, and is the General Delta transform of the unit pulse response $g(kT)$.

For linear causal time-invariant systems any discrete-time transfer function in the General Delta transform, corresponding to the operator δ, obtained by step-invariance method, converges on the continuous time model from which it has been obtained, when the sampling period T tends to zero. This can be seen in the following theorem demonstrated in [4].

Theorem 2.5 Given a linear causal time-invariant continuous time system with transfer function

$$G(s) = \frac{C(s)}{A(s)} = \frac{c_m s^m + \cdots + c_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_0}$$

with $\partial C = m \leq \partial A = n$, the following pulse transfer function, obtained by step-invariance, in the general delta form is obtained

$$\begin{aligned} G_D(\gamma) &= \tfrac{C_D(\gamma)}{A_D(\gamma)} \quad (\partial A_D = n, \partial C_D \leq n) \\ &= \tfrac{C_\epsilon(\gamma) + C_R(\gamma)}{A_D(\gamma)} \quad (\partial C_\epsilon \leq n, \partial C_R = \partial C = m) \\ &= \tfrac{c'_n\gamma^n + \cdots + c'_{m+1}\gamma^{m+1} + c'_m\gamma^m + \cdots + c'_1\gamma + c'_0}{\gamma^n + a'_{n-1}\gamma^{n-1} + \cdots + a'_0} \end{aligned} \quad (35)$$

satisfying

$$\lim_{T\to 0} A_D(\gamma) = A(\gamma) \quad (36)$$

$$\lim_{T\to 0} C_\epsilon(\gamma) = 0 \quad (37)$$

$$\lim_{T\to 0} C_R(\gamma) = C(\gamma) \quad (38)$$

where T is the sampling period and the limit refers to the coefficients of the polynomials, that is,

$$a'_{n-1} \stackrel{T\to 0}{\longrightarrow} a_{n-1}, \quad \ldots \quad , a'_0 \stackrel{T\to 0}{\longrightarrow} a_0 \quad (39)$$

$$c'_n \stackrel{T\to 0}{\longrightarrow} 0, \quad \ldots \quad , c'_{m+1} \stackrel{T\to 0}{\longrightarrow} 0 \quad (40)$$

$$c'_m \stackrel{T\to 0}{\longrightarrow} c_m, \quad \ldots \quad , c'_1 \stackrel{T\to 0}{\longrightarrow} c_1, c'_0 \stackrel{T\to 0}{\longrightarrow} c_0 \quad (41)$$

3 Numerical Advantages of the General Delta transform

Many authors have recognized the importance of the numerical questions in such fields as important as digital control ([10], [1], [11], [3]), signal processing and reconstruction ([8]) among others.

The numerical errors can be dues to several reasons. For example dues to the coefficients quantification is necessary in digital systems, in which the data are taken and manipulated by a computer and therefore represented in finite word length. It is also necessary the use of finite word length when dealing with on-line computations. A typical case of this arises from adaptive control applications, where the controller is updated on-line using an estimated model.

The sensitivity of the roots p_k of a polynomial $P(x)$ with regard to variations in the coefficient a_i is given by the following equation [3]

$$\triangle p_k \approx -\frac{p_k^{n-i}}{\prod_{\substack{j=1 \\ j\neq k}}^{n} (p_k - p_j)} \triangle a_i$$

Therefore, the numerical sensitivity in the roots due to variations in the coefficients will be very high if the roots are close. Let's now assume that P represents the denominator of the transfer function of a sampled system, in the zeta form or in the delta form. In every method of discretization based on the $\mathcal{Z}$-transform it is verified that, when the sampling period T tends to zero, the poles z_k of the sampled system go to 1, since they are of the form $z_k = e^{s_k T}$, being s_k the poles of the continuous system from which the sampled system comes from. Therefore, it is verified that

$$\triangle z_k \approx -\frac{z_k^{n-i}}{\prod_{\substack{j=1\\ j\neq k}}^{n}(z_k - z_j)} \triangle a_i \stackrel{T\to 0}{\longrightarrow} \infty \tag{42}$$

Then a fast sampling frequency, that approximates all poles to $z = 1$ will need high accuracy in the coefficients. This neighborhood to $z = 1$ where the poles of the sampled system will concentrate, due to its high numerical sensitivity will be called *dangerous zone*. This is due to the fact that small numerical errors can cause the fatality that the stable poles jump to the unstable zone and could completely change the system performance.

On the other hand, in the case of the General Delta transform, when the sampling frequency is fast, the poles γ_k of the sampled system will approach to its corresponding poles of the continuous time system, since $\gamma_k = \frac{e^{s_k T}-1}{T(n_1 - n_2 e^{s_k T})}$ (see [4]) and not to a specific value of the stability border, as it happens in the discretization methods in the zeta form.

It is also verified that the coefficients of the sampled system converge to its corresponding continuous time system coefficients when the sampling period tends to zero (theorem 2.5). Therefore, in the delta case when the sampling period T tends to zero, the sensitivity of the sampled system poles, γ_k, will depend on the sensitivity of its corresponding continuous time system poles, s_k, with respect to variations on its corresponding coefficients, a_i^c,

$$\begin{aligned}\triangle \gamma_k &\approx -\frac{\gamma_k^{n-i}}{\prod_{\substack{j=1\\ j\neq k}}^{n}(\gamma_k - \gamma_j)} \triangle a_i \\ &= \frac{\left(\frac{e^{s_k T}-1}{T(n_1-n_2 e^{s_k T})}\right)^{n-i}}{\prod_{\substack{j=1\\ j\neq k}}^{n}\left(\frac{e^{s_k T}-1}{T(n_1-n_2 e^{s_k T})} - \frac{e^{s_j T}-1}{T(n_1-n_2 e^{s_j T})}\right)} \triangle a_i \\ &\downarrow T\to 0 \\ &\frac{s_k^{n-i}}{\prod_{\substack{j=1\\ j\neq k}}^{n}(s_k - s_j)} \triangle a_i^c\end{aligned} \tag{43}$$

This makes the models in the general delta form at high sampling speeds to be more robust than the models in the zeta form.

4 Examples

In the following example it is shown that the models in the zeta form are very sensitive to numerical errors in the so called dangerous zone, whereas the models in the general delta form are much more robust. We see that in the zeta form it is necessary to have at least 24 *bits* in the coefficients quantification, in order to obtain an acceptable response. However, in the delta models, we obtain a satisfactory response with only 2 *bits*. Moreover, with 8 *bits* and $n_2 = -\frac{1}{2}$, it is obtained the same response that the corresponding without quantifications.

Let us suppose that we want to implement a sampled system whose discrete time transfer function in the zeta form is

$$H_d(z) = \frac{6.1\,10^{-8}}{z^3 - 2.9788z^2 + 2.9577122z - 0.97891214} \tag{44}$$

with poles in $z = 0.9994$, $z = 0.9922$ and $z = 0.9872$.

This transfer function, $H_d(z)$, corresponds to the discretization by means of the step invariance method of a continuous time system with a sampling period $T = 0.01$.

The corresponding pulse transfer function of the delta transform with $n_2 = -\frac{1}{2}$, is

$$H_d(\gamma_B) = \frac{(-0.0006\gamma_B + 0.0617)(1 - \frac{1}{2}0.01\gamma_B)}{\gamma_B^3 + 2.1313\gamma_B^2 + 1.1331\gamma_B + 0.0606} \tag{45}$$

that has its poles in $\gamma_B = -0.0601$, $\gamma_B = -1.2882$ and $\gamma_B = -0.7830$; and with $n_2 = 0$ is

$$H_d(\gamma_M) = \frac{0.0610}{\gamma_M^3 + 2.12\gamma_M^2 + 1.122\gamma_M + 0.06} \tag{46}$$

that has its poles in $\gamma_M = -0.0601$, $\gamma_M = -1.2802$ and $\gamma_M = -0.7797$.

In order to implement this system it is necessary to quantify the coefficients and this is why we study the quantification effects of the coefficients of the denominator polynomial in the poles of the system, in both forms z y δ.

We study the roots of the denominator polynomial of $H_d(z)$, that are obtained with the coefficients quantified for n *bits* with n from 2 to 24, round them to four digits. All computations have been made with the n digits corresponding to each

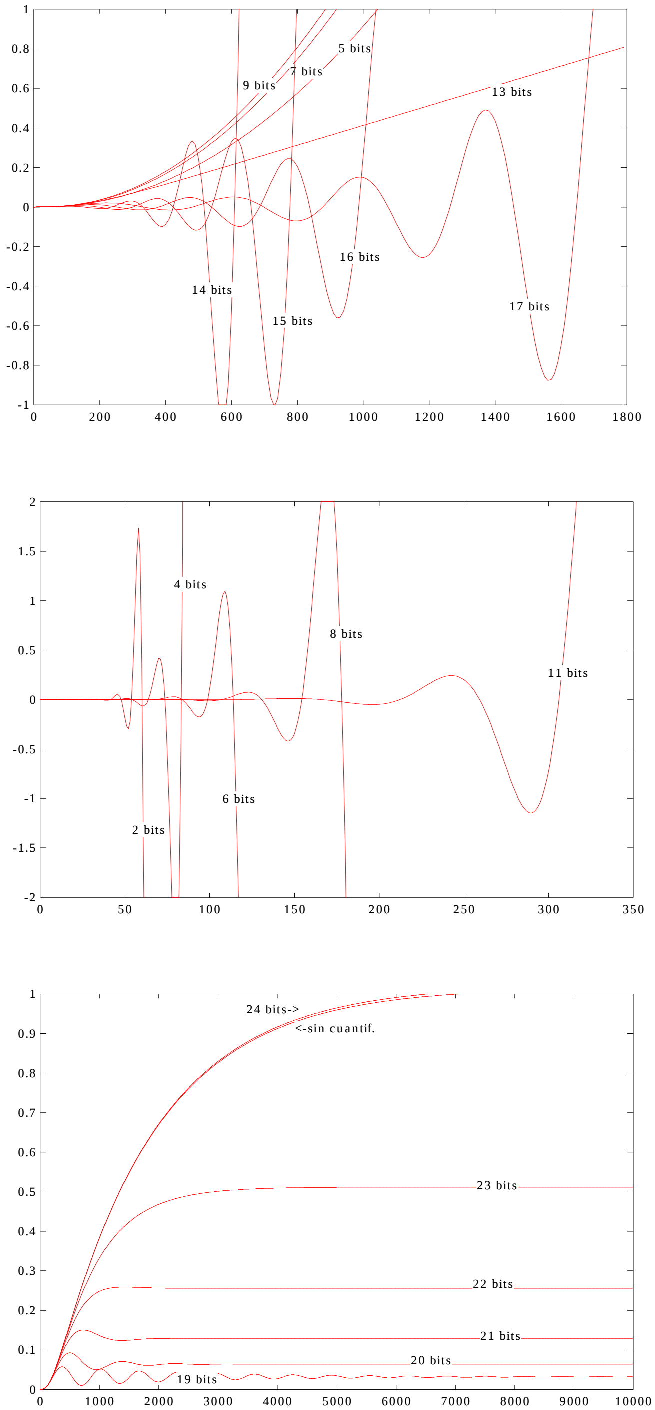

Fig. 1: Step response to the z-model

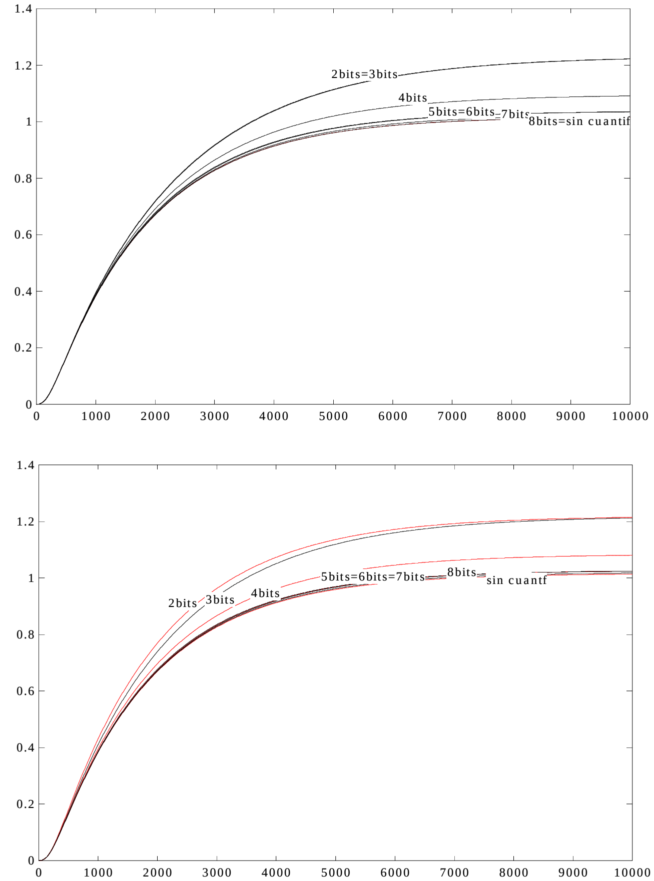

Fig. 2: Step response to the δ-models

case. It is verified that we at least need 19 $bits$ in the quantification of the coefficients in order to obtain stable poles, even though they are not exactly the same as those originals of the system. It is also verified that in order to obtain the true poles of the system, at least 24 $bits$ are necessary.

In figure 1, it is shown the step response of the system for different coefficients quantifications of the denominator polynomial of $H_d(z)$. In the first graph of the figure 1we see the response for quantifications with 5, 7, 9, 13, 14, 15, 16 y 17 $bits$; in the second graph we have the response for quantifications with 2, 4, 6, 8 y 11 $bits$; in all these cases the system is clearly unstable. In the third graph we see the behavior of the system for quantifications with 19, 20, 21, 22, 23 and 24 $bits$ and the response of the system is also shown when there are no quantifications in the coefficients of $H_d(z)$; as it can be seen the response of the quantified system is not acceptable, in comparison to the response of the true system, for a quantification of less than 24 $bits$.

In this example the transfer function $H_d(z)$ is highly sensitive with regard to the coefficients quantification of the denominator polynomial.

On the other hand we compute the roots from the coefficients of $H_d(\gamma_B)$ and $H_d(\gamma_M)$ quantified for n *bits* with n from 2 to 12, rounding them to four digits. It is verified that with only 2 *bits* in the quantification of the coefficients stable poles are obtained (see [4]),even though they are not exactly the same as those originals of the system. It is noticed that in order to obtain the true poles of the system 12 *bits* are necessary.

In figure 2 we see that the response is acceptable even with 2 *bits*. We also see, in the first graph of the figure 2, that the response of the quantified system equates to that of the system without quantifications, with only 8 *bits*, in the case $n_2 = -\frac{1}{2}$. In the case $n_2 = 0$, we see in the second graph of figure 2, that the response of the system without quantification and that of the quantified system with 8 *bits* slightly differ. As we see these discrete time models with the General Delta transform behave much better than the model $H_d(z)$.

5 Conclusion

In this work, it has been introduced an operator which, by means of a parameter, generalizes the known delta operator of discrete time. It emphasizes its properties of robustness before the variation of parameters due to quantification, without affecting the linear aspects of the discretized model. The presented analytical results guarantee the desire properties of robustness when a low sampling period is chosen. Computer simulations have been carried out showing qualitatively the differences in behavior between the proposed operator and the $\mathcal{Z}$-transform. Numerous questions, both theoretical and practical, remain open. Particularly the most adequate selection of parameter n_2, which has been studied in [4] and whose results will be presented in a future communication.

References:

[1] M.E. Ahmed and P.R. Belanger. Scaling and roundoff in fixed-point implementation of control algorithms. *IEEE Transaction on industrial electronics*, IE-31(3):228–234, August 1984.

[2] K.J. Aström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.

[3] K.J. Aström and B. Wittenmark. *Sistemas controlados por computador.* Editorial Paraninfo, Madrid, 1988.

[4] C. Torres Blanc. *Aproximación a la Unificación de Sistemas Lineales Continuos y Discretos.* F.I. - U.P.M., Madrid, 2000.

[5] G. Goodwin, R.H. Middleton, and H.V. Poor. High-speed digital signal processing and control. *IEEE Proceedings*, 80(2):238–259, February 1992.

[6] R.H. Middleton and G.C. Goodwin. Improved finite word length characteristics in digital control using delta operators. *IEEE Trans. on Automatic Control*, AC-31(11):1015–1021, November 1986.

[7] R.H. Middleton and G.C. Goodwin. *Digital Control and Estimation. A Unified Approach.* Prentice-Hal, Inc., Englewood Cliffs, New Jersey, 1990.

[8] A.V. Oppenheim and R.W. Schafer. *Digital signal processing.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[9] V. Peterka. Control of uncertain processes: applied theory and algorithms. *Kybernetika*, 22:1–102, 1986.

[10] C. Tan and B. C. McInnes. Adaptive digital control implemented using residue number systems. *IEEE Transaction on automatic control*, AC-27(2):499–502, April 1982.

[11] D. Williamson. Roundoff noise minimization and pole-zero sensitivity in fixed-point digital filters using residue feedback. *IEEE Transaction on acoustics, speech, and signal processing*, ASSP-34(5):1210–1220, October 1986.