Discretization schemes for constraint stabilization in nonlinear differential-algebraic systems

Pierluigi Di Franco, Giordano Scarciotti and Alessandro Astolfi

Abstract—In this paper the problem of simulation of differential-algebraic systems is addressed. In modelling mechanical systems the use of redundant coordinates and constraints results in differential-algebraic equations, the integration of which can lead to numerical instabilities, such as the so-called drift phenomenon. In [1] the authors have proposed a globally convergent conceptual continuous-time algorithm for the integration of constrained mechanical systems which ensures the existence of solutions and global attractivity of the solution manifold. The objective of this paper is to study the numerical implementation of the algorithm presented in [1]. In addition, the stability properties of the constrained system in the manifold are studied in both the continuous and discrete time cases. The proposed technique is illustrated by means of a simple example.

I. INTRODUCTION

The numerical analysis of differential-algebraic (DA) systems has gained great relevance in the last decades. The DA formulation is more flexible than that given by ordinary differential equations (ODEs) in modelling large-scale systems, often composed by the interconnection of multiple differential subsystems and in which the relations among the subsystems are described by algebraic constraints. For instance, constrained mechanical systems are inherently simpler to model in the framework of differential-algebraic equations (DAEs). The interest raised in the simulation of mechanical systems, together with the increasing computational power of computers, have motivated the development of different techniques for the integration of DA systems, see for example [2], [3], [4], [5] and [6].

One of the most relevant problems in integrating DAEs is the drift phenomenon. Since the solution manifold is invariant but not attractive, see [7], any constraint violation due to numerical approximations may make the solution drift away from the manifold. Stabilization methods provide a possible solution to this problem within a control theory perspective. The basic idea of such methods is to stabilize the constraint manifold by feeding back the constraint violation. Among all the methods of this class, Baumgarte’s stabilization has gained prominence because of its simplicity, see [8]. However, this method does not solve all possible numerical instabilities as recently pointed out in [1].

In [1] the authors have proposed a globally convergent conceptual continuous-time algorithm for the integration of constrained mechanical systems based on the geometric interpretation of DA systems given in [9]. This algorithm ensures global convergence of the constraint violation onto the manifold and existence of solutions. However, a discretization scheme is required for any numerical implementation of the algorithm. In addition, no results have been provided on the preservation of stability properties of the system in the manifold.

Based on the results presented in [9] and [1], the objective of this study is to investigate both the aforementioned problems. In Section II we introduce some preliminaries and recall results from [1] which are instrumental for the rest of the paper. The section is concluded with a new result regarding the stability properties for the “stabilized system” in continuous-time. In Section III we provide the main theoretical results of the paper: two discretization schemes for the algorithm proposed in [1] are presented and conditions for the preservation of the continuous-time stability properties given in [1] and in Section II are provided. In Section IV the method is validated via an example in which a simple nonlinear oscillator is integrated with one of the proposed schemes. Finally in Section V we report our conclusions.

Notation. We use standard notation. The set of real and natural numbers (including 0) are denoted, respectively, by \( \mathbb{R} \) and \( \mathbb{N} \). The symbols \( \mathbb{R}_{\geq 0} \) and \( \mathbb{N}_{\geq 0} \) indicate the sets of non-negative real numbers and non-negative integer numbers, respectively. The superscripts \( T \) and \( -T \) represent the transposition operator and the transposition of the inverse operator, respectively. \( I \) represents the identity matrix. Given a matrix \( A \), \( A < 0 \) denotes that \( A \) is negative definite and symmetric. Given a vector \( x \), the symbol \( \|x\| \) denotes its \( L^2 \) norm while the symbol \( x_i \) denotes the \( i \)-th element of \( x \). Given a map \( f : \mathbb{R}^n \to \mathbb{R} \), we use the symbol \( \frac{\partial f}{\partial x} \) to denote the row vector of partial derivatives of \( f \) with respect to the vector \( x \in \mathbb{R}^n \). The symbol \( L_f h \) denotes the Lie derivative of the function \( h \) along the vector field \( f \), while \( L_f^k h = L_f \left( L_f^{k-1} h \right) \).

II. PRELIMINARIES

In this section we introduce some definitions and give some preliminary results which are instrumental for the remaining of the paper. To begin with we recall the relation between differential-algebraic systems and a class of ordinary differential equations presented in [1]. At the end...
of the section a new result is given regarding the stability properties of the solutions of the system.

Consider a nonlinear differential-algebraic system of the form

\[ \begin{align*}
\dot{x} &= f(x) + g(x)\lambda, \\
0 &= h(x),
\end{align*} \tag{1} \]

where \( x(t) \in \mathbb{R}^n \) is the state vector, \( \lambda(t) \in \mathbb{R}^m \) is the algebraic variable, and \( f : \mathbb{R}^n \to \mathbb{R}^n, g : \mathbb{R}^n \to \mathbb{R}^{n \times m} \) and \( h : \mathbb{R}^n \to \mathbb{R}^m \) are smooth mappings. For simplicity we assume that \( m = 1 \) and that the origin is the only equilibrium point of the system. All statements are valid locally around the origin if not otherwise stated. We now recall the definition of index for a DA system.

**Definition 1.** [10] The differentiation index \( \nu \) of the differential-algebraic system (1) is equal to the minimum number of times it is necessary to differentiate (1) to find an explicit expression of \( \dot{x} \) and \( \dot{\lambda} \) as continuous functions of \( x \) and \( \lambda \).

For the DA system (1) the following proposition holds.

**Proposition 1.** [9] Consider system (1). Assume it has index \( \nu = r + 1, r \in \mathbb{N}_{\geq 0} \). Consider the mapping \( z = \Phi(x) \), with

\[ \Phi(x) = \left( h(x), L_fh(x), \ldots, L_r^{-1}h(x), \phi_r+1, \ldots, \phi_n \right)^\top. \tag{2} \]

Assume that the Jacobian of \( \Phi(x) \) is invertible and that \( \phi_i : \mathbb{R}^n \to \mathbb{R}, \) for \( i = r + 1, \ldots, n \), are such that \( L_g \phi_i(x) = 0 \). Then system (1) can be written as

\[ \begin{align*}
\xi_1 &= 0, \\
\dot{\xi}_1 &= \xi_2, \\
\dot{\xi}_2 &= \xi_3, \\
&\vdots \\
\dot{\xi}_{r-1} &= \xi_r, \\
\dot{\xi}_r &= b(\xi, \mu) + a(\xi, \mu)\lambda, \\
\dot{\mu} &= q(\xi, \mu),
\end{align*} \tag{3} \]

where \( \xi_i(t) \in \mathbb{R}, \) for \( i = 1, \ldots, r, \) \( \mu(t) \in \mathbb{R}^{n-r}, \) \( \xi = (\xi_1, \ldots, \xi_r)^\top, z = (\xi, \mu)^\top, a, b \) and \( q \) are smooth mappings and \( a(0, 0) \neq 0 \). Moreover, the solution manifold of system (3) is

\[ \mathcal{M} = \left\{ (\xi^\top, \mu^\top)^\top : \xi = 0 \right\}. \tag{4} \]

Usual index reduction techniques (see [11], [12] or [13]) would normally impose \( \dot{\xi} = 0 \) and directly integrate the differential subsystem in (3). However, since the solution of (3) is

\[ \xi(t) = \sum_{i=1}^{r} \frac{t^{r-1}}{(r-1)!}\xi(0), \]

it is clear that violation of the constraints would grow unbounded, if \( r > 1 \). This phenomenon is known as constraint “drift” and many stabilization methods have been proposed in the literature to deal with this issue, see for example [8], [14], [15] and [16]. In addition, as pointed out in [1], stabilization methods which rely on linear feedback control theory do not guarantee existence of solutions for all \( t \geq 0 \). Thus, in [1] a globally stable convergent continuous-time algorithm has been presented. The main properties of the algorithm therein are recalled in Proposition 2 and Theorem 1. Consider the following differential system

\[ \begin{align*}
\dot{\xi}_1 &= k_1(\xi, \mu)\xi_1, \\
\dot{\xi}_2 &= k_2(\xi, \mu)\xi_2, \\
&\vdots \\
\dot{\xi}_{r-1} &= k_{r-1}(\xi, \mu)\xi_{r-1}, \\
\dot{\xi}_{r} &= k_r(\xi, \mu)\xi_{r}, \\
\dot{\mu} &= q_0(\mu) + \sum_{i=1}^{r} q_i(\xi, \mu)\xi_{i},
\end{align*} \tag{5} \]

where \( \xi_i(t) \in \mathbb{R}, \) for \( i = 1, \ldots, r, \) \( \mu(t) \in \mathbb{R}^{n-r}, \) \( \xi = (\xi_1, \ldots, \xi_r)^\top, q_0 \) and \( q_i, \) for \( i = 1, \ldots, r, \) are smooth mappings and

\[ k_i(\xi, \mu) = -\frac{\delta^2}{2}||q_i(\xi, \mu)||^2 - \epsilon, \tag{6} \]

for \( i = 1, \ldots, r, \) with \( \epsilon > 0 \) and \( \delta > 0. \) For such a system the following results hold.

**Proposition 2.** [1] Consider systems (3) and (5). Suppose \( \xi(0) \) belongs to the solution manifold \( \mathcal{M}. \) Then any solution of (3) is a solution of (5) and vice versa.

**Theorem 1.** [1] Consider system (5). Assume there exists a positive definite radially unbounded function \( W(\mu) \) such that

\[ \frac{\partial W(\mu)}{\partial \mu} q_0(\mu) \leq \gamma W(\mu) + \gamma_0, \tag{7} \]

for some \( \gamma \in \mathbb{R} \) and \( \gamma_0 \in \mathbb{R}, \) and

\[ \sup_{\nu \in \mathbb{R}^{n-r}} \frac{||\frac{\partial W(\mu)}{\partial \mu}||^2}{W(\mu)} \leq \frac{1}{\bar{W}} < +\infty, \tag{8} \]

for some \( \bar{W} > 0. \) Then there exists \( \bar{\delta} > 0 \) such that for all \( \delta \geq \bar{\delta} \) in (6) the following statements hold.

1. \( \mu(t) \) and \( \xi(t) \) exist for all \( \mu(0) \in \mathbb{R}^{n-r} \) and \( \xi(0) \in \mathbb{R}^r \) and \( t \geq 0; \)
2. \( \lim_{t \to \infty} \xi_i(t) = 0, \) for \( i = 1, \ldots, r. \)

Theorem 1 ensures that the manifold \( \mathcal{M} \) is globally attractive and that the solutions of the system exist for all \( t \geq 0. \) However, it does not give any information on the stability properties of the constrained system. The first objective of this paper is to study the behaviour of the solutions in the manifold \( \mathcal{M}. \) To this end we first recall a preliminary result. Let

\[ \mu = q_0(\mu), \tag{9} \]

be the constrained dynamics, also called zero dynamics, of system (3) (equivalently of system (5)) in the manifold \( \mathcal{M}, \) i.e. for \( \xi(t) \equiv 0. \) Then the following result holds.

**Proposition 3.** [9] Suppose that the differential-algebraic
system (1) has differentiation index $\nu \leq n$. Stability properties of the zero equilibrium of system (1) are equivalent to stability properties of the zero equilibrium of the zero dynamics of system (3).

The problem we want to address in the remainder of the section is to find out under which conditions the stability properties of the zero dynamics of system (3), and according to Proposition 3 of the DA system (1), are preserved by the stabilization method in equation (5). In particular, the following result holds.

**Theorem 2.** Consider system (5). Assume (7) and (8) holds with $\gamma < 0$ and $\gamma_0 = 0$. Then there exists $\delta > 0$ such that for all $\delta \geq \delta$ in (6) the following statements hold.

1. $\mu(t)$ and $\xi(t)$ exist for all $\mu(0) \in \mathbb{R}^{n-r}$ and $\xi(0) \in \mathbb{R}^r$ and $t \geq 0$.
2. $\lim_{t \to \infty} \xi_i(t) = 0$, for all $\xi_i(0) \in \mathbb{R}$, $i = 1, \ldots, r$;
3. $\lim_{t \to \infty} \mu(t) = 0$, for all $\mu(0) \in \mathbb{R}^{n-r}$.

**Remark 1.** Observe that assumption (7) with $\gamma < 0$ and $\gamma_0 = 0$ in Theorem 3 can be rewritten as

$$W \leq \gamma W,$$

where the function $W(\mu)$ evolves along the trajectory of system (9), i.e. the zero dynamics. According to Proposition 3 and Theorem 2, if the equilibrium point of system (1) is globally asymptotically stable then also the equilibrium point of the “stabilized” system (5) is globally asymptotically stable.

### III. MAIN RESULT

In this section the problem of the discretization of system (5) is addressed. Since for a numerical implementation of any integration method a discretization scheme is required, we now study under which conditions the results obtained in [1] and in Section II of this paper for the continuous-time case are maintained when a discretization is performed. In the remainder of the paper we use the notation $x := x(t)$ and $x^+ := x(t+T)$, $\forall t \in \mathbb{N}$, where $T \in \mathbb{R}_{>0}$ is the sample period.

Consider the discrete-time system

$$\frac{\xi^+-\xi}{T} = K(\xi, \mu)\xi,$$

$$\frac{\mu^+-\mu}{T} = \varphi_T(q_0) + Q(\xi, \mu),$$

where $\mu(t) \in \mathbb{R}^{n-r}$, $\xi(t) \in \mathbb{R}^r$, $Q : \mathbb{R}^r \times \mathbb{R}^{n-r} \to \mathbb{R}^{n-r \times r}$, $Q = [q_1, \ldots, q_r]$, $\varphi_T : \mathbb{R}^{n-r} \to \mathbb{R}^{n-r}$, $\varphi_T$, $q_0$ and $q_i$, for $i = 1, \ldots, r$ are smooth mappings, and $K : \mathbb{R}^r \times \mathbb{R}^{n-r} \to \mathbb{R}^{r \times r}$. Recalling the definition of $k_i$ given in (6), let $K(\xi, \mu)$ be such that

$$K(\xi, \mu) = -\frac{\gamma^2}{2} \text{diag}(q_1^2(\xi, \mu), \ldots, q_r^2(\xi, \mu)) - \epsilon I = \text{diag}(k_1, \ldots, k_r).$$

Observe that system (11) is a discretization scheme for system (5), provided the map $\varphi_T(q_0)$ is such that

$$\lim_{T \to 0} \varphi_T(q_0) = q_0,$$

see for example [17] for conserving time-integration schemes for mechanical systems. Similarly to Theorem 1 for the continuous-time system (5), for the discrete-time system (11) the following result holds.

**Theorem 3.** Consider system (11) and assume $(\xi(0), \mu(0)) \in B$, where $B \subset \mathbb{R}^n$ is a compact set. For all $T$ such that

$$T < \frac{2}{\max_{i=1,\ldots,r} \left( \sup_{(\xi, \mu) \in B} (-k_i(\xi, \mu)) \right)}$$

the following claims hold.

1. $\lim_{t \to +\infty} \xi(t) = 0$ for all $(\xi(0), \mu(0)) \in B$ and $t \in \mathbb{N}_+$;
2. $\xi(t)$ and $\mu(t)$ exist for all $(\xi(0), \mu(0)) \in B$, $\mu(0) \in \mathbb{R}^{n-r}$ and $t \in \mathbb{N}_+$.

Note that different discretization schemes provide different properties. For instance, consider the modified discrete-time system

$$\frac{\xi^+-\xi}{T} = K(\xi, \mu)\xi,$$

$$\frac{\mu^+-\mu}{T} = \varphi_T(q_0) + Q(\xi, \mu).$$

Observe that, since

$$\lim_{T \to 0} \xi^+ = \lim_{T \to 0} \xi(t+T) = \xi(t),$$

and since (13) holds, system (15) is a discretization of system (5). The following result holds.

**Theorem 4.** Consider system (15). Then the following claims hold for any $T > 0$.

1. $\lim_{t \to +\infty} \xi(t) = 0$ for all $(\xi(0), \mu(0)) \in \mathbb{R}^{n-r}$ and $t \in \mathbb{N}_+$;
2. $\xi(t)$ and $\mu(t)$ exist for all $(\xi(0), \mu(0)) \in \mathbb{R}^{n-r}$, $\mu(0) \in \mathbb{R}^{n-r}$ and $t \in \mathbb{N}_+$.

Theorem 4 ensures that the manifold

$$\mathcal{M} = \{ (\xi^T, \mu^T)^T : \xi = 0 \}.$$

is globally attractive and that the solutions of the system exist for all $t \in \mathbb{N}_{\geq 0}$ and for any $T > 0$. We consider now the stability properties of the zero equilibrium of systems (11) and (15). The following result provides a discrete-time equivalent of Theorem 2.

**Theorem 5.** Consider system (11) and assume $(\xi(0), \mu(0)) \in B$, where $B \subset \mathbb{R}^n$ is a compact set. Suppose the assumptions of Theorem 2 hold. Let $\overline{T}$ be such that

$$\frac{1}{n-r} \sum_{i=1}^{n-r} \frac{W(\mu(t+\overline{T})) - W(\mu(t))}{\mu_i(t+\overline{T}) - \mu_i(t)} \varphi_T(\mu) \leq \gamma W(\mu),$$

(16)
and
\[
\left( \frac{1}{n-r} \sum_{i=1}^{n-r} W(\mu(t + T)) - W(\mu(t)) \right)^2 \leq W(\mu)
\]
Let \( T \leq \min(\frac{\sqrt{n-r}}{\mu_{n-r}}, \frac{\sqrt{n-r}}{\mu_{n-r}}) \) if the discretization (15) is used and \( T \leq \min(\frac{\sqrt{n-r}}{\mu_{n-r}}, \frac{\sqrt{n-r}}{\mu_{n-r}}) \) if the discretization (11) is used, where \( T \) is such that Theorem 3 holds. Then for this selection of \( T \) and for \( \delta \) and \( \epsilon \) selected as in Theorem 2, the following claims hold.

1) \( \xi(t) \) and \( \mu(t) \) exist, for all \( (\xi(0), \mu(0)) \in B \) and \( t \in \mathbb{N}_+ \);
2) \( \lim_{t \to \infty} \xi(t) = 0 \), for all \( (\xi(0), \mu(0)) \in B \) and \( t \in \mathbb{N}_+ \);
3) \( \lim_{t \to \infty} \mu(t) = 0 \), for all \( (\xi(0), \mu(0)) \in B \) and \( t \in \mathbb{N}_+ \).

Remark 2. Observe that assumption (16) can be rewritten as
\[
\frac{\Delta W}{T} = \frac{W(\mu^+) - W(\mu)}{T} \leq \gamma W(\mu),
\]
where \( W(\mu) \) is a positive definite function which evolves along the trajectories of the system
\[
\frac{\mu^+ - \mu}{T} = \varphi_T(q_0),
\]
i.e. the constrained system. Moreover, note that for \( T \to 0 \) equation (16) becomes
\[
\lim_{T \to 0} \frac{1}{n-r} \sum_{i=1}^{n-r} \frac{W(\mu(t + T)) - W(\mu(t))}{\mu(t + T) - \mu(t)} \varphi_T(\mu) = \lim_{n \to 0} \frac{\Delta W}{T} = \bar{W} \leq \gamma W.
\]
Hence, according to Remark 1, condition (16) is the discrete-time counterpart of (7). Similarly, condition (17) is the discrete-time counterpart of (8). This observation implies that for \( T \) small enough, (7) and (8) imply (16) and (17).

Observe that while the discretization scheme in (15) ensures global attractivity of the constraint manifold, for the discretization scheme (11) attraction holds only locally, see Theorems 3 and 4. However, this comes at the cost of a smaller sample time required to preserve the stability properties of the system, see the assumptions of Theorem 5. The results so far presented, together with the algorithm presented in [1], outline a procedure to simulate DA systems, as shown in the following example.

IV. EXAMPLE

Consider the nonlinear oscillator depicted in Fig. 1 with mass \( m \) and spring constant \( k \). Let \((X, Y)\) and \((\dot{X}, \dot{Y})\) be the position and the velocity coordinates of the mass \( m \), respectively. Assume that the only external forces acting on the system are the gravitational force (with \( g_a \) the gravitational acceleration), the constraint force \( \lambda \) of the floor, the friction force \( F_d = -d \dot{X} \) and the nonlinear spring force \( F_k = -kX^3 \). Such a system can be described by a set of differential-algebraic equations in the state space form (1).
has been selected according to [17] to preserve the total mechanical energy.

The data used for the simulation are \( m = 1, k = 2, d = 0.03, \epsilon = 1, \delta = 2 \) and \((x_{01}, x_{02}, x_{03}, x_{04})^T = (0.1, 0.6, 0, 0)^T\). Let

\[
\mathcal{B} = \left\{ (\xi, \mu) : -0.7 \leq \xi_1 \leq 0.7, \xi_2 \in \mathbb{R}, -0.9 \leq \mu_1 \leq 0.9, -0.9 \leq \mu_2 \leq 0.9 \right\}.
\]

For this set of parameters we have that

\[
\max_{i=1, \ldots, r} \left( \sup_{(\xi, \mu) \in \mathcal{B}} (-k_i(\xi, \mu)) \right) = 188.9. \tag{24}
\]

Note that the assumptions of Theorem 5 are satisfied for \( T = 0.001 \), with

\[
W(\mu) = \frac{1}{2} m \mu^2 + \frac{1}{4} k \mu^4,
\]

and \( \gamma = 0.00001 \).

Simulation results are shown in Figs. 2, 3, 4 and 5. As expected from Theorems 4 and 5, we observe in Fig. 2 that the constraint violation converges to zero, while from Fig. 3 we observe that the origin is an asymptotically stable equilibrium for the constrained system. Fig. 4 displays the trajectory of the system in the X-Y plane, while Fig. 5 shows the total mechanical energy which is dissipated because of the friction force acting on the system.

Consider now the case where \( d = 0 \). The results of the simulation for the same set of parameters are shown in Figs. 6, 7, 8 and 9. As expected from Theorem 4, we observe from Fig. 7 that the constraint violation converges to zero. In addition, observe from Fig. 9 that the mechanical energy is preserved at steady state, i.e. once solutions reach the manifold \( \mathcal{M} \). This observation suggests that the method proposed can be used also in the case in which the equilibrium is marginal stable, even though this has not been proved.

![Fig. 3: Time history of the coordinate X for \( d = 0.03 \).](image)

![Fig. 4: Path of the mass \( m \) on the X-Y plane for \( d = 0.03 \).](image)

![Fig. 5: Time history of the total mechanical energy for \( d = 0.03 \).](image)

V. CONCLUSIONS

We have studied the problem of integration of DA systems. In Section II we have extended the theory developed in [1] by studying the stability properties of the “stabilized” system in continuous-time. We have concluded that if the equilibrium of the constrained system is globally asymptotically stable then the stabilization method preserves this property. In Section III we have extended the results to discrete-time. We have concluded that the solution manifold remains globally attractive for one of the discretization scheme proposed for any sampling time. Moreover asymptotic stability of the equilibrium is preserved (locally). Finally, in Section IV we have demonstrated the technique with an example and discussed the results.
Fig. 6: Time history of the coordinate X for $d = 0$.

Fig. 7: Time history of the coordinate Y for $d = 0$.

Fig. 8: Path of the mass $m$ on the X-Y plane for $d = 0$.

Fig. 9: Time history of the total mechanical energy for $d = 0$.

REFERENCES


