Research Article

On a Newton-Type Method for Differential-Algebraic Equations

S. Amat,[1] M. J. Légaz,[1] and P. Pedregal[2]

[1] *Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Paseo de Alfonso XIII 52, Murcia, 30203 Cartagena, Spain*
[2] *ETSI Industriales, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain*

Correspondence should be addressed to S. Amat, sergio.amat@upct.es

This paper deals with the approximation of systems of differential-algebraic equations based on a certain error functional naturally associated with the system. In seeking to minimize the error, by using standard descent schemes, the procedure can never get stuck in local minima but will always and steadily decrease the error until getting to the solution sought. Starting with an initial approximation to the solution, we improve it by adding the solution of some associated linear problems, in such a way that the error is significantly decreased. Some numerical examples are presented to illustrate the main theoretical conclusions. We should mention that we have already explored, in some previous papers (Amat et al., in press, Amat and Pedregal, 2009, and Pedregal, 2010), this point of view for regular problems. However, the main hypotheses in these papers ask for some requirements that essentially rule out the application to singular problems. We are also preparing a much more ambitious perspective for the theoretical analysis of nonlinear DAEs based on this same approach.

1. Introduction

Differential-algebraic equations are becoming increasingly important in a lot of technical areas. They are currently the standard modeling concept in many applications such as circuit simulation, multibody dynamics, and chemical process engineering; see for instance [1–5] with no attempt to be exhaustive.

A basic concept in the analysis of differential-algebraic equations is the index. The notion of index is used in the theory of DAEs for measuring the distance from a DAE to its related ODE. The higher the index of a DAE is, the more difficulties are for its numerical solution. There are different index definitions, but for simple problems they are identical. On more complicated nonlinear and fully implicit systems they can be different (see [5] and the references therein.)

For simplicity, we focus our attention on problems of the form

$$Mx'(t) = f(x(t)) \quad \text{in } (0, T), \ x(0) = x_0, \tag{1.1}$$

where M is a given, eventually singular, matrix depending on t. More general situations can be allowed. This type of equations arises, for instance, in the functional analytic formulation of the initial value problem for the Stokes as well as for the linearized Navier-Stokes or Oseen equations [6].

For the approximation of these equations collocation-type methods are usually used. These methods are implicit, and we need to solve a nonlinear system of equations in each iteration using a Newton's type method. Given different coefficients c_i, $1 \le i \le s$, there is a (unique for h sufficiently small) polynomial of collocation $q(t)$ of degree less than or equal to s such that

$$q(t_0) = y_0, \qquad q'(t_0 + c_i h) = f(t_0 + c_i h, q(t_0 + c_i h)) \quad \text{if } 1 \le i \le s. \tag{1.2}$$

The collocation methods are defined by an approximation $y(t) \simeq q(t)$ and are equivalent to implicit RK methods of s stages

$$k_i = f\left(t_0 + c_i h, y_0 + h \sum_{j=1}^{s} a_{i,j} k_j\right),$$

$$y_1 = y_0 + h \sum_{i=1}^{s} b_i k_i, \tag{1.3}$$

for the coefficients

$$a_{i,j} = \int_0^{c_i} \prod_{l \ne j} \frac{u - c_l}{c_j - c_l} \, du,$$

$$b_i = \int_0^1 \prod_{l \ne i} \frac{u - c_l}{c_i - c_l} \, du. \tag{1.4}$$

The coefficients c_i play the role of the nodes of the quadrature formula, and the associated coefficients b_i are analogous to the weights. From (1.4), we can find implicit RK methods called Gauss of order $2s$, Radau IA and Radau IIA of order $2s - 1$, and Lobatto IIIA of order $2s - 2$. Also we can consider perturbed collocation methods like Lobato IIIC. (See [4] for more details).

A number of convergence results have been derived for these methods introducing the so-called B-convergence theory. In [7, 8] the authors extend the B-convergence theory to be valid for a class of nonautonomous weakly nonlinear stiff systems, in particular, including *the linear case*. As pointed out by the same authors, it is not clear if it is possible to cover, in a satisfactory way, highly nonlinear stiff problems, that is, problems where also the nonlinear terms are affected by large parameters. Moreover, any result should assume that, in each step, the associated nonlinear system is well approximated [5]. In particular, we should be able to

start with a good initial guess for the iterative scheme. This might be very restrictive for many stiff problems.

On the other hand, iterative methods are the typical tool to solve nonlinear systems of equations. In these schemes we compute a sequence of approximations solving associated linear problems. In this paper, we would like to introduce a new variational approach for the treatment of DAEs where we *linearize* the original equations obtaining an iterative scheme. Our ideas are based on the analysis of a certain error functional of the form

$$E(x) = \frac{1}{2} \int_0^T \left| Mx'(t) - f(x(t)) \right|^2 dx \tag{1.5}$$

to be minimized among the absolutely continuous paths $x : (0, T) \rightarrow \mathbf{R}^N$ with $x(0) = x_0$. Note that if $E(x)$ is finite for one such path x, then automatically Mx' is square integrable. This error functional is associated, in a natural way, with the Cauchy problem (1.1). Indeed, the existence of solutions for (1.1) is equivalent to the existence of minimizers for E with vanishing minimum value. This is elementary.

In this initial contribution, we want to concentrate on the approximation issue through this perspective. We will place ourselves under the appropriate hypotheses so that there are indeed solutions for (1.1), that is, there are minimizers for the error with vanishing minimum value. In addition, we would like to guarantee that the main ingredients for the iterative approximating scheme to work are valid. More explicitly, our approach for the numerical approximation of such problems relies on three main analytical hypotheses that we take for granted here.

(1) The Cauchy problem (1.1) admits a unique solution for every feasible initial condition x_0 (the definition of feasible path should depend on the index of the equation).

(2) The linearization around any feasible, absolutely continuous, path $x(t)$ with $x(0) = x_0$,

$$My'(t) - \nabla f(x(t))y(t) = f(x(t)) - Mx'(t) \quad \text{in } (0, T), \ y(0) = 0, \tag{1.6}$$

always has a unique solution, and moreover, for some constant $L > 0$ depending on M, f, x and its derivatives,

$$\|y\|_{L^\infty(0,T)}^2 \le TL \|f(x(t)) - Mx'(t)\|_{L^2(0,T)}. \tag{1.7}$$

(3) The only solution of the problem

$$\frac{d}{dt}\left(M^\top z(t)\right) + \nabla f(x(t))^\top z(t) = 0 \quad \text{in } (0, T), \ M^\top z(T) = 0, \tag{1.8}$$

is $z \equiv 0$, for every feasible, absolutely continuous, path $x(t)$ with $x(0) = x_0$.

Here the superscript $\top$ indicates transpose.

These requirements depend on the index of the equation and on some regularity on the pair $(M, \nabla f(x(t)))$. They should be more restrictive for equations with high index. More details can be found, for example, in [9, Theorem 3.9], where the authors consider DAEs transferable into standard canonical form. More precise information is outside of the scope of this paper. In any case, the equations verifying our hypotheses are, in general, a subclass of all analytically solvable systems.

In addition to the basic facts just stated on existence and uniqueness of solutions for our problems, the analysis of the approximation scheme, based on a minimization of the error functional E, requires one main basic assumption on the nonlinearity $f : \mathbf{R}^N \to \mathbf{R}^N$. It must be smooth, so that $\nabla f : \mathbf{R}^N \to \mathbf{R}^{N \times N}$ is continuous and globally Lipschitz with constant $K > 0$ ($|\nabla f| \le K$). Moreover, the main result of this paper demands a further special property on the map f: for every positive $C > 0$ and small $\epsilon > 0$, there is $D_{C,\epsilon} > 0$ so that

$$
\left| f(x + y) - f(x) - \nabla f(x) y \right| \le D_{C,\epsilon} |y|^2, \quad |x| \le C, \ |y| \le \epsilon. \tag{1.9}
$$

This regularity is somehow not surprising as our approach here is based on regularity and optimality. On the other hand, that regularity holds for most of the important problems in applications. It certainly does in all numerical tests performed in this work. Our goal here is placed on the fact that this optimization strategy may be utilized to set up approximation schemes based on the minimization of the error functional. Indeed, we provide a solid basis for this approximation procedure. One very important and appealing property of our approach states that typical minimization schemes like (steepest) descent methods will work fine as they can never get stuck in local minima and converge steadily to the solution of the problem, no matter what the initialization is.

We should mention that we have already explored, in some previous papers, this point of view. Since the initial contribution [10], we have also treated the reverse mechanism of using first discretization and then optimality [11]. The perspective of going through optimality and then discretization has already been indicated and studied in [12], though only for the steepest descent method, and without going through any further analytical foundation for the numerical procedure. However, the main hypotheses in these papers ask for some requirements that essentially rule out the application to singular problems. We will however address shortly [13] a complete treatment of DAEs with no a priori assumptions on existence and uniqueness. Rather, we will be interested in showing existence and uniqueness from scratch by examining the fundamental properties of the error functional E.

On the other hand, variational methods have been used also before in the context of ODEs. See [14, 15], where numerical integration algorithms for finite-dimensional mechanical systems that are based on discrete variational principles are proposed and studied. This is one approach to deriving and studying symplectic integrators. The starting point is Hamilton's principle and its direct discretization. In those references, some fundamental numerical methods are presented from that variational viewpoint where the model plays a prominent role.

The rest of the paper is divided in three sections. In Section 2 we introduce our variational approach and develop a convergence analysis. Section 3 introduces the numerical procedure. We analyze some numerical results in Section 4. Finally, we present the main conclusions in Section 5.

2. A Main Descent Procedure

We start with a fundamental fact for our approach.

Proposition 2.1. *Let $\overline{x}$ be a critical point for the error E. Then $\overline{x}$ is the solution of the Cauchy problem* (1.1).

Proof. The proof is elementary. Based on the smoothness and bounds assumed on the mapping f, we conclude that if $x \equiv \overline{x}$ is a critical point for the error E, then x ought to be a solution of the problem

$$-\frac{d}{dt}\left(M^{\top}(Mx'(t) - f(x(t)))\right) - \nabla f(x(t))^{\top}((Mx'(t) - f(x(t)))) = 0 \quad \text{in } (0, T),$$

$$x(0) = x_0, \quad M^{\top}(Mx'(T) - f(x(T))) = 0. \tag{2.1}$$

The vector-valued map $y(t) = Mx'(t) - f(x(t))$ is then a solution of the linear, nondegenerate problem

$$M^{\top}y'(t) + \nabla f(x(t))^{\top}y(t) = 0 \quad \text{in } (0, T), \qquad M^{\top}y(T) = 0. \tag{2.2}$$

The only solution of this problem, by our initial conditions on uniqueness of linearizations, is $y \equiv 0$, and so x is the solution of our Cauchy problem. □

On the other hand, suppose we start with an initial crude approximation $x^{(0)}$ to the solution of our basic problem (1.1). We could take $x^{(0)} = x_0$ for all t or $x^{(0)}(t) = x_0 + tf(x_0)$. We would like to improve this approximation in such a way that the error is significantly decreased. We have already pointed out that descent methods can never get stuck on anything but the solution of the problem, under global lipschitzianity hypotheses.

It is straightforward to find the Gâteaux derivative of E at a given feasible x in the direction y with $y(0) = 0$. Namely

$$E'(x)y = \int_0^T ((Mx'(t) - f(x(t))) \cdot (My'(t) - \nabla f(x(t))y(t)))dt. \tag{2.3}$$

This expression suggests a main possibility to select y from. Choose y such that

$$My'(t) - \nabla f(x(t))y(t) = f(x(t)) - Mx'(t) \quad \text{in } (0, T), \ y(0) = 0. \tag{2.4}$$

In this way, it is clear that $E'(x)y = -2E(x)$, and so the (local) decrease of the error is of the size $E(x)$. Finding y requires solving the above linear problem which is assumed to have a unique solution by our main hypotheses in the introduction. In some sense, this is like a Newton method with global convergence.

Suppose $x^{(0)}$ is a feasible path in the interval $(0, T)$ so that $x^{(0)}(0) = x_0$, $M(x^{(0)})'$ is square integrable, $|x^{(0)}(t)| \leq C$ for a fixed constant C, all $t \in (0, T)$, and the quantity

$$E\left(x^{(0)}\right) = \frac{1}{2} \int_0^T \left| M\left(x^{(0)}\right)'(t) - f\left(x^{(0)}(t)\right) \right|^2 dt \tag{2.5}$$

measures how far such $x^{(0)}$ is from being a solution of our problem.

Theorem 2.2. *For T sufficiently small, the iterative procedure $x^{(j)} = x^{(j-1)} + y^{(j)}$, starting from the above feasible $x^{(0)}$ and defining $y^{(j)}$ as the solution of the linear problem*

$$M\left(y^{(j)}\right)'(t) - \nabla f\left(x^{(j-1)}(t)\right) y^{(j)}(t) = f\left(x^{(j-1)}(t)\right) - M\left(x^{(j-1)}\right)'(t) \quad in\ (0, T),\ y^{(j)}(0) = 0, \tag{2.6}$$

converges strongly in $L^\infty(0, T)$ to the solution of (1.1).

Proof. Choose $\epsilon > 0$ and $0 < \alpha < 1$ so that

$$\frac{\epsilon}{1 - \sqrt{\alpha}} \leq C, \tag{2.7}$$

$$\left| f(z + y) - f(z) - \nabla f(z) y \right| \leq D |y|^2, \quad |y| \leq \epsilon,\ |z| \leq 2C,$$

for some constant $D > 0$ (see the main hypotheses in Section 1). We then solve for $y^{(0)}$ as the solution of the nonautonomous linear problem

$$My'(t) - \nabla f(x(t)) y(t) = f(x(t)) - Mx'(t) \quad in\ (0, T),\ y(0) = 0, \tag{2.8}$$

and pretend to update $x^{(0)}$ to $x^{(0)} + y^{(0)}$ in such a way that the error for $x^{(0)} + y^{(0)}$ be less than the error for the current iteration $x^{(0)}$. Note that

$$E\left(x^{(0)} + y^{(0)}\right) = \frac{1}{2} \int_0^T \left| f\left(x^{(0)}(t) + y^{(0)}(t)\right) - f\left(x^{(0)}(t)\right) - \nabla f\left(x^{(0)}(t)\right) y^{(0)}(t) \right|^2 dt, \tag{2.9}$$

where we have used the differential equation satisfied by $y^{(0)}$ and the definition of $E(x)$. By our assumption on f above,

$$\left| f\left(x^{(0)}(t) + y^{(0)}(t)\right) - f\left(x^{(0)}(t)\right) - \nabla f\left(x^{(0)}(t)\right) y^{(0)}(t) \right| \leq D \left| y^{(0)}(t) \right|^2, \quad t \in (0, T), \tag{2.10}$$

provided that $|y^{(0)}(t)| \leq \epsilon$. Since we know that $y^{(0)}$ is the solution of a certain linear problem, by the upper bound assumed in Section 1 on the size of these solutions,

$$\left| y^{(0)}(t) \right|^2 \leq TLE\left(x^{(0)}\right) \quad \forall t \in [0, T],\ L \in \mathbb{R}^+. \tag{2.11}$$

Assume that we select $T > 0$ so small that

$$E_{0,T}\left(x^{(0)}\right) \equiv E\left(x^{(0)}\right) \leq \frac{\epsilon^2}{TL}, \tag{2.12}$$

and then $|y^{(0)}(t)| \leq \epsilon$ for all $t \in [0, T]$. By (2.9), (2.10), and (2.11), we can write

$$E\left(x^{(0)} + y^{(0)}\right) \leq \frac{D^2}{2} \int_0^T \left|y^{(0)}(t)\right|^4 dt \leq \frac{D^2}{2} L^2 T^3 E\left(x^{(0)}\right)^2. \tag{2.13}$$

If, in addition, we demand, by making T smaller if necessary,

$$E\left(x^{(0)}\right) \leq \frac{2\alpha}{D^2 T^3 L^2}, \tag{2.14}$$

then $E(x^{(0)} + y^{(0)}) \leq \alpha E(x^{(0)})$. Moreover, for all $t \in (0, T)$,

$$\left|x^{(0)}(t) + y^{(0)}(t)\right| \leq C + \epsilon \leq 2C. \tag{2.15}$$

All these calculations form the basis of a typical induction argument, verifying

$$\left|\sum_{i=0}^{j-1} x^{(i)}(t)\right| \leq C + \epsilon \left(\sum_{i=0}^{j-2} \sqrt{\alpha^i}\right) (\leq 2C),$$

$$\left|x^{(j-1)}(t)\right| \leq \epsilon \sqrt{\alpha^{j-2}} \quad \forall t \in [0, T], \tag{2.16}$$

$$E\left(\sum_{i=0}^{j-1} x^{(i)}\right) \leq \alpha^{j-1} E\left(x^{(0)}\right) \left(\leq E\left(x^{(0)}\right)\right).$$

It is therefore clear that the sum

$$\sum_{i=0}^{\infty} x^{(i)}(t) \tag{2.17}$$

converges strongly in $L^\infty(0, T)$ to the solution of our initial Cauchy problem in a small interval $(0, T)$. □

Since the various ingredients of the problem do not depend on T, we can proceed to have a global approximation in a big interval by successively performing this analysis in intervals of appropriate small size. For instance, we can always divide a global interval $(0, T)$ into a certain number n of subintervals of small length h $(T = nh)$ with

$$\frac{E_{0,T}(x^{(0)})D^2 L^2}{2\alpha} \leq \frac{1}{h^3}, \tag{2.18}$$
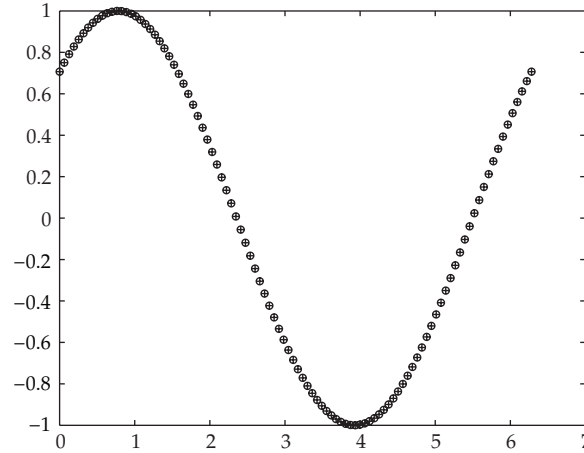
according to (2.14).

Figure 1: Index 1, $T = 2\pi$, the y-coordinate, "o"-original, "+"-approximation.
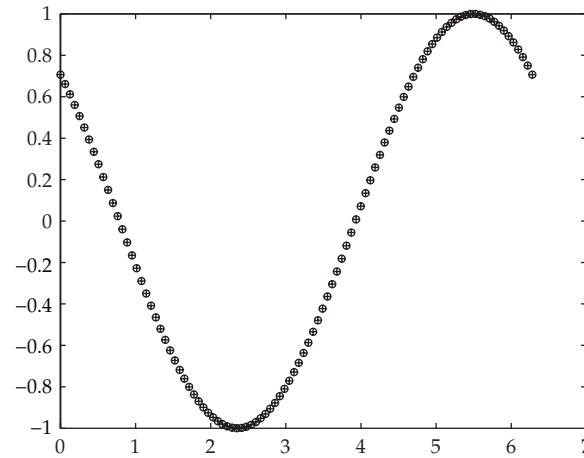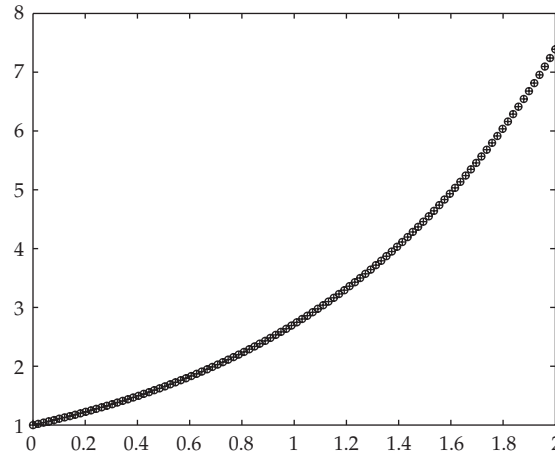

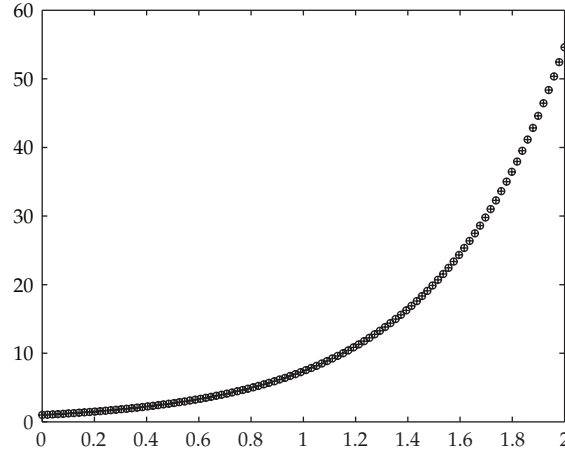**Figure 2:** Index 1, $T = 2\pi$, the z-coordinate, "o"-original, "+"-approximation.

3. Numerical Procedure

Since our optimization approach is really constructive, iterative numerical procedures are easily implementable.

(1) Start with an initial approximation $x^{(0)}(t)$ compatible with the initial conditions (e.g. $x^{(0)}(t) = x_0 + tf(x_0)$).

(2) Assume we know the approximation $x^{(j)}(t)$ in $[0, T]$.

(3) Compute its derivative $M(x^{(j)})'(t)$.

(4) Compute the auxiliar function $y^{(j)}(t)$ as the numerical solution of the problem

$$My'(t) - \nabla f\left(x^{(j)}(t)\right)y(t) = f\left(x^{(j)}(t)\right) - M\left(x^{(j)}\right)'(t) \quad \text{in } (0, T), \ y(0) = 0, \qquad (3.1)$$

Figure 3: Index 2, $T = 2$, the y_1-coordinate, "o"-original, "+"-approximation.
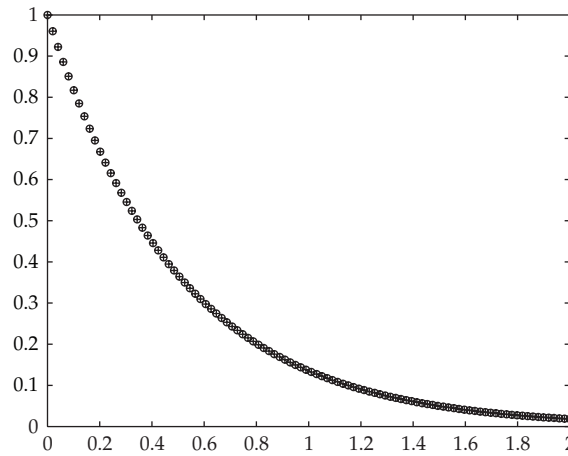
Figure 4: Index 2, $T = 2$, the y_2-coordinate, "o"-original, "+"-approximation.

by making use of a numerical scheme for DAEs with dense output (like collocation methods).

(5) Change $x^{(j)}$ to $x^{(j+1)}$ by using the update formula

$$x^{(j+1)}(t) = x^{(j)}(t) + y^{(j)}(t). \tag{3.2}$$

(6) Iterate (3), (4), and (5), until numerical convergence.

In practice, we use the stopping criterium

$$\max\left\{ \left\| y^{(j)} \right\|_\infty, \sqrt{2E(x^{(j)})} \right\} \leq TOL. \tag{3.3}$$

Figure 5: Index 2, $T = 2$, the z-coordinate, "o"-original, "+"-approximation.
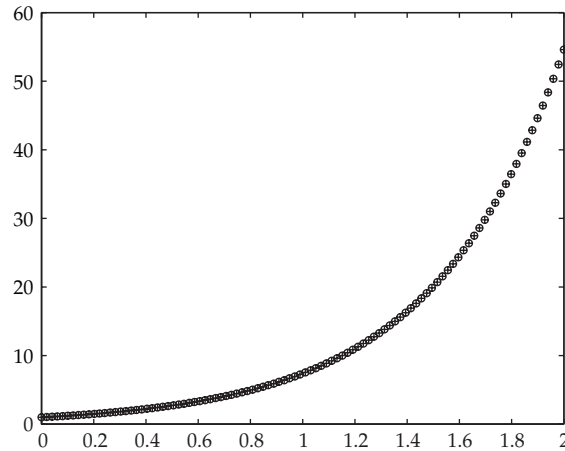
Figure 6: Index 3, $T = 2$, the y_1-coordinate, "o"-original, "+"-approximation.

In particular, one can implement, in a very easy way, this numerical procedure using a problem-solving environment like MATLAB [16].

4. Some Experiments

In this section, we approximate some problems well known in the literature for a different index [5, 17, 18]. High-order accuracy and stability are major areas of interest in this type of simulations. We do not perform an analysis of the convergence conditions imposed in the above section. We are only interested to test numerically our approach.

In our approach we only need to approximate, with at least order one, the associated linear system for $y^{(j)}$, in order to obtain the convergence of our scheme (see Theorem 2.2). The stability can be ensured by the fact that we approximate a *linear* problem using specific implicit methods [4]. This is not the case with a general nonlinear problem [19], where
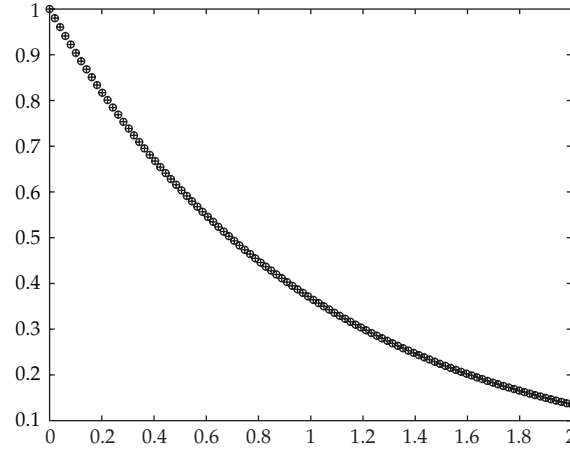
Figure 7: Index 3, $T = 2$, the y_2-coordinate, "o"-original, "+"-approximation.
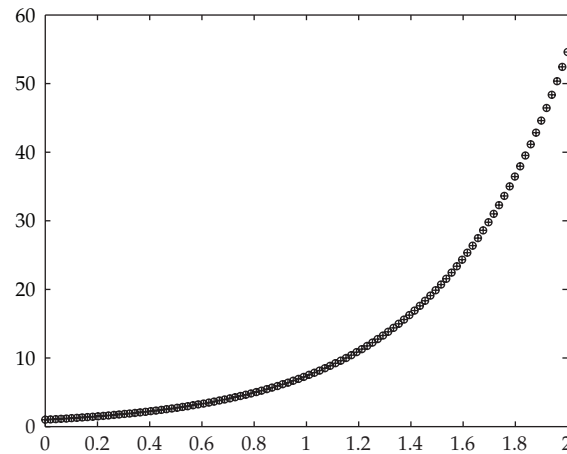

Figure 8: Index 3, $T = 2$, the z_1-coordinate, "o"-original, "+"-approximation.

we need to approximate well (with a Newton-type iterative method) the nonlinear system related to the implicitness of the scheme (see the above section). This approximation should be a difficult task due to the local (nonglobal) convergence of any iterative scheme for nonlinear problems.

In this section, we consider the convergent Lobatto IIIC method [18] valid for indexes 1–3, in order to approximate the associated linear problem for $y^{(j)}$ in each iteration. This method can be considered as a perturbation collocation method. The final error depends only on the stopping criterium. In the following examples, we stop the algorithm when

$$\max\left\{ \left\| y^{(j)} \right\|_\infty, \sqrt{2E(x^{(j)})} \right\} \leq 10^{-6} \tag{4.1}$$

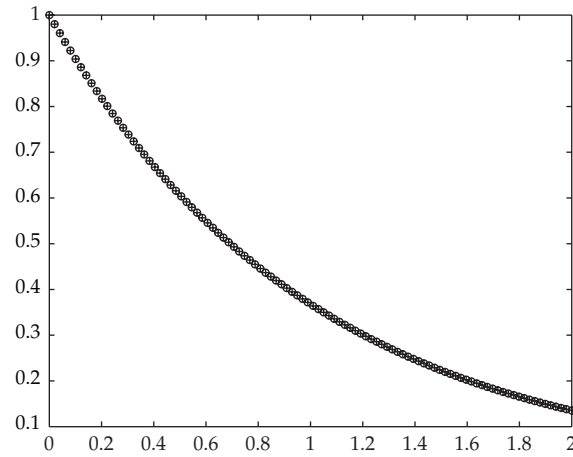and plot the solution and the approximation given by our approach.

Figure 9: Index 3, $T = 2$, the z_2-coordinate, "o"-original, "+"-approximation.
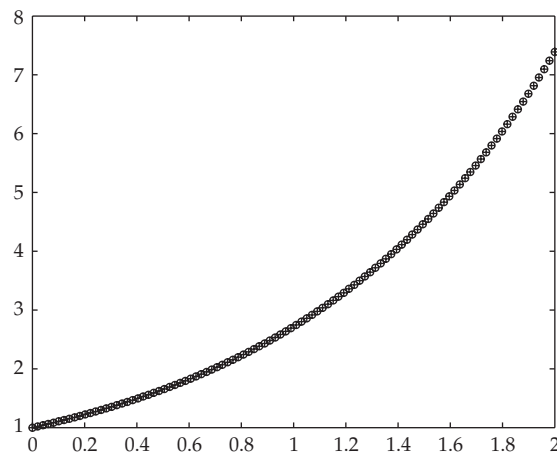

Figure 10: Index 3, $T = 2$, the u-coordinate, "o"-original, "+"-approximation.

(i) Index 1 [5]

$$
\begin{aligned}
y'(t) &= z(t), \\
y(t)^2 + z(t)^2 &= 1, \\
y(0) = z(0) &= \frac{\sqrt{2}}{2}.
\end{aligned}
\tag{4.2}
$$

The solution of this problem is $(\sin(x + \pi/4), \cos(x + \pi/4))$.

(ii) Index 2 [17]

$$y_1'(t) = \sum_{i=1}^{5} f_i(y_1(t), y_2(t), z(t)),$$

$$y_2'(t) = \sum_{i=1}^{5} g_i(y_1(t), y_2(t), z(t)),$$

$$y_1(t)^2 y_2(t) = 1,$$ (4.3)

$$y_1(0) = y_2(0) = 1,$$

$$z(0) = 1,$$

where

$$f_1(y_1(t), y_2(t), z(t))$$

$$= y_2(t) - 2y_1(t)^2 y_2(t) + y_1(t)y_2(t)^2 z(t)^2 + 2y_1(t)y_2(t)^2 - 2e^{-2t}y_1(t)y_2(t),$$

$$f_2(y_2(t), z(t)) = -y_2(t)^2 z(t) + 2y_2(t)^2 z(t)^2,$$ (4.4)

$$g_1(y_1(t), y_2(t)) = -y_1(t)^2 + y_1(t)^2 y_2(t)^2,$$

$$g_2(y_1(t), y_2(t), z(t)) = -y_1(t) + e^{-t}z(t) - 3y_2(t)^2 z(t) + z(t).$$

The solution of this problem is (e^t, e^{-2t}, e^{2t}).

(iii) Index 3 [18]

$$y_1'(t) = 2y_1(t)y_2(t)z_1(t)z_2(t),$$

$$y_2'(t) = -y_1(t)y_2(t)z_2(t)^2,$$

$$z_1'(t) = (y_1(t)y_2(t) + z_1(t)z_2(t))u(t),$$

$$z_2'(t) = -y_1(t)y_2(t)^2 z_2(t)^2 u(t),$$ (4.5)

$$y_1(t)y_2(t)^2 = 1,$$

$$y_1(0) = y_2(0) = 1,$$

$$z_1(0) = z_2(0) = 1,$$

$$u(0) = 1.$$

The solution of this problem is $(e^{2t}, e^{-t}, e^{2t}, e^{-t}, e^t)$.

In Figures 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, we compare the solution of the corresponding three problems with the approximations given by our approach. The results are very satisfactory in all cases, obtaining always the convergence to the true solution. In a first look,

the exact and computed solutions are indistinguishable, since after convergence the error is smaller than the tolerance (= 10^{-6}) used in the stopping criterium. A more systematic and careful analysis of the numerical possibilities of the method will be pursued in the future.

5. Conclusions

A new variational approach to the analysis and numerical implementation of regular ODEs has been recently introduced in [10, 20]. Because of its flexibility and simplicity, it can easily be extended to treat other types of ODEs like differential-algebraic equations (DAEs). This has been precisely the main motivation for this paper: to explore how well those ideas can be adapted to this framework. In particular, extending to this context some of the analytical results and performing various numerical tests that confirm that indeed the variational perspective is worth pursuing. One remarkable feature is that this point of view only requires to count on good numerical schemes for linear problems, and this is the reason why it fits so well in other scenarios. Because of the many good qualities of this viewpoint, it can be considered and implemented in essentially all fields where differential equations are relevant. There is, then, a long way to go.
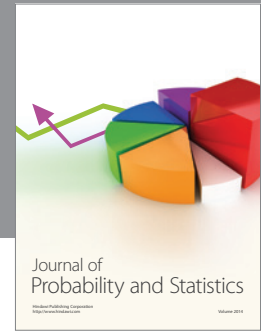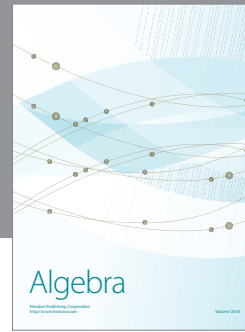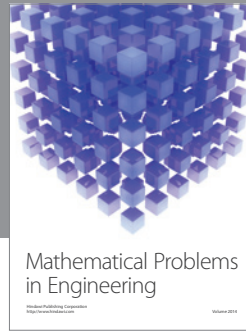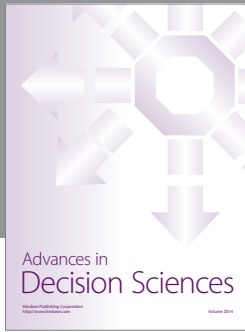
Acknowledgments

References

[1] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, SIAM, Philadelphia, Pa, USA, 1996.

[2] K. E. Brenan and L. R. Petzold, "The numerical solution of higher index differential/algebraic equations by implicit methods," *SIAM Journal on Numerical Analysis*, vol. 26, no. 4, pp. 976–996, 1989.

[3] S. L. Campbell and E. Griepentrog, "Solvability of general differential algebraic equations," *SIAM Journal on Scientific Computing*, vol. 16, no. 2, pp. 257–270, 1995.

[4] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential Algebraic Problems*, vol. 14, Springer, Berlin, Germany, 1991.

[5] R. Riaza, *Differential-Algebraic Systems*, World Scientific, Singapore, 2008.

[6] E. Emmrich and W. Mehrmann, *Analysis of Operator DiFFerential-Algebraic Equations Arising Influid Dynamics. Part I. The Finite Dimensional Case*, 2010.

[7] W. Auzinger, R. Frank, and G. Kirlinger, "An extension of *B*-convergence for Runge-Kutta methods," *Applied Numerical Mathematics*, vol. 9, no. 2, pp. 91–109, 1992.

[8] W. Auzinger, R. Frank, and G. Kirlinger, "Modern convergence theory for stiff initial value problems," *Journal of Computational and Applied Mathematics*, vol. 45, no. 1-2, pp. 5–16, 1993.

[9] T. Berger and A. Ilchmann, "On the standard canonicalform of time-varying linear DAEs," *Quarterly of Applied Mathematics*. In press.

[10] S. Amat and P. Pedregal, "A variational approach to implicit ODEs and differential inclusions," *ESAIM. Control, Optimisation and Calculus of Variations*, vol. 15, no. 1, pp. 139–148, 2009.

[11] S. Amat, D. J. López, and P. Pedregal, "An optimization approach for the numerical approximation ofdifferential equations," *Optimization*. In press.

[12] P. Pedregal, "A variational approach to dynamical systems, and its numerical simulation," *Numerical Functional Analysis and Optimization*, vol. 31, no. 7, pp. 1532–2467, 2010.

[13] S. Amat and P. Pedregal, "A constructive existence theorem for non-linearDAEs through a variational strategy," in preparation.

[14] A. Lew, J. E. Marsden, M. Ortiz, and M. West, "Variational time integrators," *International Journal for Numerical Methods in Engineering*, vol. 60, no. 1, pp. 153–212, 2004.

[15] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[16] The MathWorks, Inc. MATLAB and SIMULINK, Natick, Mass, USA.

[17] L. O. Jay, "Solution of index 2 implicit differential-algebraic equations by Lobatto Runge-Kutta methods," *BIT. Numerical Mathematics*, vol. 43, no. 1, pp. 93–106, 2003.

[18] L. C. Jay, "Convergence of Runge-Kutta methods for differential-algebraic systems of index 3," *Applied Numerical Mathematics. An IMACS Journal*, vol. 17, no. 2, pp. 97–118, 1995.

[19] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, John Wiley & Sons, 1991.

[20] S. Amat and P. Pedregal, "On an alternative approach for the analysis and numerical simulation of sti ODEs," *DCDS A*, vol. 33, no. 4, pp. 1275–1291, 2012.