

On differential singular perturbation problems: a simple variational approach

Sergio Amat¹ · M. José Legaz²

Received: 2 February 2017 / Accepted: 28 June 2017 / Published online: 28 July 2017
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2017

Abstract An alternative approach for the analysis and the numerical approximation of singular perturbation problems, using a variational framework, is presented. It is based on the natural idea of minimizing the residual of the differential equation measured in the L^2 norm. By this strategy, the approximation is based in the solution of a set of linear problems giving the descent step directions of the problem. This is the main advantage of our approach, since we can use stable and convergent methods for linear problems (without assuming the knowledge of good initial guesses used in the approximation of the associated non-linear systems necessary in Newton-type methods). Remember that for this type of problems, we should use implicit methods. We prove that our procedure can never get stuck in local minima, and the error decreases until getting to the original solution independent of the perturbation parameter. Finally, we include some numerical examples.

Keywords Variational methods · Differential singular problems · Global convergence

Mathematics Subject Classification 65L11 · 49M25

Communicated by Cristina Turner.

S. Amat: Research supported in part by Programa de Apoyo a la investigación de la fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia 19374/PI/14 and MTM2015-64382-P (MINECO/FEDER).

M. J. Legaz: Research supported in part by Programa de Apoyo a la investigación de la fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia 19374/PI/14 and MTM2015-64382-P (MINECO/FEDER).

✉ M. José Legaz
mjlegazalmansa@gmail.com

Sergio Amat
sergio.amat@upct.es

¹ Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Cartagena, Spain

² Universidad Cádiz, Cádiz, Spain

1 Introduction

We start with the singular perturbation problem

$$\begin{aligned} y'(t) &= f(y(t), z(t)), \\ \epsilon z'(t) &= g(y(t), z(t)), \quad 0 < \epsilon \ll 1, \\ y(0) &= y_0, z(0) = z_0. \end{aligned}$$

The structure of the solutions of these problems is well understood (Hairer and Wanner 1991). They are a superposition of smooth solution having an ϵ -expansion plus some rapidly decay functions. Of course, we should assume some hypotheses on the function g : $g_z(y, z)$ is invertible and his logarithmic norm is negative in a ϵ -independent neighborhood of the solution.

To approximate these equations, we should use implicit methods. For explicit methods, the discretization parameters depend on the perturbation parameters and they have not practical interest. The error analysis and the existence and uniqueness of the implicit Runge–Kutta equations are also analyzed in Hairer and Wanner (1991). The results are supported in the Newton–Kantorovich theory. In particular, we have the difficult task to find good initial guesses to approximate the nonlinear Runge–Kutta equations by Newton-type methods. In practice, this fact should be a real problem, as we will point out in the following example:

1.1 Linear vs nonlinear singular perturbation problems

The problems with nonlinear equations can appear even for very robust methods using step variable implementations. In this section, we analyze the numerical behavior, for linear and nonlinear equations, of the stiff solver *ode23s*, that we can find in the *odeset* of MATLAB (The MathWorks 2017).

We consider the linear problem (Stoer and Bulirsch 1993)

$$\begin{aligned} y_1'(t) &= \frac{\lambda_1 + \lambda_2}{2} y_1(t) + \frac{\lambda_1 - \lambda_2}{2} y_2(t), \\ \epsilon y_2'(t) &= \frac{\lambda_1 - \lambda_2}{2} y_1(t) + \frac{\lambda_1 + \lambda_2}{2} y_2(t), \end{aligned}$$

with $\lambda_i < 0$, and two (associated) nonlinear problems:

$$\begin{aligned} y_1'(t) &= \frac{\lambda_1 + \lambda_2}{2} y_1(t) + \frac{\lambda_1 - \lambda_2}{2} y_2(t) + y_1(t)y_2(t), \\ \epsilon y_2'(t) &= \frac{\lambda_1 - \lambda_2}{2} y_1(t) + \frac{\lambda_1 + \lambda_2}{2} y_2(t). \end{aligned}$$

and

$$\begin{aligned} y_1'(t) &= \frac{\lambda_1 + \lambda_2}{2} y_1(t) + \frac{\lambda_1 - \lambda_2}{2} y_2(t) + y_1(t)y_2(t), \\ \epsilon y_2'(t) &= \frac{\lambda_1 - \lambda_2}{2} y_1(t) + \frac{\lambda_1 + \lambda_2}{2} y_2(t) + 2y_1(t)y_2(t). \end{aligned}$$

We consider in all the cases $\lambda_1 = -10^{15}$, $\lambda_2 = -10^{-15}$, $\epsilon = 10^{15}$ and $(1, 1)$ as the initial conditions.

As we can see, in Fig. 1, for the linear case and for the first nonlinear perturbation, the method gives good results even for these large parameters. In this case, the nonlinear part appears only in the first equation (without ϵ).

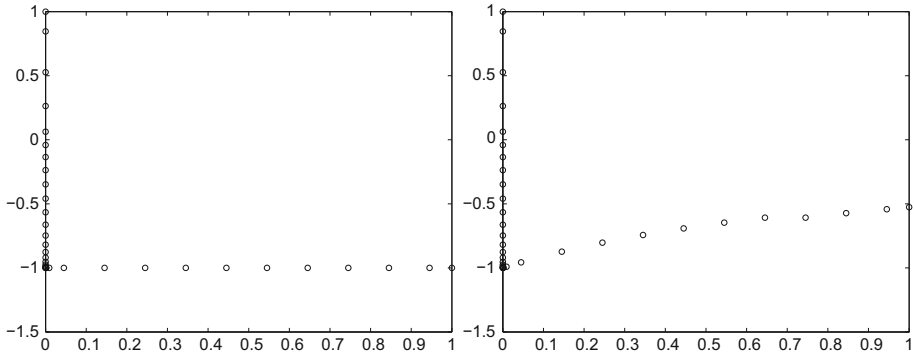


Fig. 1 *Left* second component of the linear problem, *right* second component of the first nonlinear problem

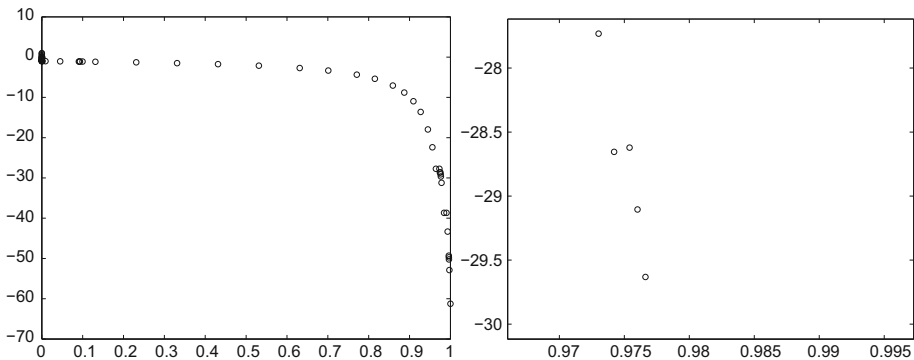


Fig. 2 *Left* second component of the second nonlinear problem, *right* a zoom of the oscillation near the end of the interval

However, some oscillations occur in the second nonlinear problem (Fig. 2). The stiffness caused by the linear part at the beginning of the interval is well integrated. The oscillations are at the end of the interval where the nonlinear part dominates. This fact indicates that the considered initial guess is outside of the basin of attraction of the used Newton-type method (Hairer and Wanner 1991).

1.2 Our variational approach

The ideas which we would like to introduce for the treatment of singular perturbation problems are based on the analysis of a certain error functional of the form:

$$E(y, z) = \frac{1}{2} \int_0^T |y'(t) - f(y(t), z(t))|^2 + |\epsilon z'(t) - g(y(t), z(t))|^2 dt,$$

to be minimized among the absolutely continuous paths $y, z : (0, T) \rightarrow \mathbf{R}^N$ with $(y(0), z(0)) = (y_0, z_0)$. Note that if $E(y, z)$ is finite for one such path (y, z) , then automatically, y' and z are square integrable.

This error functional is associated in a natural way with the original singular perturbation problem.

One main property of this functional is the absence of local minima different from the solution of the problem. Thus, the 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. In our approach, based on optimality conditions, we only need to approximate linear singular perturbation problems. In particular, the problem of having good initial guesses in the application of Newton-type methods, to approximate the associated nonlinear equations, is avoided. In fact, we will be able to ensure global convergence independent of ϵ .

We should mention that we have already explored in some previous papers this point of view. Since the initial contributions (Amat and Pedregal 2009, 2013), we have also treated the reverse mechanism of using first discretization and then optimality (Amat et al. 2014). We also have addressed, from this viewpoint, other problems like DAEs (Amat et al. 2013) or problems with retarded arguments (Amat et al. 2012).

The paper is organized as follows: in Sect. 2, we introduce our algorithm and we prove its global convergence for all $\epsilon > 0$, and in Sect. 3, we present some numerical analysis pointing out some advantages of our variational approach in comparison with the classical implicit approach.

2 Some analysis of the variational approach

The analytical part of this contribution requires a special property on the maps f and g : for every positive $C > 0$ and small $\eta > 0$, there are constants $D_{C,\eta}^f > 0$ and $D_{C,\eta}^g > 0$, so that

$$|f(y + r, z + s) - f(y, z) - \nabla f(y, z)(r, s)| \leq D_{C,\eta}^f |(r, s)|^2, \quad |(y, z)| \leq C, |(r, s)| \leq \eta,$$

and

$$|g(y + r, z + s) - g(y, z) - \nabla g(y, z)(r, s)| \leq D_{C,\eta}^g |(r, s)|^2, \quad |(y, z)| \leq C, |(r, s)| \leq \eta.$$

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. See, however, Amat and Pedregal (2009) for a similar analysis under more general assumptions. Our emphasis 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, the analytical part is oriented towards providing a solid basis for this approximation procedure.

Proposition 1 *Let (\bar{y}, \bar{z}) be a critical point for the error E . Then, (\bar{y}, \bar{z}) is the solution of the singular perturbation problem.*

Proof The proof is elementary. Based on the smoothness and bounds assumed on the mappings f and g , we conclude that if $(y, z) \equiv (\bar{y}, \bar{z})$ is a critical point for the error E , then (y, z) ought to be a solution of the problem:

$$\begin{aligned} & -\frac{d}{dt} (y'(t) - f(y(t), z(t))) - (y'(t) - f(y(t), z(t))) \nabla f(y(t), z(t)) = 0 \text{ in } (0, T), \\ & -\frac{d}{dt} (\epsilon z'(t) - g(y(t), z(t))) - (\epsilon z'(t) - g(y(t), z(t))) \nabla g(y(t), z(t)) = 0 \text{ in } (0, T), \\ & y(0) = y_0, y'(T) - f(y(T), z(T)) = 0, z(0) = z_0, \epsilon z'(T) - g(y(T), z(T)) = 0. \end{aligned}$$

The vector-valued maps $r(t) = y'(t) - f(y(t), z(t))$ and $s(t) = \epsilon z'(t) - g(y(t), z(t))$ are then solutions of the linear, nondegenerate problems:

$$\begin{aligned} r'(t) + r(t)\nabla f(y(t), z(t)) &= 0 \text{ in } (0, T), \quad r(T) = 0, \\ s'(t) + s(t)\nabla g(y(t), z(t)) &= 0 \text{ in } (0, T), \quad s(T) = 0. \end{aligned}$$

The only solutions of these problems are $r \equiv 0$ and $s \equiv 0$, and so (y, z) is the solution of the singular perturbed problem. □

It, therefore, looks like a promising strategy to search for approximations of the solution (y, z) by minimizing the error E .

On the other hand, suppose that we start with an initial crude approximation $(y_{(0)}, z_{(0)}) \equiv (y, z)$ to the solution of our original singular perturbation problem. We could take $(y_{(0)}, z_{(0)}) = (y_0, z_0)$ for all t . We would like to improve this approximation in such a way that the error is significantly decreased.

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

$$\begin{aligned} E'(y, z)(r, s) &= \int_0^T (y'(t) - f(y(t), z(t)), \epsilon z'(t) - g(y(t), z(t)))^T \\ &\quad \cdot (r'(t) - \nabla f(y(t), z(t))r(t), \epsilon s'(t) - \nabla g(y(t), z(t))s(t)) \, dt. \end{aligned}$$

This expression suggests a main possibility to select (r, s) from:

Choose (r, s) , such that

$$\begin{aligned} r'(t) - \nabla f(y(t), z(t))(r(t), s(t)) &= f(y(t), z(t)) - y'(t) \text{ in } (0, T), \quad r(0) = 0, \\ \epsilon s'(t) - \nabla g(y(t), z(t))(r(t), s(t)) &= g(y(t), z(t)) - \epsilon z'(t) \text{ in } (0, T), \quad r(0) = 0. \end{aligned}$$

In this way, it is clear that $E'(y, z)(r, s) = -2E(y, z)$, and so the (local) decrease of the error is of the size $E(y, z)$. Finding (r, s) requires only solving the above linear problem.

2.1 Strong global convergence

Suppose (y, z) are feasible paths in the interval $(0, T)$, so that $y(0) = y_0, z(0) = z_0, y', z'$ are square integrable, and the quantity

$$E(y, z) = \frac{1}{2} \int_0^T |y'(t) - f(y(t), z(t))|^2 + |\epsilon z'(t) - g(y(t), z(t))|^2 \, dt,$$

measures how far such x is from being a solution of our problem. We also assume that $|(y(t), z(t))| \leq C$ for a fixed constant C , and all $t \in (0, T)$.

Choose $\eta > 0$ and $0 < \alpha < 1$, so that $\frac{\eta}{1-\sqrt{\alpha}} \leq C$.

We then solve for (r, s) the linear problem

$$\begin{aligned} r'(t) - \nabla f(y(t), z(t))(r(t), s(t)) &= f(y(t), z(t)) - y'(t) \text{ in } (0, T), \quad r(0) = 0, \\ \epsilon s'(t) - \nabla g(y(t), z(t))(r(t), s(t)) &= g(y(t), z(t)) - \epsilon z'(t) \text{ in } (0, T), \quad s(0) = 0, \end{aligned}$$

and pretend to update (y, z) to $(y + r, z + s)$ in such a way that the error for $(y + r, z + s)$ be less than the error for the current iteration (y, z) .

Note that

$$\begin{aligned}
 E(y + r, z + s) &= \frac{1}{2} \int_0^T |y'(t) + r'(t) - f(y(t) + r(t), z(t) + s(t))|^2 \\
 &\quad + |\epsilon(z'(t) + s'(t)) - g(y(t) + r(t), z(t) + s(t))|^2 dt \\
 &= \frac{1}{2} \int_0^T |f(y(t) + r(t), z(t) + s(t)) - f(y(t), z(t)) - \nabla f(y(t), z(t))(r(t), s(t))|^2 \\
 &\quad + |g(y(t) + r(t), z(t) + s(t)) - g(y(t), z(t)) - \nabla g(y(t), z(t))(r(t), s(t))|^2 dt
 \end{aligned}$$

where we have used the differential equations satisfied by r and s .

By our assumption on f and g above,

$$|f(y + r, z + s) - f(y, z) - \nabla f(y, z)(r, s)| \leq D_{C,\eta}^f |(r, s)|^2,$$

and

$$|g(y + r, z + s) - g(y, z) - \nabla g(y, z)(r, s)| \leq D_{C,\eta}^g |(r, s)|^2,$$

provided that $|(y, z)| \leq C, |(r, s)| \leq \eta$. Let $D := \max\{D_{C,\eta}^f, D_{C,\eta}^g\}$.

Since we know that (r, s) is the solution of the above linear problem and recalling our main assumption, $g_z(y, z)$ is invertible, and his logarithmic norm is negative in a ϵ -independent neighborhood of the solution, we have the upper bounds (see Theorem 3.2 in [Hairer and Wanner 1991](#)):

$$\max\{|r(t)|^2, |s(t)|^2\} \leq TKE(y, z) \quad \text{for all } t \in [0, T], K > 0. \tag{1}$$

Notice that $(r(t), s(t))$ is the solution of a linear perturbation problem with jacobian respect to the second component equal to $g_z(y, z)$ (the same as the original problem). Thus, assuming the above hypothesis on g , we ensure the above bounds for all the associated linear problems. In particular, all these constants are independent of ϵ .

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

$$E(y, z) \leq \frac{\eta^2}{KT}, \tag{2}$$

and then $\max\{|r(t)|, |s(t)|\} \leq \eta$ for all $t \in [0, T]$.

By the previous bounds, we can write

$$E(y + r, z + s) \leq \frac{D^2}{2} \int_0^T |r(t)|^4 + |s(t)|^4 dt \leq \frac{D^2}{2} K^2 T^3 E(y, z)^2. \tag{3}$$

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

$$E(y, z) \leq \frac{2\alpha}{D^2 K^2 T^3}, \tag{4}$$

then $E(y + r, z + s) \leq \alpha E(y, z)$.

All these calculations form the basis of a typical induction argument (see [Amat and Pedregal 2013](#) for more details), giving us the following convergence result:

Theorem 1 *With the previous notation, the sequence*

$$(y^{(j)}, z^{(j)}) = (y^{(j-1)} + r^{(j)}, z^{(j-1)} + s^{(j)}),$$

where

$$\begin{aligned}
 (r^{(j)})'(t) - \nabla f(y^{(j-1)}(t), z^{(j-1)}(t))(r^{(j)}(t), s^{(j)}(t)) &= f(y^{(j-1)}(t), z^{(j-1)}(t)) - (y^{(j-1)})'(t), \\
 \epsilon(s^{(j)})'(t) - \nabla g(y^{(j-1)}(t), z^{(j-1)}(t))(r^{(j)}(t), s^{(j)}(t)) &= g(y^{(j-1)}(t), z^{(j-1)}(t)) - \epsilon(z^{(j-1)})'(t),
 \end{aligned}$$

Table 1 Nonautonomous singular perturbed problem, $\epsilon = 10^{-5}$

<i>Tol</i>	Error
10^{-3}	9.76×10^{-4}
10^{-4}	6.10×10^{-5}
10^{-5}	7.63×10^{-6}

in $(0, T)$, with $r(0) = 0, s(0) = 0$, converges strongly in $L^\infty(0, T)$ to the unique solution of the singular perturbed 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 solution 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

$$\frac{E(y, z)D^2K^2}{2\alpha} \leq \frac{1}{h^3}. \tag{5}$$

3 Numerical analysis

We pretend to provide some experimental evidence that our iterative schemes are competitive with respect to the classical approach, without pretending at this stage to provide any further numerical analysis for our approach. We consider some singular perturbation problems well known in the literature.

In practice, we use as stopping criterion that the norms of r^j and s^j be smaller than a prefixed tolerance *Tol*.

We will observe the good convergence of our approach for different sizes of ϵ and moderate sizes of the discretization parameter. As initial guess of our algorithm, we simply consider the initial condition of the problems.

In all the paper, for simplicity in the notation, we only have considered autonomous problems. However, our approach works similarly for nonautonomous singular perturbed problems.

3.1 A nonautonomous singular perturbed problem

We consider the problem (Hairer and Wanner 1991)

$$\begin{aligned} \epsilon z'(t) &= -z(t) + \cos(t), \\ z(0) &= 1, \end{aligned}$$

whose explicit solution is

$$z(t) = (1 + \epsilon^2)^{-1}(\cos(t) + \epsilon \sin(t)) + C \exp(-t/\epsilon),$$

where $C = 1 - (1 + \epsilon^2)^{-1}$.

We consider again the trapezoid method with $N = 100$ and $T = 10$. We consider different tolerances for a given ϵ and different ϵ for two particular tolerances. In Tables 1, 2, and 3, we observe a very good numerical behavior of our approach according with our global convergence result.

Table 2 Nonautonomous singular perturbed problem, $Tol = 10^{-3}$

ϵ	Error
10^{-3}	9.7650×10^{-4}
10^{-5}	9.7641×10^{-4}
10^{-10}	9.7641×10^{-4}

Table 3 Nonautonomous singular perturbed problem, $Tol = 10^{-5}$

ϵ	Error
10^{-3}	4.9494×10^{-5}
10^{-5}	7.6278×10^{-6}
10^{-10}	7.6282×10^{-6}

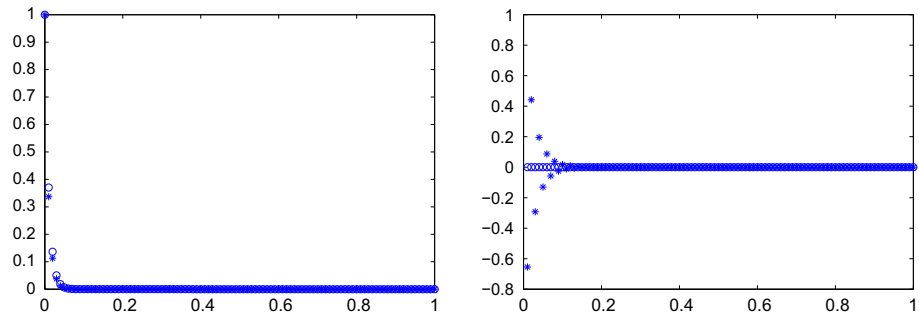


Fig. 3 Classical implementation of the trapezoid method, nonlinear test problem, ‘circle’-original, ‘asterisk’-approximation, *left* $\epsilon = 10^{-2}$, and *right* $\epsilon = 10^{-3}$

3.2 A nonlinear test problem

Finally, we consider a small nonlinear perturbation of the scalar linear test problem. Indeed, we are interesting in approximating the value $x(1)$ of the solution $x(t)$ of the problem:

$$\begin{aligned} \epsilon z'(t) &= -z(t) + 10^{-2} \cdot z^2(t), \\ z(0) &= 1, \end{aligned}$$

whose explicit solution is

$$z(t) = \frac{100}{1 + 99e^{-\frac{1}{\epsilon}t}}.$$

We consider the trapezoid method

$$z_{n+1} = z_n + \frac{h}{2}(f(z_n) + f(z_{n+1})). \tag{6}$$

This method averages the Euler and backward Euler methods, advancing the approximate solution at each step along a line whose slope is the arithmetic mean of the derivatives at its endpoints.

We consider a discretization of $N = 100$ points in $[0, 1]$. In Fig. 3, we plot the approximation using a classical implementation of the (implicit) trapezoid method, that is, solving

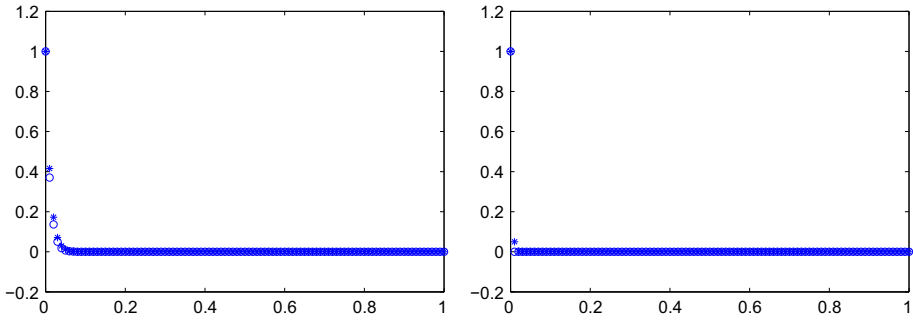


Fig. 4 New implementation of the trapezoid method, nonlinear test problem, ‘circle’-original, ‘asterisk’-approximation, *left* $\epsilon = 10^{-2}$, and *right* $\epsilon = 10^{-3}$

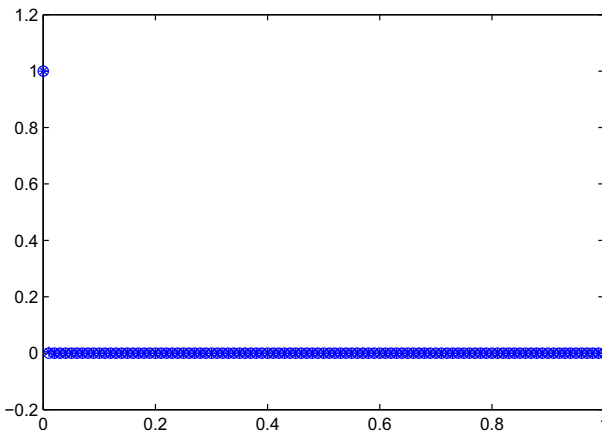


Fig. 5 New implementation of the trapezoid method, nonlinear test problem, ‘circle’-original, ‘asterisk’-approximation, and $\epsilon = 10^{-5}$

the associated nonlinear equation using Newton’s method. The method gives a bad approximation when the stiffness of the problem increases. We observe some oscillations near $t = 0$ where the algorithm is not able to find good initial guess for Newton’s method. However, looking at Figs. 4 and 5, we see that our method, starting with

$$z^0(t) = z(0) = 1$$

and taking $Tol = 10^{-4}$ in the stopping criterion of our iterative approach, converges in all the cases.

References

Amat S, Pedregal P (2009) A variational approach to implicit ODEs and differential inclusions. *ESAIM-COCV* 15(1):139–148
 Amat Sergio, Pedregal Pablo (2013) On a variational approach for the analysis and numerical simulation of ODEs. *Discret Contin Dyn Syst* 33(4):12751291
 Amat S, López DJ, Pedregal P (2014) An optimization approach for the numerical approximation of differential equations. *Optim* 63(3):337–358

- Amat S, Legaz MJ, Pedregal P (2013) Linearizing stiff delay differential equations. *Appl Math Inf Sci* 7(1):229232
- Amat S, Legaz MJ, Pedregal P (2012) On a Newton-type method for differential-algebraic equations. *J Appl Math Art.* ID 718608, 15
- Hairer E, Wanner G (1991) Solving ordinary differential equations II: stiff and differential algebraic problems. Springer-Verlag, Berlin
- Stoer J, Bulirsch R (1993) Introduction to numerical analysis, 2nd edn. Springer-Verlag, New York
- The MathWorks, Inc. (2017) MATLAB and SIMULINK, Natick, MA