CrossMark

Optimal homotopy analysis and control of error for implicitly defined fully nonlinear differential equations

Robert A. Van Gorder[1] (ID)

Abstract Implicitly defined fully nonlinear differential equations can admit solutions which have only finitely many derivatives, making their solution via analytical or numerical techniques challenging. We apply the optimal homotopy analysis method (OHAM) to the solution of implicitly defined ordinary differential equations, obtaining solutions with low error after few iterations or even one iteration of the method. This is particularly true in cases where an auxiliary nonlinear operator was employed (in contrast to the commonly used choice of an auxiliary linear operator), highlighting the need for further study on using auxiliary nonlinear operators in the HAM. Through various examples, we demonstrate that the approach is efficient for an appropriate selection of auxiliary operator and convergence control parameter.

Keywords Optimal homotopy analysis method · Implicit differential equations · Nonlinear auxiliary operator · Control of error

1 Introduction

The homotopy analysis method (HAM) is an analytical solution method which allows one to approximate the solution to a variety of problems, such as nonlinear ordinary differential equations, partial differential equations, integral equations [1–3]. The approach has been applied to many problems (see [4–8]). One reason for the popularity of the method is that it contains a convergence control parameter which allows

---

✉ Robert A. Van Gorder
Robert.VanGorder@maths.ox.ac.uk

[1]   Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK

one to adjust and control the error properties of the solution. For example, one may minimize the error or residual error of approximate solutions over this parameter, and such an approach is commonly referred to as the optimal homotopy analysis method (OHAM) [9–15]. The HAM also gives one great freedom in selecting the form of the solutions via representation of solutions [1, 16], since one has control over the type of basis functions employed in such a representation. This is tied closely to the choice of auxiliary operator used in the construction of the homotopy.

In applying the HAM, one attempts to solve a nonlinear differential equation $N[u] = 0$ by solving a related linear or nonlinear problem that is simpler and hence more amenable to analysis. Introducing an auxiliary operator, L, and convergence control parameter, h, the next step is to construct a homotopy of operators

$$H[u] = (1 - q)L[u] - hqN[u], \tag{1}$$

which is itself a differential operator. Here, $q \in [0, 1]$ is an embedding parameter, such that upon setting $q = 0$ we have $L[u] = 0$ implies $H[u] = 0$, while setting $q = 1$ gives $N[u] = 0$ implies $H[u] = 0$. Setting $H[u] \equiv 0$ identically, we have a homotopy of the operators L and N. If we then consider the more general equation $H[u] = 0$, then if a solution exists it should depend on the embedding parameter q, say $u(t) = \hat{u}(t; q)$. We should have that $\hat{u}(t; 0)$ is a solution of $L[u] = 0$, while $\hat{u}(t; 1)$ is a solution is $N[u] = 0$. If $\hat{u}(t; q)$ varies continuously in $q \in [0, 1]$, then taking q from 0 to 1 gives a map from a solution of $L[u] = 0$ to a solution of $N[u]$ (the original nonlinear problem of interest). Note that an entire literature exists for the case where $h = -1$, in which case the method is often referred to as the homotopy perturbation method (HPM); see [17–19]. However, there is no guarantee of convergence for the HPM, since the choice of $h = -1$ may not yield convergence. When h is picked in a specific way to minimize the error in the approximate solution the method is often referred to as the optimal homotopy analysis method (OHAM) [9–15], and this is the approach we shall consider.

In obtaining a homotopy solution, one often assumes that the resulting function is smooth with infinitely many derivatives existing at all points on the interior of the problem domain. However, as was recently shown [20], the HAM can also be used to construct solutions to differential equations which are not analytic over the domain. Such solutions exist when only a fixed number of derivatives exist, or when there is a sufficiently strong singularity in the domain. However, such solutions can still be physically reasonable and hence merit study. Indeed, for an nth order differential equation, a solution will in general need only have as many continuous derivatives as the order of the equation. Therefore, loss of regularity at higher order derivatives will not preclude a solution from being a physically viable solution to such a problem.

For implicit differential equations, such as those corresponding to fully nonlinear equations where the highest derivative may enter into the equation non-linearly, the solution may not be regular beyond the number of derivatives in the original equation. However, such equations do arise in various areas of science and engineering in the form of PDEs [21–23] and ODEs [24–27]. While many such equations are PDEs, traveling wave or other common reductions are often used to put them into the form of ODEs. Due to the appearance of implicitly defined differential equations in a variety of problems, it is sensible to consider the HAM for solving these equations.

However, a review of the open literature reveals that the problem of implicitly defined differential equations has not been well-studied in the context of the HAM or for other related analytic methods. This motivates us to develop the HAM for implicitly defined differential equations.

The remainder of the paper is organized as follows. In Section 2, we formulate the implicitly defined differential equations. In Section 3, we outline the application of the HAM for such problems, also discussing the optimal selection of the convergence control parameter (resulting in the OHAM). In Section 4, we consider several specific examples of implicit fully nonlinear differential equations, and apply OHAM to obtain approximate solutions. In some cases, the solutions will even be exact after few or even one iteration of the method. We discuss how, in some situations, auxiliary nonlinear operators may be preferable to selecting auxiliary linear operators. In Section 5 we give concluding remarks.

2 Formulation of the implicit problem

Let us consider the implicit ordinary differential equation

$$F(x', x, t) = 0,\tag{2}$$

subject to the initial condition

$$x(0) = a.\tag{3}$$

Here, prime denotes differentiation with respect to $t \in \mathbb{R}$. We assume that $F \in C^\infty(\mathbb{R}^2)$ with respect to x and x' (we do not make such an assumption on F with respect to t, in order to be able to consider discontinuous forcing terms), $x \in C^1(\mathbb{R})$, and $a \in \mathbb{R}$ is a constant. While we assume that F is smooth in all variables, we only assume that x is continuous with continuous first derivative. This allows our formulation to be much more general than if we were to assume that x was smooth. Under such an assumption, one could then differentiate (2) with respect to t, obtaining

$$\left(\frac{\partial F}{\partial x'}(x', x, t)\right) x'' + \left(\frac{\partial F}{\partial x}(x', x, t)\right) x' + \left(\frac{\partial F}{\partial t}(x', x, t)\right) = 0,\tag{4}$$

which is no longer an implicit equation for the highest derivative. Therefore, considering (2) is more general than considering (4).

We should also note that we can extend (2)–(3) to vector equations, viz.,

$$\mathbf{F}(\mathbf{x}', \mathbf{x}, t) = \mathbf{0},\tag{5}$$

subject to the initial condition

$$\mathbf{x}(0) = \mathbf{a}.\tag{6}$$

Here, $\mathbf{F} : \mathbb{R}^{2\xi+1} \to \mathbb{R}^\xi$, $\mathbf{F} \in C^\infty(\mathbb{R}^{2\xi})$ (with respect to $\mathbf{x}$ and $\mathbf{x}'$ but not t, again so that we may include forcing terms that may be discontinuous), $\mathbf{x} : \mathbb{R} \to \mathbb{R}^\xi$, $\mathbf{x} \in C^1(\mathbb{R})$ is a vector-valued function, $a \in \mathbb{R}^\xi$ is a constant vector, and $\xi \geq 1$ is the dimension of the system. Such a construction would then allow one to consider any differential equation with higher order derivatives. For example, the single second order differential equation

$$\sin(x'') + \cos(x') + x = t\tag{7}$$

would fit into the framework (5) as

$$\mathbf{F}(\mathbf{y}', \mathbf{y}, t) = \mathbf{0} \,, \qquad (8)$$

where $\mathbf{y} = [y_1, y_2]^T$, $y_1 = x$, $y_2 = x'$, and

$$\mathbf{F}(\mathbf{y}', \mathbf{y}, t) = \begin{bmatrix} y_1' - y_2 \\ \sin(y_2') + \cos(y_2) + y_1 - t \end{bmatrix} . \qquad (9)$$

Therefore, we can use a formulation such as (5)–(6) to study a wide variety of problems.

3 Homotopy analysis method and control of residual error for the implicit problem (2)–(3)

Let us construct the homotopy

$$\mathcal{H}[x] := (1 - q)L[x] - hq F(x', x, t) \equiv 0 \,. \qquad (10)$$

Clearly, when $q = 0$, $\mathcal{H}[x] = 0$ is equivalent to $L[x] = 0$, which is a linear problem. On the other hand, when $q = 1$, $\mathcal{H}[x] = 0$ is equivalent to $F(x', x, t) = 0$, which is the original nonlinear problem.

Let us consider the standard homotopy series solution to (10),

$$X(t; q) = \sum_{k=0}^{\infty} q^k x_k(t) \,. \qquad (11)$$

When $q = 0$, $X(t) = x_0(t)$ is the initial guess or approximation to the solution. On the other hand, if this series exists as $q \rightarrow 1^-$, then the series $X(t)$ converges to a solution $x(t)$ of the original nonlinear problem (2). Often, one will be interested in a truncated expansion of the form (11) evaluated at $q = 1$, say

$$X_K(t) = \sum_{k=0}^{K} x_k(t) \,, \qquad (12)$$

as this serves as an approximation to the true solution of (2). The approximation should gain accuracy as we increase the total number of terms included in this expansion. Theorems concerning the convergence of the series to the true solution of the initial equation can be found in [1, 3, 28–35].

In order to calculate the terms in the approximation (12), we place the series (11) into the homotopy (10),

$$(1 - q)L[X(t; q)] - hq F(X'(t; q), X(t; q), t) = 0 \,. \qquad (13)$$

Matching powers of the embedding parameter, q, we find the zeroth order deformation equation,

$$L[x_0(t)] = 0 \,, \qquad (14)$$

as well as the higher order deformation equations

$$L[x_k(t)] = L[x_{k-1}(t)] + h f_k(x_0'(t), \dots, x_{k-1}'(t), x_0(t), \dots, x_{k-1}(t), t) \,, \qquad (15)$$

for $k = 1, 2, 3, \ldots$. Here, the terms f_k are obtained through the expansion of the function $q F(X'(t; q), X(t; q), t)$ in powers of q, that is to say

$$q F \left(\sum_{k=0}^{K} x'_k(t), \sum_{k=0}^{K} x_k(t), t \right)$$
$$= q f_1(x'_0(t), x_0(t), t) + q^2 f_2(x'_0(t), x'_1(t), x_0(t), x_1(t), t) + \cdots \quad (16)$$
$$= \sum_{j=1}^{\infty} q^j f_j(x'_0(t), \ldots, x'_{j-1}(t), x_0(t), \ldots, x_{j-1}(t), t).$$

The terms in this expansion can be calculated for any given F. The first few terms are

$$f_1 = F(x'_0, x_0, t), \quad (17)$$

$$f_2 = \frac{\partial F}{\partial x'}(x'_0, x_0, t)x'_1 + \frac{\partial F}{\partial x}(x'_0, x_0, t)x_1, \quad (18)$$

$$f_3 = \frac{\partial^2 F}{\partial x' \partial x}(x'_0, x_0, t)x'_1 x_1 + \frac{1}{2}\frac{\partial^2 F}{\partial x'^2}(x'_0, x_0, t)(x'_1)^2 + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(x'_0, x_0, t)(x_1)^2$$
$$+ \frac{\partial F}{\partial x'}(x'_0, x_0, t)x'_2 + \frac{\partial F}{\partial x}(x'_0, x_0, t)x_2, \quad (19)$$

$$f_4 = \frac{1}{2}\frac{\partial^3 F}{\partial x'^2 \partial x}(x'_0, x_0, t)(x'_1)^2 x_1 + \frac{1}{2}\frac{\partial^3 F}{\partial x' \partial x^2}(x'_0, x_0, t)x'_1(x_1)^2$$
$$+ \frac{1}{6}\frac{\partial^3 F}{\partial x'^3}(x'_0, x_0, t)(x'_1)^3 + \frac{1}{6}\frac{\partial^3 F}{\partial x^3}(x'_0, x_0, t)(x_1)^3$$
$$+ \frac{\partial^2 F}{\partial x'^2}(x'_0, x_0, t)x'_1 x'_2 + \frac{\partial^2 F}{\partial x \partial x}(x'_0, x_0, t)(x'_1 x_2 + x_1 x'_2) \quad (20)$$
$$+ \frac{\partial^2 F}{\partial x^2}(x'_0, x_0, t)x_1 x_2 + \frac{\partial F}{\partial x'}(x'_0, x_0, t)x'_3 + \frac{\partial F}{\partial x}(x'_0, x_0, t)x_3.$$

Therefore, when we solve the deformation (15), we solve a linear problem for each $x_k(t)$, with the inhomogeneous term given (in general) in terms of all previously calculated terms. These equations are solved subject to the initial conditions

$$x_0(0) = a \quad \text{and} \quad x_k(0) = 0 \quad \text{for} \quad k = 1, 2, 3, \ldots. \quad (21)$$

3.1 Selection of the auxiliary linear operator L

It is clear from the form of the terms (17)–(20) that the initial approximation $x_0(t)$ will have a critical influence on the solution process. Indeed, it will be the initial approximation which molds the effects from the nonlinearity (the derivatives of F will involve only arguments of $x_0(t)$ and $x'_0(t)$). This highlights the need to find a good initial approximation for the problem.

Since our problem is first order, we will restrict our attention to first order operators, which can be put into the form

$$L[x] = x' + A(t)x. \quad (22)$$

If $A(t) = 0$, we obtain the initial approximation $x_0(t) = a$, and the full homotopy solution can be recovered in terms of powers of the variable t. However, when $A(t) \neq 0$, we should obtain

$$x_0(t) = a \exp\left(-\int_0^t A(\tau)d\tau\right).\tag{23}$$

This initial approximation can take a variety of forms, depending on the functional form of $A(t)$. It is important to pick the operator (22) so that the initial approximation is accurate. If we consider how $x_0(t)$ given in (23) acts when placed into the original nonlinear equation, we obtain

$$F(X_K'(t; h), X_K(t; h), t) = \text{Res}(t; h),\tag{24}$$

which gives the residual error. If this quantity is small, then the initial approximation is a good choice, and it is likely that relatively few additional terms will be needed to gain a sufficiently accurate homotopy solution.

Another way to view an appropriate choice of the linear operator is to consider a linearization of the original nonlinear equation at $t = 0$. Linearizing at the point $t = 0$, we find

$$\frac{\partial F}{\partial x'}(x'(0), a, 0)x' + \frac{\partial F}{\partial x}(x'(0), a, 0)x,\tag{25}$$

where we neglect constant terms in the linearization. This linearization depends on the value of x' at $t = 0$. However, this can be found provided that the algebraic equation $F(x'(0), a, 0) = 0$ can be solved. In general, this solution will be a function of the initial condition $x(0) = a$, and there may even be multiple solutions. So, writing $x'(0) = b$, and noting that $b = b(a)$ in general, we have the option to pick

$$L[x] = x' + A_0(a)x,\tag{26}$$

where

$$A_0(a) = \frac{\frac{\partial F}{\partial x}(b(a), a, 0)}{\frac{\partial F}{\partial x'}(b(a), a, 0)}\tag{27}$$

is a constant which depends on the initial condition $x(0) = a$. Of course, we assume that $\frac{\partial F}{\partial x'}(b(a), a, 0) \neq 0$; otherwise, the linearization is degenerate. For this case, the initial approximation is an exponential function,

$$x_0(t) = ae^{-A_0(a)t}.\tag{28}$$

3.2 Control of error and optimal homotopy analysis

While a proper choice for the auxiliary linear operator L is important, in order to minimize the error inherent in any truncated approximation to the true solution we should consider the role of the convergence control parameter, h. Assuming that we carry out the iterative algorithm needed to obtain a truncated solution (12), we still need to pick h appropriately. Let $X_K(t; h)$ be such a truncated homotopy solution, where we make explicit the dependence on the convergence control parameter, h. Then, if we place this approximation back into the original nonlinear equation, we obtain the residual error at each point t in the domain, given by (24). In the absence of an exact solution to compare our approximate solution with, we should attempt to

minimize this residual error, since a larger residual error implies that the approximate solution $X_k(t; h)$ is a poor fit the to original differential equation.

Therefore, we often consider the integral of squared residual errors,

$$E(h) = \frac{1}{l} \int_0^l (\text{Res}(t; h))^2 dt . \tag{29}$$

Often, it is impractical to integrate over all of $\mathbb{R}$, so we shall often consider a more restricted integral region $[0, l]$. A minimum value for $E(h)$ should exist (although it need not be unique). The so-called optimal homotopy analysis method then states that h should be chosen so as to minimize the function $E(h)$, i.e., there should be a residual error minimizing choice of h, say h^*, such that

$$\min_{h \neq 0} E(h) = E(h^*) \geq 0 . \tag{30}$$

At $h = h^*$, the residual errors are minimized over the subset of the domain $[0, l]$. Therefore, $X_K(t; h^*)$ is a residual error minimizing approximation to the true solution $x(t)$ over the interval $t \in [0, l]$. There are many examples in the literature where the success of the approximation obtained through the homotopy analysis method depends strongly on the choice of the convergence control parameter, h [9–15]. Therefore, this error minimization is essential for selecting the parameter h in the homotopy analysis method.

We should note that in rare cases where we have an exact solution to the nonlinear differential equation, we may perform a similar optimization with respect to the exact solution. Let $x_{\text{exact}}(t)$ denote a closed-form exact solution to (2). Then,

$$\text{Abs}(h) = \frac{1}{l} \int_0^l (x_{\text{exact}}(t) - X_K(t; h))^2 dt . \tag{31}$$

denotes the average squared error between the exact solution and the approximate solution over the subset $[0, l]$ of the real line $\mathbb{R}$. In such a situation where $x_{\text{exact}}(t)$ is known, h can be optimally chosen by minimizing this measure of absolute error.

4 Implementation of the method for example problems

In the previous section, we outlined the homotopy analysis method for the solution of implicit first order nonlinear initial value problems. Therefore, we are now able to use these results to present several examples of the analytical method.

4.1 A simple case where the choice of L matters

Consider the nonlinear implicit initial value problem

$$\sin^2\left(x'(t)\right) + \cos^2\left(x(t)\right) = 1 \quad \text{subject to} \quad x(0) = a . \tag{32}$$

In order to demonstrate the utility of appropriately selecting the auxiliary linear operator, we shall consider two choices for L, namely

$$L_1[x] = x' \quad \text{and} \quad L_2[x] = x' + x . \tag{33}$$

Applying the method outlined in the previous section, using L_1, we find that a five-term approximation can be iteratively calculated as

$$
\begin{aligned}
X_4(t) = a &+ \left(\frac{3h^2 - 4}{2} + 2(1 - h^2)\cos(2a) + \frac{h^2}{2}\cos(4a) \right) ht \\
&+ \frac{3}{32}((16 - 5h)\sin(2a) - 4(2 - h^2)\sin(4a) - \sin(6a))h^2 t^2 \\
&+ (1 - 3\cos(2a) + 3\cos(4a) - \cos(6a))\frac{h^3}{6}t^3 \\
&- (4\sin(2a) - 6\sin(4a) + 4\sin(6a) - \sin(8a))\frac{h^4}{64}t^4 .
\end{aligned}
\tag{34}
$$

On the other hand, if we seek a solution using the auxiliary linear operator L_2, we find that the initial approximation is given by

$$
x_0(t) = ae^{-t} ,
\tag{35}
$$

while all of the higher order terms can be shown to give zero, that is, $x_k(t) = 0$ for all $k \geq 1$. Therefore, we have the homotopy solution $X(t) = ae^{-t}$. Since this solution comes from only the initial approximation, there is no dependence on the convergence control parameter, h. It is interesting that the homotopy solution should essentially terminate after a fixed number of terms. The reason for this is clear if we consider the (32) more closely: indeed, we see that $x(t) = ae^{-t}$ is an exact solution. Therefore, with the choice of linear operator L_2, we have obtained the exact solution from the HAM. Since the exact solution is obtained, the method stops and all higher order terms are zero.

Now, while the solution obtained with the linear operator L_1 is not exact, we should determine how good the approximation is. Since we have an exact solution $x(t) = ae^{-t}$, we can compare the solution (34) directly without the need for residual errors. Let us then define the integral of the square of the absolute errors as in (31). The value of l in that equation gives the size of the interval on which we are comparing the solutions. While the domain is unrestricted, often we consider a finite region on which to calculate the error for numerical reasons. Of course, for the solution obtained using L_2, this integral is always zero. For the case where we use L_1, we should calculate

$$
\mathrm{Abs}(h) = \frac{1}{l} \int_0^l (ae^{-t} - X_4(t))^2 dt .
\tag{36}
$$

In Figs. 1 and 2, we plot this error as a function of h, for various values of l and a. As both the size of the initial condition a and the interval size l increase, the error tends to increase. As we increase the domain size, the solution loses accuracy. For such cases, we would need more terms in the homotopy solution in order to recover the lost accuracy. As the initial condition is increased, the solutions also lose accuracy, and an increase in the initial condition results in an increase in the magnitude of the solution, which results in any errors being amplified as well.

In Fig. 3, we plot the error minimizing approximation against the true solution $x(t) = ae^{-t}$. The first HAM solution is very accurate over the domain $[0, 1]$, which is where the optimization occurred. On the other hand, the solution optimized over
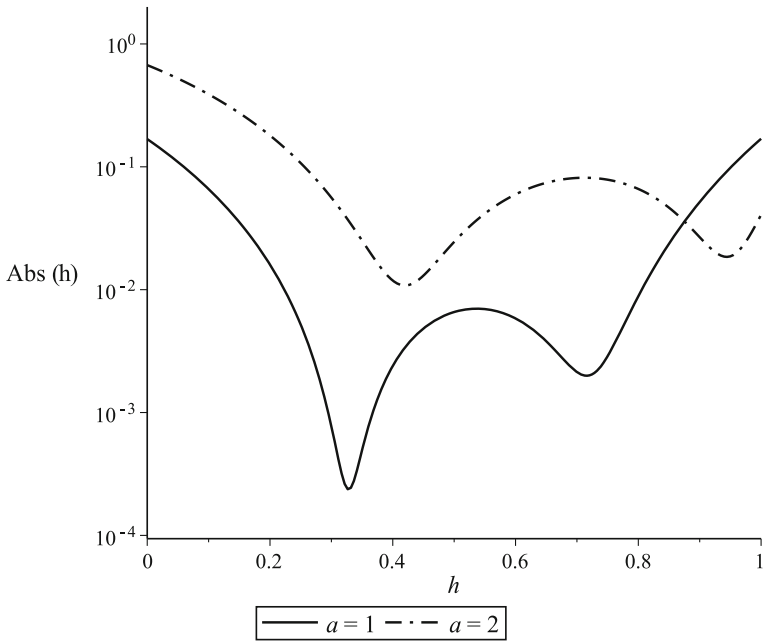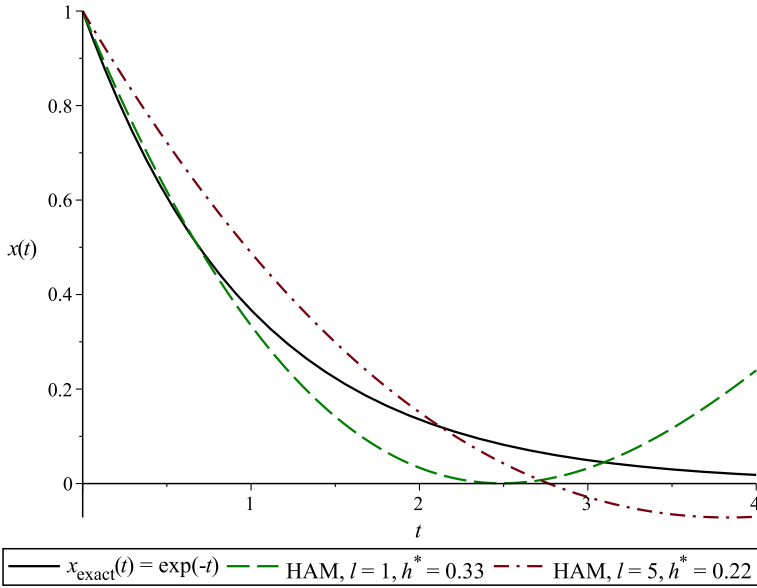
Fig. 1 Plot of the absolute errors Abs(h) given in (36) for the five-term homotopy solution (34) when compared to the exact solution $x_{\text{exact}}(t) = ae^{-t}$, for various values of a. We fix $l = 1$
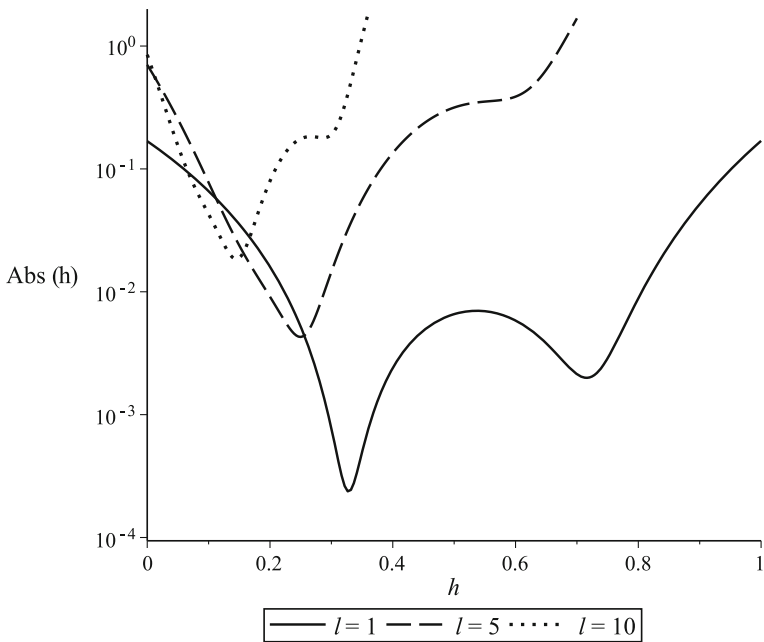
Fig. 2 Plot of the absolute errors Abs(h) given in (36) for the five term homotopy solution (34) when compared to the exact solution $x_{\text{exact}}(t) = ae^{-t}$, for various values of l. We fix $a = 1$

Fig. 3 Plot of the exact solution $x_{\text{exact}}(t) = ae^{-t}$ along with the HAM solutions with optimal convergence control parameter $h^* = 0.33$ when $l = 1$, $h^* = 0.22$ when $l = 5$

[0, 5] is not very accurate over that interval. Both of these HAM solutions correspond to the auxiliary linear operator $Lx = x'$. When we instead pick $Lx = x' + x$, we quickly recover the exact solution. This demonstrates that one should also carefully select the auxiliary linear operator, and not only the convergence control parameter, in the homotopy analysis method.

Of course, it makes sense to wonder why one choice of linear operator works much better than another. To explain this, let us return to the discussion of the linear operator from the previous section. At $t = 0$, for the present problem we obtain a linearization (25) of the form

$$2\sin(x'(0))\cos(x'(0))x' - 2\sin(a)\cos(a)x . \tag{37}$$

Yet, at $t = 0$ we must have $0 = F(x'(0), a, 0) = \sin^2(x'(0)) + \cos^2(a) - 1$, which gives $\sin^2(x'(0)) = \sin^2(a)$. We have two solutions for $x'(0)$, either $x'(0) = a$ or $x'(0) = -a$. Let us pick $x'(0) = -a$. This gives

$$A_0(a) = \frac{2\sin(-a)\cos(-a)}{-2\sin(a)\cos(a)} = 1 , \tag{38}$$

which in turn implies that

$$L[x] = x' + A_0(a)x = x' + x \tag{39}$$

corresponds exactly with the linearization of F at $t = 0$. Hence, by picking the auxiliary linear operator L to exactly correspond to the linearization of F at $t = 0$, we were quickly able to recover the exact solution using the homotopy analysis method.

4.2 Clairaut's equation with piecewise continuity

Consider next the ill-posed boundary value problem

$$U(t; 1)x'^2 + tx' - x = 0 \quad \text{subject to} \quad x(0) = 0 \quad \text{and} \quad x'(2) = 1, \quad \text{for} \quad 0 \le t \le 2. \tag{40}$$

Here, the function $U(t; \tau)$ is a step function defined by

$$U(t, \tau) = \begin{cases} 0, & t < \tau, \\ 1, & t \ge \tau. \end{cases} \tag{41}$$

Equation (40) is a form of Clairaut's equation [24] if instead U is set to unity.

This problem is ill-posed, since it is first order yet contains two boundary conditions. A continuous exact solution exists over the domain $t \in [0, 2]$, and takes the form

$$x_{\text{exact}}(t) = \begin{cases} 2t, & 0 \le t < 1, \\ 1 + t, & 1 \le t \le 2. \end{cases} \tag{42}$$

In order to solve this equation using HAM, define

$$F(x', x, t) = U(t; 1)x'^2 + tx' - x. \tag{43}$$

We shall consider each part of the solution separately. First, when $t = 0$, note that

$$\frac{\partial F}{\partial x} = -1 \quad \text{and} \quad \frac{\partial F}{\partial x'} = 0, \tag{44}$$

hence the quantity $A(0)$ is singular. Note that near $t = 0$, we can solve the equivalent problem $tx'_0 - x_0 = 0$, $x_0(0) = 0$, which gives the zeroth order approximation $x_0(t) = At$. Since this satisfies both the ODE and the boundary condition, this is the exact solution on the region where $0 \le t \le 1$, i.e., where $U(t; 1) = 0$, so we do not need to calculate higher terms as they will all be zero since we have already converged to the exact solution at the first step.

To obtain the solution on the second half of the domain, we expand F about $t = 2$. We have that $x'(2) = 1$. Meanwhile, $F(1, x, 2) = 3 - x = 0$, so $x(2) = 3$. Thus, expanding about $t = 2$, $x = 3$, $x' = 1$, we have

$$(t - 2)(x' - 1) + t - 3 - x + 4x' = O((x' - 1)^2), \tag{45}$$

or

$$(t + 2)x' - x - 1 = 0. \tag{46}$$

Define an auxiliary *nonlinear* operator by

$$L[x] = (t + 2)x' - x - 1. \tag{47}$$

Forming the homotopy

$$0 = H[x] = (1 - q)L[x] - hqN[x], \tag{48}$$

where $N[x] = U(t; 1)x'^2 + tx' - x = x'^2 + tx' - x$, the zeroth order approximation is given by the solution to

$$(t - 2)x'_0 - x_0 - 1 = 0, \tag{49}$$

$$x'_0(2) = 1. \tag{50}$$

The general solution is $x(t) = C_0(t + 2) - 1$, and using the boundary condition, we have $C_0 = 1$ giving $x_0(t) = t + 1$. However, note that $N[x_0(t)] = 0$, hence $x_0(t)$ is also an exact solution. Therefore, the homotopy series terminates after the initial term since it has converged to the exact solution after one step. Therefore, we have obtained the exact solution on the interval $1 \leq t \leq 2$.

What remains is to match the solutions on the interval $0 \leq t \leq 1$ and the interval $1 \leq t \leq 2$. To do so, note

$$A = \lim_{t \to 1^-} x_0(t) = \lim_{t \to 1^+} x_0(t) = 2 \tag{51}$$

giving $A = 2$. As such, we obtain $x(t) = 2t$ on $0 \leq t \leq 1$ and $x(t) = t + 1$ on $1 \leq t \leq 2$. This is exactly the exact solution (42).

4.3 An ODE related to the Whitney umbrella point

We consider the implicit problem

$$x'^2 - t(t - x)^2 = 0 \quad \text{subject to} \quad x(1) = 0, \quad \text{for} \quad 0 \leq t \leq 1. \tag{52}$$

This ODE is related to the Whitney umbrella point [25] and also arises in the study of slow motion of relaxation equations with one fast and two slow variables [26, 27]. Note from the form of the equation that $x'(1) = \pm 1$.

An exact solution can be given by choosing $x'(1) = -1$, but it takes the rather complicated form

$$
\begin{aligned}
x_{\text{exact}}(t) = \frac{18^{2/3}}{10} \Bigg\{ &\text{WhittakerM}\left(\frac{1}{3}, \frac{5}{6}, \frac{2}{3}\right) \exp\left(\frac{2}{3}t^{3/2} - \frac{1}{3}\right) \\
&- \sqrt{t}\,\text{WhittakerM}\left(\frac{1}{3}, \frac{5}{6}, \frac{2}{3}t^{3/2}\right) \exp\left(\frac{1}{3}t^{3/2}\right) \Bigg\},
\end{aligned} \tag{53}
$$

where WhittakerM denotes the Whittaker M function.

Note that the nonlinear operator $N[x] = x'^2 - t(t - x)^2$ for this example can be factored into a form

$$N[x] = L_-[x]L_+[x], \tag{54}$$

where $L_\pm[x] = x' \pm \sqrt{t}(t - x)$. Let us chose $L[x] = L_+[x]$ and then construct the homotopy

$$0 = H[x] = (1-q)L_+[x] - hqN[x] = (1-q)(x' - \sqrt{t}(t-x)) - hq(x'^2 - t(t-x)^2). \tag{55}$$

This represents a rare case where the homotopy operator $H[x] = 0$ itself can be solved directly. Indeed, writing

$$
\begin{aligned}
0 = H[x] &= (1-q)L_+[x] - hqN[x] \\
&= (1-q)L_+[x] - hqL_+[x]L_-[x] \\
&= L_+[x](1 - q - hqL_-[x]).
\end{aligned} \tag{56}
$$

Therefore, all of $H[x]$, $N[x]$, and $L_+[x]$ share a solution. This means that if one is able to solve $L_+[x]$, then one will have a solution to $N[x]$ and $H[x]$ without the need for iteration.

Note that $L_+[x] = 0$ gives the equation

$$x' - t^{1/2}x = -t^{3/2}, \tag{57}$$

subject to $x(1) = 0$. Solving the ODE using integrating factors, we find

$$x_{\text{exact}}(t) = C_0 \exp\left(\frac{2}{3}t^{3/2}\right) - \frac{18^{2/3}}{10}\sqrt{t}\,\text{WhittakerM}\left(\frac{1}{3}, \frac{5}{6}, \frac{2}{3}t^{3/2}\right)\exp\left(\frac{1}{3}t^{3/2}\right). \tag{58}$$

Using the boundary condition $x(1) = 0$ then allows us to choose the arbitrary constant C_0, and we recover the exact solution (53). Again, the solution to an appropriately selected auxiliary operator has given us an exact solution in one iteration.

4.4 An ODE with solutions in terms of nonlinear combinations of Lambert W functions

Let us consider the initial value problem

$$x'\exp(x') - t = 0 \quad \text{subject to} \quad x(0) = 1, \quad \text{for} \quad t \ge 0. \tag{59}$$

This solution admits a closed-form exact solution

$$x_{\text{exact}}(t) = \frac{\text{LambertW}(t)^2 - \text{LambertW}(t) + 1}{\text{LambertW}(t)}t, \tag{60}$$

where $\text{LambertW}(t)$ denotes the Lambert W function [36]. Although there is an exact solution, having an analytical solution in terms of simpler base functions may cast more light on the behavior of the solution.

Let us expand $F(x', x, t)$ about $x' = 0$, $x = 1$, $t = 0$ to get

$$x' - t = O((x')^2). \tag{61}$$

Therefore, define the operators $L[x] = x' - t$, $N[x] = x'\exp(x') - t$ and construct the homotopy

$$0 = H[x] = (1 - q)(x' - t) - hq(x'\exp(x') - t). \tag{62}$$

The zeroth order deformation equations give

$$x_0' - t = 0, \tag{63}$$

$$x_0(0) = 1, \tag{64}$$

which gives the initial approximation

$$x_0(t) = \frac{1}{2}t^2 + 1. \tag{65}$$

Generally, the higher-order deformation equations give

$$x_k' - t = f_k(t), \tag{66}$$

$$x_k(0) = 0, \tag{67}$$

where $f_k(t)$ involves the lower order terms in the homotopy series. Therefore, we obtain

$$x_k(t) = C_k + \int_0^t (s + f_k(s)) ds \,, \tag{68}$$

and using the boundary condition $x_k(0) = 0$, we have

$$x_k(t) = \int_0^t (s + f_k(s)) ds \,. \tag{69}$$

Choosing $h = -1$ and iterating through the process, we find

$$x(t) = x_0(t) + \sum_{k=1}^{\infty} x_k(t) = 1 + \frac{1}{2}t^2 - \frac{1}{3}t^3 + \frac{3}{8}t^4 - \frac{8}{15}t^5 + \frac{125}{144}t^6 + O(t^7) \,. \tag{70}$$

This is exactly the Taylor series (centered at $t = 0$) for the exact solution (60).

5 Conclusions

We have considered the application of the optimal homotopy analysis method (OHAM) to the solution of implicitly defined ordinary differential equations. Such equations may be fully nonlinear and hence may admit solutions which have only finitely many derivatives. Since the HAM solution need only be smooth in the homotopy parameter, rather than in the independent variables of the original problem, HAM allows us to obtain solutions even when higher-order derivatives of the solutions may fail to exist. Indeed, one benefit of the HAM for the solution of non-analytic solutions of differential equations was discussed recently in [20]. Furthermore, optimal selection of the convergence control parameter can aide in obtaining solutions with low error, yielding more accurate approximations.

One interesting feature of applying HAM or OHAM to the implicitly defined nonlinear ODEs was that we could sometimes obtain an optimal solution after few iterations or even one iteration of the method. This was particularly true in cases where an auxiliary nonlinear operator was employed, highlighting the need for further study on using auxiliary nonlinear operators in the HAM, in contrast to the commonly used choice of an auxiliary linear operator. The utility of using auxiliary nonlinear operators was previously discussed in [37] in the context of solution to a generalized Drinfel'd - Sokolov equation.

References

1. Liao, S.J.: Beyond perturbation: introduction to the homotopy analysis method. Chapman & Hall/CRC Press, Boca Raton (2003)

2. Liao, S.J.: On the homotopy analysis method for nonlinear problems. Appl. Math. Comput. **147**, 499–513 (2004)
3. Liao, S.J.: Homotopy analysis method in nonlinear differential equations. Springer & Higher Education Press, Heidelberg (2012)
4. Vajravelu, K., Van Gorder, R.A.: Nonlinear flow phenomena and homotopy analysis: Fluid Flow and Heat Transfer. Springer, Heidelberg (2013)
5. Tan, Y., Abbasbandy, S.: Homotopy analysis method for quadratic Riccati differential equation. Commun. Nonlinear Sci. Numer. Simul. **13**, 539–546 (2008)
6. Sajid, M., Hayat, T.: Comparison of HAM and HPM methods in nonlinear heat conduction and convection equations. Nonlinear Anal B: Real World Appl. **9**, 2296–2301 (2008)
7. Hashim, I., Abdulaziz, O., Momani, S.: Homotopy analysis method for fractional IVPs. Commun. Nonlinear Sci. Numer. Simul. **14**, 674–684 (2009)
8. Ellahi, R., Raza, M., Vafai, K.: Series solutions of non-Newtonian nanofluids with Reynolds' model and Vogel's model by means of the homotopy analysis method. Math. Comput. Model. **55**, 1876–1891 (2012)
9. Mastroberardino, A.: Homotopy analysis method applied to electrohydrodynamic flow. Commun. Nonlinear Sci. Numer. Simul. **16**, 2730–2736 (2011)
10. Liang, S., Liu, S.: An open problem on the optimality of an asymptotic solution to Duffing's nonlinear oscillation problem. Commun. Nonlinear Sci. Numer. Simul. **19**, 4189–4195 (2014)
11. Liao, S.: An optimal homotopy-analysis approach for strongly nonlinear differential equations. Commun. Nonlinear Sci. Numer. Simul. **15**, 2315–2332 (2010)
12. Mallory, K., Van Gorder, R.A.: Optimal homotopy analysis and control of error for solutions to the non-local Whitham equation. Numer. Algorithms **66**, 843–863 (2014)
13. Baxter, M., Van Gorder, R.A., Vajravelu, K.: Optimal analytic method for the nonlinear Hasegawa-Mima equation. Euro. Phys. J Plus **129**, 98 (2014)
14. Van Gorder, R.A.: Control of error in the homotopy analysis of semi-linear elliptic boundary value problems. Numer. Algorithms **61**, 613–629 (2012)
15. Abbasbandy, S., Shivanian, E., Vajravelu, K.: Mathematical properties of h-curve in the frame work of the homotopy analysis method. Commun. Nonlinear Sci. Numer. Simul. **16**, 4268–4275 (2011)
16. Van Gorder, R.A., Vajravelu, K.: On the selection of auxiliary functions, operators, and convergence control parameters in the application of the homotopy analysis method to nonlinear differential equations: a general approach. Commun. Nonlinear Sci. Numer. Simul. **14**, 4078–4089 (2009)
17. He, J.-H.: Homotopy perturbation method: a new nonlinear analytical technique. Appl. Math. Comput. **135**, 73–79 (2003)
18. Ganji, D.D.: The application of He's homotopy perturbation method to nonlinear equations arising in heat transfer. Phys. Lett. A **355**, 337–341 (2006)
19. Abbasbandy, S.: Homotopy perturbation method for quadratic Riccati differential equation and comparison with Adomian's decomposition method. Appl. Math. Comput. **172**, 485–490 (2006)
20. Van Gorder, R.A.: On the utility of the homotopy analysis method for non-analytic and global solutions to nonlinear differential equations. Numer. Algorithms **76**, 151–162 (2017)
21. Huicheng, Y., Qingjiu, Q.: Tangent interation of co-normal waves for second order full nonlinear strickly hyperbolic equations. Nonlinear Anal.: Theory Methods & Appl. **19**, 81–93 (1992)
22. Wei, G., Kirby, J.T., Grilli, S.T., Subramanya, R.: A fully nonlinear Boussinesq model for surface waves. Part 1. Highly nonlinear unsteady waves. J. Fluid Mech. **294**, 71–92 (1995)
23. Caffarelli, L., Crandall, M.G., Kocan, M., Swiech, A.: On viscosity solutions of fully nonlinear equations with measurable ingredients. Commun. Pure Appl. Math. **49**, 365–398 (1996)
24. Clairaut, A.C.: Solution de plusieurs problèmes où il s'agit de trouver des Courbes dont la propriété consiste dans une certaine relation entre leurs branches, exprimé par une Équation donnée, Histoire de l'Acadé,mie Royale des Sciences (1734), 196–215
25. Davydov, A.A., Ishikawa, G., Izumiya, S., Sun, W.Z.: Generic singularities of implicit systems of first order differential equations on the plane. Jpn. J. Math. **3**(1), 93–119 (2008)
26. Arnold, V.I.: Contact structure, relaxation oscillations and singular points of implicit differential equations. In: Global Analysis-Studies and Applications, III, Lecture Notes in Mathematics, vol. 1334, pp. 173–179. Springer (1988)
27. Davydov, A.A.: Whitney umbrella and slow-motion bifurcations of relaxation-type equations. J. Math. Sci. **126**, 1251–1258 (2005)
28. Liao, S.: Notes on the homotopy analysis method: some definitions and theorems. Commun. Nonlinear Sci. Numer. Simul. **14**, 983–997 (2009)

29. Odibat, Z.M.: A study on the convergence of homotopy analysis method. Appl. Math. Comput. **217**, 782–789 (2010)
30. Turkyilmazoglu, M.: Some issues on HPM and HAM methods: a convergence scheme. Math. Comput. Model. **53**, 1929–1936 (2011)
31. Abbasbandy, S., Hashemi, M.S., Hashim, I.: On convergence of homotopy analysis method and its application to fractional integro-differential equations. Quaest. Math. **36**, 93–105 (2013)
32. Hetmaniok, E., Nowak, I., Słota, D., Wituła, R.: Convergence and error estimation of homotopy analysis method for some type of nonlinear and linear integral equations. J. Numer. Math. **23**, 331–344 (2015)
33. Marinca, V., Herisanu, N.: Optimal homotopy asymptotic method. In: The Optimal Homotopy Asymptotic Method, pp. 9–22. Springer International Publishing (2015)
34. Turkyilmazoglu, M.: An effective approach for evaluation of the optimal convergence control parameter in the homotopy analysis method. Filomat **30**, 1633–1650 (2016)
35. Wang, Z., Zou, L., Qin, Y.: Piecewise homotopy analysis method and convergence analysis for formally well-posed initial value problems. Numer. Algorithms **76**, 393–411 (2017)
36. Corless, R.M., Gonnet, G.H., Hare, D.E., Jeffrey, D.J., Knuth, D.E.: On the LambertW function. Advan. Comput. Math. **5**(1), 329–359 (1996)
37. Sweet, E., Van Gorder, R.A.: Exponential type solutions to a generalized Drinfel'd - Sokolov equation. Physica Scripta **82**, 035006 (2010)