NUMERICAL METHODS FOR SOLVING ORDINARY DIFFERENTIAL EQUATIONS

MARIA CAMERON

Contents

1. Introduction

Ref.: John Strain, Lecture 1.

1.1. **Why do we need to study ODE solvers?**

(1) Most ODEs except for very special cases, do not have analytical solutions and need to be solved numerically.
(2) A number of commonly used solvers are implemented in high-level languages such as Matlab and Python,
    - Matlab: https://www.mathworks.com/help/matlab/math/choose-an-ode-solver.html,
    - Python, SciPy library: https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.solve_ivp.html.

    Nonetheless, there are problems where these methods are not suitable. For example, such problems are those where certain first integrals (e.g. energy, angular momentum, etc.) need to be conserved.
(3) The built-in ODE solvers contain a number of useful features such as error control, dense output, etc. We will learn what it is and how to use it.
(4) In some cases, we would like to accomplish a more elaborate task than built-in tools allow us. Then we need to be able to implement an ODE solver ourselves.
(5) The theory for ODE solvers is very enlightening. Its study will allow us to understand why methods might fail to compute a solution accurately and find a remedy for it.
(6) ODE solvers are building blocks for PDE solvers and SDE solvers.

1.2. **What will we study?** Keywords:

- Basic theory for ODE problems: *well-posedness: existence, uniqueness, stability with respect to small perturbations*;
- Basic concepts for ODE solvers: *consistency, stability, convergence, order of the method*;
- Stiff problems;
- Linear stability theory;
- Runge-Kutta methods;
- Linear Multistep methods;
- Extra features: *error control, dense output*;
- Symplectic ODE solvers.

1.3. **Types of ODE problems.** The following types of problems involving ODEs are typically considered:

- **Initial Value Problem (IVP)**, $y' = f(t,y), \quad y(t_0) = y_0$;
- **Boundary Value Problem (BVP)**, e.g. $y' = f(t,y), \quad y^1(t_0) = y_{1,0}, \quad y^1(t_1) = y_{1,1}$, where y^1 is the first component of y;
- **Optimal control problem**: $y = f(t,y,u)$, $y(0) = y_0$, and there is either the target destination $y(t_1) = y_1$ or a cost functional that needs to be minimized;
- **Inverse problem**: given a collection of solutions $y(t)$, one needs to identify the right-hand side function $f(t,y)$.

In this course, we will focus on IVP. Find the solution to a given ODE with a given initial condition:

$$y' = f(t,y), \quad y(t_0) = y_0, \tag{1}$$

on the time interval $[t_0, T]$, where $y : [t_0, T] \to \mathbb{R}^d$, $f : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$. The examples below show that the solution may fail to exist throughout $[t_0, T]$, or can be nonunique.

Example 1.

$$y' = y^2, \quad y(t_0) = y_0. \tag{2}$$

The solution is obtained by the following sequence of steps:

$$\frac{dy}{y^2} = dt, \quad -\frac{1}{y} = t - C, \quad -\frac{1}{y_0} = t_0 - C, \quad y = \frac{y_0}{1 + y_0(t_0 - t)}.$$

The solution blows up at time $t^* = t_0 + 1/y_0$. Thus, in this example, the solution to IVP (2) exists, it is unique, but it blows up at a finite time. Hence, we must choose $T < t^*$.

Example 2.

$$y' = 2\sqrt{y}, \quad y(0) = 0. \tag{3}$$

One obvious solution is $y(t) \equiv 0$. *The other solution,* $y(t) = t^2$, *is obtained by the following sequence of steps:*

$$\frac{dy}{2\sqrt{y}} = dt, \quad \sqrt{y} = t + C, \quad 0 = C, \quad y = t^2.$$

Furthermore, there is a one-parameter family of solutions

$$y(t) = \begin{cases} 0, & 0 \le t \le t_0, \\ (t - t_0)^2, & t \ge t_0, \end{cases},$$

for any $0 \le t_0 < \infty$. *Thus, in this example, the solution to IVP* (3) *exists for* $0 \le T < \infty$ *but it is not unique.*

Example 3. *Let us consider an IVP for a linear ODE:*

$$y' = Ay + g(t), \quad y(0) = y_0, \tag{4}$$

where $A \in \mathbb{R}^{d\times d}$ *is a constant matrix and* $g : \mathbb{R} \to \mathbb{R}^d$. *One can check directly that the solution to IVP* (4) *is given by*

$$y(t) = e^{tA}y_0 + \int_0^t e^{(t-s)A}g(s)ds, \tag{5}$$

where the matrix exponential e^{tA} *is defined as the fundamental solution matrix to*

$$\Psi' = A\Psi, \quad \Psi(0) = I_{d\times d}, \tag{6}$$

or the sum of the infinite series

$$e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}. \tag{7}$$

If A *is diagonalizable, i.e.,* $A = SDS^{-1}$, *then* $e^{tA} = Se^{tD}S^{-1}$.

In this example, the solution is unique, exists at all times, and is given by an explicit formula. However, this formula is not convenient for numerical evaluation unless $g \equiv 0$, d *is small, and* A *is diagonalizable. Therefore, it is still more convenient to compute the solution numerically for all practical purposes.*

1.4. **Basic theory for IVP.**

Theorem 1. *Consider IVP* (1). *Suppose* f *is a continuous function of* t *and* y *defined on the cylinder*

$$Q := \{t_0 \le t \le T, \|y - y_0\| \le r\}$$

and $\|f\| \le M$ *on* Q. *Then the IVP*

(1) *Then IVP* (1) *has a solution which exists for* $0 \le t - t_0 \le \min\left(\frac{r}{M}, T - t_0\right)$.

(2) *If, in addition, f is Lipschitz in y, i.e., for some constant L,*

$$\|f(t,y_1) - f(t,y_2)\| \le L\|y_1 - y_2\|,$$

then the solution is unique.

(3) *If in addition the Jacobian matrix*

$$Df(t,y) = \left\{\frac{\partial f^i}{\partial y^j}\right\}_{i,j=1}^{d}$$

is continuous in Q then the solution y is differentiable with respect to the initial condition y_0.

(4) *Suppose f also depends on parameters $u \in \mathbb{R}^m$ and the Jacobian matrix with respect to u,*

$$D_u f = \left\{\frac{\partial f^i}{\partial u^j}\right\}, \quad i = 1,\dots,d, \; j = 1,\dots,m,$$

exists and is continuous on Q for all u. Then the solution y is differentiable with respect to u.

Exercise 1. *Apply Theorem 1, part 1, to Example 1: fix $r \in \mathbb{R}$, define the cylinder Q, express M via r, and find the interval of time where the existence of the solution is guaranteed.*

Exercise 2. *Apply Theorem 1, parts 1 and 2, to Example 2. As in the previous exercise, find the interval where the existence of the solution is guaranteed. Then check that the function $f(y) = 2\sqrt{y}$ is not Lipschitz at $y = 0$.*

1.5. **Integral equations.** IVP (1) is equivalent to the integral equation

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))ds. \tag{8}$$

Integral equation (8) is used to construct various numerical methods including Runge-Kutta, Adams and BDF. Moreover, integral equation (8) gives a useful tool for proving Theorem 1 called the *Picard iteration*

$$y_{n+1}(t) = y_0 + \int_{t_0}^{t} f(t, y_n(s))ds. \tag{9}$$

When f is Lipschitz, it is easy to show that the Picard iterates converge uniformly to a unique continuous solution of the IVP on the interval of t specified in Theorem 1.

Exercise 3. *Compute the first few Picard iterates starting from $y(t) = 1$ for the IVP*

$$y' = y^2, \quad y(0) = 1. \tag{10}$$

Infer a general pattern and determine the interval on which the Picard iterates converge. Does it match the interval guaranteed by Theorem 1?

1.6. **Textbooks on ODE theory.**

- Witold Hurewitz, Lectures on Ordinary Differential Equations, first published in 1958 and then replicated multiple times. This is a very nice, concise and fun to read set of lectures on ODE covering a very good part of the ODE theory.
- Carmen Chicone, Ordinary Differential Equations with Applications, Springer, 1999. This is a nicely written book covering the ODE theory and its numerous applications. It is freely available online.
- E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, 1955. This is a large textbook on the ODE theory. You can look into it if you cannot find something in the two textbooks above.

2. Construction of ODE solvers

Ref: John Strain, Lecture 2, section 2. Starting from now, we will denote the exact solution by y and the numerical solution by u. The time step will be denoted by h. In addition, we will use short-hand notation replacing the time argument t_n with the subscript n: $u_n \equiv u(t_n)$, $f_n \equiv f(t_n, u_n)$, etc.

2.1. **Approach 1: Taylor expansion.** A Taylor expansion of the solution $y(t)$ at t_n yields:

$$y(t_n + h) = y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(t_n) + \dots. \tag{11}$$

Truncating this expansion at the first-order term in h we obtain the **forward Euler** method:

$$\boxed{u_{n+1} = u_n + hf(t_n, u_n).} \tag{12}$$

Do not use the forward Euler method unless there is no other choice. It is very inaccurate!

Truncating the series in (11) at the second-order term we obtain the following *second order Taylor method* which is hardly ever used:

$$\boxed{u_{n+1} = u_n + hf(t_n, u_n) + \frac{h^2}{2}\left[\frac{\partial f}{\partial t}(t_n, u_n) + [Df(t_n, u_n)]f(t_n, u_n)\right].} \tag{13}$$

While this method is reasonably accurate (commonly used PDE solvers are approximately as accurate as this method), it is not popular because it requires additional input: the time derivative and the Jacobian matrix for the right-hand side function f. There are other second-order accurate options that do not require such input.

2.2. **Integral equation approach.** Let us consider the ansatz

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(s, y(s))ds. \tag{14}$$

We will approximate the integral using various quadrature rules. Applying the *left-hand rule* we obtain the **forward Euler** method (15). The *right-hand rule* results in the **backward Euler** method:

$$\boxed{u_{n+1} = u_n + hf(t_n + 1, u_n + 1).} \tag{15}$$

While backward Euler is as inaccurate as forward Euler, backward Euler can be a reasonable choice in some cases due to its very good stability properties that we will discuss later.

The *trapezoidal rule* with u_{n+1} unknown in the right-hand side yields the method also called the **implicit trapezoidal rule**:

$$\boxed{u_{n+1} = u_n + \frac{h}{2}\left[f(t_n, u_n) + f(t_{n+1}, u_{n+1})\right].} \tag{16}$$

The implementation of this method requires the use of a nonlinear solver for u_{n+1}. This method is often used in solving PDEs. Another option is to predict u_{n+1} in the right-hand side of (16) using forward Euler. The resulting method belongs to the Runge-Kutta family and is called the **trapezoidal rule with Euler predictor**:

$$\boxed{u_{n+1} = u_n + \frac{h}{2}\left[f(t_n, u_n) + f(t_{n+1}, u_n + hf(t_n, u_n))\right].} \tag{17}$$

This method is not commonly used because contrary to the implicit trapezoidal rule, its stability properties are not good enough, and there are other much more accurate methods with stability properties similar to (17).

The *midpoint rule* gives the ansatz $u_{n+1} = u_n + hf(t_n + h/2, u_{n+1/2})$ where $u_{n+1/2}$ needs to be defined. One option is to approximate it as the half-sum of u_n and u_{n+1} resulting in the **implicit midpoint rule**:

$$\boxed{u_{n+1} = u_n + hf\left(t_n + \frac{h}{2}, \frac{u_n + u_{n+1}}{2}\right).} \tag{18}$$

Another option resulting in a different version of the implicit midpoint rule is to approximate $k \equiv f(t_n + h/2, u_{n+1/2})$ from the relationship $k = f(t_n + h/2, u_n + (h/2)k)$:

$$\boxed{k = f\left(t_n + \frac{h}{2}, u_n + \frac{h}{2}k\right), \quad u_{n+1} = u_n + hk.} \tag{19}$$

We will see later that method (19) is a good choice for integrating Hamiltonian systems because it is *symplectic*. In particular, it conserves energy.

Another option is to predict $u_{n+1/2}$ using forward Euler. The resulting method, the **midpoint rule with Euler predictor**

$$\boxed{u_{n+1} = u_n + hf\left(t_n + \frac{h}{2}, u_n + \frac{h}{2}f(t_n, u_n)\right),} \tag{20}$$

has properties similar to the trapezoidal rule with Euler predictor and hence is not used much.

A famous and commonly used method based on *Simpson's quadrature* rule

$$\int_{t_n}^{t_{n+1}} g(t)dt \approx \frac{h}{6}\left(g(t_n) + 4g(t_{n+1/2}) + g(t_{n+1})\right) \tag{21}$$

is the **four-stage fourth-order Runge Kutta method** (RK4).

$$k_1 = f(t_n, u_n), \tag{22}$$

$$k_2 = f\left(t_n + \frac{h}{2}, u_n + \frac{h}{2}k_1\right), \tag{23}$$

$$k_3 = f\left(t_n + \frac{h}{2}, u_n + \frac{h}{2}k_1\right), \tag{24}$$

$$k_4 = f(t_n + h, u_n + hk_3), \tag{25}$$

$$u_{n+1} = \frac{h}{6}\left(k_1 + 2(k_2 + k_3) + k_4\right). \tag{26}$$

2.3. **Polynomial interpolation approach.** Adams and BDF (backward differentiation formula) families of methods are constructed with the aid of polynomial interpolation using only values at grid point. These methods are often designed to have variable time step and variable order.

The **Adams-Bashforth** family is constructed by writing an interpolating polynomial through f_{n-k+1}, f_{n-k+2}, $\ldots$, f_{n-1}, and f_n, integrating this polynomial and evaluating its value at t_{n+1}. As you see, it involves extrapolation. On the other hand, it is a family of explicit methods, hence time-stepping is relatively cheap. The **Adams-Moulton** family is constructed in a similar manner except for the interpolating polynomial also uses the unknown value f_{n+1} making the methods implicit. Hence time marching requires the use of a nonlinear solver. At the same time, the Adams-Moulton methods are more accurate and have better stability properties than Adams-Bashforth.

The **BDF** family is obtained by writing an interpolating polynomial through u_{n-k+1}, $\ldots$, u_n, and u_{n+1}, differentiating it, and matching its derivative at t_{n+1} with f_{n+1}.

2.4. **Undetermined coefficients.** The method of undetermined coefficients is a very general approach for constructing ODE solvers. It works as follows. First, we choose the general form of the method that contains a number of undetermined coefficients. Second, we apply consistency and stability constraints and solve for the undetermined coefficients. This method will be used to construct the Runge-Kutta methods and will be discussed in more detail later.

3. Consistency, stability, and convergence

Consistency and stability imply convergence. This easy-to-remember claim is true for ODE (and PDE) solvers. In this section, we will define these three concepts and prove the convergence theorem for first-order methods.

3.1. **Standard assumptions about solvers for IVP.** We will start with specifying the standard assumptions about a general ODE solver. Any ODE solver, i.e., a method for solving IVP $y' = f(t,y)$, $y(t_0) = y_0$, $t \in [t_0, T]$, can be written of the form

$$u_{n+1} + a_0 u_n + \ldots + a_{k-1} u_{n-k+1} = hF(u_{n+1}, u_n, \ldots, u_{n-k+1}; t_{n+1}, f, h). \tag{27}$$

We will make the following assumption about F.

Assumption 1. *F vanishes identically whenever f does.*

Assumption 2. *F is a Lipschitz function of all u-arguments whenever f is Lipschitz. Specifically, if f is Lipschitz with respect to y, there is a constant L such that*

$$\begin{aligned} &\|F(u_{n+1}, u_n, \ldots, u_{n-k+1}; t_{n+1}, f, h) - F(v_{n+1}, v_n, \ldots, v_{n-k+1}; t_{n+1}, f, h)\| \le \\ &L\left(\|u_{n+1} - v_{n+1}\| + \|u_n - v_n\| + \ldots \|u_{n-k+1} - v_{n-k+1}\|\right). \end{aligned} \tag{28}$$

3.2. **Classification of methods for IVP.** Look at the general IVP solver given by (27).

- If $k = 1$, the method is *one-step*, otherwise it is *multi-step*.
- If F does not depend on u_{n+1}, the method in *explicit*, otherwise it is *implicit*.
- If F is a linear function of the right-hand side f, the method is *linear*, otherwise, it is *nonlinear*.

3.3. **Definition of convergence.** Starting from this section, we will shift time so that the starting time is zero. In this case, the total number of steps N and the time step h satisfy a nice relationship $Nh = T$.

Definition 1. *A method for IVP converges to the true solution if for all sufficiently smooth right-hand sides f, the numerical solution u_n satisfies*

$$\max_{0\le t_n\le T}\|u_n - y(t_n)\| \to 0 \quad \text{as} \quad h\to 0 \quad \text{and} \quad u_j \to y_j,\ 0\le j\le k-1. \tag{29}$$

The method is accurate of order p if

$$\max_{0\le t_n\le T}\|u_n - y(t_n)\| = O(h^p) + O(\|u_0-y_0\|) + \ldots + O(\|u_{k-1}-y_{k-1}\|) \tag{30}$$

as $h\to 0$ and the initial values converge.

3.4. **Consistency.**

Definition 2. *The local truncation error τ of method (27) applied to a smooth solution y of an IVP $y' = f(t,y)$, $y(0) = y_0$, is the error committed in one step starting from exact values:*

$$\tau_{n+1} = y_{n+1} + a_0 y_n + \ldots + a_{k-1}y_{n-k+1} - hF(y_{n+1}, y_n, \ldots, y_{n-k+1}; t_{n+1}, f, h). \tag{31}$$

Definition 3. *Method (27) is consistent of order p if*

$$\tau_{n+1} = O(h^{p+1}) \quad \forall 0\le n\le T/h \tag{32}$$

Note that consistency does not imply convergence. Consistency only tells us about the error committed over one step starting from the exact values.

Exercise 4. *Show that methods (16), (17), (18), (19), and (20) are consistent of order 2.*

Let us show that the trapezoidal rule with Euler predictor (17) is consistent of order 2. We need to plug in the exact solution to the ODE into the method and Taylor-expand all terms at t_n. We note that $f(t_n, y_n) = y'_n$ and

$$\begin{aligned} f(t_{n+1}, y_n + hf(t_n,y_n)) &= f(t_n,y_n) + h\frac{\partial f}{\partial t} + [Df(t_n,y_n)]f(t_n,y_n)h + O(h^2) \\ &= y'_n + hy''_n + O(h^2). \end{aligned} \tag{33}$$

Here $Df(t_n,y_n)$ is the Jacobian matrix consisting of the partial derivatives of the components of f with respect to components of y evaluated at (t_n,y_n). For brevity, we will omit subscript n. Plugging in y and (33) into (17) and Taylor-expanding y_{n+1} we obtain:

$$y + hy' + \frac{h^2}{2}y'' - y - \frac{h}{2}y' - \frac{h}{2}y' - \frac{h^2}{2}y'' = O(h^3).$$

Note that we did not check that the coefficient at h^3 does not vanish. In order to do it, we would need to write out more expansion terms in (33) which is quite tedious. We will discuss how to simplify this process later in this chapter.

Exercise 5. *Show that the three-step Adams-Bashforth method*

$$u_{n+1} = u_n + h\left(\frac{23}{12}f(t_n, u_n) - \frac{4}{3}f(t_{n-1}u_{n-1}) + \frac{5}{12}f(t_{n-2}, u_{n-2})\right) \tag{34}$$

is consistent of order 3.

Exercise 6. *Show that the three-step Adams-Moulton method*

$$u_{n+1} = u_n + h\left(\frac{9}{24}f(t_{n+1}, u_{n+1}) + \frac{19}{24}f(t_n, u_n) - \frac{5}{24}f(t_{n-1}u_{n-1}) + \frac{1}{24}f(t_{n-2}, u_{n-2})\right) \tag{35}$$

is consistent of order 4.

Exercise 7. *What is the order of consistency of the following explicit two-step method?*

$$u_{n+1} + 9u_n - 10u_{n-1} = \frac{h}{2}\left(13f(t_n, u_n) + 9f(t_{n-1}u_{n-1})\right) \tag{36}$$

3.5. **Stability.** A consistent method might not be convergent because of the way numerical errors accumulate over time steps. For example, consider the method (36). It is consistent of order 2. However, it is unstable. An easy way to see it is to apply if for solving the IVP $y' = 0$, $y(0) = a$. To initiate the method, we need two initial values $u(0) = u_0$ and $u(h) = u_1$. Since the right-hand side is zero, the method becomes the following linear recurrent relationship:

$$u_{n+1} + 9u_n - 10u_{n-1} = 0. \tag{37}$$

The general solution to (37) is $u_n = Ar_1^n + Br_2^n$ where r_1 and r_2 are the roots to the characteristic equation $r^2 + 9r - 10 = 0$, i.e., $r_1 = 1$, $r_2 = -10$. In order to obtain the constant solution $u_n = a$, we need $u_1 = u_2 = a$. Then $A = a$ and $B = 0$. If either of these values u_1 or u_2, will be slightly perturbed, the coefficient B will be nonzero and hence the solution will blow up. Note that the smaller the time step h will be, the more the solution will blow up over a fixed interval of time.

This example shows that besides requiring that the errors committed at each time step are small, they also need to accumulate in a stable manner.

Definition 4. *An IVP method is stable if and only if the numerical solution is Lipschitz with respect to perturbations of the initial values and the right-hand side. I.e., there is a Lipschitz constant S such that for any sequence of vectors δ_n, the solution v_n of the perturbed method*

$$v_{n+1} + a_0 v_n + \ldots + a_{k-1}v_{n-k+1} = hF(v_{n+1}, u_n, \ldots, v_{n-k+1}; t_{n+1}, f, h) + h\delta_{n+1} \tag{38}$$

with initial values $v_0 = u_0 + \delta_0$, ..., $v_{k-1} = u_{k-1} + \delta_{k-1}$ satisfies

$$\max_{t_0 \le t_n \le T} \|u_n - v_n\| \le S \max_{t_0 \le t_n \le T} \|\delta_n\|. \tag{39}$$

3.6. Convergence of one-step methods.

Theorem 2. *A one-step method*

$$u_{n+1} = u_n + F(u_{n+1}, u_n; t_{n+1}, h, f) \tag{40}$$

which is consistent of order p converges with order p accuracy for any IVP where f is sufficiently smooth (i.e., $f \in C^{p+1}$ in an appropriate cylinder).

This theorem implies that *a one-step method is always stable.*

Proof. The local truncation error at step $n+1$ is given by

$$\tau_{n+1} = y_{n+1} - y_n + F(y_{n+1}, y_n; t_{n+1}, h, f). \tag{41}$$

For the numerical solution u we have:

$$0 = u_{n+1} - u_n + F(u_{n+1}, u_n; t_{n+1}, h, f). \tag{42}$$

Subtracting (42) from (41) we obtain

$$\tau_{n+1} = [y_{n+1} - u_{n+1}] - [y_n - u_n] + [F(y_{n+1}, y_n; t_{n+1}, h, f) - F(u_{n+1}, u_n; t_{n+1}, h, f)]. \tag{43}$$

Introducing the notation $e_n = \|y_n - u_n\|$ and $\tau = \max_n \|\tau_n\|$, and using the fact that F is Lipschitz with respect to all u-arguments by Assumption 2, we obtain

$$e_{n+1} \le e_n + hL(e_n + e_{n+1}) + \tau. \tag{44}$$

Equation (44) can be rewritten as

$$e_{n+1} \le e_n \frac{1+hL}{1-hL} + \frac{\tau}{1-hL}. \tag{45}$$

We choose h small enough so that the coefficients in (45) are positive. By assumption, $\tau = O(h^{p+1})$. To obtain a bound for e_{n+1}, we consider the corresponding linear inhomogeneous recurrence relationship

$$x_{n+1} = a x_n + b, \quad \text{where} \quad a = \frac{1+hL}{1-hL}, \quad b = \frac{\tau}{1-hL}. \tag{46}$$

The solution to (46) is

$$x_n = a^n x_0 + \frac{a^n - 1}{a-1} b. \tag{47}$$

Since the coefficients a and b are positive, we obtain the following bound on e_n:

$$e_n \le \left(\frac{1+hL}{1-hL}\right)^n e_0 + \frac{\left(\frac{1+hL}{1-hL}\right)^n - 1}{\left(\frac{1+hL}{1-hL}\right) - 1} \frac{\tau}{1-hL}. \tag{48}$$

It is convenient to simplify (48) using the following bound:

$$a = \frac{1+hL}{1-hL} \le (1+hL)[1 + hL + (hL)^2 + (hL)^3 + \dots] \le 1 + 3hL \le e^{3hL} \quad \text{for} \quad hL < \frac{1}{3}. \tag{49}$$

Then the bound (48) can be simplified to

$$e_n \leq e^{3Lhn} e_0 + \frac{e^{3Lhn} - 1}{1 + 3hL - 1} \frac{\tau}{1 - hL}. \tag{50}$$

We observe that the denominator under τ is less than 1 and that $nh \leq hN = T$, where N is the total number of time steps. Hence, (51) becomes

$$e_n \leq e^{3LT} e_0 + \frac{e^{3LT} - 1}{3L} \frac{\tau}{h}. \tag{51}$$

Equation (51) shows that the error decays as $O(h^p) + O(e_0)$ as $h \to 0$ uniformly on the interval $[0, T]$. Hence, the method converges.

□