MAE 290B. Homework 1 Solution

Winter 2018

Problem 1

a. Central scheme

To obtain a central approximation to f'_j let us begin by writing the Taylor expansions of f_{j+1} and f_{j-1}

$$f_{j+1} = f_j + f'_j h + \frac{1}{2} f''_j h^2 + \frac{1}{6} f'''_j h^3 + ... \tag{1}$$

$$f_{j-1} = f_j - f'_j h + \frac{1}{2} f''_j h^2 - \frac{1}{6} f'''_j h^3 + ... , \tag{2}$$

substracting Eq.(1) - Eq.(2) we get

$$f_{j+1} - f_{j-1} = 2 f'_j h + f''_j h^2 + ... , \tag{3}$$

and solving for f'_j we finally obtain

$$f'_j = \frac{f_{j+1} - f_{j-1}}{2h} + O(h^2) . \tag{4}$$

Note that this is equivalent to writing $f'_j = a f_{j+1} + b f_j + c f_{j-1}$, expanding in a Taylor series and solving the system for a, b, c after cancelling all the terms but f'_j.

b. Padé scheme

The Padé approximation of f'_j uses f'_{j+1},f'_{j-1},f_j,f_{j+1} and f_{j-1}. We can then write

$$a f'_{j+1} + b f'_{j-1} + c f_{j+1} + d f_j + e f_{j-1} - f'_j = \varepsilon , \tag{5}$$

where ε is the error of the scheme. Expanding the first derivatives in a power series around f'_j and the function values in a power series around f_j, we can fill a Taylor table as given by Table 1.

To minimize ε we have to make zero as many low-order terms as possible, i.e all the possible columns starting from the left of Table 1. Having five coefficients we get five equations:

$$c + d + e = 0 \tag{6}$$

$$a + b + ch - eh - 1 = 0 \tag{7}$$

$$ah - bh + \frac{ch^2}{2} + \frac{eh^2}{2} = 0 \tag{8}$$

$$\frac{ah^2}{2} + \frac{bh^2}{2} + \frac{ch^3}{6} - \frac{eh^3}{6} = 0 \tag{9}$$

$$\frac{ah^3}{6} - \frac{bh^3}{6} + \frac{ch^4}{24} + \frac{eh^4}{24} = 0 , \tag{10}$$

(11)

Table 1: Taylor table

	f_j	f'_j	f''_j	f'''_j	f^{iv}_j	f^{v}_j
af'_{j+1}	0	a	ah	$\frac{ah^2}{2}$	$\frac{ah^3}{6}$	$\frac{ah^4}{24}$
bf'_{j-1}	0	b	$-bh$	$\frac{bh^2}{2}$	$-\frac{bh^3}{6}$	$\frac{bh^4}{24}$
cf_{j+1}	c	ch	$\frac{ch^2}{2}$	$\frac{ch^3}{6}$	$\frac{ch^4}{24}$	$\frac{ch^5}{120}$
df_j	d	0	0	0	0	0
ef_{j-1}	e	$-eh$	$\frac{eh^2}{2}$	$-\frac{eh^3}{6}$	$\frac{eh^4}{24}$	$-\frac{eh^5}{120}$
$-f'_j$	0	-1	0	0	0	0

that can be simplified and written in matrix form as

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & -1 \\ 1 & -1 & 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 1/6 & 0 & -1/6 \\ 1/6 & -1/6 & 1/24 & 0 & 1/24 \end{bmatrix} \begin{bmatrix} a \\ b \\ ch \\ d \\ eh \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} .$$

Solving the system we obtain the coefficients that minimize the error $a = b = -\frac{1}{4}$, $c = \frac{3}{4h}$, $d = 0$, $e = -\frac{3}{4h}$. Finally, we obtain

$$f'_j = -\frac{(f'_{j+1} + f'_{j-1})}{4} + \frac{3(f_{j+1} - f_{j-1})}{4h} + O(h^4) \; . \tag{12}$$

c. Modified wavenumber analysis

Let us consider the harmonic function and its exact derivative

$$f(x) = e^{ikx} \; , \tag{13}$$

$$f'(x) = ike^{ikx} \; . \tag{14}$$

Recall that $x_j = \frac{L}{N} = hj$. If we apply the central scheme of Eq.(4) we obtain

$$f'_j = \frac{f_{j+1} - f_{j-1}}{2h} = \frac{e^{ikh(j+1)} - e^{ikh(j-1)}}{2h} = e^{ikhj}\left(\frac{e^{ikh} - e^{-ikh}}{2h}\right) = i\frac{\sin(kh)}{h}e^{ikhj} \; . \tag{15}$$

Comparing Eq.(15) and Eq.(14), we can define a modified wave number $k' = \frac{\sin(kh)}{h}$.

We can follow an analogous procedure to study the accuracy of the Padé scheme.

$$\begin{aligned} f'_j &= -\frac{(f'_{j+1} + f'_{j-1})}{4} + \frac{3(f_{j+1} - f_{j-1})}{4h} = -\frac{ik'}{4}(e^{ikh(j+1)} + e^{ikh(j-1)}) + \frac{3}{4h}(e^{ikh(j+1)} - e^{ikh(j-1)}) = \\ &= i\left(-\frac{k'}{4}2\cos kh + \frac{6}{4h}\sin kh\right)e^{ikhj} \; . \end{aligned} \tag{16}$$

At this point we know that $k' = -\frac{k'}{2}\cos kh + \frac{3}{2h}\sin kh$, which implies

$$k' = \left(\frac{1}{1 + \frac{1}{2}\cos kh}\right)\frac{3}{2h}\sin kh \; . \tag{17}$$

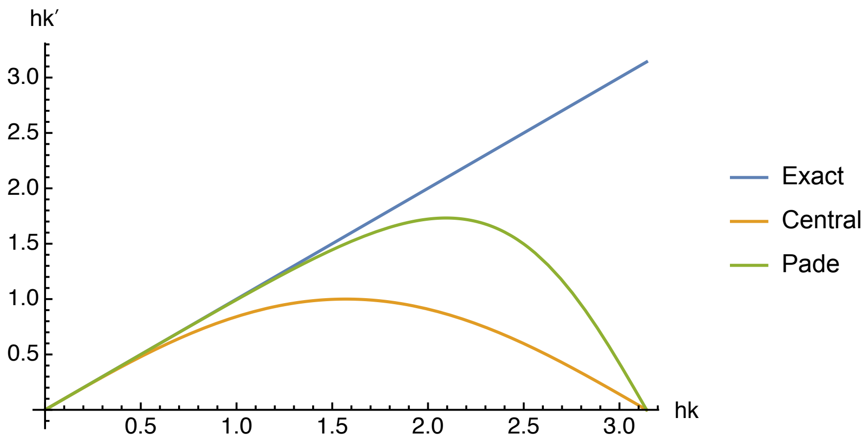

Figure 1: Accuracy of the FD scheme using modified wave number analysis. The central scheme is accurate (deviation from the ideal behavior is not visible on the figure) for approximately $kh < 0.5$ whereas the Padé scheme is accurate for $kh < 1.5$.

Problem 2

a. Stability of the θ method

Let us start the stability analysis by computing the amplification factor of the method when applied to the model problem $y' = \lambda y$ as follows.

$$\begin{aligned} y_{n+1} &= y_n + h[\theta f_{n+1} + (1-\theta) f_n] \\ y_{n+1} &= y_n + h[\theta \lambda y_{n+1} + (1-\theta)\lambda y_n] \\ y_{n+1} &= \frac{1+(1-\theta)\lambda h}{1-\theta\lambda h} y_n \\ y_{n+1} &= \left[\frac{1+(1-\theta)\lambda h}{1-\theta\lambda h}\right]^n y_0 \end{aligned} \tag{18}$$

$$\sigma \triangleq \frac{1+(1-\theta)\lambda h}{1-\theta\lambda h} \tag{19}$$

To ensure that the numerical method is stable, the modulus of the amplification factor has to be less than one, $|\sigma| \leq 1$.

- $\theta = 0$ Explicit Euler

$$\sigma = (1+\lambda h) \tag{20}$$

$$|\sigma| = |1+\lambda h| \leq 1 \tag{21}$$

$$(1+\lambda_R h)^2 + (\lambda_I h)^2 \leq 1 \tag{22}$$

which is the equation of a circle of radius one centered at $(-1, 0)$.

- $\theta = \frac{1}{2}$ Crank-Nicholson

$$\sigma = \frac{1+\frac{\lambda h}{2}}{1-\frac{\lambda h}{2}} \tag{23}$$

$$|\sigma| = \frac{\sqrt{\left(1+\frac{\lambda_R h}{2}\right)^2 + \left(\frac{\lambda_I h}{2}\right)^2}}{\sqrt{\left(1-\frac{\lambda_R h}{2}\right)^2 + \left(\frac{\lambda_I h}{2}\right)^2}} \leq 1 \tag{24}$$

which always holds for $\lambda_R \leq 0$.

- $\theta = 1$ Implicit Euler

$$\sigma = \frac{1}{1-\lambda h} \tag{25}$$

$$|\sigma| = \left|\frac{1}{1-\lambda h}\right| \leq 1 \tag{26}$$

$$(1-\lambda_R h)^2 + (\lambda_I h)^2 \geq 1 \tag{27}$$

which is the outer region of a circle of radius one centered at $(1, 0)$.

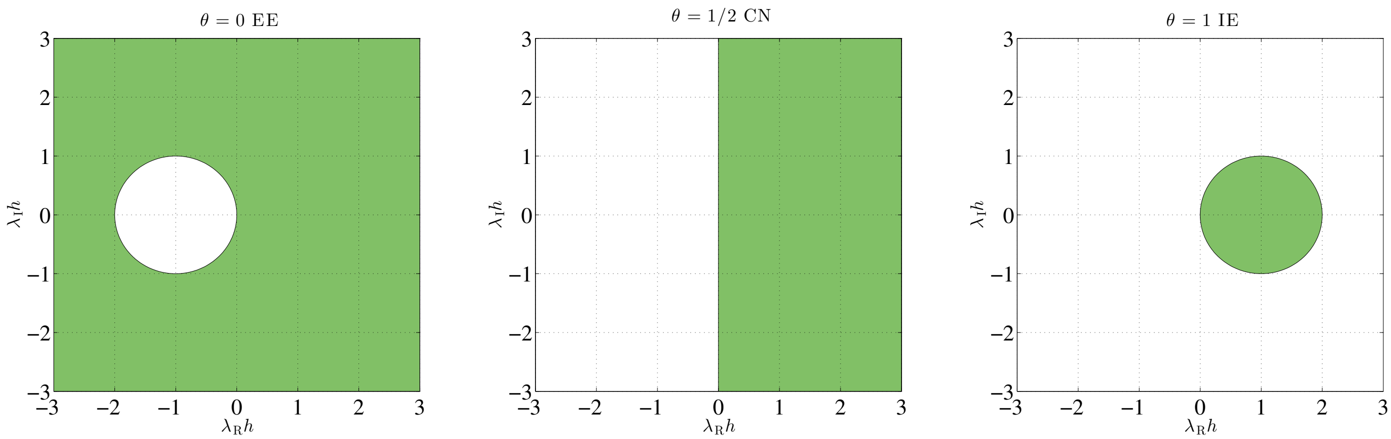

Figure 2: Stability diagram of the method for various θ. Stable regions are plotted in white.

b. Implementation

The exact solution of the IVP is $y(t) = e^{-t/\tau}(\cos(wt) + i\sin(wt))$. The implementation of the numerical solution is done in Matlab as follows.

```
%% Theta method time marching

tau = 200/0.25 ; omega = 0.25;  lambda = (-1/tau + i*omega); %physical parameters

dt=0.01; theta=3/4; T=1000; time=0:dt:T; %numerical parameters


f = @(t) exp(-t./tau).*cos(omega.*t); sol = f(time); %exact real solution

y=zeros(length(time),1);

y(1)=1; %initial condition

for ii=1:length(y)-1 %time marching
    y(ii+1) = y(ii)*(1+dt*(1-theta)*lambda)/(1-dt*theta*lambda);
end
```

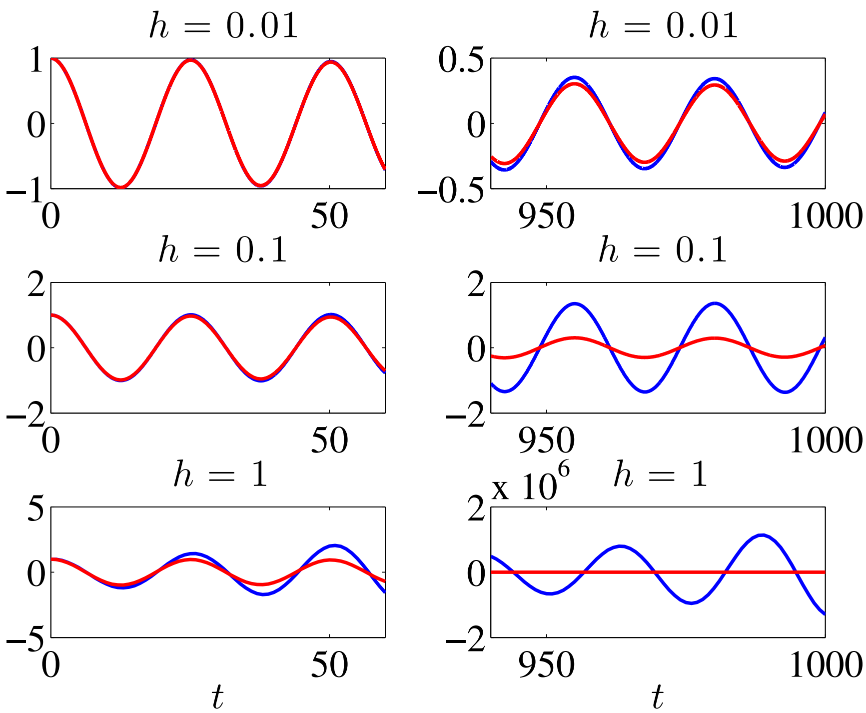

Figure 3: Comparison of the exact (red) and numerical (blue) solution for different h and $\theta = 1/4$.

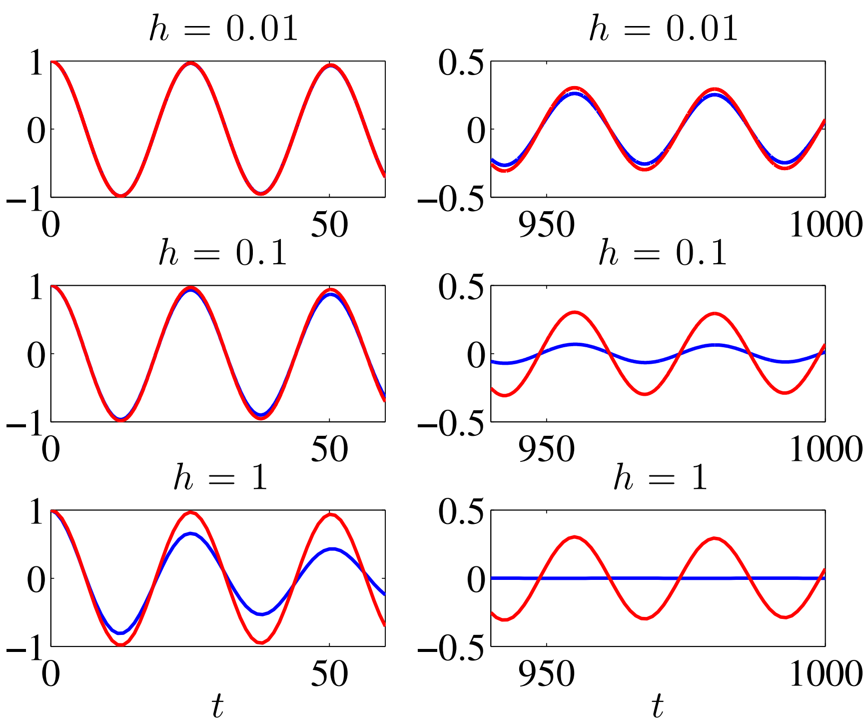

Figure 4: Comparison of the exact (red) and numerical (blue) solution for different h and $\theta = 3/4$.

Looking at the solutions Fig.(3) and Fig.(4), we clearly observe how increasing h reduces the accuracy and the convergence. Increasing the step increases both the amplitude and the phase error. The stability of the method increases with θ. The solution with $\theta = 1/4$ and $h = 1$ diverges whereas the solution with $\theta = 3/4$ and $h = 1$ remains bounded. Note that although $\theta = 3/4$ leads to a bounded solution at long time, the result is inaccurate.

c. Accuracy

After assuming that the characteristic decay time tends to infinity, the eigenvalue of the problem becomes $\lambda = iw$. We start by substituting $\lambda = iw$ and $\theta = 3/4$ in the expression for the amplification factor of our method, Eq.(19), to get

$$\sigma = \frac{1 + \frac{1}{4}iwh}{1 - \frac{3}{4}iwh} \, . \tag{28}$$

To obtain the amplification error (AE) and the phase error (PE) we have to compute the difference between the amplification and the phase of the amplification factor and the amplification and the phase of the exact solution.

When $\lambda \to \infty$ the exact solution becomes

$$y(t) = \cos(wt) + i \sin(wt) \, , \tag{29}$$

which has an oscillatory response with amplitude one and phase wh[1].

To obtain the amplitude of the numerical solution, we need to compute $|\sigma|$. To obtain the phase, we need $\arg(\sigma)$.

$$|\sigma| = \left| \frac{1 + \frac{1}{4}iwh}{1 - \frac{3}{4}iwh} \right| = \frac{|1 + \frac{1}{4}iwh|}{|1 - \frac{3}{4}iwh|} = \frac{\sqrt{1 + (\frac{1}{4}wh)^2}}{\sqrt{1 + (\frac{3}{4}wh)^2}} = 1 - \frac{1}{4}(wh)^2 + O((wh)^4) \tag{30}$$

$$\begin{aligned} \arg(\sigma) = \arg\left(\frac{1 + \frac{1}{4}iwh}{1 - \frac{3}{4}iwh}\right) &= \arg\left(1 + \frac{1}{4}iwh\right) - \arg\left(1 - \frac{3}{4}iwh\right) = \arctan\left(\frac{1}{4}wh\right) - \arctan\left(-\frac{3}{4}wh\right) \\ &= wh - \frac{7}{48}(wh)^3 + O((wh)^4) \end{aligned} \tag{31}$$

Note that these simplifications use a Taylor expansion around $wh = 0$ assuming that $wh << 1$. *Wolfram Mathematica* was used to simplify the expressions.

At this point obtaining AE and PE is straightforward.

$$AE \triangleq |\sigma| - 1 = 1 - \frac{1}{4}(wh)^2 - 1 = -\frac{1}{4}(wh)^2 \tag{32}$$

$$PE \triangleq \arg(\sigma) - wh = wh - \frac{7}{48}(wh)^3 - wh = -\frac{7}{48}(wh)^3 \tag{33}$$

We observe in Fig.(4) that, as expected, the amplitude error is bigger than the phase error. The analytical results match the simulation since the numerical solution amplitude is less than the exact solution amplitude $(AE < 0)$ and the numerical solution falls behind the exact solution, i.e phase lag $(PE < 0)$.

[1] Note that the exact solution is $e^{iwt} = e^{iwnh} = [e^{iwh}]^n$ thus the phase written in terms of h instead of t is wh.

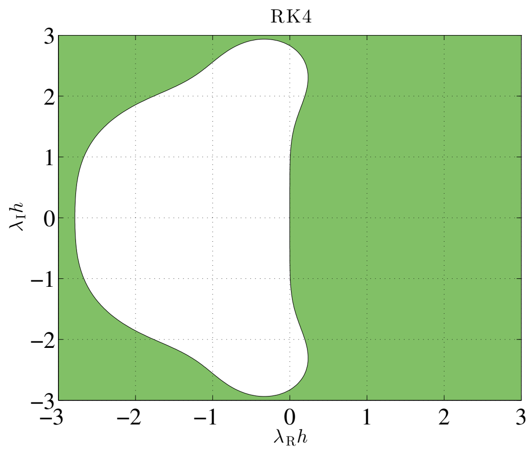

Figure 5: Stability region of RK4. Stable region plotted in white.

Problem 3

a. Stability of RK4

Since RK4 is fourth-order accurate, the amplification factor of the method when applied to the model problem, $y' = \lambda y$, has to be of the form

$$\sigma = 1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \frac{1}{6}(\lambda h)^3 + \frac{1}{24}(\lambda h)^4 \, . \tag{34}$$

To obtain the stability region we have to solve the nonlinear equation

$$1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \frac{1}{6}(\lambda h)^3 + \frac{1}{24}(\lambda h)^4 \; = e^{i\theta} \tag{35}$$

for every θ. Attached is a piece of code that computes directly the curve for which $|\sigma| = 1$. The solution is shown in Fig.(5). The stability limit for $\lambda \in \mathbb{R}$ is $h < 2.76/|\lambda|$.

```
%% Stability region of RK4

%Calculation
x=linspace(-3,3,200);
y=linspace(-3,3,200);
[X,Y] = meshgrid(x,y);
points    = X + i*Y;
sigma = abs(1+points+1/2.*points.^2+1/6.*points.^3+1/24.*points.^4);
%Plot contour with sigma = 1
contourf(X,Y,sigma,[1 1])

%Plot options
grid on
colormap summer
t=title('RK4');
%Use text formatting language Latex to get nice axis labels!
set(t,'Interpreter','Latex','fontsize',18);
l  =  xlabel('$\lambda_{\rm{R}} h $');
set(l,'Interpreter','Latex','fontsize',18);
s  =  ylabel('$\lambda_{\rm{I}} h $');
set(s,'Interpreter','Latex','fontsize',18);
```

b. Implementation

The implementation of RK4 and the solutions are shown below. Notice how much better the accuracy of RK4 is compared to the previous methods. The given choices of $h = 0.01$, 0.1 and 1 show little error. An additional case with h as large as 10 shows inaccuracy.

```
%% RK4 time marching

tau = 200/0.25
omega = 0.25
lambda = (-1/tau + i*omega)

dt=0.01
T=1000

f = @(t) exp(-t./tau).*cos(omega.*t);
time_sol=0:0.01:T;
sol = f(time_sol);

b1=1/6; b2=1/3; b3=1/3; b4=1/6;
a21=1/2; a31=0; a32=1/2;
a41=0; a42=0; a43=1;

time=0:dt:T;

y=zeros(length(time),1);

y(1)=1;

for ii=1:length(y)-1

    f1 = lambda*y(ii);
    f2 = lambda*(y(ii)+a21*dt*f1);
    f3 = lambda*(y(ii)+a31*dt*f1+a32*dt*f2);
    f4 = lambda*(y(ii)+a41*dt*f1+a42*dt*f2+a43*dt*f3);
    y(ii+1)=y(ii)+dt*(b1*f1+b2*f2+b3*f3+b4*f4);

end

%Plot
subplot(4,2,1)
h1=plot(time(1:60/dt+1),y(1:60/dt+1),time_sol(1:60/dt+1),sol(1:60/dt+1));
set(h1(2),'Color','r','LineStyle','--','LineWidth',1);
set(gca,'FontSize',10,'FontName','Times New Roman');
axis([0 60 -1 1]);
title(['$h=',num2str(dt),'$'])
```

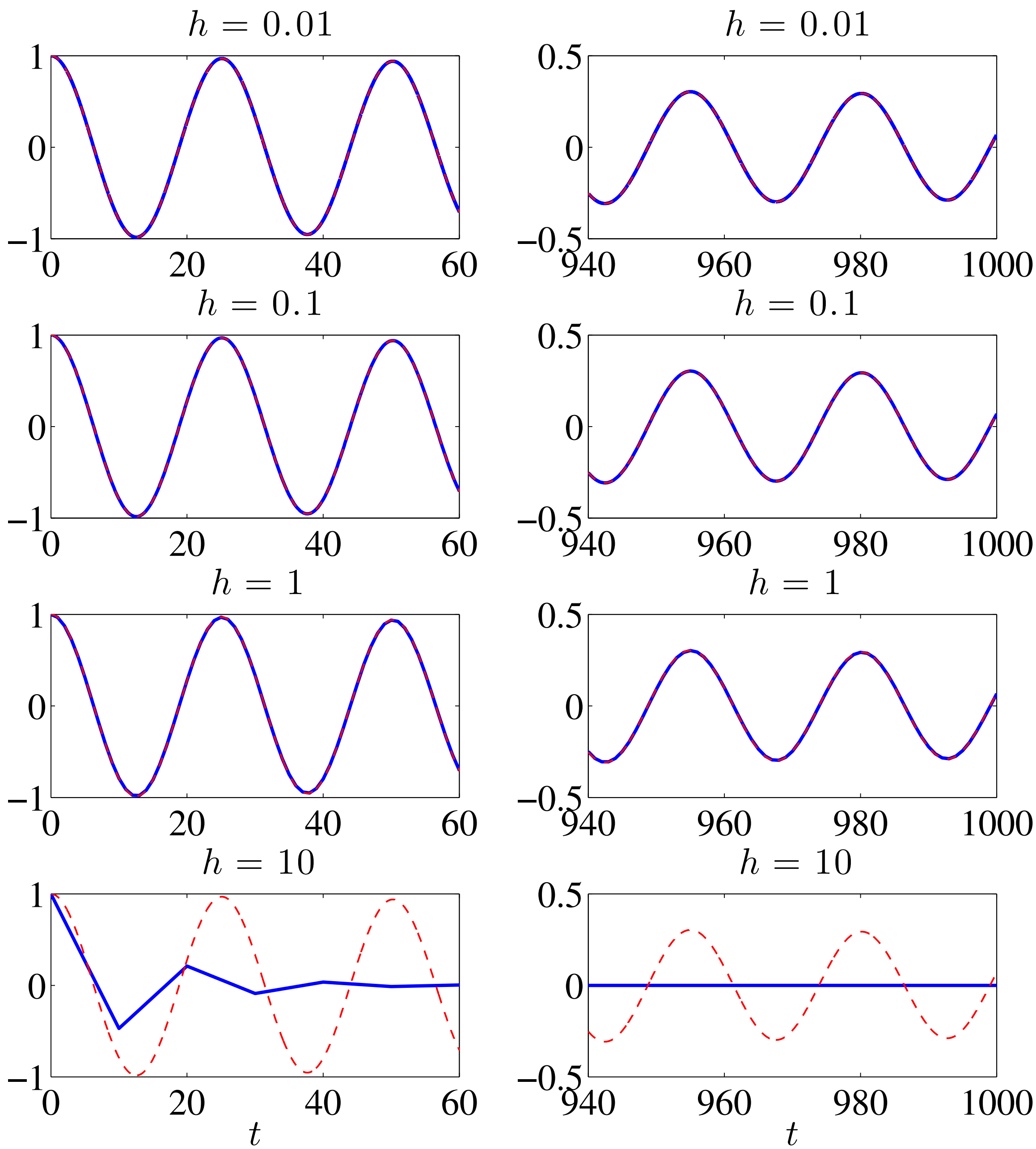

Figure 6: Comparison of the exact (red) and numerical (blue) solution for different h.