Spectral Method Linear PDE Solver

Danny Ober-Reynolds
doberrey@asu.edu

Alexander Reynolds
abreynolds@asu.edu

ARIZONA STATE UNIVERSITY

Honors Project
MAT 425: Numerical Analysis II
Dr. Sébastien Motsch
Spring 2016

1 Introduction

Consider the second order linear PDE of a periodic function u given by

$$\begin{cases} u_t = au_{xx} + bu_x + cu & 0 \le x < \ell, t \ge 0 \\ u(0,t) = u(\ell,t) & t \ge 0 \\ u(x,0) = u_0(t) & t = 0. \end{cases} \tag{1}$$

This PDE has a number of physical analogues, although since u is only a function of a single variable x the reality of the models is quite limited. Regardless, interpretations can give an intuition to the problem.

If $b = c = 0$, then this equation is the heat equation with periodic boundary conditions. Such an instance could occur if one was interested in how heat is diffused through a ring, or any other metal loop. The initial condition $u_0(x)$ would then describe the current concentration of heat in the ring, and the diffusion coefficient a would be a property of the metal, revealing how quickly the ring would come to a uniform temperature. If the term $c > 0$, then this analogy could be extended to that of a ring with an initial heat distribution $u_0(x)$ and diffusivity a inside of an environment with a rising temperature.

If $a = c = 0$, then this yields the transport equation with periodic boundary conditions. For an example, consider a lazy river ride at a water park. The inflatable rafts in the river could be viewed as particles which are transported along with the water that do not diffuse; in this case $u_0(x)$ would describe the concentration of inflatable rafts at different points along the river. If $a > 0$, then there is a mix of diffusion and transport. The raft analogy breaks down here a bit due to diffusion, so consider the similar example of concentrations of chlorine in the lazy river. In this case, $u_0(x)$ would describe the initial concentration at patches along the one-dimensional river, and with diffusion and transport the chlorine would be spread throughout and moved along the river. If $c > 0$ then this would suggest that the chlorine levels rising in the river throughout time.

The imagination of such scenarios that this equation can describe all share some common terms, so it will be useful to state these explicitly. The term au_{xx} represents *diffusion*, bu_x represents *transport*, and cu represents *growth* with a, b, and c the diffusion, transport, and growth coefficient respectively. Since a is the diffusion coefficient, $a \ge 0$ (otherwise the quantity of interest gets concentrated to a singularity); if $a = 0$ then there is no smoothing of the initial condition over time. If $a > 0$ then the initial condition gets smoothed over time, more speedily as a increases. If $b > 0$ then the function will be shifting to the left (or anti-clockwise) as time increases, and if $b < 0$ then the function will be shifting to the right (or

clockwise). The coefficient c simply determines whether the function is increasing or decreasing as time increases over the whole domain.

As a second order linear equation, this PDE could be solved using finite difference methods. However, to explore a topic outside of the normal curriculum, we opted to solve this problem using spectral methods. Spectral methods can give a natural way of viewing the problem, and can lead to great accuracy with minimal computation in certain cases. A very brief background to Fourier analysis will be presented to develop the idea behind spectral methods.

2 Theory

The motivation for spectral methods is no different than the general idea of a transform in the solution of differential equations. The transform allows to lift the function into a space with simpler derivatives, or obvious explicit solutions, or easier numerical schemes. In the case of Fourier transforms, the function is transformed into spectral space, and the time evolution of each mode is studied instead of the entire function.

The Fourier transform takes a periodic piecewise-smooth function and breaks it up into infinitely many simple sin and cos functions, which when summed in series yield back the original function. However, this is clearly infeasible to calculate. A useful property of the Fourier transform shows that the amplitudes of the n-th wave is inversely proportional to n. Thus computationally, we truncate the Fourier series to only give the first $2N+1$ modes while still maintaining close accuracy to the original function, especially in smooth regions.

Definition 2.1 (Fourier Series)**.** The $2N+1$-th *Fourier partial sum* of a function f is a harmonic approximation to f given by

$$S_N f(x) = \sum_{k=-N}^{N} \hat{f}_k e^{ikx}, \tag{2}$$

where the *Fourier coefficients* $\hat{f}_k$ are given by the *Fourier transform* of f, viz.,

$$\hat{f}_k = \int_{-\pi}^{\pi} f(x) e^{-ikx} \, \mathrm{d}x. \tag{3}$$

Note that the Fourier coefficients $\hat{f}_k$ are complex numbers for each value of k, so they can be represented in the form $\hat{f}_k = A_k e^{i\theta_k}$ where $A_k \geq 0$ is the amplitude and $e^{i\theta_k}$ of the Fourier coefficient. The Fourier coefficients thus describe the amplitude and the phase of each frequency component (corresponding to k) making up the function $f(x)$.

The Fourier transform has a particularly awesome property involving derivatives that allow for effortless derivative computation in the spectral domain.

Proposition 2.1 (Fourier Derivative Property)**.** *If $\hat{f}_k$ is the Fourier transform of a piecewise-smooth periodic function f, then $\widehat{f^{(n)}}_k = (ik)^n \hat{f}_k$.*

Proof. Using the (continuous) inverse transform,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}_\omega e^{i\omega x} \, \mathrm{d}\omega$$

it is easy to see that

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}_\omega e^{i\omega x} \, \mathrm{d}\omega \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (i\omega) \hat{f}_\omega e^{i\omega x} \, \mathrm{d}\omega. \qquad \square$$

Thus, the Fourier transform looks like a good way to compute derivatives instead of using finite difference methods. The finite difference methods compute derivatives using local information; we infer the slopes from a neighborhood around each point. Spectral methods are instead global; the derivatives are computed using modes which extend throughout the entire domain. This idea will be applied to solve Equation 1.

Taking the Fourier transform of both sides of Equation 1 and using the derivative relationship from Proposition 2.1 yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{u}_k = a(ik)^2\hat{u} + b(ik)\hat{u} + c\hat{u} = (-ak^2 + ibk + c)\hat{u}_k, \tag{4}$$

a simple linear ODE of $\hat{u}_k$ for each k. That is, the Fourier transform gives a system of independent linear ODEs, one for each value of k.

The interpretation of the problem has now changed. Instead of looking for how $u(x,t)$ changes over time, the concern now is in how $\hat{u}_k$ changes over time for each k. As mentioned in the beginning of the section, these $\hat{u}_k$ are Fourier coefficients and thus simply complex numbers, each with an amplitude and phase shift. Studying how $\hat{u}_k$ changes with time is studying how the k-th mode is varying in amplitude and in phase shift over time. The solution to the differential equation corresponding to each k describe how every mode is scaled and shifted until the final time. Then, taking the inverse transform of the modes after they have been scaled and shifted to the final time position, the evolution of the original function is obtained. In a way, this is similar to the cases when a PDE is turned into a system of ODEs, each representing the path of a particle.

Figure 1 shows this natural interpretation. The first plot shows many different Fourier modes extended in physical space which add up to give an approximation of the initial condition $u_0(x) = \mathrm{sq}(x)$, a unit pulse which is *on* over $[\frac{\pi}{2}, \frac{3\pi}{2})$, viz.

$$\mathrm{sq}(x) := \begin{cases} 1 & \frac{\pi}{2} \le x < \frac{3\pi}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The time evolution of the PDE is then shown in the second plot, and the third plot shows the modes resulting in the waveform at the final time step. Notice that there are very few modes in the final solution; the others decayed rapidly and are overlapping on the line riding 0. Further, notice that these modes are shifted due to the transport coefficient being non-zero. This fully exemplifies the benefits of working with the time evolution of the Fourier coefficients (and thus the modes). For a PDE with diffusion, the number of Fourier modes drops quickly.

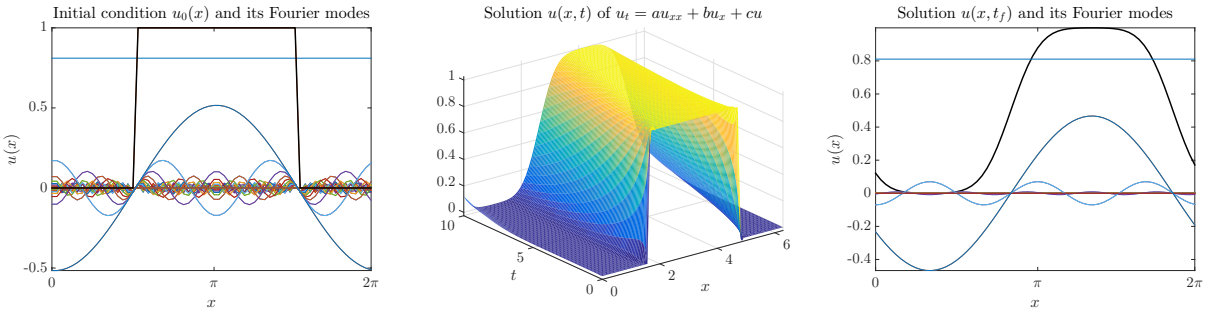


Figure 1: The first plot shows the initial condition $u_0(x) = \mathrm{sq}(x)$ and 0 elsewhere with its associated Fourier modes. The second plot shows the time evolution of u satisfying $u_t = au_{xx} + bu_x + cu$ where $a = 0.01$, $b = -0.1$, and $c = 0$ with the initial condition $u_0(x)$. The final plot shows the solution $u(x,t_f)$ at the final time step and its associated Fourier modes.

0	0	0	0	0
1/2	1/2	0	0	0
1/2	0	1/2	0	0
1	0	0	1	0
	1/6	1/3	1/3	1/6

Table 1: Butcher Tableau for fourth-order Runge-Kutta (RK4)

However, spectral methods have their problems as well. While global data give great approximations of smooth functions, discontinuous functions have local features which global data do not capture well. As seen in Figure 1, a discontinuous function can still behave well, but the accuracy is not nearly as good as starting with a smooth initial condition. This will be discussed explicitly in the later section on error and stability.

3 Numerics

While there is an analytic solution for the linear ODEs when given values of a, b, and c, a numerical program allows to quickly compute the approximate solutions for any values. The numerical method that was used is included at the end of this paper. The set up and algorithm for the numerical scheme will be discussed in this section.

First, the periodic domain is discretized with N steps, and the initial condition $u_0(x)$ is evaluated at these points to obtain the sequence

$$\{u_0(x_n)\}_{n=1}^{N} \tag{6}$$

The Fourier transformation is numerically estimated using a discrete Fourier transformation (DFT)

$$\hat{u}_k = \int_0^{2\pi} u(x)e^{-ikx}\,\mathrm{d}x \approx \sum_{n=1}^{N} u(x_n)e^{-i\frac{2\pi(k-1)}{N}(n-1)}, \quad 0 \le k \le N-1 \tag{7}$$

resulting in N wavenumbers $k = 0, 1, \ldots, N-1$. The DFT is computed exactly and quickly by the built-in MATLAB `fft` function.

Using that the Fourier transform is linear with the result from Proposition 2.1, the coefficients are transformed into a sequence indexed by k,

$$\left\{\alpha_k \coloneqq (-ak^2 + ibk + c)\right\}_{k=0}^{N-1}. \tag{8}$$

As in Equation 4, this produce a sequence of independent ODEs indexed by k

$$\left\{\frac{\mathrm{d}}{\mathrm{d}t}\hat{u}_k = \alpha_k \hat{u}_k\right\}_{k=0}^{N-1} \tag{9}$$

which can each be solved for each k.

To solve each ODE, time is discretized $\{t_s\}_{s=0}^{f}$ and the evolution of the Fourier coefficient $\hat{u}_k$ is solved using a fourth order Runge-Kutta Method (RK4) given by the Butcher Tableau in Table 1. Note that other methods could be used here as well, but RK4 was chosen due to the high accuracy even when some of the ODEs become more stiff.

Applying RK4 to the N ODEs yields a solution at each time step, giving a matrix holding the Fourier coefficient $\hat{u}_k$ at time t

$$\left\{\hat{u}_{k,t}\right\}_{k=0,t=t_0}^{N-1,t_f} \tag{10}$$

This is a sequence of modes at each time step. Applying the inverse discrete transform with the MATLAB command `ifft` at each time step then builds a function from the associated modes. The result is thus a matrix

$$\{u_{x,t}\}_{x=0,t=t_0}^{\ell,t_f} \tag{11}$$

of the solution at each value x and each time t until t_f.

4 Error and Stability

There is some concern for stability in the solution of ODEs in Fourier space. If a large number of modes is chosen, the coefficient $\alpha(k)$ from Equation 8 will be large for the larger values of k, and thus the system from Equation 9 will have a number of stiff ODEs. If the time steps are too large, the RK4 method will be numerically unstable. A simple constraint on the time step however can guarantee stability, as for a linear ODE

$$|\alpha(k)|\Delta t < 1$$

ensures that the method will be stable. When the diffusion coefficient $a > 0$, the function is smoothed as time goes on, requiring less modes to approximate it. Thus the interpretation for the stiffness is simply that as more modes are added, the higher modes decay more quickly to zero. As shown in Figure 1, at the final time there are only a few modes with non-negligible amplitudes, so starting with more only means that more of them will decay quickly.

There are three sources of error. The first occurs in estimating the Fourier transformation of the initial function u_0 using the discrete Fourier transformation. This is an approximation of the actual integral-based Fourier transformation, but is known to be very small especially when k is reasonably large. For smooth functions, `fft` is especially accurate. However, for discontinuous initial functions like the pulse or sawtooth functions, the Gibb's phenomenon presents itself around the discontinuities. The Gibb's phenomenon is the observation of oscillations around the points of discontinuities in Fourier approximations. While the oscillations can get relatively high in amplitude, the oscillations are narrow enough that the total source of L_2 error still remains fairly low. Additionally, even if a perfect estimation were possible of a discontinuous function like a pulse from Fourier modes, with any small timestep $t_\epsilon > 0$ after the initial t_0 the function becomes smooth so long as the diffusion coefficient $a > 0$.

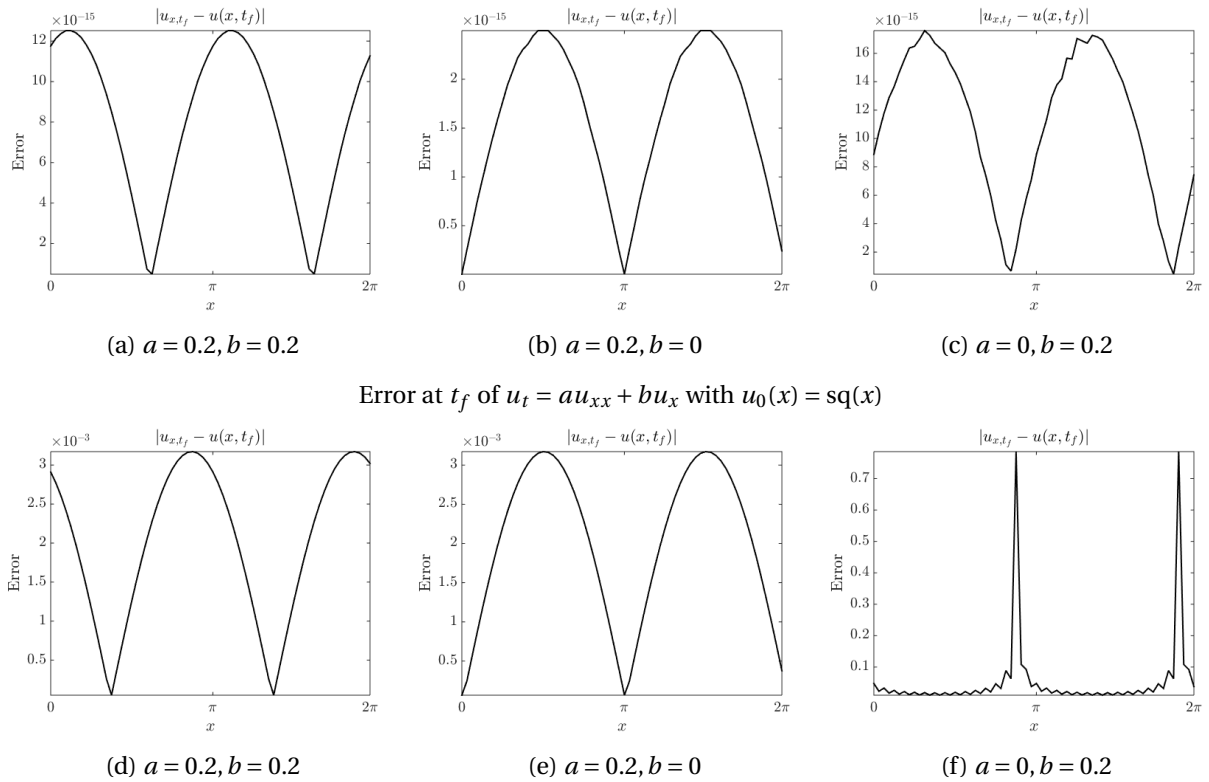
A second source of error is introduced when solving the system of ODEs in Fourier space. By choosing RK4, a method of order 4, the error is quite small.

Lastly, there is some error with transforming our solution back to physical space from Fourier space using the inverse fast Fourier transformation.

Unfortunately these errors are not simply added as in other numerical methods, but compounded throughout the process in a less straightforward way. In order to analyze the error, the estimated solutions can be compared to 'exact' solutions and with different values of N, corresponding to a different number of Fourier modes. First,

The error plots in Figure 2 show the error at the final timestep t_f for two initial functions, a smooth sin curve and a discontinuous pulse. The error was computed by comparing the solution with a high number of modes (256) and a very small timestep $\Delta t = 0.001$ to a more realistic numerical model with a smaller number of modes (64) and a larger timestep $\Delta t = 0.005$.

These error plots show a substantial amount about the method. First, the equations with $u_0(x) = \sin(x)$ give extreme precision. There is virtually no error at the locations where $\sin(x) = 0$, and otherwise machine-precision error. This is to be expected, since a sin function can be approximated from just a single Fourier mode. The only reason why there is any error is simply due to the discretization of physical space at non-exact real numbers (numerical estimations of fractions of π).



(a) $a = 0.2, b = 0.2$ (b) $a = 0.2, b = 0$ (c) $a = 0, b = 0.2$

(d) $a = 0.2, b = 0.2$ (e) $a = 0.2, b = 0$ (f) $a = 0, b = 0.2$

Figure 2: Errors at the final time step $t_f = 10$ for $u_t = au_{xx} + bu_x$. The first column has $a = 0.2, b = 0.2$, the second has $a = 0.2, b = 0$, and the third has $a = 0, b = 0.2$. Thus the columns refer to the error of the solutions for a diffusive transport equation, a diffusion equation, and a transport equation. The top row shows the error for $u_0(x) = \sin(x)$ while the bottom row shows the error for $u_0(x) = \mathrm{sq}(x)$.

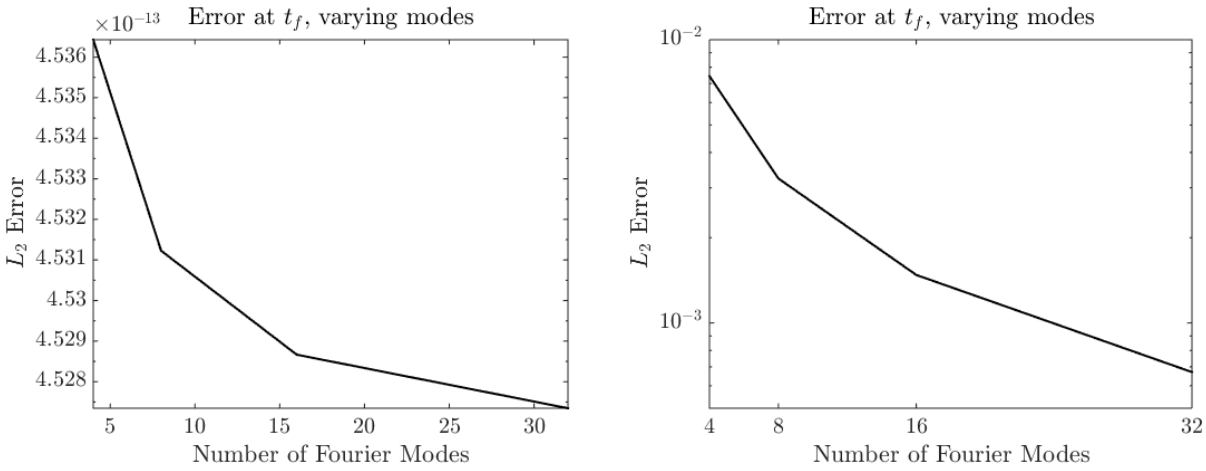
For the initial condition of the discontinuous square wave $\mathrm{sq}(x)$, with diffusion the methods show a pretty strong amount of accuracy with errors on the order of 10^{-3}. However, one particular error plot stands out: the transported square with no diffusion shows very large (almost order 1) error. This is to be expected, though, when comparing two square approximations with different numbers of Fourier modes. With a lower number of Fourier modes, the slopes around the discontinuities will be less steep than the slopes from a higher number of Fourier modes. This error graph would be roughly identical to that of the difference at the initial time step then. Furthermore, the L_2 error (which corresponds to the area under the curve), arguably more important for our purposes would still be quite small due to the narrowness of these errors.

In order to better analyze the error, the estimated solutions from varying values of N were compared to exact solutions corresponding to a high number of Fourier modes (256). The exact solutions for the ODEs in Fourier space are given by a simple exponential since they have the form as in Equation 9. The approximate solutions were then computed for $N = 4, 8, 16, 32$ modes and using RK4 with a timestep $\Delta t = 0.01$. The L_2 error given by

$$L_2 \text{ Error} := \sqrt{\sum (u_{x,t_f} - u(x,t_f))^2 \Delta x} \tag{12}$$

where u_{x,t_f} is the numerical approximation from the spectral method and using RK4 to estimate the ODEs, and $u(x,t_f)$ is the exact solution given by the exponential in Fourier space. This L_2 error was

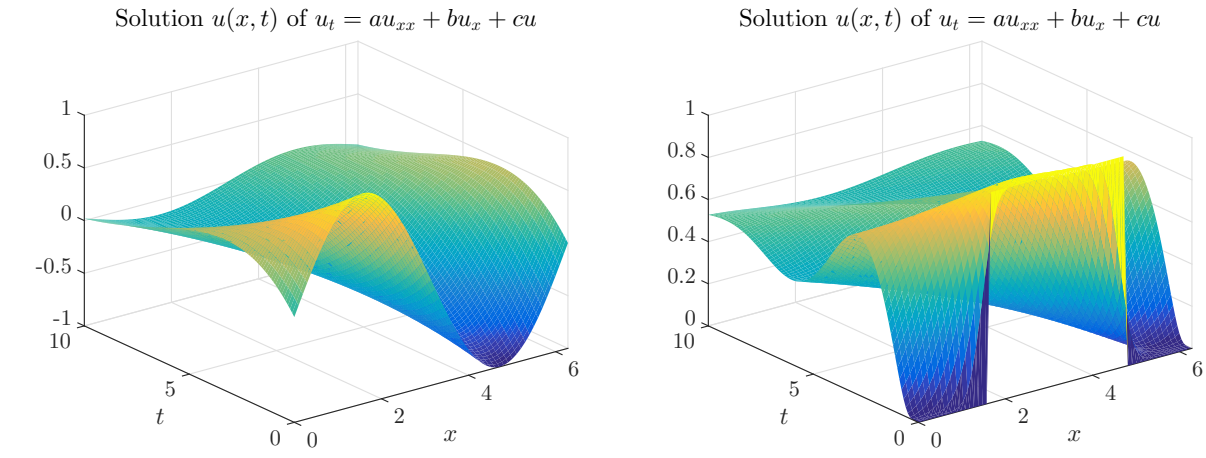
then computed and plotted over the number of Fourier modes to see how the error decreases. The error plots are logarithmic in the vertical axes, showing better-than-exponential decay of error with increased Fourier modes. As before, two plots are generated, one with the initial condition of $u_0(x) = \sin(x)$, the other with the initial condition of $u_0(x) = \mathrm{sq}(x)$.



(a) $u_0(x) = \sin(x)$ (b) $u_0(x) = \mathrm{sq}(x)$

Figure 3: L_2 error at t_f for $u_t = 0.5u_{xx} + 0.2u_x$.

To wrap up this project, the numerical scheme is demonstrated on two linear PDEs with the initial conditions that have been used in this paper. The code for this project can be accessed and downloaded from `http://reynoldsalexander.com`.



(a) $a = 0.7, b = 0.3, c = 0.5$ with $u_0(x) = \sin(x)$ (b) $a = 0.3, b = 1.0, c = 0$ with $u_0(x) = \mathrm{sq}(x)$

Figure 4: Two examples of the numerical method showing the time evolution of a smooth initial function with diffusive transport and growth, and of a discontinuous initial function with diffusive transport.