NUMERICAL SIMULATIONS OF THE ENERGY-STABLE SCHEME FOR SWIFT-HOHENBERG EQUATION

by

Jun ZHOU*

School of Mathematics and Statistics, Yangtze Normal University, Chongqing, China

Original scientific paper
https://doi.org/10.2298/TSCI180515080Z

A collocation Fourier scheme for Swift-Hohenberg equation based on the convex splitting idea is implemented. To ensure an efficient numerical computation, we propose a general framework with linear iteration algorithm to solve the non-linear coupled equations which arise with the semi-implicit scheme. Following the contraction mapping theorem, we present a detailed convergence analysis for the linear iteration algorithm. Various numerical simulations, including verification of accuracy, dissipative property of discrete energy and pattern formation, are presented to demonstrate the efficiency and the robustness of proposed method.

Key words: Swift-Hohenberg equation, energy stability, collocation Fourier method, linear iteration

Introduction

The Swift-Hohenberg (SH) equation is an important physical model and can be applied in roll patterns in Rayleigh-Benard convection in [1], chemical reactions [2], and liquid crystal displays [3, 4].

The SH equation is given by the $L^2$ gradient flow:

$$\frac{\partial u}{\partial t} = -u + \gamma u - \epsilon^2 \Delta u - \delta^2 \Delta^2 u$$

on $\Omega = [0, L_x] \times [0, L_y]$, and $\gamma$, $\delta^2$, and $\epsilon^2$ are positive real constants. Here we only consider the periodic boundary value condition.

Regarding to the numerical consideration, it is well known that the numerical treatment of gradient flow poses certain challenges owing to the physical complexity. Energy stability of the proposed numerical scheme is necessary for SH equation. For example, the linearized schemes are given by Cheng and Warren [5] with three undetermined weights and Backofen et al. [6] with a directly linearized formula $(\phi^{n+1})^t \approx 3(\phi^t)^{n+1} - 2(\phi^t)^{n+1}$. More extensive applications of energy-stable or energy conservative method to a wide class of physical models also are available. See the related works for wave equations [7, 8], the phase field crystal equation [9] and the Cahn-Hilliard equation [10-12], etc.

In general, for most convex splitting numerical works, a local spatial discretization, such as the finite difference or finite element approximation is often used, because some highly efficient non-linear, for example, the steepest descent method [13], solvers can be borrowed readily. But a spatial approximation with a global nature, such as spectral or collocation spec-

*Author’s e-mail: flzjzkln@126.com
A numerical method for SH equation, is very difficult. The key reason is that, the convex splitting scheme usually treats the non-linear term implicitly, since the non-linear part corresponds to the convex part of the Lyapunov energy functional. Whereas, for collocation spectral method, the major advantage is that it is easier to implement, and very efficient due to the fast Fourier transform. In this paper, a first-order convex splitting scheme, with a collocation Fourier method for the SH eq. (1) is proposed. To implement the non-linear numerical scheme, a linear iteration algorithm is introduced and the corresponding contraction mapping property under a given condition is shown.

The numerical Scheme

Collocation Fourier spectral discretization space

Assume that $L_x = N_x h_x$, $L_y = N_y h_y$, for some mesh sizes $h_x, h_y > 0$ and some positive integers $N_x$ and $N_y$. For simplicity of presentation, we use a square domain, i.e., $L_x = L_y = 1$, and a uniform mesh size $h_x = h_y$, $N_x = N_y = N$. We will always assume that $N = 2K + 1$ is always odd. All the variables are evaluated at the regular numerical grid $(x_i, y_j)$, with $x_i = i h_x$, $y_j = j h_y$, $0 \leq i, j \leq N$.

For a periodic function, $f$, over the given 2-D numerical grid, its discrete Fourier expansion is given:

$$f_{i,j} = \sum_{l,m=-K}^{K} \hat{f}_{l,m} \exp[2\pi i(lx_i + my_j)]$$

its collocation Fourier spectral approximations to first and second order partial derivatives are given:

$$(D_x f)_{i,j} = \sum_{l,m=-K}^{K} (2\pi i l) \hat{f}_{l,m} \exp[2\pi i(lx_i + my_j)]$$

$$(D_x^2 f)_{i,j} = \sum_{l,m=-K}^{K} (4\pi^2 l^2) \hat{f}_{l,m} \exp[2\pi i(lx_i + my_j)]$$

and the corresponding collocation spectral differentiations in the $y$-direction can be defined in the same way.

We introduce the $\ell^p$ and $\ell^p$, $\ell \leq p < \infty$ norms for a grid function:

$$\| f \|_{\ell^p} := \max_{0 \leq i,j \leq N-1} | f_{i,j} |, \quad \| f \|_{\ell^p} := (h^2 \sum_{i,j=0}^{N-1} | f_{i,j} |^p)^{1/p}$$

The following Lemma will play important role in the contraction mapping analysis for our algorithm.

Lemma 1. Suppose $\gamma_1, \gamma_2 > 0$. For any periodic grid function $f$, we have:

$$2C_\alpha \alpha \| f \|_2^2 \leq \gamma_1 \| f \|_2^2 + \gamma_2 \| \Delta_x f \|_2^2$$

where $\alpha = -2\gamma_2 + [2\gamma_2 (\gamma_1 + \gamma_2)]^{1/2}$.

Proof. For the proof of eq. (5), a discrete version of Sobolev embedding from $H^2$ (the discrete case of $H^2$ space) into $\ell^4$, we have to utilize the grid function, $f$, with its discrete Fourier expansion defined as eq. (2). An application of discrete Parseval equality gives:

$$\| f \|_2^2 = \sum_{l,m=-K}^{K} |\hat{f}_{l,m}|^2, \quad \nabla_x f \|_2^2 = \sum_{l,m=-K}^{K} |\hat{\lambda}_{x,m}|^2 |\hat{f}_{l,m}|^2, \quad \Delta_x f \|_2^2 = \sum_{l,m=-K}^{K} |\hat{\lambda}_{x,m}^2 |\hat{f}_{l,m}|^2$$

with $\lambda_{x,m} = 4\pi^2(F + m^2)$. Then, letting:
\( \gamma_1 = \alpha + \beta, \quad \alpha > 0, \quad \beta > 0 \) \hfill (7)

we have:

\[
\gamma_1 \| f \|_2^2 + \gamma_2 \| \Delta_x f \|_2^2 = \alpha \| f \|_2^2 + \sum_{l,m=-K}^{K} (\beta + \gamma_2 \lambda_{lm}^2) \| \hat{f}_{lm} \|^2 \geq \alpha \| f \|_2^2 + 2\beta \| \Delta_x f \|_2^2
\]

Next, we choose values of \( \alpha \) and \( \beta \) with \( \alpha = 2(|\beta\gamma_2|^{1/2}) \). It follows from eq. (8):

\[
\alpha = -2\gamma_2 + 2\sqrt{\gamma_2(\gamma_1 + \gamma_2)} \quad \beta = \gamma_1 + 2\gamma_2 - 2\sqrt{\gamma_2(\gamma_1 + \gamma_2)} \hfill (9)
\]

Therefore, from eq. (9), follows:

\[
\gamma_1 \| f \|_2^2 + \gamma_2 \| \Delta_x f \|_2^2 \geq \alpha \| f \|_2^2 + \alpha \| \nabla_x f \|_2^2 = \alpha \| f \|_2^2 \geq C_0 \| f \|_2^2 \hfill (10)
\]

in which a discrete Sobolev embedding from \( H^1 \) to \( \ell^4 \) is used. This completes the proof.

**Semi-implicit numerical scheme**

We propose the following fully discrete first-order (in time) Fourier collocation spectral Scheme for SH eq. (1):

\[
\frac{u^{n+1} - u^n}{\Delta t} = -(u^{n+1})^3 + \gamma u^n - \varepsilon^2 \Delta_x u^n - \delta^2 \Delta_x^2 u^{n+1} \hfill (11)
\]

where \( \Delta t \) is the discrete time step and \( u^n \) denotes the time-discrete approximation of \( u(t, n\Delta t) \), \( n = 0, 1, \ldots, M = \lceil T/\Delta t \rceil \), and \( T \) – the given final time.

**Lemma 2.** Assume that \( u^n \) is the numerical solution of eq. (11). Then, the following estimates are valid:

\[
\| u^n \|_2 \leq C_1, \quad \| u^n \|_2 \leq C_2, \quad \| \Delta_x u^n \|_2 \leq C_3, \quad \| u^n \|_\infty \leq C_4
\]

\hfill (12)

for \( n = 1, 2, \ldots, M \).

**Linear iteration algorithm**

In this section, we mainly discuss an efficient algorithm for solving eq. (11). Firstly, we note that the scheme (11) can be reformulated as a closed equation for \( u^{n+1} \):

\[
\left( \frac{1}{\Delta t} + \delta^2 \Delta_x^2 \right) u^{n+1} = \left( \frac{1}{\Delta t} + \gamma - \varepsilon^2 \Delta_x \right) u^n - (u^{n+1})^3 \hfill (13)
\]

or, equivalently:

\[
(\mathcal{I} + \Delta t \delta^2 \Delta_x^2) u^{n+1} = (\mathcal{I} + \gamma - \varepsilon^2 \Delta_x) u^n - (u^{n+1})^3 \hfill (14)
\]

where \( \mathcal{I} \) denotes the identity operator. Also, define a linear operator \( \mathcal{A} \) and the value \( f_n \) at the \( n^{\text{th}} \) level:

\[
\mathcal{A} := \mathcal{I} + \Delta t \delta^2 \Delta_x^2, \quad f_n := [(1 + \gamma \Delta t) \mathcal{I} - \Delta t \varepsilon^2 \Delta_x] u^n \hfill (15)
\]

Then, eq. (14) can be simplified:

\[
\mathcal{A} u^{n+1} = f_n - \Delta t (u^{n+1})^3 \hfill (16)
\]

Obviously, the non-linear part in this equation is treated implicitly. To overcome the difficulty associated with the implicit treatment of the non-linear term, a linear solver is necessary, and we propose the following linear iteration algorithm:
\[ Au^{n+1(k+1)} = f_n - \Delta t(u^{n+1(k)})^3 \]  

(17)

in which \( u^{n+1(k)} \) corresponds to the numerical solution at the \( k \)th iteration.

The following theorem gives an affirmative answer for the convergence of such an iteration algorithm.

**Theorem 1.** The linear eq. (17) is a contraction mapping in the discrete norm, provided that:

\[ \Delta t < \frac{C_0^2 \delta^2}{5C_1^2(5C_1 + 2C_0 \delta^2)} \]  

(18)

with the positive constants \( C_0, C_1 \), which are given in eqs. (5) and (12), respectively.

**Proof.** Let \( u^{n+1} \) be the unique periodic solution to eq. (14) and define the iteration error at each stage via:

\[ e^k := u^{n+1(k)} - u^{n+1} \]  

(19)

Subtracting eq. (17) from eq. (16) yields the following discrete equation:

\[ A e^{k+1} = -\Delta t[(u^{n+1(k)})^3 - (u^{n+1})^3] \]  

(20)

Taking a discrete inner product of eq. (20) with \( e^{k+1} \), and using integration by parts leads to:

\[ < A e^{k+1}, e^{k+1} > = \Delta t \delta^2 \Delta_x e^{k+1}, e^{k+1} > + \Delta t \delta^2 \Delta_x e^{k+1}, e^{k+1} > + \Delta t \delta^2 \Delta_x e^{k+1}, e^{k+1} > = \| e^{k+1} \|^2 + \Delta t \delta^2 \| \Delta_x e^{k+1} \|^2 \]

(21)

in which a discrete Holder inequality was applied at the last step. To proceed the non-linear analysis, we use the induction method to handle the bound of error function \( e^k \). In general, we choose \( u^{n+1(0)} = u^* \) as the initial value. By the preliminary bound (12) in Lemma 2, we can get an estimate for \( e^0 \) in the \( \| \cdot \| \) norm:

\[ \| e^0 \| \leq 2C_1 \]  

(22)

and make the following the inductive assumption at the iteration stage \( k \):

\[ \| e^k \| \leq 2C_1 \]  

(23)

With this assumption, a bound of \( u^{n+1(k)} \) in the \( \| \cdot \| \) norm is derived:

\[ \| u^{n+1(k)} \| \leq \| e^k + u^{n+1} \| \leq \| e^k \| + \| u^{n+1} \| \leq 3C_1 \]  

(24)

Going back to eq. (21), we consider the term \( \| e^{k+1} \|^2 + \Delta t \delta^2 \| \Delta_x e^{k+1} \|^2 \). According to Lemma 1, taking \( \gamma_1 = 1 \) and \( \gamma_2 = \Delta t \delta^2 \), we have:

\[ \| e^{k+1} \|^2 + \Delta t \delta^2 \| \Delta_x e^{k+1} \|^2 \geq 2C_0 \left[ -2\Delta t \delta^2 + 2\sqrt{\Delta t \delta^2 (\Delta t \delta^2 + 1)} \right] \| e^{k+1} \|^2 \]  

(25)

which indicates:

\[ 2C_0 \left[ -2\Delta t \delta^2 + 2\sqrt{\Delta t \delta^2 (\Delta t \delta^2 + 1)} \right] \| e^{k+1} \|^2 \leq \Delta t \| u^{n+1(k)} \|^2 + \| u^{n+1} \|^2 \| e^k \| \| e^{k+1} \| \]  

(26)

Hence, it follows from eq. (26) that:

\[ C_0 \left[ \sqrt{\Delta t \delta^2 (\Delta t \delta^2 + 1)} - \Delta t \delta^2 \right] \| e^{k+1} \| \leq 5C_1^2 \Delta t \| e^k \| \]  

(27)

and we arrive at the estimate:
\[ || e^{x+1} ||_1 \leq \frac{5C_i^2 \Delta t}{C_t (\sqrt{\Delta t \delta^2 (\Delta t \delta^2 + 1)} - \Delta t \delta^2)} || e^x ||_1 \]  (28)

As a result, a contraction is assured under the condition that:

\[ \frac{5C_i^2 \Delta t}{C_t (\sqrt{\Delta t \delta^2 (\Delta t \delta^2 + 1)} - \Delta t \delta^2)} < 1 \]  (29)

that is:

\[ \Delta t < \frac{C_t \delta^2}{5C_i^2 (5C_i^2 + 2C_t \delta^2)} \]  (30)

Clearly, we are justified in our a priori assumption of eq. (23), since:

\[ || e^{x+1} ||_1 || e^x ||_1 \leq 2C_i \]  (31)

provided that condition of eq. (31) is enforced, which completes the proof of Theorem 1.

**Numerical experiments**

In this section, we present some numerical experiments to verify the theoretical results obtained in the previous sections, particularly the accuracy test, the energy stability and the pattern formation.

**Accuracy test**

An example provides the numerical evidence for our numerical scheme being first-order accurate in time and spectral accurate in space. The set-up of this accuracy test is based on that presented [11, 14]. We solve the 1-D SH equation on the domain \( \Omega = [0.32] \). The initial condition is defined:

\[ u(x) = 0.07 - 0.02 \cos \left( \frac{2\pi(x-12)}{32} \right) + 0.0171 \cos \left( \frac{2\pi(x+10)}{32} \right) - 0.0085 \sin \left( \frac{4\pi x}{32} \right) \]  (32)

For exploring the temporal accuracy in time, we fix spatial resolution as \( N = 2048 \) so that the numerical error is dominated mainly by the temporal ones. We computed a reference solution at time \( T = 1 \) using a time step \( \Delta t = 1/2^m \) and assume that this space time discretization is fine enough as to suppose that the reference solution is exact. With a sequence of time step sizes \( \Delta t = 0.1, 0.005, 0.025, 0.0125, \) and \( 0.00625 \), we also compute the numerical errors at \( T = 1 \) for the parameter group (I): \( \gamma = 0.001, \varepsilon = 1, \delta^2 = 1 \), and the parameter group (II): \( \gamma = 0.025, \varepsilon = 1, \delta^2 = 10 \), respectively. The results are presented in table 1 in which a first-order in time accuracy is shown clearly.

**Table 1. The \( || \cdot ||_1 \) errors and convergence orders of the numerical solution \( T = 1 \) at for different parameter group \( \gamma, \varepsilon^2, \) and \( \delta^2 \)**

| \( \Delta t \) | \( \gamma = 0.001, \varepsilon = 1, \delta^2 = 1 \) | Order | \( \gamma = 0.025, \varepsilon = 1, \delta^2 = 10 \) | Order |
|---|---|---|---|
| 0.1 | 6.7072556e-6 | – | 1.3791082e-4 | – |
| 0.05 | 3.3186000e-6 | 1.015147 | 6.8618141e-5 | 1.007073 |
| 0.025 | 1.6374393e-6 | 1.019133 | 3.3978888e-5 | 1.013951 |
| 0.0125 | 8.0141592e-7 | 1.030818 | 1.6648254e-5 | 1.029267 |
| 0.00625 | 3.8417922e-7 | 1.060771 | 7.9846732e-6 | 1.060065 |
To demonstrate the accuracy in space, we take $\Delta t = 10^{-4}$ so that the temporal numerical error is negligible. Also, we computed a reference solution at time $T = 5$ with $N = 2048$ and assume that this space time discretization is fine enough as to suppose that the reference solution is exact. With grid sizes $N = 8, 16, 32, 64$, and 128, we solve eq. (11) up to for two different parameter groups given in previous subsection, respectively. The discrete norm $\|\cdot\|_2$ of numerical errors at $T = 5$ is given in Tab. 2, which shows the spatial spectral accuracy is verified. Noting that when $N$ increases from 64 to 128, there have so little difference between errors. The main reason is that the value of $N$ is so large that the numerical errors are dominated by the temporal discretization.

Table 2. The $\|\cdot\|_2$ errors of spectral accuracy for the numerical solution at $T = 5$

<table>
<thead>
<tr>
<th>$\Delta t$</th>
<th>$\gamma = 0.001, \varepsilon^2 = 1, \delta^2 = 1$</th>
<th>$\gamma = 0.025, \varepsilon^2 = 1, \delta^2 = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>9.21122e-05</td>
<td>1.246072e-06</td>
</tr>
<tr>
<td>16</td>
<td>1.37439e-06</td>
<td>1.103986e-12</td>
</tr>
<tr>
<td>32</td>
<td>2.67306e-13</td>
<td>2.376737e-14</td>
</tr>
<tr>
<td>64</td>
<td>5.71089e-14</td>
<td>3.670676e-14</td>
</tr>
<tr>
<td>128</td>
<td>7.14353e-14</td>
<td>4.489314e-14</td>
</tr>
</tbody>
</table>

In this subsection, we present some numerical tests to support the theoretical analysis for the proposed linear iteration algorithm of eq. (14). We fix $\gamma = 1$, $\varepsilon^2 = 1$, $N = 256$, and $\Delta t = 0.01$. Different values of the coefficient $\delta^2$ and the final time are used to compare the times for the linear iteration. By setting $10^{-9}$ as the tolerance of iteration error, we present the detailed results in Tab. 3. From Tab. 3, it is clear that the linear iteration error reaches a saturation after certain iteration stages. We observe that the time for the linear iteration increases with a decreasing value of $\delta^2$, which in turn implies that numerical implementation of the linear iteration algorithm of eq. (14) becomes more challenging with a smaller diffusion coefficient $\delta^2$. This result matches with our theoretical analysis in the proof of Theorem 1.

Table 3. Iterations of the propose algorithm at different final time with different $\delta^2$

<table>
<thead>
<tr>
<th>$\delta^2$</th>
<th>$T = 4$</th>
<th>$T = 5$</th>
<th>$T = 6$</th>
<th>$T = 7$</th>
<th>$T = 8$</th>
<th>$T = 9$</th>
<th>$T = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>0.1</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>0.01</td>
<td>28</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>23</td>
</tr>
</tbody>
</table>

Pattern formation

Localized structures, which can be described by real order parameter equations in the form of SH type of models, remain of great interest in the pattern-formation community. In fact, Many recent experimental and theoretical studies have focused exclusively on localized pattern formation [15]. From SH eq. (1), one expects to observe the single stripe becomes unstable by the emergence of undulations.

Here the parameters are $\gamma = 1$, $\varepsilon^2 = 2$, $\delta^2 = 1$. Take $\Omega = [0.40]^2$ and fix the spatial resolution as $N = 512$ and the time step as $\Delta t = 0.1$. The initial condition is a constant state in which we embed a curvy vertical stripe with the phase variable taking the value $u = 1$. The initial pattern evolves developing horizontal fingers that might bifurcate, see fig. 1.

The fig. 1(b) illustrates the manifestation of this undulations under the consideration of an infinitely long rod-like structure, to avoid border effects. Later, this undulated stripe is replaced by the emergence of facets that form a zigzag structure. However, the higher non-linear terms control the evolution of the single stripe, then the dynamics of initial zigzag is replaced by the growth of undulations without saturation as it is depicted in fig. 2. The numerical results are consistent with the experiments on this topic [14].
Conclusion

In this paper we simulated an energy-stable first-order numerical scheme for the SH equation with the Fourier pseudospectral approximation in space. Since the energy stability has to be considered, it leads to the non-linearity of the derived scheme. Although the global nature of the Fourier pseudo-spectral scheme makes a direct non-linear solver not feasible, we introduced a linear iteration algorithm and its contraction mapping property in a discrete $\| \cdot \|_4$ norm is also justified at a theoretical level. Moreover, various experiments, such as accuracy tests, efficiency of the iteration and pattern deformation are given to demonstrate efficiency and robustness of the proposed scheme.

References


