



139
035
THS

1
2006

LIBRARY
Michigan State
University

This is to certify that the
dissertation entitled

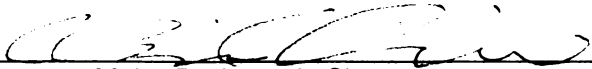
A fully explicit optimal two-stage numerical method
for solving reaction-diffusion-chemotaxis systems

presented by

Jui-Ling Yu

has been accepted towards fulfillment
of the requirements for the

PhD degree in Mathematics


Major Professor's Signature

5-6-05

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

ABSTRACT

A FULLY EXPLICIT OPTIMAL TWO-STAGE NUMERICAL SCHEME FOR SOLVING REACTION-DIFFUSION-CHEMOTAXIS SYSTEMS

By

Jui-Ling Yu

Jui-Ling Yu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2005

ABSTRACT

A fully explicit optimal two-stage numerical scheme for solving
reaction-diffusion-chemotaxis systems

By

Jui-Ling Yu

Reaction-diffusion-chemotaxis systems have been proved to be fairly accurate mathematical models for many pattern formation problems in chemistry and biology. These systems are important for computer simulations of the patterns, parameter estimations as well as analysis of the biological properties. In order to solve reaction-diffusion-chemotaxis systems, efficient and reliable numerical algorithms are essential for the pattern generations. In this thesis, a general reaction-diffusion-chemotaxis system is considered and specific numerical issues are discussed. We propose a fully explicit discretization combined with a variable optimal time step strategy for solving the reaction-diffusion-chemotaxis system. Theorems about stability and convergence of the algorithm are given to show that the algorithm is highly stable and efficient. Numerical experiment results are given for one testing problem and one real experimental problem.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and thanks to my advisor, Professor Chi-Chia Chiu for her constant and patient help, encouragement, and excellent advice. I would also like to thank Professors Michael Frazier, Mogun Tang, Chang-Yi Wang, and Farhad Jaberi for their time and valuable suggestions.

Thanks for people in the Lansing Buddhism Association, who offer their time and space for me to learn Buddha's teaching.

Thanks for **To my parents, brother, sister and my grandparents** - Chu Lee, Charles MacCluer who give me valuable suggestions and support.

Thanks for my friends, Cherry Lee, Yi-Chia Lin, Elif Sehin, Jaegwi Go, Hsin-Chuan Wei, Kuerak Chung, Jung-Chun Liu, EunJeong Kim, Steve Gargola, Derek Lorek, Danyu Zhu, Mei-Yu Tsai, Aloka, Chi-Chieh Chiang's family who helped me in many different ways.

Thanks for my lovely cousins who supported and comforted me when grandpa passed away.

Thanks for my cousin Maggie Lin.

Thanks for my students, who make me become a better teacher.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and thanks to my advisor, Professor ChiChia Chiu for her constant and patient help, encouragement, and excellent advice. I would also like to thank Professors Michael Frazier, Moxun Tang, Chang-Yi Wang, and Farhad Jaberl for their time and valuable suggestions.

Thanks for people in the Lansing Buddhism Association, who offer their time and space for me to learn Buddha's teaching.

Thanks for my teachers and professors Wei-Yueh Lin, Wei-Eihn Kuan, Min-Chu Lee, Charles MacCluer who give me valuable suggestions and support.

Thanks for my friends, Cherry Lee, Yi-Chia Lin, Elif Seckin, Jaegwi Go, Hsiu-Chuan Wei, Kuerak Chung, Jung-Chun Liu, EunJeong Kim, Steve Gagola, Derek Lorek, Danyu Zhu, Mei-Yu Tsai, Aloka, Chin-Chieh Chiang's family who helped me in many different ways.

Thanks for my lovely cousins who supported and comforted me when grandpa passed away.

Thanks for my cousin Maggie Lin.

Thanks for my students, who make me become a better teacher.

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES	viii
Introduction	1
1 A fully explicit discretization	7
2 An optimal two-stage Runge-Kutta method	13
2.1 Stability issue	15
3 An optimal two-stage numerical scheme	20
4 Stability and Convergence	27
5 Numerical Experiments	31
5.1 1D testing problem	31
5.1.1 Results for 1D experiment	37
5.2 Simulation of Adler's experiment	38
5.2.1 Chemotaxis phenomenon	39
5.2.2 Experiment results	41
5.2.3 Mathematical Model for pattern formation in <i>E.coli</i> experiment	46
5.2.4 The coefficient function	50
5.2.5 Intuitive explanation of pattern formation mechanism	51
5.2.6 An optimal two-stage numerical scheme	51
5.2.7 Results of numerical experiment	59
6 Conclusion	63
BIBLIOGRAPHY	65

5.8	The picture shows the movement of bacteria when the concentration of galactose is below the oxygen.	61
5.9	Relation of the amount of oxygen and the distance of the bands of bacteria.	62

LIST OF FIGURES

1.1	Method of Line	9
2.1	The solution of the MinMax problem.	18
5.1	Error between numerical simulation and analytic solution.	38
5.2	Photographs showing bands of <i>E. coli</i> in a capillary tube. In all the experiments reported here, capillary tubes were filled with a liquid medium, inoculated at one end with 200000 to 2000000 bacteria, and then closed at the ends with plug of agar and clay, all according to procedure described in full elsewhere. The tube were incubated horizontally at 37C. The origin, which is turbid because of the bacteria that have not moved out, is visible at the left, then the second band of bacteria, then the first band. Plugs at the ends are not shown. The concentration of galactose was 0.00025 mole per liter.	41
5.3	Utilization of oxygen by bands of bacteria in 0.00025 <i>M</i> galactose. Oxygen was measured polargraphically by inserting an oxygen needle electrode into the capillary tube in 4-mm steps when the first and second bands were visible where shown by arrows.	43
5.4	Utilization of 0.00025 <i>M</i> galactose. The tube contained C14- galactose (0.00025 mole/liter and 1000000 counts per minute per milliliter). At 4 or 8 hours, when the first and second bands were visible where shown by arrows, the tube was fractionated into ten compartments, each 8 mm long. the contents of each compartment were chromatographed on paper, with n-butanol, acetic acid, and water (12:3:5) as the solvent, and the radioactivity in the galactose region of the chromatogram was measured in a paper-strip counter.	44
5.5	Utilization of oxygen by bands of bacteria in 0.000005 <i>M</i> galactose. Measurements were made as in Figure 5.3 and Figure 5.4. Utilization of 0.000005 <i>M</i> galactose. The tube contained C14-galactose (0.000005 molde/liter and 200000 counts per minute per milliliter). At 4.75 hours when the first and second bands were visible where shown by the arrows.	45
5.6	The relation between the amount of galactose and the distance of the bands of bacteria.	59
5.7	This set of pictures simulates the movement of bands of bacteria when the concentration of galactose is higher than oxygen.	60

- 5.8 The picture shows the movement of bacteria when the concentration of galactose is below the oxygen. 61
- 5.9 Relation of the amount of oxygen and the distance of the bands of bacteria. 62

LIST OF TABLES

- 5.1 Number of iterations 37

LIST OF TABLES

5.1 Number of iterations.	37
-----------------------------------	----

A significant number of researches in molecular biology were done by Julius Adler and his co-workers around 1956. He tested the chemotaxis behavior of bacteria by using *Escherichia coli* and *Salmonella typhimurium* bacteria. His strategy of studying the chemotaxis phenomenon by using the simplest organism has provided the proof

Introduction

the mechanism of synthesis and control of biology macromolecules. In this thesis, we studied one of Adler's papers on the chemotaxis behavior of the *E. coli* bacteria because "the vast knowledge of its biochemistry and One of the important branches in mathematical biology is the study of pattern formation. In 1952, Alan Turing proposed the reaction-diffusion model as the chemical mechanism for biological patterns [25]. The idea was that the interaction between the passive diffusion and chemical reaction will involve into a spatiotemporal pattern.

movement during which the cells are reoriented to a new direction. The movement of Moreover, the motion of the biological cells is often influenced by the environment. This influence refers to chemotaxis, which is often known as directed random movement. The presence of chemotaxis provides a general tendency for bacteria to move toward the chemotaxattractant [17].

The study of chemotaxis phenomenon among bacteria began in the nineteenth century. In 1884, Wilhelm Pfeffer, Engelmann, and their associates discovered that bacteria move preferentially toward a concentration of oxygen, minerals and organic nutrients. These researches demonstrated chemotaxis by observing whether bacteria in a suspension gathered near or away a gas bubble which contains a certain amount of oxygen [4]. In 1893, Beijerinck discovered that bacteria move toward oxygen macroscopically. He concluded that the bacteria search for "a certain optimum concentration of oxygen." Later, Sherris, Preston, and Shoesmith and Baracchini and Sherris confirmed the results.

A significant number of researches in molecular biology were done by Julius Adler and his co-workers around 1966. He tested the chemotaxis behavior of bacteria by using *Escherichia coli* and *Salmonella typhimurium* bacteria. His strategy of studying the chemotaxis phenomenon by using the simplest organism has provided the proof for and inspired the study in the mechanism of synthesis and control of biology macromolecules. In this thesis, we studied one of Adler's papers on the chemotaxis behavior of the *E.coli* bacteria because "the vast knowledge of its biochemistry and genetics could be brought to bear on the problem. [18]"

The mobility of the *E.coli* and many other bacteria depends on swimming and tumbling. Swimming is a near straight run. Tumbling is a very rapid and jerky movement during which the cells are reoriented to a new direction. The movement of swimming (lasts for seconds) is interrupted by tumbling (lasts for tenths of seconds). Cells respond to the temporal and spatial changes in the environment by modulating the probability of tumbling. When cells encounter a favorable environment, they reduce the frequency of tumbling so that the swimming time in a favorable direction is increased. Therefore, a net migration along the desired direction is formed. That is chemotaxis [8].

The motion of swimming and tumbling is controlled by bacteria flagella which consists of a long filament. It is a left handed helix. The flagella turns either clockwise or counterclockwise as a propeller. The clockwise motion producing a pushing force results in the motion of swimming. The counterclockwise motion resulting in deformation of the filament produces tumbling [8].

Bacteria sense the change in the local concentration of chemicals through

membrane receptors binding chemical's molecules. The cells estimate the change in the gradient of chemicals by measuring the change in the receptor's occupation [16]. It results in a shift in the clockwise and counterclockwise motion in the flagella motor. The addition of an attractant results in the suppression of tumbling. The suppression time may last for a few seconds or minutes. It depends on the strength of the chemotaxis [8].

The general form of the reaction-diffusion-chemotaxis system is

$$\begin{aligned}\frac{\partial u}{\partial t} &= D_u \Delta u - \nabla(u\chi(a)\nabla a) + f(x, y, t) \\ \frac{\partial a}{\partial t} &= \nabla \cdot D_a \nabla a + g(a, u)\end{aligned}$$

$u(x, y, t)$, $a(x, y, t)$ represent the concentration of cells and the concentration of chemoattractant substrate respectively. $D_u \Delta u$ is the diffusion term or random walk, and which mainly characterize the movement of swimming. $D_u > 0$ is the constant diffusivity describing the speed of bacteria. $\nabla(u\chi(a)\nabla a)$ describes the biased random movement which is induced by external chemicals. It has been proven that the nonlinearity in the reaction term is a necessary condition for the pattern to be formed. No general analytical solutions are available at the present time. Thus, a numerical simulation provides a comprehensive understanding of the reaction-diffusion-chemotaxis systems and the mechanism of the underlying biology [17].

For the past years, numerical methods for reaction-diffusion systems have been well-studied: for example, the ADI type of scheme, the finite element method, monotone iterative methods, and the multigrid method. However, the development of numerical methods for solving reaction-diffusion-chemotaxis systems have not been thoroughly studied due to the additional nonlinear chemotaxis term [26]. Furthermore, most of the time, numerical methods are too complicated to be used

in practice. The idea of this thesis is to construct a numerical method, which is not only reliable but also easy to implement so that it can be applied to complex systems.

Many numerical methods have been studied to solve the PDEs: for example, the fully implicit method, the ADI method, and the explicit method [3, 23]. The explicit method is especially easy to implement and does not have to invert the matrix resulting from the implicit method. One of the approaches to get the explicit scheme is to use the method of line, which discretizes the spatial and temporal variables separably, changes PDEs to an ODE system, and solves the ODE system at each time level [1]. In this thesis, the standard finite difference and the upwind scheme are proposed for spatial discretization, and the spatial matrix obtained from it are sparse matrices. They are less expensive to integrate than the full matrix during the numerical simulation [11].

The convergence rate of the numerical solution strongly depends on the distribution of the eigenvalues of the spatial discretization. If the matrix is symmetric and positive definite, many efficient iterative methods can be achieved to obtain fast convergence. The spatial matrix we obtained from the finite difference scheme is not symmetric due to the chemotaxis term and the boundary condition. Therefore, it is necessary to develop an efficient numerical scheme to treat the nonsymmetric case.

The development of numerical methods for solving the ODE system are well-studied: for example, Euler's method and the n^{th} -stage Runge-Kutta method. The explicit Runge-Kutta method has advantages of low storage, high accuracy, flexible stability region and easy implementation [11]. However, since the method is explicit, the time step is often restrictive. It has become an unfavorable feature of the explicit

scheme. There are many research papers addressing the issue of improving the convergence rate for Runge-kutta method with different stages [11, 12, 13].

Since the spatial matrix of the ODE system is time dependent in the chemotaxis system, the stability region will vary by time. Since the size of the matrix is very large and the matrix is not symmetric, it is not possible to find the eigenvalues analytically, and using the existing numerical methods to estimate the distribution of eigenvalues are too expensive. Therefore, the time step has to be forced to be very small to satisfy all the stability conditions for the "time dependent" ODE system. In this thesis, we develop an efficient numerical method. The analytic form of explicit time can be easily calculated and the time step is locally optimal.

We study the properties of a particular Runge-Kutta method with $n=2$ for choosing the optimal time step. Since for n^{th} -stage Runge-Kutta method, the stability function is a n th degree polynomial function, the analysis for the MinMax problem becomes complicated and the analytic form of the optimal time step does not seem possible. Furthermore, introducing more stages does not guarantee the efficiency of the CPU time [12].

The generation of patterns depends on the distribution of eigenvalues of the ODE system. The ODE system contains eigenvalues with positive and negative real parts. We have to take these two parts into consideration during the numerical simulation. The stability issue has to be addressed since we want to make sure that patterns which are generated from the computation simulation are derived from the original PDEs, not due to numerical instability.

We propose an optimal two-stage method to solve the ODE system with time dependent spatial matrix. The method keeps the advantage of the two-stage Runge-Kutta method for its higher accuracy and easy implementation. In addition, the time steps are enlarged, also we can arrive at the explicit formulation for the time step. The numerical experiments show that the optimal two-stage scheme is accurate and converges faster than the standard two-stage Runge-Kutta method for the time evolution problem.

In Chapter 1, the fully explicit discretization for the spatial variables is discussed. The method of line and the upwind scheme for the chemotaxis term are the main topics of this chapter. The method we used to find the optimal time step for the standard two-stage Runge-Kutta method will be addressed in Chapter 2. In this chapter, we mainly discuss the method to solve a MinMax problem. In Chapter 3, a fully explicit optimal two-stage scheme is proposed to solve the ODE system. The consistency of the numerical method is investigated. Stability and Convergency are addressed in Chapter 4. Finally, the results of numerical experiment on the comparison between the standard two-stage Runge-Kutta method and our optimal two-stage scheme will be shown. A real problem will be tested against our numerical method.

CHAPTER 1

A fully explicit discretization

Many interesting patterns happen as bacteria consume nutrients, proliferate, perform random walks and the biased movement towards high concentration of chemicals [4, 10, 16, 21]. The time evolution of chemicals is described by reaction-diffusion-chemotaxis equations [20]. In these models, the microorganisms are simulating through 2D density, and a reaction-diffusion-chemotaxis equation of this density characterized their time evolution. The equation is coupled with other reaction-diffusion equations, since the chemotaxis substrate diffuses and produces by itself [16, 20].

The general reaction-diffusion-chemotaxis systems have the following form:

$$\begin{aligned}\frac{\partial u}{\partial t} &= D_u \Delta u - \nabla(u\chi(a)\nabla a) + f(x, y, t) \\ \frac{\partial a}{\partial t} &= \nabla \cdot D_a \nabla a + g(a, u).\end{aligned}$$

$u(x, y, t)$, $a(x, y, t)$ represent the concentration of cells and the concentration of chemoattractant substrate respectively. $D_u \Delta u$ is the diffusion term or random walk, and which mainly characterize the movement of swimming. $D_u > 0$ is the constant

diffusivity describing the speed of bacteria. $\nabla(u\chi(a)\nabla a)$ describes the biased random movement which is induced by external chemicals. The movement is biased either along the gradient direction or in the opposite direction of the gradient [16]. $\chi(a) \geq 0$ is the chemotaxis parameter function describing the speed of chemotaxis, $f(x, y, t)$ is the reaction term for cells, which indicate the growth rate of cells.

For simplicity, we use a 1D reaction-diffusion-chemotaxis equation to demonstrate the analysis of our numerical scheme. We also assume boundaries are periodic with period ℓ , and the initial value is known,

$$\frac{\partial u}{\partial t} = d_u u_{xx} - (u\chi(a)a_x)_x + f(x, t), x \in [0, \ell], \quad (1.1)$$

the first-order upwind scheme for the gradient term. The assumption of using an upwind scheme [19] is crucial in the development of our numerical method. It ensures the stability of the numerical scheme.

The equation (1.1) is generally nonlinear so the analytical solution is difficult to solve. Therefore, numerical simulation becomes crucial to investigating the mechanism of the underlying biology. To solve this equation numerically, a very common method is to apply the method of lines [15] which discretizes spatial and temporal variables separately, then turns the PDE into an ODE system and solves the ODE system at each time level.

where $(qu_x)_x$ is approximated by the second-order centered difference

Take a uniform mesh of $\Omega = [0, \ell]$ with mesh size $h = \ell/N$. Let $x_i \in \Omega$ be a grid point, then $x_{i+1} = x_i + h$. A time step of size $t^{n+1} - t^n$ will be denoted as Δt , and we adopt the standard notation $u_i^n \approx u(x_i, t^n)$, etc. We use the standard second-order centered difference to approximate the Laplace operator, and

where

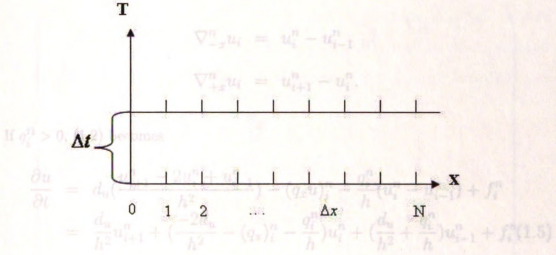


Figure 1.1: Method of Line

the first-order upwind scheme for the gradient operator. The assumption of using an upwind scheme [19] is crucial in the development of our numerical method. It ensures the stability of the numerical scheme.

Let $u^n = [u_i^n]$ and $f^n = [f_i^n]$ be a vector obtained by the usual ordering. Then

(1.5) can be represented by the matrix form as follows:

For simplicity, let us denote $\chi(a)\nabla a$ by $q(t)$. Then,

$$\begin{aligned} \frac{\partial u}{\partial t} &= d_u u_{xx} - (q(t)u)_x + f(x, t) \\ &= d_u \left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \right) - (q_x u + q u_x)_i^n + f_i^n, \end{aligned} \quad (1.2)$$

where $(q u_x)_i^n$ is approximated by the first-order upwind scheme, i.e.

$$(q u_x)_i^n \approx \frac{q_i^n + |q_i^n|}{2h} \nabla_{-x}^n u_i \quad (q_i^n > 0) \quad (1.3)$$

$$+ \frac{q_i^n - |q_i^n|}{2h} \nabla_{+x}^n u_i, \quad (q_i^n < 0) \quad (1.4)$$

where

$$\nabla_{-x}^n u_i = u_i^n - u_{i-1}^n$$

$$\nabla_{+x}^n u_i = u_{i+1}^n - u_i^n.$$

If $q_i^n > 0$, (1.2) becomes

$$\begin{aligned} \frac{\partial u}{\partial t} &= d_u \left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \right) - (q_x u)_i^n - \frac{q_i^n}{h} (u_i^n - u_{i-1}^n) + f_i^n \\ &= \frac{d_u}{h^2} u_{i+1}^n + \left(\frac{-2d_u}{h^2} - (q_x)_i^n - \frac{q_i^n}{h} \right) u_i^n + \left(\frac{d_u}{h^2} + \frac{q_i^n}{h} \right) u_{i-1}^n + f_i^n \end{aligned} \quad (1.5)$$

The periodic boundary condition can be approximated by

$$\begin{cases} u_{N-1} = u_{-1} \\ u_{N+1} = u_1 \end{cases}$$

Let $u^n = [u_i^n]$ and $f^n = [f_i^n]$ be a vector obtained by the usual ordering. Then

(1.5) can be represented by the matrix form as follows:

$$\frac{du}{dt} = M(t^n)u^n + f^n. \quad (1.6)$$

Let $\omega_i = \frac{-2d_u}{h^2} - (q_x)_i^n - \frac{q_i^n}{h}$ $i = 0, 1, 2, \dots, N$, then $M(t^n) =$

step constraint. Our idea is to adjust the appropriate time step so that the spectral radius of the spatial matrix in the ODE system is minimized. Therefore, optimal convergence can be achieved [11, 13]. This leads to the solution of the MinMax problem $\frac{d_y}{h^2} + \frac{q_0^n}{h}$. Detailed analysis of the MinMax problems will be addressed on Chapter Two.

$$\begin{pmatrix} \omega_0 & \frac{d_y}{h^2} & & & \frac{d_y}{h^2} + \frac{q_{-1}^n}{h} \\ \frac{d_y}{h^2} + \frac{q_0^n}{h} & \omega_1 & \frac{d_y}{h^2} & & \\ & \frac{d_y}{h^2} + \frac{q_1^n}{h} & \omega_2 & \frac{d_y}{h^2} & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \frac{d_y}{h^2} & 0 & \frac{d_y}{h^2} + \frac{q_{N-1}^n}{h} & \omega_N \end{pmatrix}$$

a $(N+1) \times (N+1)$ matrix. Observe that the spatial matrix is time dependent.

By the method of line, we have turned the PDE into an ODE system. Many numerical methods have been studied to approximate the temporal derivative in the ODE system. For example, Euler's method, and the n^{th} stage Runge-Kutta method [15]. Among all of these numerical schemes, the explicit Runge-Kutta method has been widely applied. It has the advantages of high accuracy, flexible stability region, low storage, and ease of implementation [11]. However, step size has been limited due to the nature of the explicit method. Due to this fact, we are forced to choose a smaller step size than is needed to guarantee the convergence of the numerical scheme. Oftentimes, the restricted time step increase the CPU time which becomes an unfavorable feature of the explicit scheme. Many researchers have been working on improving the performance of explicit Runge-Kutta method by releasing the time

If the we set the reaction vector as zero, the system becomes

$$\frac{du}{dt} = Mu. \quad (2.2)$$

CHAPTER 2

Large Runge-Kutta integration is often (2.2), we get

$$u^{(n)} = u^0 + \Delta t M u^{(n-1)} \quad (2.3a)$$

An optimal two-stage Runge-Kutta method

The parameter Δt usually stands for the time step. Then we write (2.3) as a

standard iterative method. It is usually written as

For the n th stage explicit RK method, the stability function $R(z)$ is a function with degree n . When $n \geq 3$, the formulation of $R(z)$ becomes complicated and the analytic form of the optimal time step does not seem possible. Furthermore, it is too expensive to update the time step by using the modified projected Lagrangian Algorithm and introducing more stages of the optimal RK method does not guarantee the efficiency of the CPU time [12]. Therefore, We focus on applying the optimal two-stage RK method to solve the chemotaxis system under the stability consideration.

The main purpose of this section is to obtain the optimal time step for the standard two-stage Runge-Kutta integration. First of all, let us assume the spatial discretization matrix M and the reaction vector in the ODE system is time independent.

$$\frac{du}{dt} = Mu + f. \quad (2.1)$$

The generation of pattern

ODE system. The system consists

If we set the reaction vector as zero, the system becomes
Therefore, the stability issue has to be addressed.

$$\frac{du}{dt} = Mu. \quad (2.2)$$

Let us use the standard two-stage Runge-Kutta integration to solve (2.2), we get

2.1 Stability issue

$$u^{(0)} = u^n + \Delta t Mu^n \quad (2.3a)$$

Let us consider the next

$$u^{n+1} = u^n + \frac{\Delta t}{2}(Mu^n + Mu^{(0)}). \quad (2.3b)$$

$$\frac{du}{dt} = Mu^n + f^n. \quad (2.7)$$

The parameter Δt usually stands for the time step. When we treat (2.3) as a
where M is a constant matrix, f^n is a constant vector.
standard iterative method, it is actually a one-step method:

Let v^0 denote the perturbation of the initial value u^0 . Then the solution vector v^{n+1} after $n+1$ iterations and the exact solution u^{n+1} are given by the following:

$$\begin{aligned} u^{n+1} &= u^n + \frac{\Delta t}{2}Mu^n + \frac{\Delta t}{2}M(u^n + \Delta tMu^n) \\ &= (I + \Delta tM + \frac{\Delta t^2}{2}M^2)u^n. \end{aligned} \quad (2.4)$$

Let P denote the amplification factor, i.e., (2.8)

$$P(\Delta t) = I + \Delta tM + \frac{\Delta t^2}{2}M^2, \quad (2.5)$$

Let $u^{n+1} = v^{n+1} + u^{n+1}$ denote the exact solution after $n+1$ iterations,
we have
then (2.4) can be written as

$$u^{n+1} = P(\Delta t)u^n \quad (2.10)$$

From equation (2.10), we know (2.6)

$$u^{n+1} = P^{n+1}(\Delta t)u^0.$$

of the amplification factor is how fast $P(\Delta t)$ grows. The amplification will

The generation of patterns depends on the distribution of eigenvalues of the
ODE system. The system contains eigenvalues with positive and negative real parts.

We have to take care of these two parts during the numerical simulation processes. Therefore, the stability issue has to be addressed.

2.1 Stability issue

Let us consider the nonhomogeneous ODE system,

$$\frac{du}{dt} = Mu^n + f^n, \quad (2.7)$$

Note that the characteristic polynomial $\lambda(\Delta t, \mu)$ is a function of time step and where M is a constant matrix, f^n is a constant vector.

Let v^0 denote the perturbation of the initial condition u^0 . Then the solution vector v^{n+1} after n th iterations and the exact solution u^{n+1} satisfy the following:

$$u^{n+1} = P(\Delta t)u^n + (1 + \Delta t + \frac{\Delta t^2}{2})f^n \quad (2.8)$$

$$v^{n+1} = P(\Delta t)v^n + (1 + \Delta t + \frac{\Delta t^2}{2})f^n. \quad (2.9)$$

Let $\epsilon^{n+1} = v^{n+1} - u^{n+1}$ denotes the perturbation error after $n+1$ iterations,

we have

$$\epsilon^{n+1} = p(\Delta t)\epsilon^n = p^{n+1}(\Delta t)\epsilon^0. \quad (2.10)$$

From equation (2.10), we know that the error ϵ^{n+1} will be bounded if the norm of the amplification factor is less than one. Nevertheless, the iterative procedure will converge, if and only if, the spectral radius of the amplification factor is less than a unity (Issacson and Keller 1966). Moreover, the convergence speed of the iterative

procedure depends on the size of the spectral radius; if the size of the spectral radius is small, the iterative procedure will converge faster. This gives rise to formation of the MinMax problem.

The characteristic polynomial of the amplification factor is

$$\lambda(\Delta t, \mu) = 1 + \Delta t\mu + \frac{\Delta t^2}{2}\mu^2,$$

where μ is an eigenvalue of M .

Note that the characteristic polynomial $\lambda(\Delta t, \mu)$ is a function of time step and eigenvalues.

Definition 1 Let $\mu = a + bi$. Define $f(a, b, \Delta t) = |\lambda(\Delta t, \mu)|^2$.

Theorem 2 If b is bounded, then $f(a, b, \Delta t) = f_a^2(\Delta t) = (1 + a\Delta t + \frac{1}{2}a^2\Delta t^2)^2 + O(\Delta t^3)$.

Proof.

$$\begin{aligned} f(a, b, \Delta t) &= |1 + \Delta t(a + bi) + \frac{\Delta t^2}{2}(a + bi)^2|^2 \\ &= |1 + \Delta t(a + bi) + \frac{\Delta t^2}{2}(a^2 - b^2 + 2abi)|^2 \\ &= [(1 + a\Delta t + \frac{\Delta t^2}{2}(a^2 - b^2))]^2 + (\Delta tb + \Delta t^2 ab)^2 \\ &= (1 + a\Delta t)^2 + \Delta t^2(a^2 - b^2)(1 + a\Delta t) + \frac{\Delta t^4}{4}(a^2 - b^2)^2 \\ &\quad + b^2\Delta t^2 + O(\Delta t^3) \\ &= (1 + a\Delta t)^2 + a^2\Delta t^2(1 + a\Delta t) - b^2\Delta t^2 + \frac{\Delta t^4}{4}a^4 \end{aligned}$$

(2.11)

Proof. The proof is as follows. The truncation error is a polynomial. \square

$$+ b^2 \Delta t^2 + O(\Delta t^3) + O(\Delta t^4)$$

$$= (1 + a\Delta t + \frac{1}{2}a^2\Delta t^2)^2 + O(\Delta t^3) \quad (2.12)$$

Theorem 3 shows that the stability region is the set of z such that $|1 + az + \frac{1}{2}a^2z^2| \geq 1$ and smallest real part of the eigenvalues is λ .

\square

Since the truncation error of the standard two-stage Runge-Kutta method is $O(\Delta t^3)$, the higher order terms appeared in can be ignored. Moreover, the $\min_{\Delta t} \max_a f(a, b, \Delta t)$ can be simplified as $\min_{\Delta t} \max_a f_a(\Delta t)$ since we are only interested in the case where $\frac{1}{2} \leq f_a(\Delta t) < 1$ when $a < 0$. It is worth noting that the optimal two-stage scheme matches well with the theoretical analysis in the following sense: $e^{a\Delta t}$ is numerically represented by $f_a(\Delta t) = 1 + a\Delta t + \frac{1}{2}a^2\Delta t^2$. If $a > 0$, the accuracy of the numerical solution is the only criteria for us to choose the time step. We have to choose Δt small enough so that $f_a(\Delta t)$ will accurately integrate e^{aT} . The stability issue is not important here, since e^{aT} is growing and $a\Delta t$ will not lie in the stability region. If $a < 0$, we have to consider both accuracy and stability of the numerical scheme. However, they often imply each other. When $a < 0$, $f_a(\Delta t)$ corresponds to a decay mode. We have to choose the right time step so that $f_a(\Delta t)$ is less than a unity. In other words, the choice of the time step has to restrict $a\Delta t$ to stay in the stability region so that $f_a(\Delta t)$ is a correct integration of e^{aT} [15].

Lemma 3 For $a < 0$, f_a has the following properties:

- (a) $f_a(0) = 1$ and $f_a(-\frac{2}{a}) = 1$.
- (b) $f'_a(0) = a$.
- (c) $f_a(\Delta t) \geq 0$, $\Delta t > 0$, and $f_a(\Delta t) < 1$, $\Delta t \in (0, -2/a)$.
- (d) $\min_{\Delta t \in R} f_a(\Delta t)$ occurs at $\Delta t = -1/a$.
- (e) $\min_{\Delta t \in R} f_a(\Delta t) = 1/2$.

Proof. The proof is straightforward using the properties of the second-degree polynomials. \square

Theorem 3 shows that the optimal time step can be determined by the largest and smallest real part of the eigenvalues of the ODE system.

Theorem 4 *If $a < 0$ and $-a_{\max} < a < -a_{\min}$, where $a_{\min} > 0$ and $a_{\max} > 0$ are constant then $\min_{\Delta t} \max_{a \in [-a_{\max}, -a_{\min}]} |f_a(\Delta t)|^2$ is obtained at $\Delta t_{\text{opt}} = \frac{2}{a_{\min} + a_{\max}}$.*

Proof. $f'_a(\Delta t) = a + a^2 \Delta t = 0$. Let $f'_a(\Delta t) = 0$, get $\Delta t = -\frac{1}{a}$.

Moreover, $f(-\frac{1}{a}) = \frac{1}{2} > 0$.

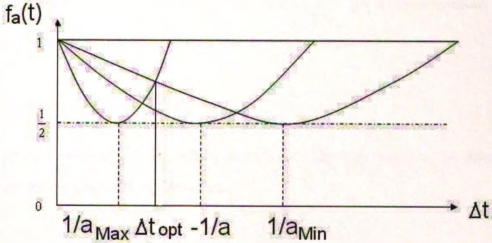


Figure 2.1: The solution of the MinMax problem.

From Lemma (3) and the graph, it is clear that the optimal time step will be obtained if

$$f_{-a_{\min}}(\Delta t_{\text{opt}}) = f_{-a_{\max}}(\Delta t_{\text{opt}}).$$

i.e.

$$1 - a_{\min} \Delta t_{\text{opt}} + \frac{1}{2} a_{\min}^2 \Delta t_{\text{opt}}^2 = 1 - a_{\max} \Delta t_{\text{opt}} + \frac{1}{2} a_{\max}^2 \Delta t_{\text{opt}}^2$$

$$\frac{1}{2} (a_{\max}^2 - a_{\min}^2) \Delta t_{\text{opt}} = a_{\max} - a_{\min}.$$

which implies

$$\Delta t_{\text{opt}} = \frac{2}{a_{\min} + a_{\max}}.$$

□
scheme

In this section, we will consider the numerical solution of systems with time dependent coefficients.

An optimal two-stage scheme will be discussed in this section.

Let us consider the following two-stage scheme:

$$\begin{aligned} u^{(1)} &= u^n + \tau \Phi(u^n, u^n) \\ u^{n+1} &= u^{(1)} + \tau \Psi(u^n, u^{(1)}) \end{aligned}$$

There are two parameters τ and θ in the scheme. The two-stage scheme is consistent if

being $O(\Delta t^3)$. Meanwhile, the time step will be locally optimal according to the analysis in Chapter 2. Treating (3.2) as a standard iterative method, the two-stage scheme becomes

CHAPTER 3

$$\begin{aligned} u^{n+1} &= u^n + \alpha_2 \Delta t (M(t^n) + M(t^{n+1})) [u^n + \alpha_1 \Delta t (M(t^n) + M(t^{n+1})) u^n] \\ &= [I + \alpha_2 \Delta t (M(t^n) + M(t^{n+1})) + \alpha_1 \alpha_2 \Delta t^2 (M(t^n) + M(t^{n+1}))^2] u^n. \end{aligned}$$

An optimal two-stage numerical

Since $M(t^{n+1}) = M(t^n) + \Delta t M'(t^n) + O(\Delta t^2)$,

scheme

$$\begin{aligned} u^{n+1} &= [I + \alpha_2 \Delta t (M(t^n) + M(t^n + \Delta t M'(t^n) + O(\Delta t^2))) \\ &\quad + \alpha_1 \alpha_2 \Delta t^2 (M(t^n) + M(t^n + \Delta t M'(t^n) + O(\Delta t^2)))^2] u^n \end{aligned}$$

In this section, we will extend ideas from chapter 2 to the nonhomogeneous ODE systems with time dependent spatial matrix, i.e.,

$$u'(t) = M(t)u(t) + f(t). \quad (3.3)$$

Moreover, Taylor expansion $\frac{du}{dt} = M(t)u + f(t). \quad (3.1)$

An optimal two-stage scheme will be constructed. The consistency of the numerical scheme will be discussed in this section.

Let us consider the following two-stage scheme for solving (3.1)

$$u^{(1)} = u^n + \alpha_1 \Delta t [M(t^n) + M(t^{n+1})] u^n \quad (3.2a)$$

$$u^{n+1} = u^n + \alpha_2 \Delta t [M(t^n) + M(t^{n+1})] u^{(1)}. \quad (3.2b)$$

There are two parameters α_1 and α_2 . We are looking for α_1 and α_2 so that the two-stage scheme is consistent in time and space, i.e., having the truncation error

being $O(\Delta t^3)$. Meanwhile, the time step will be locally optimal according to the analysis in Chapter 2. Treating (3.2) as a standard iterative method, the two-stage scheme becomes

$$\begin{aligned} u^{n+1} &= u^n + \alpha_2 \Delta t (M(t^n) + M(t^{n+1})) [u^n + \alpha_1 \Delta t (M(t^n) + M(t^{n+1})) u^n] \\ &= [I + \alpha_2 \Delta t (M(t^n) + M(t^{n+1})) + \alpha_1 \alpha_2 \Delta t^2 (M(t^n) + M(t^{n+1}))^2] u^n. \end{aligned}$$

Therefore, we can set up the optimal two-stage scheme as the following:

Since $M(t^{n+1}) = M(t^n) + \Delta t M'(t^n) + O(\Delta t^2)$,

$$\begin{aligned} u^{n+1} &= [I + \alpha_2 \Delta t (M(t^n) + M(t^n) + \Delta t M'(t^n) + O(\Delta t^2))] \\ &\quad + \alpha_1 \alpha_2 \Delta t^2 (M(t^n) + M(t^n) + \Delta t M'(t^n) + O(\Delta t^2))^2] u^n \\ &= [I + 2\alpha_2 \Delta t M(t^n) + \alpha_2 \Delta t^2 M'(t^n) \\ &\quad + 4\alpha_1 \alpha_2 \Delta t^2 M(t^n)^2 + O(\Delta t^3)] u^n. \end{aligned} \quad (3.3)$$

Moreover, Taylor expansion of u at time t^n is

$$\begin{aligned} u(t^{n+1}) &= u(t^n) + u'(t^n) \Delta t + \frac{u''(t^n)}{2} \Delta t^2 + O(\Delta t^3) \\ &= u(t^n) + \Delta t M(t^n) u(t^n) + \frac{\Delta t^2}{2} (M'(t^n) u(t^n) \\ &\quad + M(t^n) u'(t^n)) + O(\Delta t^3) \end{aligned}$$

Proof.

$$\begin{aligned} &= u(t^n) + \Delta t M(t^n) u(t^n) + \frac{\Delta t^2}{2} (M'(t^n) u(t^n) \\ &\quad + M(t^n)^2 u(t^n)) + O(\Delta t^3) \\ &= [I + \Delta t M(t^n) + \frac{\Delta t^2}{2} M'(t^n) \\ &\quad + \frac{\Delta t^2}{2} M(t^n)^2] u(t^n) + O(\Delta t^3). \end{aligned} \quad (3.4)$$

Matching coefficients from (3.3),(3.4) get

$$\alpha_1 = \frac{1}{4},$$

and

$$\alpha_2 = \frac{1}{2}.$$

Therefore, we can set up the optimal two-stage scheme as the following:

$$u^{(1)} = u^n + \frac{1}{4}\Delta t[M(t^n) + M(t^{n+1})]u^n \quad (3.5a)$$

$$u^{n+1} = u^n + \frac{1}{2}\Delta t[M(t^n) + M(t^{n+1})]u^{(1)}. \quad (3.5b)$$

Note that the coefficient of the Euler's predictor is 1, and the coefficient of Trapezoid corrector is $\frac{1}{2}$ [15]. (3.5) is not a standard two-stage Runge-Kutta method. The consistence of the numerical scheme will be derived in Theorem 5, which shows that the difference between the differential equations and the two-stage scheme we presented here can be controlled. Because the analysis for the 2D case is similar to the case for 1D, for simplicity, the 1D model equation is presented in the proof.

Theorem 5 (Convergence) Assume that $\|M^n + M^{n+1}\| \leq \alpha$ for all n , let $\|e^n\| = \|u^n - u(t^n)\|$, and $(n+1)\Delta t \leq T$. If $\|e^0\| = 0$ then $\|e^n\| \leq O(h) + O(\Delta t^2)$.

Proof.

The general reaction-diffusion-chemotaxis equation is

$$\frac{\partial u}{\partial t} = d_u u_{xx} - (q(x)u)_x + f(x, t). \quad (3.6)$$

Since the upwind scheme is of order one in space, the finite difference discretization

gives

$$\begin{aligned}
 \frac{\partial u}{\partial t} &= d_u \left(\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \right) - \frac{q_i^n - q_{i-1}^n}{h} u_i \\
 &\quad - \frac{q_i^n + |q_i^n|}{2h} \nabla_{-x}^n u_i - \frac{q_i^n - |q_i^n|}{2h} \nabla_{+x}^n u_i + f_i^n + O(h) \\
 &= \frac{d_u}{h^2} u_{i+1}^n + \left(\frac{-2d_u}{h^2} \right. \\
 &\quad \left. - \frac{q_i^n - q_{i-1}^n}{h} - \frac{q_i^n + |q_i^n|}{2h} \nabla_{-x}^n - \frac{q_i^n - |q_i^n|}{2h} \nabla_{+x}^n \right) u_i + \frac{d_u}{h^2} u_{i-1}^n + f_i^n + O(h)
 \end{aligned}$$

where $i = 0, 1, 2, \dots, N$.

Therefore,

$$\frac{\partial u}{\partial t} = M(t^n)u^n + f^n + O(h). \quad (3.7)$$

Applying the two-stage scheme to (3.1), we get

$$u^{(1)} = u^n + \frac{1}{4} \Delta t [(M(t^n) + M(t^{n+1}))u^n + f(t^n) + f(t^{n+1})] \quad (3.8)$$

$$u^{n+1} = u^n + \frac{1}{2} \Delta t [(M(t^n) + M(t^{n+1}))u^{(1)} + f(t^n) + f(t^{n+1})] \quad (3.9)$$

Viewing two-stage scheme as a standard iterative method, we get

$$\begin{aligned}
 u^{n+1} &= [I + \Delta t \left(\frac{M^n + M^{n+1}}{2} \right) + \frac{\Delta t}{2} \left(\frac{M^n + M^{n+1}}{2} \right)^2] u_n \\
 &\quad + \frac{\Delta t}{2} [(f^n + f^{n+1}) + \frac{\Delta t}{4} (M^n + M^{n+1})(f^n + f^{n+1})].
 \end{aligned}$$

The Taylor expansion of u at t^n is

$$u(t^{n+1}) = u(t^n) + u'(t^n) \Delta t + \frac{u''(t^n)}{2} \Delta t^2 + O(\Delta t^3)$$

Assuming there is no rounding,

$$\begin{aligned}
f^n = f(t^n) &= u(t^n) + \Delta t(M(t^n)u(t^n) + f(t^n) + O(h)) \\
&+ \frac{\Delta t^2}{2}(M'(t^n)u(t^n) + M(t^n)^2 f(t^n) + M(t^n)f(t^n) \\
&+ f(t^n)' + O(h)) + O(\Delta t^3) \\
&+ \frac{\Delta t^2}{2}M^n f^n + \Delta t f^n + \frac{\Delta t^2}{2}f_n' + \Delta t O(h) + \frac{\Delta t^2}{2}O(h)
\end{aligned}$$

Since $M(t^n) = M(t^{n+1}) - M'(t^n)\Delta t + O(\Delta t^2)$,

$$\begin{aligned}
u(t^{n+1}) &= u(t^n) + \frac{\Delta t}{2}M(t^n)u(t^n) + \frac{\Delta t}{2}(M(t^{n+1}) - M'(t^n)\Delta t \\
&+ O(\Delta t^2))u(t^n) + \Delta t f^n + \Delta t O(h) + \frac{\Delta t^2}{2}(M'(t^n)u(t^n) \\
&+ M(t^n)^2 u(t^n) + M(t^n)f^n + (f^n)' + O(h)) + O(\Delta t^3) \\
&= u(t^n) + \frac{\Delta t}{2}(M(t^n) + M(t^{n+1}))u(t^n) - \frac{\Delta t^2}{2}M(t^n)'u(t^n) \\
&+ O(\Delta t^3) + \Delta t f^n + \Delta t O(h) + \frac{\Delta t^2}{2}(M'(t^n)u(t^n) + M(t^n)^2 u(t^n)) \\
&+ \frac{\Delta t^2}{2}M(t^n)f^n + \frac{\Delta t^2}{2}(f^n)' + \frac{\Delta t^2}{2}O(h) + O(\Delta t^3) \\
&= u(t^n) + \frac{\Delta t}{2}(M(t^n) + M(t^{n+1}))u(t^n) + \frac{\Delta t^2}{2}M(t^n)^2 u(t^n) + O(\Delta t^3) \\
&+ \frac{\Delta t^2}{2}M(t^n)f^n + \Delta t f^n + \Delta t O(h) + \frac{\Delta t^2}{2}O(h) + \frac{\Delta t^2}{2}(f^n)' \\
&= u(t^n) + \frac{\Delta t}{2}(M(t^n) + M(t^{n+1}))u(t^n) + \frac{\Delta t^2}{2}\left(\frac{M(t^n)}{2} \right. \\
&+ \frac{M(t^{n+1})}{2} - \frac{M'(t^n)}{2}\Delta t + O(\Delta t^2))u(t^n) + O(\Delta t^3) \\
&+ \frac{\Delta t^2}{2}M(t^n)f^n + \Delta t f^n + \Delta t O(h) + \frac{\Delta t^2}{2}O(h) + \frac{\Delta t^2}{2}(f^n)' \\
&= (I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M(t^n) + M(t^{n+1})^2)u(t^n) + O(\Delta t^3) \\
&+ \frac{\Delta t^2}{2}M(t^n)f^n + \Delta t f^n + \frac{\Delta t^2}{2}(f^n)' + \Delta t O(h) + \frac{\Delta t^2}{2}O(h).
\end{aligned}$$

Assuming there is no rounding error, i.e.,

$f^n = f(t^n)$, $M(t^n) = M^n$, and $e^{n+1} = u(t^{n+1}) - u^n + 1$, then

$$\begin{aligned}
e^{n+1} &= [I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M^n + M^{n+1})^2]e^n + O(\Delta t^3) \\
&+ \frac{\Delta t^2}{2}M^n f^n + \Delta t f^n + \frac{\Delta t^2}{2}f'_n + \Delta t O(h) + \frac{\Delta t^2}{2}O(h) \\
&- \frac{\Delta t}{2}[f^n + f^{n+1} + \frac{\Delta t}{4}(M^n + M^{n+1})(f^n + f^{n+1})] \\
&= [I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M^n + M^{n+1})^2]e^n + O(\Delta t^3) + \Delta t O(h) \\
&+ \frac{\Delta t^2}{2}O(h) + \Delta t f^n + \frac{\Delta t^2}{2}(f^n)' - \frac{\Delta t}{2}(f^n + f^{n+1}) \\
&+ \frac{\Delta t^2}{2}M^n f^n - \frac{\Delta t^2}{8}(M^n + M^{n+1})(f^n + f^{n+1}) \\
&= [I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M^n + M^{n+1})^2]e^n + O(\Delta t^3) + \Delta t O(h) \\
&+ \Delta t f^n + \frac{\Delta t^2}{2}(f^n)' - \frac{\Delta t}{2}f^n - \frac{\Delta t}{2}(f^n + (f^n)'\Delta t + O(\Delta t^2)) \\
&+ \frac{\Delta t^2}{2}M^n f^n - \frac{\Delta t^2}{8}(2M^n + O(\Delta t))(2f^n + O(\Delta t)) \\
&= [I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M^n + M^{n+1})^2]e^n + O(\Delta t^3) + \Delta t O(h) \\
&= [I + \frac{\Delta t}{2}(M^n + M^{n+1}) + \frac{\Delta t^2}{8}(M^n + M^{n+1})^2]^{n+1}e^0 \\
&+ (n+1)\Delta t(O(h) + O(\Delta t^2)).
\end{aligned}$$

Since $(n+1)\Delta t \leq T$ and $e^0 = 0$, it implies

$$\|e^n\| \leq O(h) + O(\Delta t^2).$$

This completes the proof. \square

Let $J_n = \frac{M(t^n) + M(t^{n+1})}{2}$, then the two-stage scheme for the homogeneous problem can be written as

$$u^{n+1} = (I + \Delta t J_n + \frac{1}{2} \Delta t^2 J_n^2) u^n.$$

It follows from the analysis in chapter 2 that the optimal time step will be obtained if we set $\Delta t_{opt} = \frac{2}{a_{min} + a_{max}}$. Let $\{a_i\}_{i=1}$ be the real parts of the eigenvalues of J_n , then

Stability and Convergence

$$a_{min} = \min_{a_i} < 0 \{a_i\}$$

$$a_{max} = \max_{a_i} > 0 \{a_i\}.$$

In this section, we will investigate the stability and convergence of the optimal two-stage scheme.

Observe that the matrix $I + \Delta t J_n + \frac{\Delta t^2}{2} J_n^2$ depends on time. The stability region varies as the numerical scheme is marching on time. In practice, the optimal time step will be updated according to the distribution of the eigenvalues of the matrix $I + \Delta t J_n + \frac{\Delta t^2}{2} J_n^2$ at each time level. We call the two-stage scheme an optimal scheme in this sense.

Proof. Let $\|e^n\| = \|u(t_n) - u^{(2)}\|$. From the proof in Theorem 3, we get

$$\begin{aligned} \|e^{n+1}\| &\leq \|I + \frac{\Delta t}{2}(M_n + M_{n+1}) + \frac{\Delta t^2}{8}(M_n + M_{n+1})^2\| \|e^n\| + O(\Delta t^3) \\ &\quad + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2) \|e^n\| + O(\Delta t^3) + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^2 \|e^{n-1}\| \\ &\quad + (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)(O(\Delta t^3) + \Delta t O(h)) + O(\Delta t^3) + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n+1} \|e^0\| \end{aligned}$$

CHAPTER 4

Stability and Convergence

In this section, we will investigate the stability and convergence of the optimal two-stage scheme.

Theorem 6 (Convergence) Assume that $\|M^k + M^{k+1}\| \leq \alpha$ for all $k + 1 \leq N$, ($t_N \leq T, t_{N+1} > T$) and $t_n = n\Delta t \leq T$. If $\|u(t_0) - u^0\| = 0$, then $\|u(t_n) - u^n\| \leq C(T)(O(h) + O(\Delta t^2))$.

Proof. Let $\|e^n\| = \|u(t_n) - u^n\|$. From the proof in Theorem 3, we get

$$\begin{aligned} \|e^{n+1}\| &\leq \|I + \frac{\Delta t}{2}(M_n + M_{n+1}) + \frac{\Delta t^2}{8}(M_n + M_{n+1})^2\| \|e^n\| + O(\Delta t^3) \\ &\quad + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2) \|e^n\| + O(\Delta t^3) + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2) \|e^{n-1}\| \\ &\quad + (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)(O(\Delta t^3) + \Delta t O(h)) + O(\Delta t^3) + \Delta t O(h) \\ &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n+1} \|e^0\| \end{aligned}$$

$$\begin{aligned}
& + \left[\left(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2\right)^n + \left(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2\right)^{n-1} \right. \\
& + \left. \cdots + 1 \right] (O(\Delta t^3) + \Delta t O(h)).
\end{aligned}$$

Since $\|e^0\| = 0$, and

$$\begin{aligned}
1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2 &= 1 + \left(\frac{\alpha}{2} + \frac{\Delta t}{8}\alpha^2\right)\Delta t \\
&\leq 1 + \beta\Delta t,
\end{aligned}$$

therefore,

$$\left(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2\right)^{n+1} \leq e^{\beta T}.$$

Moreover, since

$$\frac{\alpha}{2} + \frac{\alpha^2}{8}\Delta t \geq \frac{\alpha}{2},$$

it implies that

$$\frac{1}{\frac{\alpha}{2} + \frac{\alpha^2}{8}\Delta t} \leq \frac{2}{\alpha}.$$

We conclude that

$$\begin{aligned}
\|e^{n+1}\| &= \frac{\left(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2\right)^{n+1} - 1}{\frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2} (O(\Delta t^3) + \Delta t O(h)). \\
&\leq \frac{2}{\alpha} e^{\beta T} (O(\Delta t^2) + O(h)) \\
&\leq C(T) (O(\Delta t^2) + O(h)).
\end{aligned}$$

□

Lax's equivalence theorem says that if we approximate a linear initial-value problem by a linear finite-difference method that satisfies the consistency condition,

then stability is necessary and sufficient condition for convergence [23]. Since the reaction-diffusion-chemotaxis equation is nonlinear, it does not satisfy Lax's condition. Therefore, we have to investigate the stability issue.

Theorem 7 (Stability) Assume that $\|M_n + M_{n+1}\| \leq \alpha$ for all $n + 1 \leq N$, let $|e^n| = \max_l |u_l - u(t_l)|$, $0 \leq l \leq n$ for all n such that $n\Delta t \leq T$. If $|e^0| \leq \varepsilon$ then $|e^n| \leq C_1(T)\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h))$, where $C = C(T)$ is a constant which depends on T , but not Δt and Δx .

Proof. It follows from the proof of convergence, we have

$$\begin{aligned}
\|e^{n+1}\| &\leq \|I + \frac{\Delta t}{2}(M_n + M_{n+1}) \\
&\quad + \frac{\Delta t^2}{8}(M_n + M_{n+1})^2\| \|e^n\| + O(\Delta t^3) + \Delta t O(h) \\
&\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2) \|e^{n-1}\| \\
&\quad + (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)(O(\Delta t^3) + \Delta t O(h)) \\
&\quad + O(\Delta t^3) + \Delta t O(h) \\
&\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n+1}\varepsilon + [(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^n \\
&\quad + (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n-1} + \dots + 1](O(\Delta t^3) + \Delta t O(h)).
\end{aligned}$$

Since

$$(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^n + (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n-1} + \dots + 1$$

$$= \frac{(1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n+1} - 1}{\frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2} \leq C(T), \text{ and,}$$

$$\begin{aligned} 1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2 &= 1 + (\frac{\alpha}{2} + \frac{\Delta t}{8}\alpha^2)\Delta t \\ &\leq 1 + \beta\Delta t, \end{aligned}$$

therefore,

$$\begin{aligned} \| e^{n+1} \| &\leq (1 + \frac{\Delta t}{2}\alpha + \frac{\Delta t^2}{8}\alpha^2)^{n+1}\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)) \\ &\leq (1 + \beta\Delta t)^{n+1}\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)) \\ &\leq e^{\beta(\Delta t)(n+1)}\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)) \\ &\leq e^{\beta T}\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)) \\ &\leq e^{\beta T}\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)) \\ &\leq C_1(T)\varepsilon + C(T)(O(\Delta t^3) + \Delta t O(h)), \end{aligned}$$

where $C_1(T) = e^{\beta T}$ and $C(T) = 2/\alpha e^{\beta T}$. \square

CHAPTER 5

Numerical Experiments

In this chapter, we use a 1D testing equation to test the efficiency of the two-stage numerical scheme. The results of a comparison between the numerical and analytical solutions will be given. Furthermore, we will show that the optimal two-stage scheme converges with the analytical solution for the time evolution problem while the standard two-stage Runge-Kutta method converges only when the time step is small enough. Moreover, the optimal two-stage scheme converges much faster than the standard two-stage Runge-Kutta method in the region of convergence. Later in this chapter, we will test the numerical scheme against a realistic problem.

5.1 1D testing problem

Let us consider the following testing equation

$$\frac{du}{dt} = u_{xx} - (txu)_x + f, \quad (5.1)$$

where $f = (tx - 1)e^{-t}\cos(x - t) + te^{-t}\sin(x - t)$.

Assume $x \in [0, 2\pi]$, $t > 0$, and the boundary is periodic.

The analytical solution of (5.1) is given by

$$u(x, t) = e^{-t} \sin(x - t).$$

We solve the equation by the two-stage scheme. We use the first-order upwind method to approximate the chemotaxis term, and the standard second-order centered difference to approximate the diffusion term. Since tx is positive for $t > 0$, the chemotaxis is approximated by the backward difference operator. This leads to the following setting:

Let $u_i(t) \approx u(x_i, t)$, then

$$\begin{aligned} \frac{du}{dt} &= \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} - tu_i - tx_i \left(\frac{u_i - u_{i-1}}{h} \right) + f(x_i, t) \\ &= \frac{1}{h^2} u_{i+1} + \left(\frac{-2}{h^2} - t - \frac{tx_i}{h} \right) u_i + \left(\frac{1}{h^2} + \frac{tx_i}{h} \right) u_{i-1} + f(x_i, t), \end{aligned}$$

where $i = 0, 1, 2, \dots, N$.

The boundary condition is

$$u_{-1} = u_{N-1}.$$

Therefore, we can set up the ODE system as the following :

$$\frac{du}{dt} = M(t^n)u + f^n,$$

where $M(t^n) =$

$$\begin{pmatrix} \tau_0^n & \frac{1}{h^2} + \frac{t^n x_1}{h} & & \frac{1}{h^2} - \frac{t^n}{h} & & \\ \frac{1}{h^2} + \frac{t^n x_0}{h} & \tau_1^n & \frac{1}{h^2} & & & \\ & \frac{1}{h^2} + \frac{t^n x_1}{h} & \tau_2^n & \frac{1}{h^2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots \\ & & & & & \ddots \\ & & & & & \frac{1}{h^2} & \frac{1}{h^2} + \frac{t^n x_{N-1}}{h} & \tau_N^n \end{pmatrix}$$

where $\tau_i^n = \frac{-2}{h^2} - t^n - \frac{t^n x_i}{h}, i = 0, 1, 2, \dots, N$.

The matrix $\overline{M^n} = \frac{M(t^n) + M(t^{n+1})}{2}$ can be evaluated by changing the time level from t^n to t^{n+1} .

Since the matrix size is usually very large and entries of the matrix vary according to different time levels, it is not possible to find all the eigenvalues in analytic form. The existing numerical algorithms, e.g. the power method, are too expensive to estimate the eigenvalues. Moreover, we only need the extreme eigenvalues to calculate the adapted time step. The Gerschgorin Theorem from linear algebra will be used for this purpose.

Theorem 8 (Gerschgorin) *The union of all discs $K_i := \{\mu \in \mathbb{C} \mid |\mu - a_{ii}| \leq \sum_{k=1, k \neq i}^n |a_{ik}|\}$ contains all eigenvalues of the $n \times n$ matrix $A = [a_{ik}]$.*

Proof. See the reference [23]. \square

The estimation of the eigenvalues for the first row and the rest of the rows will be discussed separately, since they have different structures which are due to the effects of boundary conditions.

For the i th row, where $i = 1, 2, \dots, N$, the diagonal entries are

$$a_{ii} = \frac{\tau_i^n + \tau_1^{n+1}}{2} = \frac{-2}{h^2} - \left(\frac{t^n + t^{n+1}}{2}\right) - \frac{x_i(t^n + t^{n+1})}{2h}.$$

The absolute sum of the off diagonal entries for each row is

$$\sum_{j=0, j \neq i}^N |a_{ij}| = \frac{2}{h^2} + \frac{x_i(t^n + t^{n+1})}{2h}.$$

Gerschgorin's Theorem implies all the eigenvalues of \overline{M} are contained in the union of

$$|\lambda - \left(\frac{-2}{h^2} - \left(\frac{t^n + t^{n+1}}{2}\right) - \left(\frac{x_i t^n + t^{n+1}}{2h}\right)\right)| \leq \frac{2}{h^2} + \frac{x_i(t^n + t^{n+1})}{2h}.$$

The smallest and largest real part of the eigenvalues satisfy

$$-\left(\frac{t^n + t^{n+1}}{2}\right) \leq \lambda \leq \frac{-4}{h^2} - \left(\frac{t^n + t^{n+1}}{2}\right) - \frac{x_i(t^n + t^{n+1})}{h}$$

For the extreme case,

$$-a_{\min} = \frac{-4}{h^2} - \left(\frac{t^n + t^{n+1}}{2}\right) - \frac{x_i(t^n + t^{n+1})}{h}$$

,

$$-a_{\max} = -\left(\frac{t^n + t^{n+1}}{2}\right).$$

Since $-a_{\max} = -(\frac{t^n + t^{n+1}}{2}) < 0$, it shows that the largest eigenvalues are on the left hand side of the complex plane. Therefore, we only have to consider the case for $-a_{\max} < 0$. By theorem 4 and

$$\begin{aligned}\Delta t_{row0} &= \frac{1}{\frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2})} > \Delta t_{row1} = \frac{1}{\frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2}) + h\frac{(t^n + t^{n+1})}{h}} > \\ \Delta t_{row2} &= \frac{1}{\frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2}) + 2h\frac{(t^n + t^{n+1})}{h}} > \dots \\ &> \Delta t_{rowN} = \frac{1}{\frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2}) + \pi\frac{(t^n + t^{n+1})}{h}},\end{aligned}$$

the time step is

$$\Delta t_{opt} = \frac{2}{a_{\min} + a_{\max}} = \frac{1}{\frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2}) + \pi\frac{(t^n + t^{n+1})}{h}}. \quad (5.2)$$

For the first row of the matrix \overline{M} , $i = 0$,

$$\begin{aligned}a_{00} &= -\frac{2}{h^2} - (\frac{t^n + t^{n+1}}{2}), \\ \sum_{j=1}^N |a_{0,j}| &= \frac{1}{h^2} + \frac{t^n + t^{n+1}}{2} + \left| \frac{1}{h^2} - (\frac{t^n + t^{n+1}}{2h}) \right|.\end{aligned}$$

The location of eigenvalues for this row is

$$\left| \lambda + \frac{2}{h^2} + (\frac{t^n + t^{n+1}}{2}) \right| \leq \frac{1}{h^2} + (\frac{t^n + t^{n+1}}{2}) + \left| \frac{1}{h^2} - (\frac{t^n + t^{n+1}}{2h}) \right|.$$

In this case,

$$\frac{-3}{h^2} - (t^n + t^{n+1}) - \left| \frac{1}{h^2} - \left(\frac{t^n + t^{n+1}}{2h} \right) \right| \leq \lambda \leq -\frac{1}{h^2} + \left| \frac{1}{h^2} - \left(\frac{t^n + t^{n+1}}{2h} \right) \right|.$$

The lower bound is

$$-a_{\min} = \frac{-3}{h^2} - (t^n + t^{n+1}) - \left| \frac{1}{h^2} - \left(\frac{t^n + t^{n+1}}{2h} \right) \right|.$$

The upper bound is

$$-a_{\max} = -\frac{1}{h^2} + \left| \frac{1}{h^2} - \left(\frac{t^n + t^{n+1}}{2h} \right) \right|.$$

When $-a_{\max} \geq 0$, we set $-a_{\max} = 0$. Therefore, the time step is separated into two cases as follows:

$$\Delta t = \begin{cases} \frac{\frac{1}{4 + (t^n + t^{n+1})(4h^2 + h)}}{4h^2} & \text{if } \frac{2}{h^2} \leq \frac{t^n + t^{n+1}}{2} \\ \frac{\frac{1}{\frac{2}{h^2} + \frac{t^n + t^{n+1}}{2}}}{\frac{2}{h^2} + \frac{t^n + t^{n+1}}{2}} & \text{if } \frac{2}{h^2} > \frac{t^n + t^{n+1}}{2} \end{cases} \quad (5.3)$$

Since time step in (5.3) is greater than time step in (5.2), the optimal time step is

$$\Delta t_{\text{opt}} = \frac{1}{\frac{2}{h^2} + \left(\frac{t^n + t^{n+1}}{2} \right) + \pi \frac{(t^n + t^{n+1})}{2}}. \quad (5.4)$$

5.1.1 Results for 1D experiment

Method / Time	1	2	3	4	5
Optimal two stage	284	637	1060	1553	2116
Standard Runge-Kutta $\Delta t = 0.002$	500	1000	1500	2000	2500
Standard Runge-Kutta $\Delta t = 0.004$	250	Div	Div	Div	Div

Table 5.1: Number of iterations.

This table is the number of iterations with different methods at different time. The table says that the standard two-stage Runge-Kutta method will converge when the time step is small enough. However, the optimal two-stage method converges faster than the Runge-Kutta method when both methods converge. When the time step for the Runge-Kutta method is increased, it will diverge.

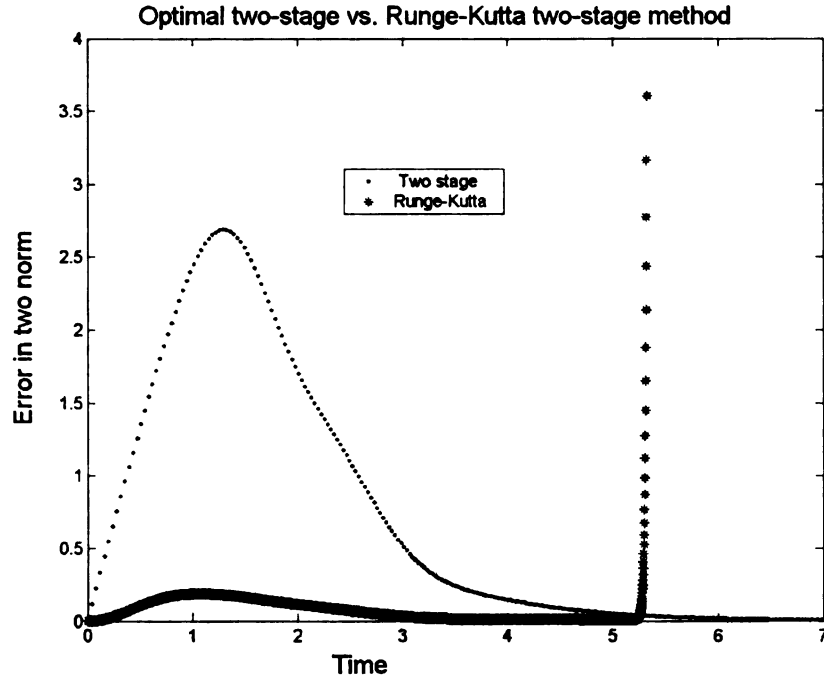


Figure 5.1: Error between numerical simulation and analytic solution.

This picture shows the relative error between numerical simulation and analytic solution. While the two-stage Runge-Kutta method blows up in a short time, the optimal two-stage method is still very accurate. It shows that the optimal two-stage scheme not only is accurate but also converges faster than the two-stage Runge-Kutta method.

5.2 Simulation of Adler's experiment

In this section, we will test the optimal two-stage scheme with a realistic problem.

5.2.1 Chemotaxis phenomenon

The study of chemotactic behavior of bacteria originated in the nineteenth century. In 1884, Pfeffer discovered the chemotaxis phenomenon in bacteria by observing a capillary tube which contained an attractant and motile bacteria. The bacteria accumulated near the higher concentration of the attractant. In 1901, Rothert, Jennings and Crosby found that bacteria often avoided the region of lower density in a capillary tube. In 1972, H. C. Berg and D. A. Brown analyzed the microscope motion of the bacteria by comparing the quantitative run-tumble analysis between a wild type, a nonchemotactic mutant and an uncoordinated mutant. They discovered that bacteria responded to the external environment by modulating their pattern of motion shifting between swimming and tumbling: swimming is a nearly straight run, tumbling is a very rapid and jerky movement which redirects the cell to a new random direction. Usually, the motion of swimming is followed by a brief tumbling. When bacteria encounter a favorable external environment, they decrease the tumbling probability and extend the time for swimming in the preferred directions. It results in a bias in the random walk and net movement in the gradient of the attractant. That is chemotaxis [8, 16].

The swimming and tumbling modes are modulated by the bacteria flagella. Each one consists a long filament, a left handed helix of protein subunits and a basal end attached to the cell membrane. The basal end serves as a turning wheel. It turns the filament as a propeller. The motion of swimming is produced by the anticlockwise rotation of the left-handed helix. It results in a forward movement. The clockwise rotation gives rise to a backward motion, producing tumbling.

Bacteria sense the change of the external environment via membrane receptors

binding the chemical substrate. The cells reflect the change in the concentration of chemicals by the relative number of occupied binding sites, not by the absolute concentration of the substrate. “The receptor proteins perform as reporter molecules, which sending information about the current concentration of ligand to a central mechanism (tumble regular). The tumble regular then creates signal to rule the direction of the rotation of flagellar rotary motors” [8].

It is worthwhile to mention that we can derive the chemotaxis coefficient from the proportion between the occupied (denoted by N_o) and free receptors (denoted by N_f). The ratio of the occupied receptors and the total number of receptors is $\frac{N_o}{N_f + N_o}$. Let τ_0 and τ_f represent the mean time of a receptor occupation and the mean time when the receptor is free. Note that τ_f is inversely proportional to the concentration of the substrate and τ_0 is a constant which is related to the process of internal cellular. Thus

$$\frac{N_o}{N_f + N_o} = \frac{\tau_0}{\tau_f + \tau_0} = \frac{C}{K + C} \quad (5.5)$$

and

$$\frac{\partial}{\partial x} \left(\frac{N_o}{N_f + N_o} \right) = \frac{K}{(K + C)^2} \frac{\partial C}{\partial x}, \quad (5.6)$$

where $K \equiv \frac{(C\tau_f)}{\tau_0}$ is a constant, and $\frac{K}{(K + C)^2}$ is known as the “receptor law”. The receptor law says that the chemotactic reponse vanishes when all of the receptors have been occupied [16].

We tested our numerical scheme by using Julius Adler’s series of work [4, 5, 6, 22] in the quantitative investigations into the chemotaxis. His work has inspired and

influenced researchers in the area of molecular biology and mathematical biology, especially in the field of pattern formation. We will briefly describe his experimental work and the results of one of his papers in 1966 before demonstrating the numerical results.

5.2.2 Experiment results

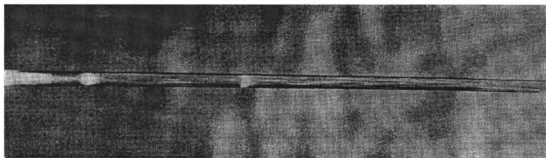


Figure 5.2: Photographs showing bands of *E.coli* in a capillary tube. In all the experiments reported here, capillary tubes were filled with a liquid medium, inoculated at one end with 200000 to 2000000 bacteria, and then closed at the ends with plug of agar and clay, all according to procedure described in full elsewhere. The tube were incubated horizontally at 37C. The origin, which is turbid because of the bacteria that have not moved out, is visible at the left, then the second band of bacteria, then the first band. Plugs at the ends are not shown. The concentration of galactose was 0.00025 mole per liter.

Julius Adler and his associates use *E.Coli* in most of his experiments. Since “the vast knowledge of its biochemistry and genetics could be brought to bear on the problem.” It is an inhabitant in the human or animal intestine and it is nonpathogenic in normal environments. *E.coli* is a mobile bacteria, it moves by propelling itself with flagella around cells. *E.Coli* is known to be moving preferentially toward oxygen.

E. Coli were placed at one end of a closed capillary tube. The other end of the capillary tube was closed with plugs of agar and clay. It contained different amounts of oxygen and an energy source. Two clear and visible bands of bacteria with a constant speed moved from the end after adding bacteria into the tube. Some cells still remained at the end of the tube. If the amount of dissolved oxygen was lower than the energy source (the amount of oxygen is not enough to oxidize all the energy source), then the first band traveled along, consuming all of the oxygen. The second band would move out when the cells were able to deplete the energy source anerobically. On the other hand, when there is enough oxygen for the cells to oxidize the energy source, the first band totally exhausted the energy source aerobically. Under these circumstances, the second band used up the remaining oxygen to oxidize the endogenous energy source. In this situation, the bacteria in the first band generated a steeper gradient along the concentration of the energy source than the gradient in the concentration of oxygen. Bacteria then moved preferentially toward the higher concentration of the energy source. Bacteria in the second band created gradients in the concentration of oxygen and galactose. Bacteria therefore migrated toward the higher concentration of oxygen [4, 18, 24].

These two crowded regions of bacteria are the effect of chemotaxis. Since if random walk or diffusion is the only transport mechanism for the movement of bacteria, the distribution of bacteria density should look like a linear function which falls from high to low density.

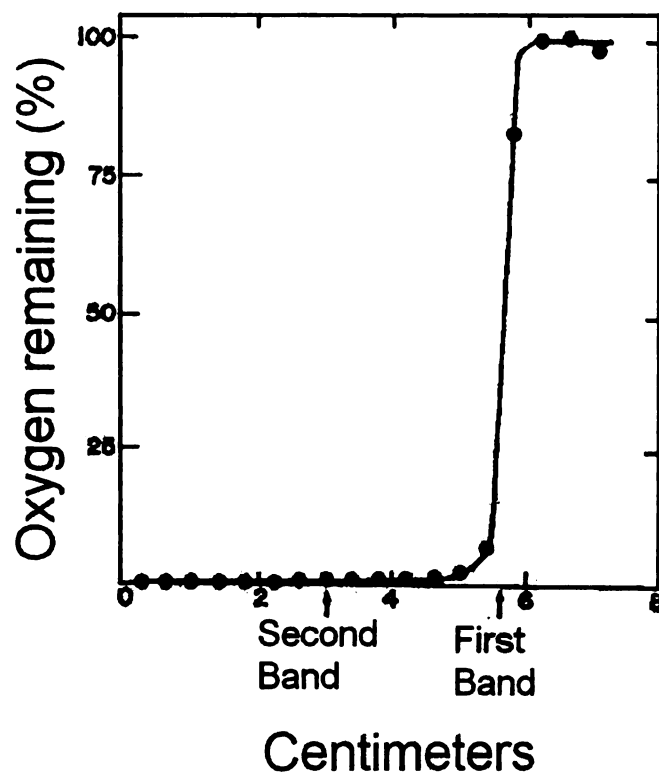


Figure 5.3: Utilization of oxygen by bands of bacteria in 0.00025 *M* galactose. Oxygen was measured polarographically by inserting an oxygen needle electrode into the capillary tube in 4-mm steps when the first and second bands were visible where shown by arrows.

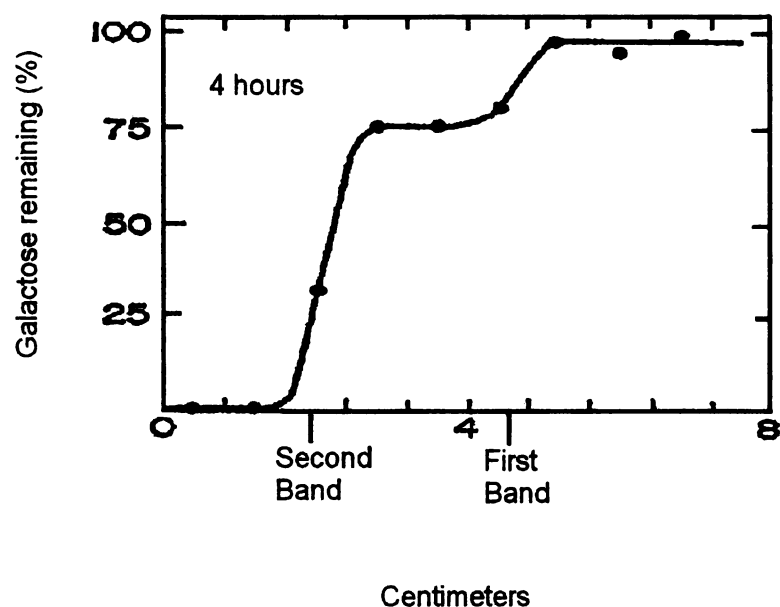


Figure 5.4: Utilization of 0.00025 *M* galactose. The tube contained C14- galactose (0.00025 mole/liter and 1000000 counts per minute per milliliter). At 4 or 8 hours, when the first and second bands were visible where shown by arrows, the tube was factionated into ten compartments, each 8 mm long. the contents of each compartment were chromatographed on paper, with *n*-butanol, acetic acid, and water (12:3:5) as the solvent, and the radioactivity in the galactose region of the chromatogram was measured in a paper-strip counter.

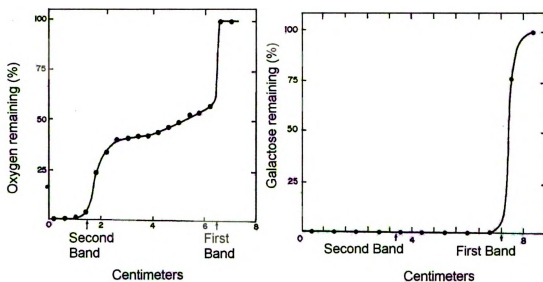


Figure 5.5: Utilization of oxygen by bands of bacteria in 0.000005 *M* galactose. Measurements were made as in Figure 5.3 and Figure 5.4. Utilization of 0.000005 *M* galactose. The tube contained C14-galactose (0.000005 molde/liter and 200000 counts per minute per milliliter). At 4.75 hours when the first and second bands were visible where shown by the arrows.

5.2.3 Mathematical Model for pattern formation in *E.coli* experiment

We deal with a continuous, deterministic model to describe bacteria growth. The characteristics of the model we describe are all two dimensional. First, we will explain the following notation and its units [.]

- u : bacteria density : $[u]=\text{cells}/\text{cm}^3$.
- g : concentration of galactose : $[g]=\text{mmole}/\text{cm}^3$.
- s : concentration of oxygen : $[s]=\text{mmole}/\text{cm}^3$.
- t : time : $[t]=\text{hours}$.
- x : distance along the tube (peri dish) : $[x]=\text{cm}$.
- $\chi_1(g)$ = chemotactic sensitivity for galactose : $[\chi_1(s)]=\text{cm}^2/\text{hour}$.
- $\chi_2(s)$ = chemotactic sensitivity for oxygen : $[\chi_2(g)]=\text{cm}^2/\text{hour}$.
- $D_u > 0$: bacteria motility : $[D_u]=\text{cm}^2/\text{hr}$
- $D_g > 0$, and $D_s > 0$: constant diffusivities : $[D]=\text{cm}^2/\text{hr}$.
- $Y_g > 0$: bacteria yield constant, which gives the yield of bacteria per galactose taken up : $[Y_g]=\text{mass u}/\text{mass g}$.
- $Y_s > 0$: bacteria yield constant, which gives the yield of bacteria per oxygen taken up : $[Y_s]=\text{mass u}/\text{mass s}$.

For a closed region, the time rate of change of the bacteria concentration should be equal to the bacteria flux across the boundary of the closed region. Assume that

reproduction is not negligible. This gives

$$\frac{\partial u}{\partial t} = -\frac{\partial J}{\partial x} + \text{"Birth"}, \quad (5.7)$$

where $J = J(x, t)$ is the flux of the bacteria. It consists of two parts : diffusion and chemotaxis.

The diffusion or random walk is given by

$$-D_u \nabla u. \quad (5.8)$$

The chemotaxis flux is

$$u\chi_1(g)\nabla g + u\chi_2(s)\nabla s, \quad (5.9)$$

where $\chi_1(g)$ and $\chi_2(s)$ are the chemotactic sensitivity to the galactose and oxygen respectively. They represent the strength of chemoattractant. The presence of u in the above equation is because the bacteria flux is the number of bacteria crossing a unit area per unit time. Therefore, increasing the number of bacteria will increase the bacteria flux. In the case of repellent, the flux will have a positive sign [16].

Here we use the Keller-Segel model with $p=2$ or the receptor law [18].

$$\chi_{KS}(s) = \frac{C_0}{(C_1 + s)^2} \quad (5.10)$$

where C_0 and C_1 are positive constants. Note that as $\lim_{s \rightarrow \infty} \chi_{KS}(s) = 0$, which means that bacteria will prefer to stay where they are if there is plenty of chemoattractant.

Thus, the total bacteria flux J is:

$$J = u\chi_1(g)\nabla g + u\chi_2(s)\nabla s - D_u\nabla u. \quad (5.11)$$

For the growth term, we adapt Monod's law

$$f_1(g) = \frac{V_1 g}{K_1 + g}, \quad (5.12)$$

where $V_1 > 0$ and $K_1 > 0$ are constants.

If the amount of nutrient (galactose, in this case) is restricted, bacteria will consume as much as food as they can. When the supply of nutrient is abundant, the consumption rate is independent of the concentration of nutrient. We say the growth rate is saturated with respect to the nutrient [18]. Under this circumstance, the reproduction of bacteria only depends on the bacteria density and the eating rate per bacteria is V_1 .

Together with (5.11),(5.12), the bacteria equation u is

$$u_t = D_u\Delta u - \nabla(u\chi_1(g)\nabla g) - \nabla(u\chi_2(s)\nabla s) + f_1(g)u, \quad (5.13)$$

where $\Delta = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$ is the standard Laplace operator, which simulates the diffusion process. $\nabla = \partial/\partial x_1 + \partial/\partial x_2$ is the gradient operator.

The model also contains nutrient (galactose) and oxygen. Nutrient was taken up by bacteria at the same rate as bacteria reproduction. This suggests that the nutrient consumption (following Monod's law) has a sigmoid characteristic. Meanwhile,

the nutrient also diffuses itself. A similar assumption also holds for oxygen. However, oxygen does not contribute to the growth of bacteria. Hence, the equations for nutrient and chemotaxis are:

$$g_t = D_g \Delta g - f_1(g)u/Y_g, \quad (5.14)$$

$$s_t = D_s \Delta s - f_2(s)u/Y_s. \quad (5.15)$$

where $Y_g > 0, Y_s > 0$ are the bacteria yield constants, Y_g gives the yield of bacteria per nutrient taken up. Y_s gives the yield of bacteria per oxygen taken up.

According to Adler's experiment, the ends of the capillary tube are closed. No substrate flows through the boundary. Therefore, the normal derivative on the boundary is zero. i.e,

$$\frac{\partial u}{\partial n}|_{\partial\Omega} = \frac{\partial g}{\partial n}|_{\partial\Omega} = \frac{\partial s}{\partial n}|_{\partial\Omega} = 0. \quad (5.16)$$

Moreover, the initial condition for bacteria is :

$$u(x, 0) = \begin{cases} u_0(x) & \text{if } \|x\| \leq \ell \\ 0 & \text{Otherwise.} \end{cases}$$

which approximates the initial inoculum of the bacteria in the center of the agar plate (in the two dimensional case).

Galactose and oxygen are uniformly distributed at the beginning. The initial conditions fitting Adler's experiment are

$$g(x, 0) = g_0(x), \quad (5.17)$$

$$s(x, 0) = s_0(x). \quad (5.18)$$

5.2.4 The coefficient function

To complete the model equations, we have to estimate the parameter function. It is crucial to have an accurate parameter estimation to ensure the presentation of patterns. The original paper does not contain the quantitative data of the experiment, so we use ChiChia Chiu's paper [14] as a reference to guess the correct order for the parameter function. The values of parameters are chosen as follows.

- $D_u = 0.001, D_g = 0.005, D_s = 0.033$
- $u_0(x) = 5 \times 10^{-6}$ if $\|x\| \leq \ell$, where $\ell = 7.5 \times 10^{-2}$
- $s_0(x) = 4 \times 10^{-5}, g_0(x) = 2.64 \times 10^{-5}$
- $Y_g = 0.001, Y_s = 0.0001$
- $CK_1 = 0.02, DK_1 = 2 \times 10^{-6}$
- $CK_2 = 0.04, DK_2 = 3.3 \times 10^{-5}$
- $V_1 = 0.35, K_1 = 4 \times 10^{-6}$
- $V_2 = 0.6, K_1 = 5.5 \times 10^{-5}$
- $\Omega = [-R, R] \times [-R, R]$, where $R = 1$

These parameters come from experiments and literature. In the numerical simulation, we fix all parameters except $u_0(x), g_0(x)$ in order to accommodate Adler's experiment.

5.2.5 Intuitive explanation of pattern formation mechanism

The nutrient (galactose) and chemoattractant (galactose, oxygen) will exhaust as time goes to infinity. This implies that the growth of bacteria and the effect of chemotaxis will eventually become zeros. In the end, the only term which remains in effect is the diffusion of bacteria. The solution become spatially homogeneous. This is not an interesting phenomenon in pattern formation. Thus, we have to investigate a dynamic solution [20].

Note that galactose not only plays a role as an energy source but also as chemoattractant. Therefore, galactose and oxygen both serve as chemoattractant. They both direct the movement of bacteria. However, oxygen does not contribute to the growth of bacteria. The bacteria take up suger and oxygen, creating a gradient of nutrient and chemoattractant. They move preferentially toward higher concentration of substrates and generate a population there. However, the diffusion of the bacteria has a dissipated effect, which flattens out the population peak. The speed of diffusion and the eating speed of the bactera for these chemicals are not identical. As a consequence, two sharp and visible bands migrate out.

5.2.6 An optimal two-stage numerical scheme

Let $\Omega = [0, \ell]^2$ with mesh size $h = 2\ell/N$. Let $x_{ij} \in \Omega$ be a grid point, then $x_{i+1,1} = x_{i1} + h$ and $x_{1,j+1} = x_{1j} + h$, etc. The time step is denoted as Δt , and $u_{ij} \approx u_{x_{ij}}$.

Let us solve the equations for the nutrient concentration, g , the attractant con-

centration, s , using the optimal two stage scheme:

$$\begin{aligned}
\frac{\partial g}{\partial t} &= D_g \left(\frac{g_{i+1,j} - 2g_{i,j} + g_{i-1,j}}{h^2} \right) + D_g \left(\frac{g_{i,j+1} - 2g_{i,j} + g_{i,j-1}}{h^2} \right) \\
&- f_2(g_{ij}u_{ij})/Y_g \\
&= \frac{D_g}{h^2} g_{i+1,j} + \frac{D_g}{h^2} g_{i,j+1} - \frac{4D_g}{h^2} g_{i,j} + \frac{D_g}{h^2} g_{i-1,j} + \frac{D_g}{h^2} g_{i,j-1} \\
&- f_2(g_{ij}u_{ij})/Y_g.
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
\frac{\partial s}{\partial t} &= D_s \left(\frac{s_{i+1,j} - 2s_{i,j} + s_{i-1,j}}{h^2} \right) + D_s \left(\frac{s_{i,j+1} - 2s_{i,j} + s_{i,j-1}}{h^2} \right) \\
&- f_1(s_{ij}u_{ij})/Y_s \\
&= \frac{D_s}{h^2} s_{i+1,j} + \frac{D_s}{h^2} s_{i,j+1} - \frac{4D_s}{h^2} s_{i,j} + \frac{D_s}{h^2} s_{i-1,j} + \frac{D_s}{h^2} s_{i,j-1} \\
&- f_1(s_{ij}u_{ij})/Y_s.
\end{aligned} \tag{5.20}$$

The boundary conditions are

$$g_{-1,j} = g_{1,j}, g_{N+1,j} = g_{N-1,j}, g_{i,-1} = g_{i,1}, \text{ and } g_{-1,N+1} = g_{i,N-1},$$

$$s_{-1,j} = s_{1,j}, s_{N+1,j} = s_{N-1,j}, s_{i,-1} = s_{i,1}, \text{ and } s_{-1,N+1} = s_{i,N-1},$$

where $i, j = 0, 1, 2, \dots, N$.

Now, we can investigate the time step of the ODE system $\partial./\partial t = M(t) + f(x, t)$ for g and s .

From (5.19) and the boundary, the diagonal entries of g matrix and the summand of the off diagonal entries are

$$\frac{-4D_g}{h^2}, \frac{4D_g}{h^2}$$

respectively. By the Gerschgorin Theorem, we get

$$|\lambda + \frac{4D_g}{h^2}| \leq \frac{4D_g}{h^2}.$$

For $i = 0, 1, 2, \dots, N$. Therefore, the real part of the extreme eigenvalues of \overline{M} for g are contained in

$$[\frac{-8D_g}{h^2}, 0].$$

Hence, the optimal time step for g is

$$\Delta t_g = \frac{h^2}{4D_g}. \quad (5.21)$$

Similarly, the optimal time step for s -equation is

$$\Delta t_s = \frac{h^2}{4D_s}. \quad (5.22)$$

In the following section, we present optimal two-stage scheme for solving the reaction-diffusion-chemotaxis equation.

The equation for u has the following form

$$u_t = D_u \Delta u - \nabla(u\chi_1(s)\nabla s) - \nabla(u\chi_2(g)\nabla g) + f_1(g)u. \quad (5.23)$$

The optimal two-stage scheme for u is defined by the following:

$$\begin{aligned} \frac{\partial u}{\partial t} &= D_u \frac{\Delta_{xh} u_{ij}}{h^2} + D_u \frac{\Delta_{yh} u_{ij}}{h^2} \\ &- \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{2h} \nabla_{xh}^- \right)_{ij} u_{ij} \\ &- \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{2h} \nabla_{xh}^+ \right)_{ij} u_{ij} \end{aligned} \quad (5.24)$$

$$\begin{aligned}
& - \left(\frac{\chi_1(s) \nabla_y s + |\chi_1(s) \nabla_y s|}{2h} \nabla_{yh}^- \right)_{ij} u_{ij} \\
& - \left(\frac{\chi_1(s) \nabla_y s - |\chi_1(s) \nabla_y s|}{2h} \nabla_{yh}^+ \right)_{ij} u_{ij} \\
& - \left(\frac{\chi_2(g) \nabla_x g + |\chi_2(g) \nabla_x g|}{2h} \nabla_{xh}^- \right)_{ij} u_{ij} \\
& - \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{2h} \nabla_{xh}^+ \right)_{ij} u_{ij} \\
& - \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{2h} \nabla_{yh}^- \right)_{ij} u_{ij} \\
& - \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{2h} \nabla_{yh}^+ \right)_{ij} u_{ij} \\
& + f_1(g_{ij}) u_{ij}
\end{aligned} \tag{5.25}$$

where ∇_{xh}^- and ∇_{yh}^- are first order backward difference operators such that

$$(\nabla_{xh}^- + \nabla_{yh}^-) u_{ij} = (u_{ij} - u_{i-1j}) + (u_{ij} - u_{ij-1}),$$

and ∇_{xh}^+ and ∇_{yh}^+ are first order forward difference operators such that

$$(\nabla_{xh}^+ + \nabla_{yh}^+) u_{ij} = (u_{i+1j} - u_{ij}) + (u_{ij+1} - u_{ij}).$$

∇_x , ∇_y are the centered difference operators

$$\nabla_x = u_{i+1j} - u_{i-1j},$$

$$\nabla_y = u_{ij+1} - u_{ij-1}.$$

Note

$$\Delta_{xh} u_{ij} = \nabla_{xh}^+ u_{ij} - \nabla_{xh}^- u_{ij}.$$

and $(\chi_1(s))_{ij} = \chi_1(s_{ij})$.

The zero flux boundary conditions for u are approximated by

$$u_{-1j} = u_{1j}, u_{N+1j} = u_{N-1j}, u_{i-1} = u_{i1}, \text{ and } u_{-1,N+1} = u_{i,N-1}.$$

Therefore, (5.24) can be written as

$$\begin{aligned} \frac{\partial u}{\partial t} &= D_u \frac{u_{i+1j} - 2u_{ij} + u_{i-1j}}{h^2} + D_u \frac{u_{ij+1} - 2u_{ij} + u_{ij-1}}{h^2} \\ &- \left(\frac{\chi_1(s) \nabla_x s + |\chi_1(s) \nabla_x s|}{2h} \right)_{ij} (u_{ij} - u_{i-1j}) \\ &- \left(\frac{\chi_1(s) \nabla_x s - |\chi_1(s) \nabla_x s|}{2h} \right)_{ij} (u_{i+1j} - u_{ij}) \\ &- \left(\frac{\chi_1(s) \nabla_y s + |\chi_1(s) \nabla_y s|}{2h} \right)_{ij} (u_{ij} - u_{ij-1}) \\ &- \left(\frac{\chi_1(s) \nabla_y s - |\chi_1(s) \nabla_y s|}{2h} \right)_{ij} (u_{ij+1} - u_{ij}) \\ &- \left(\frac{\chi_2(g) \nabla_x g + |\chi_2(g) \nabla_x g|}{2h} \right)_{ij} (u_{ij} - u_{i-1j}) \\ &- \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{2h} \right)_{ij} (u_{i+1j} - u_{ij}) \\ &- \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{2h} \right)_{ij} (u_{ij} - u_{ij-1}) \\ &- \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{2h} \right)_{ij} (u_{ij+1} - u_{ij}) \\ &+ f_1(g_{ij}) u_{ij}. \end{aligned}$$

That is,

$$\begin{aligned} \frac{\partial u}{\partial t} &= \left[\frac{D_u}{h^2} + \left(\frac{\chi_1(s) \nabla_x s + |\chi_1(s) \nabla_x s|}{2h} \right)_{ij} \right. \\ &\quad \left. + \left(\frac{\chi_2(g) \nabla_x g + |\chi_2(g) \nabla_x g|}{2h} \right)_{ij} \right] u_{i-1j} \\ &+ \left[\frac{D_u}{h^2} + \left(\frac{\chi_1(s) \nabla_y s + |\chi_1(s) \nabla_y s|}{2h} \right)_{ij} \right. \\ &\quad \left. + \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{2h} \right)_{ij} \right] u_{ij-1} \\ &+ \left[\frac{-4D_u}{h^2} - \left(\frac{\chi_1(s) \nabla_x s + |\chi_1(s) \nabla_x s|}{2h} \right)_{ij} \right. \end{aligned}$$

$$\begin{aligned}
& - \left(\frac{\chi_1(x)\nabla_y s + |\chi_1(s)\nabla_y s|}{2h} \right)_{ij} \\
& + \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{2h} \right)_{ij} + \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{2h} \right)_{ij} \\
& - \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{2h} \right)_{ij} - \left(\frac{\chi_2(g)\nabla_y g + |\chi_2(g)\nabla_y g|}{2h} \right)_{ij} \\
& + \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{2h} \right)_{ij} + \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{2h} \right)_{ij}] u_{ij} \\
& + \left[\frac{D_u}{h^2} - \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{2h} \right)_{ij} \right. \\
& \left. - \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{2h} \right)_{ij} \right] u_{i+1j} \\
& + \left[\frac{D_u}{h^2} - \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{2h} \right)_{ij} \right. \\
& \left. - \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{2h} \right)_{ij} \right] u_{ij+1},
\end{aligned}$$

where $i, j = 0, 1, 2, \dots, N$.

Let $(\bullet)^n$ denote the time level. Then the diagonal entries of the spatial matrix $\bar{M} = \frac{M(t^n) + M(t^{n+1})}{2}$ of the ODE system for u is

$$\begin{aligned}
a_{ii} = & \frac{-4D_u}{h^2} - \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^n - \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^n - \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_2(g)\nabla_y g + |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g)\nabla_y g + |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^{n+1}.
\end{aligned}$$

The summand of the absolute value of the off diagonal entries for each row is

$$\begin{aligned}
\sum_{j-1, j \neq i} |a_{ij}| &= \frac{4D_u}{h^2} - \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^n \\
&\quad - \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{4h} \right)_{ij}^{n+1} \\
&\quad - \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^n - \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{4h} \right)_{ij}^{n+1} \\
&\quad - \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g)\nabla_x g - |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{4h} \right)_{ij}^{n+1} \\
&\quad - \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g)\nabla_y g - |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_2(g)\nabla_y g + |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g)\nabla_y g + |\chi_2(g)\nabla_y g|}{4h} \right)_{ij}^{n+1}.
\end{aligned}$$

Therefore, the estimated largest and the smallest eigenvalues can be found by the

following. The smallest eigenvalue of the spatial matrix \overline{M} is

$$\begin{aligned}
-a_{\min} &= \frac{-8D_u}{h^2} - \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{2h} \right)_{ij}^n \\
&\quad - \left(\frac{\chi_1(s)\nabla_x s + |\chi_1(s)\nabla_x s|}{2h} \right)_{ij}^{n+1} \\
&\quad - \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{2h} \right)_{ij}^n - \left(\frac{\chi_1(s)\nabla_y s + |\chi_1(s)\nabla_y s|}{2h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{2h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_x s - |\chi_1(s)\nabla_x s|}{2h} \right)_{ij}^{n+1} \\
&\quad + \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{2h} \right)_{ij}^n + \left(\frac{\chi_1(s)\nabla_y s - |\chi_1(s)\nabla_y s|}{2h} \right)_{ij}^{n+1} \\
&\quad - \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{2h} \right)_{ij}^n - \left(\frac{\chi_2(g)\nabla_x g + |\chi_2(g)\nabla_x g|}{2h} \right)_{ij}^{n+1}
\end{aligned}$$

$$\begin{aligned}
& - \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{2h} \right)_{ij}^n - \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{2h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{2h} \right)_{ij}^n + \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{2h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{2h} \right)_{ij}^n + \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{2h} \right)_{ij}^{n+1}.
\end{aligned}$$

Note that $-a_{\min} < 0$. The largest eigenvalue of the spatial matrix \overline{M} is

$$-a_{\max} = 0.$$

Therefore, the time step for b system is

$$\begin{aligned}
\Delta t_b = & 1 / \left(\frac{4D_u}{h^2} + \left(\frac{\chi_1(s) \nabla_x s + |\chi_1(s) \nabla_x s|}{4h} \right)_{ij}^n \right. \\
& + \left(\frac{\chi_1(s) \nabla_x s + |\chi_1(s) \nabla_x s|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_1(s) \nabla_y s + |\chi_1(s) \nabla_y s|}{4h} \right)_{ij}^n + \left(\frac{\chi_1(s) \nabla_y s + |\chi_1(s) \nabla_y s|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_1(s) \nabla_x s - |\chi_1(s) \nabla_x s|}{4h} \right)_{ij}^n - \left(\frac{\chi_1(s) \nabla_x s - |\chi_1(s) \nabla_x s|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_1(s) \nabla_y s - |\chi_1(s) \nabla_y s|}{4h} \right)_{ij}^n - \left(\frac{\chi_1(s) \nabla_y s - |\chi_1(s) \nabla_y s|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g) \nabla_x g + |\chi_2(g) \nabla_x g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g) \nabla_x g + |\chi_2(g) \nabla_x g|}{4h} \right)_{ij}^{n+1} \\
& + \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{4h} \right)_{ij}^n + \left(\frac{\chi_2(g) \nabla_y g + |\chi_2(g) \nabla_y g|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g) \nabla_x g - |\chi_2(g) \nabla_x g|}{4h} \right)_{ij}^{n+1} \\
& - \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{4h} \right)_{ij}^n - \left(\frac{\chi_2(g) \nabla_y g - |\chi_2(g) \nabla_y g|}{4h} \right)_{ij}^{n+1} \Big).
\end{aligned}$$

The optimal time step for the our model equation is

$$\min \{ \Delta t_b, \Delta t_g, \Delta t_s \}. \quad (5.26)$$

5.2.7 Results of numerical experiment

In this section, the results of the numerical experiments will be demonstrated to compare with Adler's experimental results [4].

This picture is taken from Adler's experiment. The explanation of this picture will

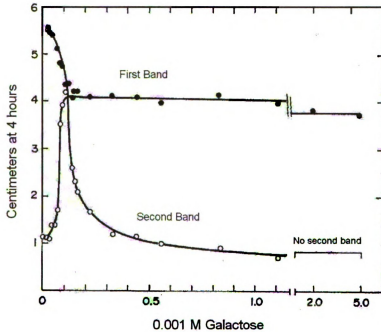


Figure 5.6: The relation between the amount of galactose and the distance of the bands of bacteria.

be described upon each graph of numerical simulation. We will use 2D pictures to simulate the 1D phenomenon. The reason is, first, Adler's experiment can be done in the agar plate. Secondly, most of patterns are shown by using 2D pictures. We want to make sure that our numerical method is suitable for general patterns.

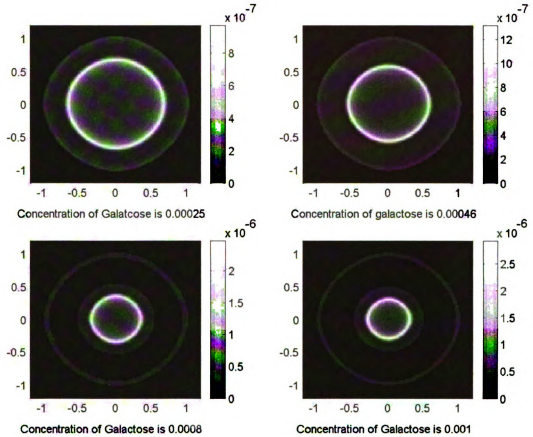


Figure 5.7: This set of pictures simulates the movement of bands of bacteria when the concentration of galactose is higher than oxygen.

The amount of galactose and oxygen varied during the experiment. When the concentration of galactose was above a certain level, the second band moved less far from the origin as the concentration of galactose was increased. It showed that the band would not proceed unless the bacteria ingested all the galactose. On the other hand, the distance for the first band did not vary. Probably, the consumption rate of oxygen was saturated.

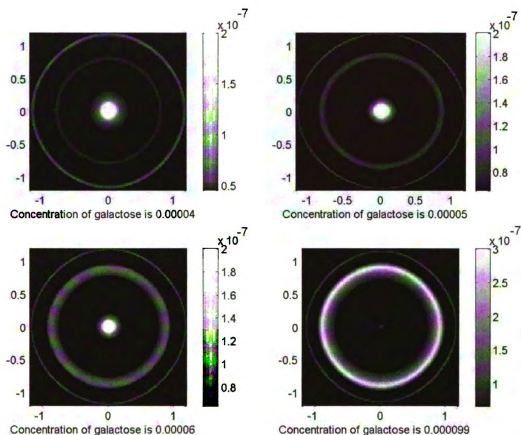


Figure 5.8: The picture shows the movement of bacteria when the concentration of galactose is below the oxygen.

The experiment also tested the movement of bacteria when the concentration of galactose was under a certain amount. In this case, increasing the concentration of galactose slowed down the movement of the first band and facilitated the movement of the second band. Since bacteria in the first traveled along and consumed all galactose, it would take more time for the bacteria to consume nutrient when the amount of galactose was increased. At the same time, the consumption of oxygen was also increased for the first band of bacteria. Therefore, less oxygen was left for the bacteria in second band. Bacteria in the second band would then needed less time to use it up.

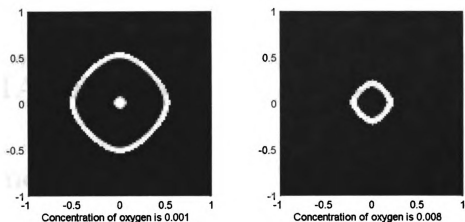


Figure 5.9: Relation of the amount of oxygen and the distance of the bands of bacteria.

When the concentration of galactose was too high, only the first band appeared during the experiment. The experiment of different oxygen levels was tested under this condition. It was observed that the higher the concentration of oxygen, the slower the band moved. The condition was similar to the galactose, the bacteria moved out only when they used up all the surrounding oxygen. Therefore, increasing the concentration of oxygen would decrease the speed of the band.

CHAPTER 6

Conclusion

This thesis develops a new numerical method to solve the reaction-diffusion-chemotaxis systems. This optimal two-stage method is to improve the performance of the explicit two-stage Runge-Kutta method for the time evolution problems. It retains the advantage of the explicit two-stage Runge-Kutta method of simple, easy implementation, and higher accuracy. In addition, the explicit formulation of the time step can be calculated. The time step is adapted and locally optimal. The consistency, convergence and the stability issues have been proved. The comparison between the optimal two-stage scheme and the two-stage Runge-Kutta method has been made. Table 1 and Figure 1 show that the optimal two-stage scheme is accurate and converges faster than the RK method for the time evolution problems. A real problem is also used to test out numerical method. Mathematical equations are constructed to model Adler's experiment in which patterns of two bands emerge. The numerical simulation and the experiment results are well matched.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] D. A. Anderson, J. C. Tannehill, R. H. Pletcher, *Computational fluid mechanics and heat transfer*, Series in computational fluid mechanics and heat transfer, Hemisphere Publishing Corporation.
- [2] VanBuskirk AM, Burlingham WJ, Jankowska-Gan E, Chin T, Kusaka S, Geissler F, Pelletier RP, Orosz, *Human allograft acceptance is associated with immune regulation*, J Clin Invest, Vol 106, 145-155, 2000.
- [3] U. Ascher, S. Ruuth, R. Spiteri, *Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. Special issue on time integration*. Appl. Numer. Math., 25, 151-167, 1997.
- [4] Julius Adler, *Chemotaxis in Bacteria*, Science, New Series, Vol 153, No.3737. Aug. 12, 1966, pp.708-716.
- [5] J. Adler and B. Templeton, *The effect of environmental conditions on the motility of Escherichia coli*, J. Gen. Microbiol. 46:175-184, 1967.
- [6] O. Baracchini and J. Sherris, J. Pathol.Bacteriol, 77, 565, 1959.
- [7] E. O. Budrene, H. C. Berg, *Complex patterns formed by motile cells of Escherichia coli*, Nature, 349, 630 - 633, 1991.
- [8] *Biology of the Chemotactic Response*, edited by J.M. Lackie, Cambridge University Press, Cambridge, 1986.
- [9] H. C. Berg, *Random Walks in Biology*, Princeton University Press, Princeton, NJ, 1989.
- [10] E. O. Budriené, A. A. Polezhaev, and M. O. Ptitsyn, *Mathematical modeling of intercellular regulation causing the formation of spatial structures in bacterial colonies*, J. Theor. Biol. 135, 323-341, 1988.

- [11] Chichia Chiu, *Optimal one-stage and two-stage schemes for steady state solutions of hyperbolic equations*, Applied Numerical Mathematics, 11, 1993, 475-496.
- [12] Mei-Qin Chen, Chichia Chiu, *Optimal m-stage Runge-Kutta Schemes for steady-State Solutions of hyperbolic equations*, Numerical methods for partial differential equations, 9, 643-666, 1993.
- [13] Chichia Chiu, David A. Kopriva, *An optimal Runge-Kutta method for steady state solutions of hyperbolic systems*, SIAM J.NUMER. ANAL, Vol 29, No2, pp 425-438, 1992.
- [14] Chichia Chiu, Frank Hoppensteadt, *Mathematical models and simulations of bacterial growth and chemotaxis in diffusion gradient chamber*. J. Math. Biol, Springer-Verlag, 2000.
- [15] G. W. Gear, *Numerical initial value problems in Ordinary Differential Equations*, Prentice-Hall, Inc. 1971.
- [16] Eshel Ben-Jacob, Inon Cohen, Ido Golding, Yonathan Kozlovsky, *Modeling branching and chiral colonial patterning of lubricating bacteria*, Mathematical models for biological pattern formation : Frontiers in Application of mathematics, edited by Philip K. Maini, Hans G. Othmer, The IMA Volumes in Mathematics and its applications Vol. 121, Springer-Verlag, 2001.
- [17] P. Grindrod, *The theory and applications of reaction-diffusion equations: patterns and waves, 2nd Ed*, Oxford University Press Inc., New York, 1996.
- [18] E. Keller, L. Segel, *Model for chemotaxis*, Jour. Theor. Biol.,30, 225-234, 1971a.
- [19] K. W. Morton, *Numerical Solution of convection-diffusion problems*, Applied Mathematics and Mathematical Computation 12, Chapman & Hall.
- [20] J. D. Murray, *Mathematical Biology. 3rd edition*, Springer-Verlag.
- [21] M. R. Myerscough, P. K. Maini, K. J. Painter, *Pattern formation in a generalized chemotactic model*, Bulletin of Mathematical Biology, 60, 1-26, 1998.
- [22] J. C. Sherris, N. W. Preston, J. G. Shoesmith, J. Gen. Microbiol. 16, 86, 1957.
- [23] G. D. Smith, *Numerical solution of partial differential equations: finite difference methods - 3rd Ed*. Oxford applied mathematics and computing solutions.
- [24] Scribner, T., Segel, L., Rogers, E., *A numerical study of the formation and propagation of traveling bands of chemotactic bacteria*, J. Theor. Biol, 189-219, 1974.

- [25] Turing, A., *The chemical basis of morphogenesis*. Proc. R. Soc. Lond. 237: 37-72, 1952.
- [26] Rebecca Tyson, L. G. Stern, Randall J. LeVeque, *Fractional step methods applied to a chemotaxis model*, J. Math. Biol. 41,455-475, 2000.
- [27] R. C. Tyson, S. Lubkin, J. Murray, *Model and analysis of chemotactic bacterial patterns in a liquid medium*, J. Math. Biol, 38, 359-375, 1999.