Stability and multiscale analysis of an advective lattice Boltzmann scheme

Martin Rheinländer

FB Mathematik und Statistik, Universität Konstanz,
Postfach D194, 78457 Konstanz, Germany
E-mail: martin.rheinlaender@uni-konstanz.de

Abstract: In this paper we consider a two-population lattice-Boltzmann algorithm to approximate the advection equation. First, the stability of this model algorithm is examined. The analysis is based on the analytic computation of the spectrum pertaining to the evolution matrix. After proving a necessary stability condition, the stability of the evolution matrix is shown, which is related to the CFL-condition. We use the model algorithm to demonstrate that formal stability criteria based on a multiscale expansion may fail to predict instability.

Keywords: lattice-Boltzmann; advection; eigenvalues; spectrum; stability; CFL-condition; convergence; asymptotic analysis; multiscale expansion.

Biographical notes: Martin Rheinländer graduated in Physics at the University of Essen in 1999. He then continued his studies by completing a *Diplôme d' Études Approfondies* in numerical analysis (University Paris VI). After spending some years at the Fraunhofer ITWM Kaiserslautern, where he became also acquainted with lattice Boltzmann methods, he started a dissertation in this field. In 2007, he obtained his Doctoral Degree (Dr. rer.nat.) in Mathematics from the University of Constance. Currently, he works in the group of Professor M. Junk participating in research and teaching.

1 Introduction

It is well known that consistency analysis plays a crucial role in understanding numerical schemes discretising differential equations. Especially in the context of lattice-Boltzmann schemes the relation between the numeric algorithm and the macroscopic target equation (e.g., Navier-Stokes etc.) is not obvious. Traditionally, this connection is established by the Chapman-Enskog expansion (Frisch et al., 1987; Wolf-Gladrow, 2000). Alternatively, the consistency analysis has been performed by analysing finite difference stencils extracted from equivalent moment systems (Junk, 2001) or recently by equivalent partial differential equations (Dubois, 2007).

Classical lattice-Boltzmann schemes work on uniform grids strictly coupling space and time step with each other. So, they are basically characterised by a single, scalar discretisation parameter h (grid spacing). This suggests to approach the consistency analysis by a regular asymptotic expansion with respect to h . Various papers introduce and apply this idea (Junk et al., 2005; Junk and Yang, 2005a, 2005c; Rheinländer, 2005). The cited works illuminate also, how this ansatz can be used to improve a given algorithm, which makes it attractive as a designing tool for the development of more powerful schemes.

Moreover, the regular expansion ansatz can be refined by taking additional time scales into account. By this way it is possible to resolve initial layers (Caiazzo, 2005) and to understand the long time behaviour (Junk and Rheinländer, 2008).

The general convergence theory (Lax equivalence theorem) requires consistency and *stability* (LeVeque, 1992). Whereas consistency analysis involves mainly elementary techniques like Taylor expansion, that may be applied in a quasi mechanical manner, the inspection of stability necessitates quite different arguments and often represents a real challenge.

This paper is an attempt to fathom whether it is possible to examine stability by asymptotic expansions just as in the case of consistency. Concretely, the paper is inspired by the following background-question: Can multiscale expansions be used to find precise stability criteria for finite difference schemes? For the specific example of the advection equation discretised with the upwind method, a positive answer has been given in Junk and Yang (2004). However this is not universally true as we demonstrate with an exemplary lattice-Boltzmann model algorithm, for which a twoscale expansion indicates stability while the method is actually unstable. For this we resume the twoscale expansion derived in the companion paper (Junk and Rheinländer, 2008), that is devoted to a profound

consistency and long-term behaviour analysis of the model algorithm.

Another motivation is to gain further thorough insights into the intrinsic workings of lattice-Boltzmann algorithms. A less complex algorithm, whose investigation is largely tractable with analytic methods, is well suited for this purpose.

The paper is organised in the following way: In Section 2 the model algorithm is presented and transformed in matrix notation. The spectrum of the evolution matrix is computed in Section 3. This helps to establish a necessary stability condition in Section 4. Furthermore the connection to the CFL condition is mentioned. Section 5 addresses the stability with respect to the ℓ_2 -norm, which is the main result from the mathematical point of view. This is followed in Section 6 by a comparison regarding the obtained stability results and the possible conclusions that can be drawn in this respect from the twoscale expansion. The paper ends with a glance on the relation between the multiscale expansion and the eigenvalues.

Finally, let us mention that there are still only few papers about rigorous stability results in the lattice-Boltzmann context. In some earlier papers like Lallemand and Luo (2000) and Sterling and Chen (1996) certain stability characteristics are computed in dependence of the wave number etc. These works are focused on the case of the D2Q9 lattice-Boltzmann algorithm for the Stokes and Navier-Stokes equation. A mathematical approach discussing stability and convergence of this algorithm for the case of the linear Stokes equation is found in Junk and Yang (2005b) and Junk and Yong (2006).

2 Model algorithm in matrix form

Let us briefly present the object of investigation – the lattice-Boltzmann model algorithm. Adopting the notation introduced in Junk and Rheinländer (2008) we denote by $F : \mathcal{T}_h \times \mathcal{X}_h \rightarrow \mathbb{R}^2$ the two-component population function that is defined on the temporal grid $\mathcal{T}_h = \{0, h, 2h, 3h, \dots\}$ and the spatial mesh $\mathcal{X}_h$. The latter one discretises the spatial domain $[0, 1]$ by setting $\mathcal{X}_h = \{\frac{1}{2}h, \frac{3}{2}h, \dots, \frac{2N-1}{2}h\}$ or alternatively $\mathcal{X}_h = \{0, h, \dots, (N-1)h\}$ with $h = \frac{1}{N}$ for some $N \in \mathbb{N}$. Note that the spatial and temporal increments are coupled via $\Delta x = \Delta t = h$.

As usual, the algorithm is defined by the difference equation

$$F_j(t+h, x+s_j h) = F_j(t, x) + [JF(t, x)]_j, \quad j=1, 2. \quad (1)$$

The arguments are restricted to the nodes of an underlying time-space grid, i.e., $(t, x) \in \mathcal{T}_h \times \mathcal{X}_h$. F_1, F_2 represent the distribution functions (populations) for imaginary left- and right-moving particles. The simple BGK collision operator J is given by the formula

$$JF = \omega(E(U) - F), \quad (2)$$

involving the *equilibrium* $E(U)$. Usually, this is considered as a function of the mass moment U :

$$E(U) := \frac{1}{2}(1 + as)U \quad \text{with} \quad U = F_1 + F_2. \quad (3)$$

Yet, it can also be seen as an operator (2×2 matrix) acting on the two-component population function F . Besides the relaxation parameter ω (collision frequency), that appears here as a purely algorithmic parameter, the advection velocity a occurs in the equilibrium. Its ‘macroscopic’ meaning becomes clear quite shortly.

As we want to consider only grids with finitely many nodes (indexed by $\{1, \dots, N\}$ from left to right) Equation (1) must be complemented by boundary conditions for the trailing populations in the left- and rightmost nodes. This is done by periodic boundary conditions where the leftmost node becomes the right neighbour of the rightmost one and vice versa. In formulae:

$$\begin{aligned} F_2(t+h, x_1) &= F_2(t, x_N) + [JF(t, x_N)]_2, \\ F_1(t+h, x_N) &= F_1(t, x_1) + [JF(t, x_1)]_1. \end{aligned}$$

In Junk and Rheinländer (2008) it was shown by examples and a consistency analysis that the mass moment U approximates the solution of the advection equation

$$\partial_t v + a \partial_x v = 0. \quad (4)$$

The corresponding initial value problem on $[0, 1]$ with periodic boundary conditions is referred to as the *target problem*. The initialisation of lattice-Boltzmann algorithm is an issue of its own. The most simple possibility is to set $F(0, x) = E(v_0(x))$ where v_0 denotes the initial value of the target problem.

Now let us cast the algorithm (1) in matrix form. Each iteration (evolution step) consists of two substeps. The first one is referred to as the *collision* executed as purely nodal operation. For this, in each node the collision products $\tilde{F}$ are computed from the current populations:

$$\begin{aligned} \tilde{F}(t, x) &:= (I + J)F(t, x) \\ \Leftrightarrow \begin{pmatrix} \tilde{F}_1(t, x) \\ \tilde{F}_2(t, x) \end{pmatrix} &:= \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} F_1(t, x) \\ F_2(t, x) \end{pmatrix}. \end{aligned}$$

Reading off the coefficients of the 2×2 collision block from Equations (1)–(3) one finds

$$\begin{aligned} \alpha &:= 1 - \frac{1}{2}\omega - \frac{1}{2}\omega a, & \beta &:= \frac{1}{2}\omega - \frac{1}{2}\omega a, \\ \gamma &:= \frac{1}{2}\omega + \frac{1}{2}\omega a, & \delta &:= 1 - \frac{1}{2}\omega + \frac{1}{2}\omega a. \end{aligned} \quad (5)$$

The iteration is completed by the *propagation* or *transport* of the collision products to their corresponding destination nodes. This second substep is non-nodal for the basic operation of copying a collision product involves two different nodes simultaneously.

If we assume, for the purpose of illustration, a grid with four nodes as shown in Figure 1 and consider all eight populations assembled in a single column vector,

Figure 1 A left and right moving population is located in each grid node (dashed crosses). The suggested numbering of the populations is rather convenient as it results in a simple, block-structured matrix representing the transport operator. The curved arrows indicate the propagation of the collision products. We assume *periodic* boundary conditions to handle the transport of the collision products indexed by 1 and 8

the collision and transport step can be performed by applying the following matrices:

$$\mathbf{C} = \left(\begin{array}{cccc|cccc} \alpha & \cdot & \cdot & \cdot & \beta & \cdot & \cdot & \cdot \\ \cdot & \alpha & \cdot & \cdot & \cdot & \beta & \cdot & \cdot \\ \cdot & \cdot & \alpha & \cdot & \cdot & \cdot & \beta & \cdot \\ \cdot & \cdot & \cdot & \alpha & \cdot & \cdot & \cdot & \beta \\ \hline \gamma & \cdot & \cdot & \cdot & \delta & \cdot & \cdot & \cdot \\ \cdot & \gamma & \cdot & \cdot & \cdot & \delta & \cdot & \cdot \\ \cdot & \cdot & \gamma & \cdot & \cdot & \cdot & \delta & \cdot \\ \cdot & \cdot & \cdot & \gamma & \cdot & \cdot & \cdot & \delta \end{array} \right),$$

$$\mathbf{T} = \left(\begin{array}{cccc|cccc} \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{array} \right).$$

Due to the special indexation of the populations both matrices reveal a clear block structure. In particular, we recognise that the collision matrix $\mathbf{C}$ is composed of four diagonal matrices that are multiples of the 4×4 identity matrix. On the other hand the transport matrix $\mathbf{T}$ adopts a block-diagonal form. The upper block $\mathbf{L}$ acts on a column vector like an up-shift which corresponds to a *left-shift* if the vector components are plotted vs. their indices. Similarly, the effect of the lower block $\mathbf{R}$ is a down-shift or *right-shift* (see Fig. 1).

The evolution matrix is the product of $\mathbf{C}$ and $\mathbf{T}$ where the apparently inverse position of the factors just reflects the order of the substeps.

$$\mathbf{E} = \mathbf{T} \cdot \mathbf{C} = \begin{pmatrix} \mathbf{L} & 0 \\ 0 & \mathbf{R} \end{pmatrix} \begin{pmatrix} \alpha \mathbf{I} & \beta \mathbf{I} \\ \gamma \mathbf{I} & \delta \mathbf{I} \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{L} & \beta \mathbf{L} \\ \gamma \mathbf{R} & \delta \mathbf{R} \end{pmatrix}. \quad (6)$$

So, the evolution matrix $\mathbf{E}$ becomes a 2×2 block matrix with four $N \times N$ blocks. By means of $\mathbf{E}$, the lattice-Boltzmann algorithm (1) is brought in the general form

$$\mathbf{F}(kh + h) = \mathbf{E} \mathbf{F}(kh) = \mathbf{E}^{k+1} \mathbf{F}(0). \quad (7)$$

Here $\mathbf{F}(kh)$ denotes a vector containing the populations of the whole grid – suitably arranged as components – at the k th iteration corresponding to the time $t = kh$. Thus the algorithm fits in the class of explicit schemes taking up the form of a simple recursion, that is solved by the iterative application of the evolution matrix.

3 Spectrum of the evolution matrix

The evolution of linear recursions is essentially determined by the spectrum (set of eigenvalues) of the evolution matrix. Consequently, we start our investigation of stability by examining the eigenvalues of $\mathbf{E}$ defined in Equation (6).

Traditionally this is done by applying the von Neumann analysis in the following way: Exploiting the analogy between difference and differential equations with constant coefficients and periodic boundary conditions, an exponential ansatz is chosen for the eigenvectors. This is inserted into Equation (1) and the supplementing boundary conditions. Eventually this yields equations equivalent to Equations (10) and (11). In the following presentation we execute essentially the same steps but prefer to stay in the framework of matrix notation.

Since $\mathbf{E}$ is composed of the $N \times N$ shift matrices $\mathbf{L}$ and $\mathbf{R} = \mathbf{L}^\top = \mathbf{L}^{-1}$, it is reasonable to first look at their eigenvalues and eigenvectors. By elementary computations, it is found that the spectrum is given by the N unit roots of N th order. Setting

$$w := e^{\frac{2\pi i}{N}} \quad \text{and} \quad \mathbf{v}_m := (w^{\ell m})_{\ell \in \{0, \dots, N-1\}}$$

we obtain for $m \in \{0, \dots, N-1\}$ the following eigenvector equations:

$$\mathbf{L} \mathbf{v}_m = w^m \mathbf{v}_m \quad \text{and} \quad \mathbf{R} \mathbf{v}_m = w^{-m} \mathbf{v}_m \equiv \bar{w}^m \mathbf{v}_m.$$

Hence $\mathbf{v}_m$ is simultaneously eigenvector of $\mathbf{L}$ and $\mathbf{R}$ associated to the eigenvalues w^m and its complex conjugate $\bar{w}^m$ respectively.

It is quite conceivable that an eigenvector $\mathbf{V}$ of $\mathbf{E}$ is closely connected to those of the shift-matrices. This motivates the ansatz

$$\mathbf{V} := \begin{pmatrix} \mathbf{v}_m \\ a \mathbf{v}_m \end{pmatrix} \quad (8)$$

with some unknown amplitude a . Inserting (8) into the eigenvalue equation where λ denotes the eigenvalue to be determined as well, we obtain the following equation by exploiting directly the properties of $\mathbf{v}_m$:

$$\begin{pmatrix} \alpha \mathbf{L} & \beta \mathbf{L} \\ \gamma \mathbf{R} & \delta \mathbf{R} \end{pmatrix} \begin{pmatrix} \mathbf{v}_m \\ a \mathbf{v}_m \end{pmatrix} = \begin{pmatrix} \alpha w^m \mathbf{v}_m + a \beta w^m \mathbf{v}_m \\ \gamma \bar{w}^m \mathbf{v}_m + a \delta \bar{w}^m \mathbf{v}_m \end{pmatrix} \stackrel{!}{=} \lambda \begin{pmatrix} \mathbf{v}_m \\ a \mathbf{v}_m \end{pmatrix}. \quad (9)$$

Comparison of coefficients with respect to $\mathbf{v}_m$ yields two equations¹

$$\alpha w^m + a\beta w^m = \lambda \quad (10)$$

$$\gamma \bar{w}^m + a\delta \bar{w}^m = \lambda a \quad (11)$$

that must be used to determine a and λ . For the moment we are primarily interested in the eigenvalues. Multiplying Equation (10) by λ , inserting Equation (11) for $a\lambda$ and using Equation (10) to substitute $a\beta w^m$ to eliminate a , finally yields the following quadratic equation

$$\lambda^2 - (\alpha w^m + \delta \bar{w}^m)\lambda + \alpha\delta - \beta\gamma = 0 \quad (12)$$

that fixes the eigenvalue λ . Now the question arises: Are all eigenvalues of the evolution matrix $\mathbf{E}$ obtained as roots of Equation (12) for $m \in \{0, \dots, N-1\}$? To pose the question in other words, we could also ask whether all eigenvectors fit to the ansatz (8).

The question could be answered affirmatively if one manages to verify that the ansatz leads to $2N$ different eigenvalues or that it produces $2N$ linearly independent eigenvectors (eigenbasis). Here we proceed by another reasoning to ensure that the ansatz really captures all eigenvalues. For this we define the unitary $N \times N$ matrix²

$$\mathbf{F} := \frac{1}{\sqrt{N}} [\bar{\mathbf{v}}_0, \dots, \bar{\mathbf{v}}_{N-1}]$$

whose columns are given by the complex conjugated $\mathbf{v}_m$'s. Due to its symmetry, the inverse of $\mathbf{F}$ is just its complex conjugated, i.e., $\mathbf{F}^{-1} = \bar{\mathbf{F}}$. As being composed of the eigenvectors belonging to $\mathbf{L}, \mathbf{R}$, it is no surprise that $\mathbf{F}$ and $\mathbf{F}^{-1}$ diagonalise the shift-matrices. More precisely one finds

$$\begin{aligned} \mathbf{F}\mathbf{L}\mathbf{F}^{-1} &= \mathbf{W} \quad \Leftrightarrow \quad \mathbf{F}^{-1}\mathbf{L}\mathbf{F} = \bar{\mathbf{W}} \\ \mathbf{F}^{-1}\mathbf{R}\mathbf{F} &= \mathbf{W} \quad \Leftrightarrow \quad \mathbf{F}\mathbf{R}\mathbf{F}^{-1} = \bar{\mathbf{W}} \end{aligned}$$

with $\mathbf{W} := \text{diag}(1, w, w^2, \dots, w^{N-1})$. Employing these relations, $\mathbf{E}$ is converted by the following similarity transformation

$$\begin{pmatrix} \mathbf{F} & 0 \\ 0 & \mathbf{F} \end{pmatrix} \begin{pmatrix} \alpha\mathbf{L} & \beta\mathbf{L} \\ \gamma\mathbf{R} & \delta\mathbf{R} \end{pmatrix} \begin{pmatrix} \mathbf{F}^{-1} & 0 \\ 0 & \mathbf{F}^{-1} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{W} & \beta\mathbf{W} \\ \gamma\bar{\mathbf{W}} & \delta\bar{\mathbf{W}} \end{pmatrix} =: \hat{\mathbf{E}}$$

into a matrix $\hat{\mathbf{E}}$ whose four blocks are all diagonal $N \times N$ matrices. By means of a further unitary similarity transformation with an appropriate $2N \times 2N$ permutation matrix $\mathbf{P}$ (note $\mathbf{P}^\top = \mathbf{P}^{-1}$) swapping lines from the left and columns from the right we define

$$\tilde{\mathbf{E}} := \mathbf{P}\hat{\mathbf{E}}\mathbf{P}^\top. \quad (13)$$

$\tilde{\mathbf{E}}$ attains the form of a block-diagonal matrix with 2×2 blocks on its diagonal. Recall that the determinant and a fortiori the characteristic polynomial remain invariant under similarity transformations. Furthermore the determinant of a block-diagonal matrix is equal to the product of determinants pertaining to each block. Hence

the characteristic polynomial of the evolution matrix becomes

$$\begin{aligned} \chi_{\mathbf{E}}(\lambda) &= \prod_{m=0}^{N-1} \det \begin{pmatrix} \alpha w^m - \lambda & \beta w^m \\ \gamma \bar{w}^m & \delta \bar{w}^m - \lambda \end{pmatrix} \\ &= \prod_{m=0}^{N-1} [\lambda^2 - (\alpha w^m + \delta \bar{w}^m)\lambda + \alpha\delta - \beta\gamma]. \end{aligned} \quad (14)$$

In fact, this shows, that by solving the N quadratic equations given in Equation (12) for $m \in \{0, \dots, N-1\}$ no eigenvalue will evade.

If we write $w^m = e^{i\phi}$ with $\phi = \frac{2\pi m}{N}$, the factors of the characteristic polynomial (14) may be considered as the same parameterised polynomial differing only in the value of the angle ϕ . Therefore let us set

$$\begin{aligned} \chi(\lambda, \phi) &:= \lambda^2 - (\alpha e^{i\phi} + \delta e^{-i\phi})\lambda + \alpha\delta - \beta\gamma \\ &= \lambda^2 + [(\omega - 2)\cos(\phi) + i\omega a \sin(\phi)]\lambda + 1 - \omega, \end{aligned} \quad (15)$$

where we have reinserted Equation (5) to obtain Equation (15). The set of zeros of $\chi(\lambda, \phi)$,

$$\mathfrak{S}(a, \omega) := \{\lambda \in \mathbb{C} \mid \exists \phi \in [0, 2\pi) \text{ with } \chi(\lambda, \phi) = 0\}, \quad (16)$$

where the parameter ϕ varies inside the period length of $e^{i\phi}$, forms curves in the complex plane. For fixed a and ω the set $\mathfrak{S}(a, \omega)$ contains all eigenvalues of the evolution matrix independently of the grid spacing h or the number of grid nodes N . With increasing N the eigenvalues sample $\mathfrak{S}(a, \omega)$ more and more densely; that is why we call it the *spectral limit set*. Its importance for stability investigations derives from this condensation property. The spectral limit set lies within the closed unit disk if and only if $\rho(\mathbf{E}) \leq 1$ independently³ of h . Thus, $\mathfrak{S}(\omega, a) \subset \overline{D_1(0)}$ is a necessary condition for stability.

Figures 2 and 3 illustrate how some spectral limit sets look like for fixed ω and varying a . In certain cases like $a = 0$ or $|a| = 1$ the spectral limit set becomes geometrically simple and can be described by line segments and circles.

The zeros of $\chi(\lambda, \phi)$ are analytically obtained by the standard formula for solving quadratic equations.

$$\begin{aligned} \lambda_{\pm}(\phi) &= -\frac{1}{2} [(\omega - 2)\cos(\phi) + i\omega a \sin(\phi)] \\ &\quad \pm \sqrt{\frac{1}{4} [(\omega - 2)\cos(\phi) + i\omega a \sin(\phi)]^2 - (1 - \omega)}. \end{aligned} \quad (17)$$

As the radicand is generally complex-valued, it is not immediately obvious to estimate the moduli of the zeros and thus to find out conditions on the parameters a, ω providing stability.

Remark on the spectral branches and the eigenvectors

It is obvious to associate with λ_- and λ_+ two spectral branches.⁴ Eigenvectors with constant populations belong to the so-called *sink points* that are obtained for $\phi = 0$

Figure 2 Sequence of spectral limit sets: underrelaxed case with $\omega = 0.6$. The radius of the small circle is $|1 - \omega|$ (also in Fig. 3). For $a = 0$ the two branches coincide on a circle of radius $\sqrt{1 - \omega}$. The markers indicate eigenvalues associated with a grid of 20 nodes. Dots mark the λ_- -branch while circles indicate the λ_+ -branch (also below)

with $\lambda_+(0) = 1$ and $\lambda_-(0) = 1 - \omega$. Each sink point is accompanied by a *source point*. Both of them lie on the real line where the source point is in some cases symmetrically

situated to its corresponding sink point with regard to the imaginary axis. Eigenvectors of high spatial frequency (shortwave modes) refer to the eigenvalues near the source

Figure 3 Sequence of spectral limit sets: overrelaxed case with $\omega = 1.5$. In the plots of the upper row, sink- and source-points belonging to the same branch are not symmetric with respect to the imaginary axis

points while those with low spatial frequency (longwave modes) belong to eigenvalues around the sink points. Moving along a branch from eigenvalue to eigenvalue towards the sink point, the spatial oscillations of the associated eigenvectors decreases by 1. When the number of grid points is increased, it seems as if new eigenvalues are generated near the source points where eigenvectors reveal higher spatial frequencies than existing before. The other eigenvalues are apparently shifted in direction of the sink points.

4 The CFL-condition and stability

The discretisation of (hyperbolic) PDEs necessitates the observation of the CFL-condition being indispensable for convergence. It requires the domain of dependence of the discretisation method to contain the domain of dependence of the PDE (see also LeVeque, 1992). For a three-point stencil difference scheme solving the one-dimensional advection equation the CFL-condition translates into

$$\left| a \frac{\Delta t}{\Delta x} \right| \leq 1. \quad (18)$$

The inequality admits an easy interpretation, namely that the grid speed $\Delta x/\Delta t$ must not be inferior to the advection speed $|a|$.

Remarkably, the CFL-condition is equivalent with stability in the case of standard finite-difference schemes like Upwind, Lax-Friedrich or Lax-Wendroff. For the lattice-Boltzmann algorithm of Section 2 a similar equivalence holds as shown in Theorem 2. Before turning to this, we verify that $|a| \leq 1$ is necessary for stability which is obtained from Equation (18) by setting $\Delta t = \Delta x = h$.

Theorem 1: *The spectral limit set is entirely contained in the closed unit disk, i.e., $\mathfrak{S}(\omega, a) \subset \overline{D_1(0)}$ if and only if $0 \leq \omega \leq 2$ and $-1 \leq a \leq 1$ or $\omega = 0$.*

If $\mathfrak{S}(\omega, a) \not\subset \overline{D_1(0)}$, then for all but finitely many grids the corresponding evolution matrix has an eigenvalue whose distance from $\overline{D_1(0)}$ must be greater than a constant minimal distance since the eigenvalues fill $\mathfrak{S}(\omega, a)$ densely. Hence the evolution matrix admits an eigenvalue greater than 1, so that $\|\mathbf{E}^n\|$ cannot be bounded independently of $n \in \mathbb{N}$ whatever matrix norm $\|\cdot\|$ is chosen.

The proof of Theorem 1 is furnished by the following lemmas distinguishing several cases by decomposing the $a\omega$ -plane as depicted in Figure 4. While Lemmas 1 and 2 characterise the area of instability, Lemma 3 demarcates the stability domain where $\rho(\mathbf{E}) \leq 1$. The verification of this main lemma is based on the theorem of *Rouché* ($\rightarrow$ appendix) from complex analysis. Finally, each of the Lemmas 4–6 deals with parts of the boundary of the stability domain.

Lemma 1: *Suppose $\omega < 0$ or $\omega > 2$. Then $\mathfrak{S}(\omega, a) \not\subset \overline{D_1(0)}$.*

Figure 4 The stability domain of $\mathbf{E}$ in the $a\omega$ -plane. The grey shading and line types visualises the case differentiation to prove Theorem 1. As the collision matrix corresponds to the identity if ω is set to 0, the whole line $\omega = 0$ belongs – apart from the closed rectangle $[-1, 1] \times [0, 2]$ – to the stability domain too

Proof: The constant term of the quadratic polynomial $\chi(\phi, \lambda)$ is $1 - \omega$. According to *Vieta's* rule the zeros of $\lambda \mapsto \chi(\phi, \lambda)$ satisfy $\lambda_-(\phi) \cdot \lambda_+(\phi) = 1 - \omega$. Now the hypothesis entails $|1 - \omega| > 1$ which implies that the modulus of at least one zero is greater than 1. Then this zero lies outside the closed unit disk. $\square$

Lemma 2: *Let $0 < \omega \leq 2$ and $|a| > 1$. Then $\mathfrak{S}(\omega, a) \not\subset \overline{D_1(0)}$.*

Proof: Assume $a > 1$. The assertion is proved by showing $|\lambda_+(\frac{\pi}{2})| > 1$. Setting $\phi = \frac{\pi}{2}$ in Equation (17) yields:

$$\lambda_+\left(\frac{\pi}{2}\right) = \frac{1}{2}\omega a i + \sqrt{\omega - 1 - \frac{1}{4}\omega^2 a^2}.$$

A small computation confirms that the root is imaginary since the radicand is negative for arbitrary ω and $|a| > 1$:

$$\omega - 1 - \frac{1}{4}\omega^2 a^2 = -\left(1 - \frac{1}{2}\omega\right)^2 - \frac{1}{4}\omega^2 \overbrace{(a^2 - 1)}^{>0} < 0.$$

Hence $\lambda_+(\frac{\pi}{2}) = i\left(\frac{1}{2}\omega a + \sqrt{1 + \frac{1}{4}\omega^2 a^2 - \omega}\right)$. Finally, this leads to the estimate

$$\begin{aligned} \left| \lambda_+\left(\frac{\pi}{2}\right) \right| &= \frac{1}{2}\omega a + \sqrt{1 + \frac{1}{4}\omega^2 a^2 - \omega} \\ &> \frac{1}{2}\omega + \sqrt{1 + \frac{1}{4}\omega^2 - \omega} \\ &= \frac{1}{2}\omega + \left(1 - \frac{1}{2}\omega\right) = 1, \end{aligned}$$

where we have used the hypothesis $0 < \omega \leq 2$ and $a > 1$. Similarly one can show that $\lambda_-(\frac{\pi}{2})$ is located outside the unit circle if $a < -1$. $\square$

Lemma 3: If $0 < \omega < 2$ and $|a| < 1$ then $\mathfrak{S}(\omega, a) \subset \overline{D_1(0)}$.

Proof: We have to show that, under the above hypothesis, the roots of Equation (15) lie inside the closed unit disk independently of $\phi \in [0, 2\pi)$. For this, we distinguish two cases. The first one with $\phi \in \{0, \pi\}$ can be treated directly:

$$\begin{aligned}\chi(\lambda, 0) &= \lambda^2 + (\omega - 2)\lambda + (1 - \omega), \\ \chi(\lambda, \pi) &= \lambda^2 + (2 - \omega)\lambda + (1 - \omega).\end{aligned}$$

Both polynomials do not depend on a . It is easily seen that the zeros of $\chi(\cdot, 0)$ are 1 and $1 - \omega$, while the zeros of $\chi(\cdot, \pi)$ are -1 and $\omega - 1$. So, they are located inside the closed unit disk if $0 < \omega < 2$.

The second case concerns the general situation $\phi \in [0, 2\pi) \setminus \{0, \pi\} = (0, \pi) \cup (\pi, 2\pi)$, which is handled by using the theorem of *Rouché* ($\rightarrow$ appendix). To apply it, we need a function of comparison ψ , that must have as many zeros inside the *open* unit disk as we claim that χ has. So ψ is required to have two zeros. Moreover it must satisfy the following estimate on the unit circle:

$$|\chi(z) - \psi(z)| < |\psi(z)| \quad \text{for all } z \in \partial D_1(0).$$

We choose for ψ the polynomial χ itself *but* setting $\omega = 1$:

$$\psi(\lambda, \phi) := \lambda^2 + [-\cos(\phi) + ia \sin(\phi)]\lambda.$$

The zeros of ψ are 0 and $\cos(\phi) - ia \sin(\phi)$. If $\phi \neq 0$ and $\phi \neq \pi$ then $\cos(\phi) \neq 1$ and $\sin(\phi) \neq 0$. This implies⁵ that both zeros of ψ are located inside the *open* unit disk assuming the hypothesis $|a| < 1$.

Let us now verify the required estimate. We start with the right hand side and replace z by $e^{i\theta} \in \partial D_1(0)$ with $\theta \in [0, 2\pi)$. Drawing the factor $(\omega - 1)e^{i\theta}$ out of the modulus, using $|e^{i\theta}| = 1$ and observing that the absolute value does not change under complex conjugation, we get:

$$\begin{aligned}|\chi(e^{i\theta}, \phi) - \psi(e^{i\theta}, \phi)| &= |(\omega - 2)\cos(\phi)e^{i\theta} + \cos(\phi)e^{i\theta} \\ &\quad + ia(\omega - 1)\sin(\phi)e^{i\theta} + (1 - \omega)| \\ &= |\omega - 1| |\cos(\phi)e^{i\theta} + ia \sin(\phi)e^{i\theta} - 1| \\ &= |\omega - 1| |e^{i\theta}| |\cos(\phi) + ia \sin(\phi) - e^{-i\theta}| \\ &= |\omega - 1| |\overline{\cos(\phi) + ia \sin(\phi) - e^{-i\theta}}| \\ &= |\omega - 1| |\cos(\phi) - ia \sin(\phi) - e^{i\theta}|.\end{aligned}$$

For the left hand side we find by factoring out $-e^{i\theta}$:

$$\begin{aligned}|\psi(e^{i\theta}, \phi)| &= |e^{2i\theta} + (-\cos(\phi) + ia \sin(\phi))e^{i\theta}| \\ &= |-e^{i\theta}| |-e^{i\theta} - (-\cos(\phi) + ia \sin(\phi))| \\ &= |\cos(\phi) - ia \sin(\phi) - e^{i\theta}|.\end{aligned}$$

So we can conclude:

$$\begin{aligned}|\chi(e^{i\theta}, \phi) - \psi(e^{i\theta}, \phi)| &= |\omega - 1| |\cos(\phi) - ia \sin(\phi) - e^{i\theta}| \\ &< |\cos(\phi) - ia \sin(\phi) - e^{i\theta}| = |\psi(e^{i\theta}, \phi)|.\end{aligned}$$

The strict inequality holds, indeed, because the assumption $\omega \in (0, 2)$ implies $|1 - \omega| < 1$. Furthermore $|\cos(\phi) - ia \sin(\phi) - e^{i\theta}|$ cannot be zero, since $e^{i\theta}$ is on the unit circle whereas $\cos(\phi) - ia \sin(\phi)$ lies in its interior as justified above. Thus, all assumptions of *Rouché's* theorem are fulfilled which gives eventually the assertion. $\square$

Lemma 4: Assume $0 \leq \omega \leq 2$ and $|a| = 1$. Then $\mathfrak{S}(\omega, a) \subset \overline{D_1(0)}$.

Proof: Setting $a = 1$ in Equation (15), it is to show that the zeros of the polynomial

$$\chi(\lambda, \phi) = \lambda^2 + [(\omega - 2)\cos(\phi) + i\omega \sin(\phi)]\lambda + (1 - \omega)$$

lie in the closed unit disk. It is easy to check that $e^{-i\phi}$ is a root of χ . By *Vieta's* rule the product of the zeros must be equal to the coefficient of the constant term. Hence the other zero must be given by $(1 - \omega)e^{i\phi}$. So both of the zeros are located inside the closed unit disk under the hypothesis. The other case $a = -1$ is treated likewise. $\square$

Lemma 5: Suppose $\omega = 0$. Then $\mathfrak{S}(\omega, a) \subset \partial D_1(0) \subset \overline{D_1(0)}$.

Proof: In this case the collision matrix becomes the identity. Hence the evolution matrix is just the transport matrix. So the spectrum lies on the unit circle. $\square$

Lemma 6: Suppose $\omega = 2$ and $|a| \leq 1$. Then $\mathfrak{S}(\omega, a) \subset \partial D_1(0) \subset \overline{D_1(0)}$.

Proof: In this case Equation (17) simplifies to $\lambda_{\pm}(\phi) = -ia \sin(\phi) \pm \sqrt{1 - a^2 \sin^2(\phi)}$. As long as $|a| \leq 1$ the root is real, and we get $|\lambda_{\pm}(\phi)|^2 = a^2 \sin^2(\phi) + 1 - a^2 \sin^2(\phi) = 1$. $\square$

5 Stability in the ℓ_2 -norm

Theorem 1 already gives the necessary fundamental condition for stability. If $\mathbf{E}$ is diagonalisable⁶ ($\rightarrow$ appendix, Corollary 1) and $\rho(\mathbf{E}) \leq 1$ the powers of the evolution matrix remain bounded. In this section we show that the bound is independent of the underlying grid which defines the size (dimension) of the evolution matrix. It must be noted that this statement depends substantially⁷ on the norm. An illustration for this is provided by Figure 5.

Theorem 2: The lattice-Boltzmann algorithm described in Section 2 with the evolution matrix $\mathbf{E}$ defined in Equation (6) is stable with respect to the ℓ_2 -norm if and only if $0 \leq \omega \leq 2$ and $-1 \leq a \leq 1$, or $\omega = 0$.

So, under this condition on ω and a , there exists a constant $C > 0$ being *independent* of the grid spacing (time step) h such that $\|\mathbf{E}^n\|_2 < C$ for all $n \in \mathbb{N}_0$.

In contrast, Theorem 1 just guarantees the existence of such a constant which might depend on h entailing

Figure 5 The upper diagram visualises the maximum-norm $\|\mathbf{E}^n\|_\infty$ over the iteration index $n \in \{0, \dots, 200\}$. For clarity only every fifth value is plotted, which lets the curves appear a little bit smoother. The curves refer to grids of 20, 40, 80, 160 and 320 nodes with $\omega = 2$ and $a = 0.8$. Analogously, the lower diagram has to be interpreted except that it refers to $\|\mathbf{E}^n\|_2$. Evidently, as far as the two-norm $\|\cdot\|_2$ is concerned, one realises a common bound for the powers of the evolution matrices associated to different grids; regarding the $\|\cdot\|_\infty$ -norm this bound seems to be lacking

that $\|\mathbf{E}^n\|_2$ is uniformly bounded with respect to n but not necessarily with respect to h . This subtle difference sometimes gives rise to confusion since, from a practical point of view, the asymptotic behaviour does not matter that much as the algorithm can be run only on a finite number of grids.

Before we start with the proof let us mention that the matrix $\|\cdot\|_2$ -norm is induced by the vectorial $\|\cdot\|_2$ -norm.

$$\begin{aligned} \|M\|_2^2 &:= \sup_{\|x\|=1} \|Mx\|_2^2 = \sup_{\|x\|=1} \langle Mx, Mx \rangle \\ &= \sup_{\|x\|=1} \langle MM^*x, x \rangle = \rho(MM^*). \end{aligned}$$

From this ensues that the (matricial) $\|\cdot\|_2$ -norm keeps invariant under unitary similarity transformations as unitary matrices generate the isometries of the scalar product.

Proof: Defining

$$B(\phi) := \begin{pmatrix} \alpha e^{i\phi} & \beta e^{i\phi} \\ \gamma e^{-i\phi} & \delta e^{-i\phi} \end{pmatrix} \quad (19)$$

it was established in Section 3 (see Eq. (13)) that the evolution matrix $\mathbf{E}$ is similar to the block-diagonal matrix

$$\tilde{\mathbf{E}} = \begin{pmatrix} B(0) & & & \\ & B(\frac{2\pi}{N}) & & \\ & & \ddots & \\ & & & B(\frac{2\pi(N-1)}{N}) \end{pmatrix} \quad (20)$$

by a unitary transformation. Moreover all powers of $\mathbf{E}$ are unitarily similar to those of $\tilde{\mathbf{E}}$ too. This implies

$\|\mathbf{E}^n\|_2 = \|\tilde{\mathbf{E}}^n\|_2$. Evaluating the norm by means of Lemma 7, we obtain

$$\begin{aligned} \|\mathbf{E}^n\|_2 = \|\tilde{\mathbf{E}}^n\|_2 &= \max_{m \in \{0, \dots, N-1\}} \left\| B\left(\frac{2\pi m}{N}\right)^n \right\|_2 \\ &\leq \sup_{\phi \in [0, 2\pi]} \|B(\phi)^n\|_2. \end{aligned} \quad (21)$$

Thanks to the estimate at right, we get rid of the high dimensional matrix $\tilde{\mathbf{E}}$ and the problem is formulated in terms of the family of 2×2 block-matrices $B(\phi)$. To finish the proof we have to bound the right hand side; for this two cases $a = 0$ and $a \neq 0$ are distinguished.

Case 1: $a = 0$

In this situation we directly calculate $\|B(\phi)\|_2^2 = \rho(B(\phi)B(\phi)^*)$. The computation

$$\begin{aligned} &\det(B(\phi)B(\phi)^* - \lambda I) \\ &= \det \begin{pmatrix} \alpha^2 + \beta^2 - \lambda & (\alpha\gamma + \beta\delta)e^{i\phi} \\ (\alpha\gamma + \beta\delta)e^{-i\phi} & \gamma^2 + \delta^2 - \lambda \end{pmatrix} \\ &= \lambda^2 - (\alpha^2 + \beta^2 + \gamma^2 + \delta^2)\lambda + (\alpha\delta - \beta\gamma)^2 \\ &= \lambda^2 - (2 - 2\omega + \omega^2)\lambda + (1 - \omega)^2 \end{aligned}$$

yields the eigenvalues 1 and $(1 - \omega)^2$ pertaining to $B(\phi)B(\phi)^*$. Independently of ϕ we find $\|B(\phi)\|_2 = 1$ for $\omega \in [0, 2]$. Hence it follows with the standard estimate

$$\sup_{\phi \in [0, 2\pi]} \|B(\phi)^n\|_2 \leq \sup_{\phi \in [0, 2\pi]} \|B(\phi)\|_2^n \leq 1.$$

Case 2: $a \neq 0$

Here the direct argument does not work although $\|B(\phi)\|_2$ is also independent of ϕ . Actually we have $\|B(\phi)\|_2 > 1$

whenever the advection velocity does not vanish. Therefore let us introduce the following sequence of positive functions for $n \in \mathbb{N}$:

$$f_n : [0, 2\pi] \rightarrow \mathbb{R}, \quad f_n(\phi) := \|B^n(\phi)\|_2.$$

As $B(\phi)$ depends continuously on ϕ and due to the continuity of the matrix multiplication and the matricial $\|\cdot\|_2$ -norm, the functions f_n are continuous too.

According to Lemma 9 the matrices $B(\phi)$ with $\phi \in [0, 2\pi]$ are diagonalisable if $a \neq 0$. From Theorem 1 follows furthermore: $\rho(B(\phi)) \leq 1$. Hence Lemma 8 ($\rightarrow$ appendix) guarantees that arbitrary powers of $B(\phi)$ remain bounded.

This shows that the function sequence $(f_n)_{n \in \mathbb{N}}$ is *pointwise bounded*, i.e.,

$$\exists C_\phi > 0 : \forall n \in \mathbb{N} : f_n(\phi) < C_\phi.$$

So we can apply the *principle of uniform boundedness* ($\rightarrow$ appendix) ensuring that the sequence $(f_n)_{n \in \mathbb{N}}$ is even *locally bounded*. Eventually, the compactness of the domain of definition, which is the interval $[0, 2\pi]$, yields global boundedness. So the right hand side of Equation (21) must be bounded. $\square$

The case $a = 0$ has been treated separately, because proving the diagonalisability of $B(\phi)$ (Lemma 9) needs some more effort. Observe that Case 2 can be made more explicit, by writing $B(\phi)^n = W(\phi)D(\phi)^nW(\phi)^{-1}$ where the diagonal matrix $D(\phi)$ contains only elements of modulus less or equal 1.

6 Multiscale expansion and stability

Let us briefly recall the results of Junk and Rheinländer (2008) concerning the twoscale expansion with a linear and quadratic time $t_1 = nh$ and $t_2 = nh^2$. The idea is to approximate the population function F of the lattice-Boltzmann algorithm in the following way:

$$F(nh, x) \approx f^{(0)}(nh, nh^2, x) + hf^{(1)}(nh, nh^2, x) + \dots \quad (22)$$

The asymptotic order functions $f^{(0)}, f^{(1)}, \dots$ are defined such that the order of the residual (with respect to h) becomes as large as possible. In contrast to a regular expansion with a single time, the order functions do not come out in a cogent manner. Therefore additional conditions may be imposed simplifying the calculation as much as possible without loosing the flexibility of the ansatz. So the subsequent PDEs are extracted, which determine the evolution of the mass moment $u^{(0)}$ associated to $f^{(0)} = \frac{1}{2}(1 + as)u^{(0)}$.

$$\partial_{t_1} u^{(0)}(t_1, t_2, x) + a \partial_x u^{(0)}(t_1, t_2, x) = 0 \quad (23)$$

$$\partial_{t_2} u^{(0)}(t_1, t_2, x) - \mu \partial_x^2 u^{(0)}(t_1, t_2, x) = 0 \quad (24)$$

with

$$\mu = \left(\frac{1}{\omega} - \frac{1}{2} \right) (1 - a^2). \quad (25)$$

The advection Equation (23) is already found using the regular expansion. It describes the most conspicuous behaviour of $u^{(0)}$ because t_1 evolves faster than t_2 . In opposition, the diffusion equation accounts for a less striking process, that becomes only noticeable after an observation over rather many iterations.

The sign of diffusivity μ plays a crucial role. In case of $\mu \geq 0$ the solutions of the diffusion equation, being damped and smoothed, exhibit the typical decently behaving dynamics. The situation changes dramatically if μ attains negative values. This case is referred to as the *backward diffusion equation*; the associated initial value problem is known to be ill-posed.⁸

Whenever the parameters a and ω are set such that μ becomes negative the twoscale expansion signals via Equation (24) that there might arise some trouble with the stability of the algorithm. So we are led to the assumption that the domain of stability is given by the subset $S_\mu := \{(a, \omega) \in \mathbb{R}^2 : \mu(a, \omega) = (\frac{1}{\omega} - \frac{1}{2})(1 - a^2) \geq 0\}$ of the $a\omega$ -plane sketched in Figure 6. Although this guess is not completely right, because it contains the actual domain of stability as a proper subset, it must be emphasised that it is only based on simple formal asymptotic computations and some elementary knowledge about the diffusion equation.

Figure 6 The stability domain S_μ as predicted by the twoscale expansion, suggesting that the four closed quadrants – each one sharing its corner with the central square – belong to the stability domain as well. Compare with Figure 4

Still the question remains open, why exactly the analysis of the long-term behaviour should contribute to understand stability properties. If the parameters a, ω are chosen outside of the actual stability domain, the algorithm blows up after a certain number of iterations that is almost

Figure 7 Left: spectrum and spectral limit set of the evolution operator for $\omega = 1.5$, $a = 1.2$ and 64 grid nodes. Right: mass moments of two eigenvectors plotted over the grid. The sinusoidal curve is associated with the highlighted eigenvalue near 1, while the other refers to the eigenvalue on the positive imaginary axis. If the fourth eigenvalue from 1 is considered, the curve interpolating the corresponding mass moment does not change its shape under grid refinement. In contrast, the other mass moment referring to the eigenvalue on the imaginary axis would display more oscillations

Figure 8 Another case with $\omega = 2.1$ and $a = 1.5$ in analogy to Figure 7. The sinusoidal curve (above) represents the mass moment of the eigenvector pertaining to the highlighted eigenvalue on the interior branch λ_+ . The curve below refers to the highlighted eigenvalue on the exterior branch (dotted) corresponding to λ_- . Observe that the imaginary part of the mass moments vanishes identically

independent of the grid spacing h . This insinuates rather that the instable behaviour strikes very fast, more precisely in the discrete time scale. In the light of this observation the chosen twoscale ansatz with its additional slow time appears especially inappropriate, even contradictory.

Throwing a glance at the spectra in Figures 7 and 8, reveals that the occurrence of instabilities must be attributed to the eigenvalues of maximal modulus located on the imaginary axis. Their associated eigenvectors are slightly excited because of rounding errors if not due to the initialisation itself. By application of the evolution matrix they undergo the strongest amplification such that they quickly become disturbing and lead finally to the ‘explosion’.

Since the spectral limit set depends only on a and ω , the position of the ‘bad’ eigenvalues, most responsible for instability, does practically not change under grid refinement. Furthermore, as these eigenvalues are located half-way between the source and sink points, the corresponding eigenvectors are strongly oscillatory with half of the maximal spatial frequency (see Fig. 7 right). As the frequency grows with the number of

grid nodes, the algorithm uncovers its discrete nature here, that eludes a description by smooth⁹ functions. However, those are used to construct the asymptotic order functions $f^{(0)}(t_1, t_2, x)$ etc. Restricted to the grid, they are represented essentially as linear combinations of the ‘smooth-looking’ eigenvectors (longwave modes, cf. Fig. 7 right above) pertaining to the eigenvalues near 1. It seems to be a trivial consequence that it is not possible to grasp the full behaviour of the evolution matrix by just analysing its action on a small subspace.

Nevertheless it is possible to gain certain information by this ansatz, if, figuratively speaking, the eigenvalues near 1 are drawn out of the closed unit disk $\overline{D_1(0)}$ by those which are farther outside already. This situation occurs if μ is negative (see Fig. 7). The eigenvectors pertaining to those eigenvalues are only faintly amplified by the evolution matrix so that their blow-up happens after comparatively many iterations. Moreover, the eigenvalues whose eigenvectors do not exceed a fixed maximal spatial frequency converge to 1 as h tends to 0. Therefore the amplification rate of the associated eigenvectors decreases under grid refinement and this entails that

their blow-up takes place after more and more iterations. The quadratic time variable $t_2 = nh^2$ in Equation (22) accounts for this phenomenon. What we actually sense by the twoscale expansion is not the early blow-up of the ‘rough’ eigenvectors but the blow-up of the ‘smooth’ ones occurring much later.

Figure 8 depicts a situation where the asymptotic approach is outwitted, which happens if $|a| > 1$ and $\omega \notin [0, 2]$ such that $\mu > 0$. Now the spectral branch referring to λ_+ and containing 1 lies within $\overline{D_1(0)}$. Instead the other branch has completely left $\overline{D_1(0)}$. Despite of μ being positive the algorithm cannot be stable. Why does the prediction based on the twoscale expansion miss this instability? The reason is simple: even those eigenvectors are very oscillating that belong to the eigenvalues being scarcely outside $\overline{D_1(0)}$ and close to 1 (confer Fig. 8 right below). Using only smooth ansatz functions, the expansion (22) is blind for them. If rounding errors were avoided and if the initialisation involved only eigenvectors associated to the λ_+ -branch, the algorithm would run stably, however.

Accordance of multiscale expansion and expansion of eigenvalues near 1

In order to elucidate the relation between the multiscale expansion and the spectrum of the evolution matrix we return to Equation (17) which gives an analytic expression for the eigenvalues setting $\phi \in \{\frac{2\pi m}{N} : m \in 0, \dots, N-1\}$. Substituting $1/N$ by h we recognise that the eigenvalues $\lambda_+(2\pi mh)$ converge to 1 for h tending to 0 and fixed m . Performing a Taylor expansion around $h = 0$

$$\begin{aligned} \lambda_+(2\pi mh) &= 1 - 2\pi i m a h - 2\pi^2 m^2 a^2 h^2 \\ &\quad - 4\pi^2 m^2 \left(\frac{1}{\omega} - \frac{1}{2} \right) (1 - a^2) h^2 + O(h^3) \end{aligned} \quad (26)$$

gives an approximation of certain¹⁰ eigenvalues near 1. The coefficients of first and second order turn out to be the negative eigenvalues¹¹ of the advection operator $a\partial_x$ and the diffusion operator $-\mu\partial_x^2 = -(\frac{1}{\omega} - \frac{1}{2})(1 - a^2)\partial_x^2$ (on $[0, 1]$ with periodic boundary conditions) appearing in Equations (23) and (24). The additional term in the second order can be made clear by looking at the eigenvalues of the evolution operators related to the advection and diffusion equation and associated to a time interval Δt .

$$\left. \begin{aligned} \text{adv. } e^{-2\pi i m a \Delta t} &= 1 - 2\pi i m a \Delta t \\ &\quad - 2\pi^2 m^2 a^2 \Delta t^2 + O(\Delta t^3) \\ \text{diff. } e^{-4\pi^2 m^2 \mu \Delta t} &= 1 - 4\pi^2 m^2 \mu \Delta t + O(\Delta t^2) \end{aligned} \right\} \quad (27)$$

Apparently, Equation (26) is obtained by setting $\Delta t = h$ in Equation (27)₁ and $\Delta t = h^2$ in Equation (27)₂, multiplying both truncated series and keeping only terms of order less than $O(h^3)$. So the expansion of the eigenvalues Equation (26) corroborates the results of the multiscale expansion:

- The effect of the lattice-Boltzmann evolution matrix (on a certain subspace) corresponds to an intertwining

of an advection and diffusion process where the first one is dominant.

- There is an inherent separation of time scales: the advection takes place in the linear (fast) time scale while the diffusion evolves in the quadratic (slow) time scale.

7 Conclusion and outlook

Let us briefly pass in review the main points of the sections, which might also be useful in other situations beyond the context of the analysed D1Q2 scheme:

- Formal asymptotic expansions like the regular and multiscale expansions in Junk and Rheinländer (2008) are rigorously justified by stability results.
- The eigenvalues of the lattice-Boltzmann evolution operator have been computed. They are shown to lie within the closed unit disk under a CFL-condition. To obtain norm stability, additional considerations are required. This subtle point has been emphasised.
- The CFL-condition is found as stability condition for the advective D1Q2 lattice-Boltzmann scheme. This parallels the situation for various finite difference methods (upwind, Lax-Friedrichs, Lax-Wendroff). It should be remarked, that the stability condition is a special property of the D1Q2 scheme and does not hold for analogous D1Q3 schemes¹² revealing a smaller domain of stability.
- Formal asymptotic methods can give indications about stability but they must be distrusted generally. The reason for this is the appearance of unstable eigenvectors which cannot be described by smooth functions. Therefore they escape the chosen ansatz.

Furthermore, instabilities can only be discovered if the eigenvalues near 1 “belonging to the same branch” are slightly outside the closed unit disk. Multiscale expansions can be helpful to detect those eigenvalues whose distance to the closed unit disk is of smaller magnitude than the discretisation parameter.

In a forthcoming publication the considered lattice-Boltzmann D1Q2 algorithm provides a basis to investigate the Chapman-Enskog expansion. Two aspects are of interest:

- the relation to other expansion approaches like the regular expansion and the multiscale expansion as presented in Rheinländer (2007)
- stability properties of the higher order equations (Burnett and super Burnett level) obtained from the Chapman-Enskog expansion.

Furthermore we remark that Theorem 1 cannot be extended to analogue D1Q3 schemes which differ from the presented D1Q2 scheme essentially by a rest population.

It turns out that in this case another condition is necessary being more restrictive than the CFL-condition (Rheinländer, 2007). For two-dimensional algorithms a similar behaviour is expected which will be investigated in a future project.

Acknowledgement

The author thanks Professor M. Junk from the Universität Konstanz for his motivating and stimulating interaction and his continual interest in this work.

References

- Caiazzo, A. (2005) ‘Analysis of lattice-Boltzmann initialization routines’, *J. Stat. Phys.*, Vol. 121, pp.37–48.
- Deuflhard, P. and Bornemann, F. (2002) *Scientific Computing with Ordinary Differential Equations*, Springer, New York.
- Dubois, F. (2007) ‘Equivalent partial differential equations of a lattice Boltzmann scheme’, *Computers and Mathematics with Applications*, doi: 10.1016/j.camwa.2007.08.003.
- Frisch, U., d’Humières, Hasslacher, B., Lallemand, P., Pomeau, Y. and Rivet, J.P. (1987) ‘Lattice gas hydrodynamics in two and three dimensions’, *Complex Systems*, Vol. 1, pp.649–707.
- Junk, M. (2001) ‘A finite difference interpretation of the lattice-Boltzmann method’, *Numer. Methods Partial Differ. Equations*, Vol. 17, pp.383–402.
- Junk, M., Klar, A. and Luo, L.S. (2005) ‘Asymptotic analysis of the lattice Boltzmann equation’, *Journal Comp. Phys.*, Vol. 210, pp.676–704.
- Junk, M. and Rheinländer, M. (2008) ‘Regular and multiscale expansions of a lattice Boltzmann method’, *Prog. Comp. Fluid Dyn.*, Vol. x, No. x, pp.xxx–xxx. (this issue)
- Junk, M. and Yang, Z. (2004) ‘Asymptotic analysis of finite difference methods’, *Appl. Math. Comput.*, Vol. 158, pp.267–301.
- Junk, M. and Yang, Z. (2005a) ‘Asymptotic analysis of lattice Boltzmann boundary conditions’, *J. Stat. Phys.*, Vol. 121, pp.3–35.
- Junk, M. and Yang, Z. (2005b) *Computer and Mathematics with Applications*, doi: 10.1016/j.camwa.2007.08.002.
- Junk, M. and Yang, Z. (2005c) ‘One-point boundary condition for the lattice Boltzmann model’, *Phy. Rev. E*, Vol. 72, p.066701.
- Junk, M. and Yong, W.A. (2006) *Weighted l2 Stability of the Lattice-Boltzmann Method*, submitted to SINUM.
- Lallemand, P. and Luo, L.S. (2000) ‘Theory of the lattice-Boltzmann method: dispersion, dissipation, isotropy, Galilean invariance, and stability’, *Physical Review E*, Vol. 61, No. 6, pp.6546–6562.
- Lang, S. (1993) *Complex Analysis*, 3rd ed., Springer, New York.
- LeVeque, R. (1992) *Numerical Methods for Conservation Laws*, Birkhäuser, Basel.
- Rheinländer, M. (2005) ‘A consistent grid coupling method for lattice-Boltzmann schemes’, *J. Stat. Phys.*, Vol. 121, pp.49–74.
- Rheinländer, M. (2007) *Analysis of Lattice Boltzmann Methods*, PhD Thesis, Universität Konstanz, Konstanz, <http://www.ub.uni-konstanz.de/kops/volltexte/2007/3635/>
- Sohrab, H.H. (2003) *Basic Real Analysis*, Birkhäuser, Boston.
- Sterling, J.D. and Chen, S. (1996) ‘Stability analysis of the lattice-Boltzmann method’, *Journal of Computational Physics*, Vol. 123, pp.196–206.
- Wolf-Gladrow, D.A. (2000) *Lattice-Gas Cellular Automata and Lattice Boltzmann Models*, Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, New York.

Notes

- ¹Observe that the system is not linear in the unknown variables a and λ because of the right hand side of Equation (11).
- ² $\mathbf{F}$ is a specific Vandermonde matrix known as the *Fourier* matrix, which is used to define the discrete Fourier transformation in $\mathbb{C}^N$.
- ³It is imaginable that $\text{spec}(\mathbf{E}) \subset \overline{D_1(0)}$ although $\mathcal{S}(\omega, a) \not\subset \overline{D_1(0)}$ if $\mathbf{E}$ belongs to a relatively coarse grid such that its eigenvalues sample $\mathcal{S}(\omega, a)$ coarsely too.
- ⁴In order to obtain connected curves as natural branches (see Figs. 2 and 3) it might be necessary to compute the complex root not in the usual way by slitting the complex plane with a ray in positive or negative direction of the real axis but using a tilted ray.
- ⁵Geometrically the zero $\cos(\phi) - ia \sin(\phi)$ describes an ellipse for $\phi \in [0, 2\pi)$ parameterised in (counter) clockwise sense if a is positive(negative). Its major semiaxis is 1 along the real axis while its minor semiaxis in imaginary direction has length $|a|$. Evidently, the ellipse is totally contained in the closed unit disk if and only if $|a| \leq 1$.
- ⁶If $\mathbf{E}$ were not diagonalisable, i.e., if the Jordan blocks were not simple, then $\|\mathbf{E}^k\|$ increased polynomially.
- ⁷It is true that in finite dimensional vector spaces ($\mathbb{R}^n$) all norms are equivalent. But the coefficients occurring in mutual estimates depend on the dimension of the space. Therefore stability with respect to one norm need not be equivalent with stability in another norm.
- ⁸Formally the associated evolution should be inverse to that of the diffusion equation. Therefore its effect would be roughing instead of smoothing. However, the example of an Gaussian bell function, evolving in finite time into a δ -distribution-like peak, gives clear evidence that the associated IVP generally permits no solution in the standard function spaces. Thus the twoscale expansion breaks down for $\mu < 0$, which implies that it is no more adequate to approximate the population function $\mathbf{F}$.
- ⁹One may object why, actually, it is reasonable to analyse discrete algorithms by means of smooth functions, if they can not fully mimic discrete behavior. There are two reasons: First, in many cases one is mostly interested in the properties of discrete algorithms that are not typically discrete. Second, one has generally a better feeling for smooth functions than for irregular, discrete functions.
- ¹⁰Bear in mind that there might be other eigenvalues close to 1 (e.g., on the branch associated with λ_-) which are not approximated by the above expansion.
- ¹¹It is quickly checked that $x \mapsto e^{2\pi m i x}$ with $m \in \mathbb{Z}$ represents a basis of eigenfunctions respecting periodic boundary conditions.
- ¹²The D1Q2 scheme can be considered as a special case of the D1Q3 scheme with a trivial, decoupled rest population.

Appendix

For the sake of completeness we give a statement of the two theorems that we have essentially used in Sections 4 and 5. Proofs can be found in almost every relevant textbook. Moreover three lemmas are provided used in the proof of Theorem 2.

Theorem 3 (Rouché's Theorem): *Let γ be a closed path null-homologous in $G \subset \mathbb{C}$ and assume that γ has an interior. Let f, g be holomorphic (analytic) on G , and*

$$|f(\zeta) - g(\zeta)| < |f(\zeta)| \quad \text{for all } \zeta \text{ on } \gamma.$$

Then f and g have the same number of zeros in the interior of γ .

A proof is presented by Lang (1993, p. 181).

Theorem 4 (Osgood's Theorem—principle of uniform boundedness): *Let (M, d) be a complete metric space and $\mathcal{F} \subset \mathcal{C}(M, \mathbb{R})$. Suppose that, for each $x \in M$, there exists a constant $C_x > 0$ such that $\forall f \in \mathcal{F} : |f(x)| \leq C_x$. Then there exists a nonempty open ball $B \subset M$ and a constant $C > 0$ such that $\forall f \in \mathcal{F}, \forall x \in B : |f(x)| \leq C$.*

For the proof we refer to Sohrab (2003, p. 184).

Lemma 7: *Let $M = \text{diag}(B_1, \dots, B_N)$ be a block-diagonal matrix. Then $\|M\|_2 = \max_k \|B_k\|_2$.*

Proof: The assertion is due to the fact that the $\|\cdot\|_2$ -norm is connected to the eigenvalues of MM^* , which can be computed block-wise:

$$\begin{aligned} \|M\|_2^2 &= \rho(MM^*) \\ &= \max \{ |\lambda| : \det(MM^* - \lambda) = 0 \} \\ &= \max \left\{ |\lambda| : \prod_k \det(B_k B_k^* - \lambda) = 0 \right\} \\ &= \max_k \max \{ |\lambda| : \det(B_k B_k^* - \lambda) = 0 \} \\ &= \max_k \rho(B_k B_k^*) = \max_k \|B_k\|_2^2. \end{aligned} \quad \square$$

Observe that other induced matrix norms (e.g., $\|\cdot\|_1$ or $\|\cdot\|_\infty$) behave analogically.

Lemma 8: *Let M be a diagonalisable matrix and assume $\rho(M) \leq 1$. Then there is for any matrix norm $\|\cdot\|$ a constant $C > 0$ such that: $\forall n \in \mathbb{N} : \|M^n\| \leq C$.*

A proof of a slightly more general assertion is contained in Deuffhard and Bornemann (2002, Theorem 3.33, p. 108).

Lemma 9: *The 2×2 matrix $B(\phi)$ defined in Equation (19) is diagonalisable for $a \neq 0$.*

Proof: It is enough to show that the characteristic polynomial, i.e., $\chi(\lambda, \phi)$ has two different zeros since this implies immediately the existence of two different eigenvectors forming an eigenbasis.

Case 1: $\phi \notin \{0, \pi\}$ and $\omega \neq 0$.

Under this condition the coefficient of the linear term is complex-valued while the constant term $1 - \omega$ is real. From Vieta's rule we can conclude that at least one zero is properly complex, say $z_1 = r_1 e^{i\theta}$ with $r_1 > 0$ and $\theta \in (-\pi, 0) \cup (0, \pi)$. Furthermore we get $z_2 = r_2 e^{-i\theta}$ with $r_2 > 0$ for $\omega \neq 1$ or $r_2 = 0$ for $\omega = 1$ such that the product is real, more exactly $z_1 z_2 = 1 - \omega$. So we may conclude $z_1 \neq z_2$.

Case 2: $\phi \in \{0, \pi\}$ or $\omega = 0$.

In these cases the characteristic polynomial and thus the eigenvalues do not depend on a . An explicit computation yields two different eigenvalues 1 and $\omega + 1$ for $\omega \neq 0$ and $\phi = 0$. Similarly, -1 and $\omega - 1$ are eigenvalues if $\phi = \pi$ and $\omega \neq 0$. If only the second condition is satisfied, then the eigenvalues are $e^{i\phi}$ and $e^{-i\phi}$. Finally, $B(\phi)$ is automatically diagonal (positive/negative identity) if both conditions are fulfilled. $\square$

For $a = 0$ the discriminant¹³ of the characteristic polynomial is strictly positive for $\omega > 1$ and implies the existence of two eigenvalues and eigenvectors. However, if $\omega \leq 1$, then ω and ϕ can be chosen such that the discriminant may vanish. In this case it is not clear a priori whether $B(\phi)$ has two linearly independent eigenvectors.

The lemma implies immediately that the transformed evolution matrix $\tilde{\mathbf{E}}$ defined in Equation (20) is diagonalisable as well. The similarity between $\tilde{\mathbf{E}}$ and $\mathbf{E}$ entails the subsequent corollary.

Corollary 1: *The evolution matrix $\mathbf{E}$ defined in Equation (6) is diagonalisable if $a \neq 0$.*