

Thomas Bridges · Jonathan Pennant · Sergey Zelik

Degenerate hyperbolic conservation laws with dissipation: reduction to and validity of a class of Burgers-type equations

May 30, 2014

Abstract A conservation law is said to be degenerate or critical if the Jacobian of the flux vector evaluated on a constant state has a zero eigenvalue. In this paper, it is proved that a degenerate conservation law with dissipation will generate dynamics on a long time scale that resembles Burger's dynamics. The case of k -fold degeneracy is also treated, and it is shown that it leads to a reduction to a quadratically coupled k -fold system of Burgers-type equations. Validity of the reduction and existence for the reduced system is proved in the class of uniformly local spaces, thereby capturing both finite and infinite energy solutions. The theory is applied to some examples, from stratified shallow-water hydrodynamics, that model the birth of hydraulic jumps.

1 Introduction

Consider a conservation law with dissipative regularization in one unbounded space dimension,

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{D}\mathbf{U}_{xx}, \quad x \in \mathbb{R}, \quad \mathbf{U} \in \mathbb{R}^n, \quad (1.1)$$

where \mathbf{U} is the state variable, $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the flux vector which is assumed to be smooth and \mathbf{D} is a constant $n \times n$ matrix satisfying

$$\mathbf{D} + \mathbf{D}^T > 0. \quad (1.2)$$

Let $\mathbf{U}_0 \in \mathbb{R}^n$ be any constant vector. Then clearly \mathbf{U}_0 is a solution of (1.1). These solutions are sometimes called *uniform states* in applications, and they form an n -parameter family of solutions. A particular constant solution $\mathbf{U}_0 := \mathbf{p} \in \mathbb{R}^n$ is said to be *critical* if $\mathbf{D}\mathbf{F}(\mathbf{p})$, the Jacobian of the flux vector, is singular

$$\det[\mathbf{D}\mathbf{F}(\mathbf{p})] = 0. \quad (1.3)$$

The flux vector is assumed to be hyperbolic at criticality: $\mathbf{D}\mathbf{F}(\mathbf{p})$ has only real eigenvalues and is diagonalizable. The set of all $\mathbf{p} \in \mathbb{R}^n$ satisfying (1.3) defines, generically, a hypersurface in \mathbb{R}^n . This hypersurface is called the *criticality surface*. In compressible fluid flow criticality corresponds to the case where the Mach number of a constant flow is unity, and in shallow water hydrodynamics, criticality corresponds to the case where the Froude number of the uniform flow is unity.

The purpose of this paper is to investigate the nonlinear problem when the system is *near critical*. It is found that the dynamics splits into a part governed by Burgers equation, or a system of Burgers-type equations, and a part which is enslaved to the Burgers dynamics.

Burgers equation has been used as a simplified model for compressible flow (e.g. Chapter 4 of WHITAM [29]), it has been widely used as model equation in acoustics (e.g. CRIGHTON [10]), and it is a model for metastability (e.g. BECK & WAYNE [4]). Two coupled Burgers equations arise as a model for polydisperse sedimentation (e.g. ESIPOV [14]).

Here, we are interested in showing that Burgers equation or coupled Burgers equation emerges as a universal model from a larger set of PDEs. Burger's equation has been shown to be a universal model equation in other contexts: in the complex Ginzburg-Landau equation (BERNOFF [6]), for channel flow

governed by the Navier-Stokes equations (HEWITT & HALL [18]) and for phase dynamics in reaction-diffusion equations (DOELMAN ET AL. [11]). In all three cases, a single Burgers equation is generated by modulation of spatially periodic waves. In the case of DOELMAN ET AL. [11] a rigorous validity proof of the reduction to Burgers equation is given.

The emergence of Burger's equation in this paper is due to a different mechanism. The Burgers dynamics is generated by the geometry of the flux vector, evaluated on spatially uniform states, and the basic state is not modulated.

The dimension of the reduced system of Burgers-type equations is dependent on the rank deficiency of $D\mathbf{F}(\mathbf{p})$. We will treat in detail the two most interesting cases, where $D\mathbf{F}(\mathbf{p})$ has a simple zero eigenvalue, in which case a single Burgers equation governs the dynamics, and secondly when $D\mathbf{F}(\mathbf{p})$ has a double semisimple zero eigenvalue in which case a coupled pair of Burgers equations governs the dynamics.

The basic idea of the formal reduction to Burgers equation can be seen by considering the simplified case when $\mathbf{D} = \nu\mathbf{I}$, for some positive constant ν , and $D\mathbf{F}(\mathbf{p})$ has a simple zero eigenvalue with right and left eigenvectors

$$D\mathbf{F}(\mathbf{p})\boldsymbol{\xi} = 0 \quad \text{and} \quad \boldsymbol{\eta}^T D\mathbf{F}(\mathbf{p}) = 0, \quad \langle \boldsymbol{\eta}, \boldsymbol{\xi} \rangle = 1. \quad (1.4)$$

Here and throughout $\langle \cdot, \cdot \rangle$ is the standard inner product on \mathbb{R}^n . Let ε be a measure of the distance between a constant state \mathbf{U}_0 and a point \mathbf{p} on the criticality surface. Express $\mathbf{U}(x, t)$ as a perturbation from the point \mathbf{p}

$$\mathbf{U}(x, t, \varepsilon) = \mathbf{p} + \varepsilon u(x, t, \varepsilon)\boldsymbol{\xi} + \varepsilon^2 \mathbf{V}(x, t, \varepsilon) \quad \text{with} \quad \langle \boldsymbol{\eta}, \mathbf{V} \rangle = 0.$$

Substitution into (1.1) splits it into two equations

$$\begin{aligned} u_t + \frac{1}{\varepsilon} \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + \varepsilon u\boldsymbol{\xi} + \varepsilon^2 \mathbf{V}) \rangle_x &= \nu u_{xx} \\ \mathbf{V}_t + \frac{1}{\varepsilon^2} (\mathbf{P}\mathbf{F}(\mathbf{p} + \varepsilon u\boldsymbol{\xi} + \varepsilon^2 \mathbf{V}))_x &= \nu \mathbf{V}_{xx}, \end{aligned} \quad (1.5)$$

where \mathbf{P} is a projection onto the complement of the kernel of $D\mathbf{F}(\mathbf{p})$. The second term in the first equation simplifies to

$$\frac{1}{\varepsilon} \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + \varepsilon u\boldsymbol{\xi} + \varepsilon^2 \mathbf{V}) \rangle_x = \varepsilon \langle \boldsymbol{\eta}, D^2\mathbf{F}(\mathbf{p})(\boldsymbol{\xi}, \boldsymbol{\xi}) \rangle u u_x + \mathcal{O}(\varepsilon^2),$$

using (1.4), reducing the first equation to a Burgers equation to leading order

$$u_t + \varepsilon \kappa u u_x = \nu u_{xx} + \dots, \quad (1.6)$$

with

$$\kappa = \left. \frac{d^2}{ds^2} \right|_{s=0} \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + s\boldsymbol{\xi}) \rangle. \quad (1.7)$$

By introducing the scaling

$$X = \varepsilon x \quad \text{and} \quad T = \varepsilon^2 t \quad (1.8)$$

into (1.6) it becomes ε -independent to leading order

$$u_T + \kappa u u_X = \nu u_{XX}, \quad (1.9)$$

suggesting that it is natural in both the formal and rigorous reduction to work with scaled time and space variables.

The parameter κ is associated with the nonlinearity. In order to have a non-trivial nonlinearity, it will be assumed that $\kappa \neq 0$, which is a condition on the flux vector. The expression (1.7) is in fact the *intrinsic second derivative* of the flux vector. It is intrinsic since it is evaluated on the kernel of the first derivative. This terminology comes from singularity theory where it was introduced by PORTEOUS [25] and is now widely used (e.g. ARNOLD ET AL. [1], GOLUBITSKY & GUILLEMIN [17]). In §8 a geometric interpretation of the condition $\kappa \neq 0$ is given.

If the second equation in (1.5) is well defined in the limit $\varepsilon \rightarrow 0$, and can be solved for \mathbf{V} as a function of u – a strategy similar to a center-manifold reduction – then one might expect the dynamics to be governed to leading order by the Burgers equation (1.9).

The above formal argument can be extended to the case where $D\mathbf{F}(\mathbf{p})$ has an k -fold ($k < n$) semisimple zero eigenvalue, leading to a k -fold coupled system of Burgers equations,

$$\mathbf{u}_T + \mathbf{Q}(\mathbf{u})_X = \widehat{\mathbf{D}}\mathbf{u}_{XX}, \quad \mathbf{u} \in \mathbb{R}^k, \quad (1.10)$$

in scaled variables, which we call Q -Burgers. The reduced dissipation matrix has entries

$$\nu_{i,j} := \widehat{\mathbf{D}}_{i,j} = \langle \boldsymbol{\eta}_i, \mathbf{D}\boldsymbol{\xi}_j \rangle, \quad (1.11)$$

where $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_k\}$ and $\{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_k\}$ are the right and left eigenvectors respectively of $\mathbf{D}\mathbf{F}(\mathbf{p})$. The reduced flux vector $\mathbf{Q}(\mathbf{u})$ is a homogeneous quadratic function. For example, in the case $k = 2$, with $\mathbf{u} = (u_1, u_2)$,

$$\mathbf{Q}(\mathbf{u}) = \begin{pmatrix} \frac{1}{2}\Gamma_{11}^1 u_1^2 + \Gamma_{12}^1 u_1 u_2 + \frac{1}{2}\Gamma_{22}^1 u_2^2 \\ \frac{1}{2}\Gamma_{11}^2 u_1^2 + \Gamma_{12}^2 u_1 u_2 + \frac{1}{2}\Gamma_{22}^2 u_2^2 \end{pmatrix},$$

where the Γ_{ij}^l are components of the intrinsic second derivative of the flux vector generalizing (1.7). The geometry of the flux vector is presented in §8. The formal argument reducing (1.1) to Q -Burgers when \mathbf{p} lies in a criticality surface is given in §2.

The main aim of this paper is to prove that the above formal argument makes sense, and to identify appropriate function spaces in which it makes sense. The choice of function space is dictated by the requirement that it is defined for functions on the real line and it captures both finite energy and infinite energy solutions. An example of the latter is fronts which are asymptotic to distinct nonzero constants as $x \rightarrow \pm\infty$. The classical Sobolev spaces based on $L^2(\mathbb{R})$ are inadequate and can be naturally replaced by the so-called *uniformly local spaces*. These spaces contain constant and non-decaying functions, as well as finite energy solutions and are widely used in the analysis of PDEs in unbounded domains. These spaces have been used in the validity argument for reduction of modulated wavetrains to Burger's equation (cf. Chapter 5 of DOELMAN ET AL. [11]), and have been widely used in the study of attractors of parabolic PDEs (e.g. [23] and references therein).

Uniformly local spaces are defined as follows. Let $B_x^1 = \{y \in \mathbb{R} : |y - x| \leq 1\}$ and let $\mathcal{D}'(\mathbb{R})$ be the space of distributions on \mathbb{R} . Then

$$L_b^2(\mathbb{R}) = \left\{ \mathbf{u} \in \mathcal{D}'(\mathbb{R}) : \sup_{x_0 \in \mathbb{R}} \left(\int_{B_{x_0}^1} |\mathbf{u}(x)|^2 dx \right) < \infty \right\}, \quad (1.12)$$

with norm

$$\|\mathbf{u}\|_{L_b^2} = \left(\sup_{x_0} \int_{B_{x_0}^1} |\mathbf{u}|^2 dx \right)^{1/2} = \sup_{x_0 \in \mathbb{R}} \|\mathbf{u}\|_{L^2([x_0-1, x_0+1])} \quad (1.13)$$

and the uniformly local Sobolev spaces $W_b^{l,2}(\mathbb{R})$ are naturally defined as the spaces of distributions whose derivatives up to order l belong to $L_b^2(\mathbb{R})$. The essential properties of these spaces have been developed by MIELKE & SCHNEIDER [22], ZELIK [30] and EFENDIEV & ZELIK [12,13]. A comprehensive review of the properties of these spaces is given in the encyclopedia article of MIRANVILLE & ZELIK [23].

The main result of the paper is the long-time validity of the Q -Burgers reduction (1.10) for system (1.1) in the uniformly local spaces. Namely, it will be shown (see Theorem 6.1 and Corollary 6.1) that, for every sufficiently smooth solution $\mathbf{u} = (u_1, \dots, u_k)$ of the Q -Burgers system (1.10) on the time interval $T \in [0, \mathcal{T}]$, there exists $\varepsilon_0 = \varepsilon(\mathbf{u}, \mathcal{T}) > 0$ such that, for every $\varepsilon < \varepsilon_0$, there exists a solution $\mathbf{U}_\varepsilon(t, x)$ of the initial problem (1.1) such that

$$\left\| \mathbf{U}_\varepsilon(t, x) - \mathbf{p} - \varepsilon \sum_{j=1}^k u_j(\varepsilon^2 t, \varepsilon x) \boldsymbol{\xi}_j \right\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{3/2}, \quad t \in [0, \mathcal{T}/\varepsilon^2], \quad (1.14)$$

where C is independent of ε . Solutions of Q -Burgers stay close to solutions of the original problem for a time interval tending to infinity as $\varepsilon \rightarrow 0$. These *validity* results are in contrast to *asymptotic* results which work precisely with a neighborhood of $t = \infty$ (e.g. SZEPESSY & ZUMBRUN [27] and references therein). In [27] the large-time asymptotic behavior of a weak rarefaction wave, for a system of the form (1.1), is analyzed and it is proved that the perturbed rarefaction data converge in the limit $t \rightarrow \infty$ to an approximate ‘‘Burgers’’ rarefaction wave.

Another approach which leads to a Burgers-type reduction is to use center-manifold theory on the steady problem, and then an Evans function analysis on the linear stability problem (e.g. FREISTÜHLER & SZMOLYAN [16] and PLAZA & ZUMBRUN [16]). In this approach, one studies the steady problem (1.1) relative to a moving frame, which reduces to an ODE,

$$\mathbf{D}\mathbf{U}_x = \mathbf{F}(\mathbf{U}) - c\mathbf{U}, \quad \mathbf{U} \in \mathbb{R}^n,$$

(modulo a vector-valued constant of integration). When $[\mathbf{DF}(\mathbf{U}) - c\mathbf{I}]$ has rank $n - k$, then a center-manifold reduction can be used to reduce it to an ODE, which is steady Burgers (or steady multi-fold Burgers). Denote the front solution by $\widehat{\mathbf{U}}$. The linearization of (1.1) about this solution with the spectral ansatz, $\partial_t \mapsto \lambda$, is

$$\lambda \mathbf{U} - c \mathbf{U}_x - (\mathbf{DF}(\widehat{\mathbf{U}})\mathbf{U})_x = \mathbf{D}\mathbf{U}_{xx}.$$

Since $\widehat{\mathbf{U}}$ tends to a constant as $x \rightarrow \pm\infty$, this system can be reformulated as a λ -dependent system of ODEs, of dimension $2n$, in standard form for an Evans function analysis. The advantage of this approach for Burgers shocks, or multi-fold Burgers shocks, is that rigorous results on linear stability can be obtained. On the other hand, the reduction to time-dependent Burgers, or Q -Burgers, invites analysis of *all* solutions of (1.10), with the proviso that their relevance for (1.1) can only be guaranteed for a time interval of order ε^{-2} .

There are three steps in the proof of validity in this paper. The first step is to prove that the reduced Q -Burgers (1.10) is well posed in L_b^2 for some time interval $[0, \mathcal{T}]$. In the case $k = 1$, it follows from a maximum principle argument that $\mathcal{T} = \infty$. When $k > 1$ it is not in general true that solutions of Q -Burgers exist globally in time. However, we are able to prove global existence of Q -Burgers for any k when $\mathbf{Q}(\mathbf{u})$ is a gradient function. In the case $k = 2$ we can appeal to a result of SCHAEFFER & SHEARER [26] that assures that $\mathbf{Q}(\mathbf{u})$ is a gradient function if $\mathbf{D}\mathbf{Q}$ is hyperbolic. The proof of regularity, uniqueness and global existence for Q -Burgers is presented in §4.

The second stage of the proof is to derive estimates in L_b^2 for the class of linear parabolic PDEs

$$\mathbf{u}_t + \frac{1}{\varepsilon} \mathbf{A}\mathbf{u}_x - \mathbf{D}\mathbf{u}_{xx} = \mathbf{h}(t), \quad (1.15)$$

for given \mathbf{h} , where \mathbf{A} is a constant symmetrizable matrix. Systems of this type arise when analyzing the second equation of (1.5). The regularity properties of the solutions of (1.15) are presented in §5.

The third stage is the validity proof. Take initial data for (1.1) and take the projection of this initial data onto the kernel of $\mathbf{DF}(\mathbf{p})$ as initial data for Q -Burgers. Look at the distance between these two solutions over time, measured in a uniformly local norm. In §6 it is proved that the distance between these two solutions is of order $\varepsilon^{3/2}$ for a time interval which tends to infinity as $\varepsilon \rightarrow 0$, confirming the estimate (1.14).

In §9 and §10 the theory is applied to examples from shallow water hydrodynamics, and it is shown there that the reduction to Burgers is associated with the generation of hydraulic jumps. We also show how to modify the theory if dissipation does not act on all components; e.g. when one of the equations in (1.1) is a pure conservation law.

2 Formal reduction to Burgers equations

Let ε be a small parameter that measures the distance from the criticality surface. Suppose that $\mathbf{DF}(\mathbf{p})$ has a k -fold semisimple zero eigenvalue with

$$\text{Ker}(\mathbf{DF}(\mathbf{p})) = \text{span}\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_k\} \quad \text{and} \quad \text{Ker}(\mathbf{DF}(\mathbf{p})^T) = \text{span}\{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_k\},$$

with the standard normalization $\langle \boldsymbol{\eta}_i, \boldsymbol{\xi}_j \rangle = \delta_{i,j}$.

Introduce the Burgers scaling (1.8) into the full equation (1.1),

$$\mathbf{U}_T + \frac{1}{\varepsilon} \mathbf{F}(\mathbf{U})_X = \mathbf{D}\mathbf{U}_{XX}, \quad (2.1)$$

and express $\mathbf{U}(x, t)$ in (1.1) in terms of functions of the scaled variables,

$$\mathbf{U}(x, t, \varepsilon) = \mathbf{p} + \varepsilon \sum_{j=1}^k u_j(X, T, \varepsilon) \boldsymbol{\xi}_j + \varepsilon^2 \mathbf{V}(X, T, \varepsilon), \quad (2.2)$$

with

$$\langle \boldsymbol{\eta}_j, \mathbf{V}(X, T, \varepsilon) \rangle = 0, \quad \text{for } j = 1, \dots, k.$$

The space \mathbb{R}^n has been split into the kernel of $\mathbf{DF}(\mathbf{p})$ and its complement. Let \mathbf{P} denote the projection onto the complement of the kernel of $\mathbf{DF}(\mathbf{p})$.

Substitute (2.2) into (1.1)

$$\begin{aligned}\mathbf{U}_t &= \varepsilon^3 \sum_{j=1}^k \frac{\partial u_j}{\partial T} \boldsymbol{\xi}_j + \varepsilon^4 \mathbf{V}_T \\ \mathbf{F}(\mathbf{U})_x &= \varepsilon^3 \mathbf{D}\mathbf{F}(\mathbf{p})\mathbf{V}_X + \frac{1}{2}\varepsilon^3 \mathbf{D}^2\mathbf{F}(\mathbf{p}) \left(\sum_{j=1}^k u_j \boldsymbol{\xi}_j + \varepsilon \mathbf{V}, \sum_{j=1}^k u_j \boldsymbol{\xi}_j + \varepsilon \mathbf{V} \right)_X + \dots \\ \mathbf{U}_{xx} &= \varepsilon^3 \left(\sum_{j=1}^k \frac{\partial^2 u_j}{\partial X^2} \boldsymbol{\xi}_j + \varepsilon \mathbf{V}_{XX} \right).\end{aligned}$$

Split into two parts according to the splitting of \mathbb{R}^n

$$\frac{\partial u_\ell}{\partial T} + \frac{\partial}{\partial X} \left(\frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \Gamma_{ij}^\ell u_i u_j \right) = \sum_{j=1}^k \nu_{\ell j} \frac{\partial^2 u_j}{\partial X^2} + \varepsilon \mathbf{S}_\ell, \quad \ell = 1, \dots, k, \quad (2.3)$$

with $\nu_{i,j}$ defined in (1.11), and

$$\varepsilon V_T + \mathbf{P}\mathbf{D}\mathbf{F}(\mathbf{p})V_X + \frac{1}{2}\mathbf{P}\mathbf{D}^2\mathbf{F}(\mathbf{p}) \left(\sum_{j=1}^k u_j \boldsymbol{\xi}_j, \sum_{j=1}^k u_j \boldsymbol{\xi}_j \right)_X = \mathbf{P}\mathbf{D} \left(\sum_{j=1}^k \boldsymbol{\xi}_j \frac{\partial^2 u_j}{\partial X^2} \right) + \varepsilon \mathbf{P}\mathbf{D}\mathbf{V}_{XX} + \dots,$$

or

$$\frac{d}{dX} \left[\mathbf{P}\mathbf{D}\mathbf{F}(\mathbf{p})\mathbf{V} + \frac{1}{2}\mathbf{P}\mathbf{D}^2\mathbf{F}(\mathbf{p}) \left(\sum_{j=1}^k u_j \boldsymbol{\xi}_j, \sum_{j=1}^k u_j \boldsymbol{\xi}_j \right) - \mathbf{P}\mathbf{D} \left(\sum_{j=1}^k \boldsymbol{\xi}_j \frac{\partial u_j}{\partial X} \right) \right] = \varepsilon \mathbf{T}, \quad (2.4)$$

where \mathbf{S}_ℓ , $\ell = 1, \dots, k$, and \mathbf{T} are remainder terms.

The leading order term in the first set of equations (2.3) is Q -Burgers (1.10). The complementary equation (2.4) appears to be solvable for \mathbf{V} as a function of u_1, \dots, u_k .

Consider a perturbation expansion for u_1, \dots, u_k and \mathbf{V} ,

$$\begin{aligned}u_\ell(X, T, \varepsilon) &= u_\ell^0(X, T) + \varepsilon u_\ell^1(X, T) + \mathcal{O}(\varepsilon^2), \quad \ell = 1, \dots, k, \\ \mathbf{V}(X, T, \varepsilon) &= \mathbf{V}_0(X, T) + \varepsilon \mathbf{V}_1(X, T) + \mathcal{O}(\varepsilon^2),\end{aligned}$$

Substitution shows that u_l^0, \dots, u_k^0 satisfy Q -Burgers and \mathbf{V}_0 can be determined as a function of u_1^0, \dots, u_k^0 ,

$$\mathbf{V}_0(X, T) = [\mathbf{P}\mathbf{D}\mathbf{F}(\mathbf{p})]^{-1} \left(-\frac{1}{2}\mathbf{P}\mathbf{D}^2\mathbf{F}(\mathbf{p}) \left(\sum_{j=1}^k u_j^0 \boldsymbol{\xi}_j, \sum_{j=1}^k u_j^0 \boldsymbol{\xi}_j \right) + \mathbf{P}\mathbf{D} \left(\sum_{j=1}^k \boldsymbol{\xi}_j \frac{\partial u_j^0}{\partial X} \right) \right) + \mathbf{V}_0^{(0)}(T), \quad (2.5)$$

where $\mathbf{V}_0^{(0)}(T)$ is an arbitrary function of T . The invertibility of $\mathbf{P}\mathbf{D}\mathbf{F}(\mathbf{p})$ follows since \mathbf{P} is projection onto the complement of the kernel of $\mathbf{D}\mathbf{F}(\mathbf{p})$. The successive terms in the series for u_ℓ and \mathbf{V} then satisfy linear inhomogeneous PDEs. Formally, each term in the series can be determined.

3 Towards a proof of validity

In this section we attempt to justify the formal argument in §2. Let $\mathbf{U}(X, T, \varepsilon)$ be an exact solution of (1.1) in the scaled form (2.1) and define

$$\mathbf{z}(X, T, \varepsilon) = \frac{1}{\varepsilon} (\mathbf{U}(X, T, \varepsilon) - \mathbf{p}). \quad (3.1)$$

Then $\mathbf{z}(X, T, \varepsilon)$ satisfies

$$\mathbf{z}_T + \frac{1}{\varepsilon^2} \mathbf{F}(\mathbf{p} + \varepsilon \mathbf{z})_X = \mathbf{D}\mathbf{z}_{XX},$$

or, after expanding \mathbf{F} in a Taylor series,

$$\mathbf{z}_T + \varepsilon^{-1} \left(\mathbf{D}\mathbf{F}(\mathbf{p})\mathbf{z} + \frac{1}{2}\varepsilon \mathbf{D}^2\mathbf{F}(\mathbf{p})(\mathbf{z}, \mathbf{z}) + \varepsilon^2 \bar{\Phi}_\varepsilon(\mathbf{z}) \right)_X = \mathbf{D}\mathbf{z}_{XX}, \quad (3.2)$$

where the remainder function $\Phi_\varepsilon(\mathbf{z})$ remains regular as $\varepsilon \rightarrow 0$.

Now split \mathbf{z} into a leading order part and a remainder

$$\mathbf{z}(X, T, \varepsilon) = \mathbf{w}(X, T, \varepsilon) + \mathbf{R}(X, T, \varepsilon), \quad (3.3)$$

with

$$\mathbf{w}(X, T, \varepsilon) = \mathbf{w}_0(X, T) + \varepsilon \mathbf{V}_0(X, T). \quad (3.4)$$

The perturbation $\mathbf{V}_0(X, T)$ is the leading order term in the asymptotic expansion (2.5). The first term in (3.4) is

$$\mathbf{w}_0(X, T) := \sum_{j=1}^k u_j(X, T) \boldsymbol{\xi}_j, \quad (3.5)$$

where $(u_1(X, T), \dots, u_k(X, T))$ is an exact associated solution of the Q -Burgers equation (1.10) with initial data

$$\mathbf{w}_0(X, 0) = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (\mathbf{I} - \mathbf{P})(\mathbf{U}(X, 0, \varepsilon) - \mathbf{p}). \quad (3.6)$$

Using Q -Burgers and the equation for \mathbf{V} in (2.4), the vector-valued function \mathbf{w} satisfies

$$\mathbf{w}_T + \frac{1}{\varepsilon} \left(\mathbf{D}\mathbf{F}(\mathbf{p})\mathbf{w} + \frac{1}{2} \varepsilon \mathbf{D}^2 \mathbf{F}(\mathbf{p})(\mathbf{w}, \mathbf{w}) \right)_X - \mathbf{D} \mathbf{w}_{XX} = \varepsilon \mathbf{h}, \quad (3.7)$$

where the function

$$\mathbf{h} := \mathbf{h}(\partial_T \mathbf{w}_0, \partial_X \mathbf{w}_0, \partial_X^2 \mathbf{w}_0, \partial_X^3 \mathbf{w}_0, \partial_T \partial_X \mathbf{w}_0),$$

is a cubic polynomial determined by the solution $(u_1, \dots, u_k)(X, T)$ of the Q -Burgers equation (1.10). Moreover, it is linear with respect to the higher derivatives $\partial_T \partial_X \mathbf{w}_0$ and $\partial_X^3 \mathbf{w}_0$:

$$\mathbf{h} = \mathbf{h}_0(\partial_T \mathbf{w}_0, \partial_X \mathbf{w}_0, \partial_X^2 \mathbf{w}_0) + \mathbf{C}_1 \partial_T \partial_X \mathbf{w}_0 + \mathbf{C}_2 \partial_X^3 \mathbf{w}_0, \quad (3.8)$$

where \mathbf{C}_1 and \mathbf{C}_2 are constant matrices depending only on \mathbf{D} , $\mathbf{D}\mathbf{F}(\mathbf{p})$ and $\mathbf{D}^2 \mathbf{F}(\mathbf{p})$.

Thus, the difference \mathbf{R} solves

$$\mathbf{R}_T + \varepsilon^{-1} \mathbf{D}\mathbf{F}(\mathbf{p})\mathbf{R}_X - \mathbf{D}\mathbf{R}_{XX} = \mathbf{H}, \quad (3.9)$$

with

$$\mathbf{H} := \varepsilon \mathbf{h} - \partial_X (\varepsilon \Phi_\varepsilon(\mathbf{z})) - \frac{1}{2} \mathbf{D}^2 \mathbf{F}(\mathbf{p})(\mathbf{R}, 2\mathbf{w} + \mathbf{R}). \quad (3.10)$$

Equation (3.9) is the key equation that needs to be understood. The validity proof reduces to showing that the remainder function $\mathbf{R}(X, T, \varepsilon)$ is well behaved. To show this we need two results: first we need to show that the solution $\mathbf{w}_0(X, T)$ of Q -Burgers is well-behaved and secondly we need to show that solutions of the linear parabolic equation (3.9) are well-defined in the uniformly local spaces. These questions are made precise and addressed in the next sections. The first question is addressed in §4 and the second question is addressed in §5.

4 The reduced Q -Burgers system

In this section, we study the global well-posedness and regularity of Q -Burgers type systems without remainder (1.10). Consider the Q -Burgers system in the form

$$\partial_t \mathbf{u} = \mathbf{E} \partial_x^2 \mathbf{u} + \partial_x \mathbf{Q}(\mathbf{u}), \quad x \in \mathbb{R}, \quad \mathbf{u}|_{t=0} = \mathbf{u}_0, \quad (4.1)$$

where the k -component vector-valued function $\mathbf{u}(x, t)$ is the k -component vector-valued unknown function, \mathbf{E} is a given constant matrix satisfying the uniform parabolicity assumption

$$\operatorname{Re} \sigma(\mathbf{E}) > 0 \quad (4.2)$$

and $\mathbf{Q} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is a given nonlinearity satisfying

$$|\mathbf{D}\mathbf{Q}(\mathbf{u})| \leq C(1 + |\mathbf{u}|) \quad \text{for all } \mathbf{u} \in \mathbb{R}^k, \quad (4.3)$$

for some constant C (throughout of the paper the constants whose concrete values are not important will be denoted by the letter C and this notation will not change even if the constant changes).

We are interested in weak solutions of problem (4.1) in the uniformly local spaces. To be more precise, a function $\mathbf{u} = \mathbf{u}(t, x)$ is a weak solution of (4.1) on the time interval $t \in [0, T]$ if

$$\mathbf{u} \in L^\infty([0, T], L_b^2(\mathbb{R})) \cap C([0, T], L_{loc}^2(\mathbb{R})), \quad \mathbf{u}_x \in L_b^2([0, T] \times \mathbb{R}) \quad (4.4)$$

and satisfies (4.1) in the sense of distributions.

The section is organized as follows. In the first subsection, we briefly discuss the key technical tools associated with weighted and uniformly local Sobolev spaces which will be used throughout of the paper. In the second subsection, we verify the uniqueness of a weak solution and its additional regularity. Finally, in the third subsection, we establish the global existence of weak solutions under additional assumptions on \mathbf{Q} and \mathbf{E} .

Note also that the critical growth of the nonlinearity \mathbf{Q} in the class of energy solutions is cubic, so most of the results of this section remain true for the nonlinearities \mathbf{Q} of a cubic growth. However, since in all our applications \mathbf{Q} is a quadratic polynomial, we impose condition (4.3) and restrict the possible growth of \mathbf{Q} to be at most quadratic.

4.1 Weighted and uniformly local Sobolev spaces

A key feature of the analysis is the intermediate step of proving existence in weighted spaces. In this subsection, the key properties needed in the analysis are recorded, see [23] for the more detailed exposition. We start with introducing the class of suitable weight functions.

A function $\phi \in L_{loc}^\infty(\mathbb{R})$ is called a weight function with exponential rate of growth ($\nu > 0$) if the conditions

$$\phi(x) > 0 \quad \text{and} \quad \phi(x + y) \leq C e^{\nu|x|} \phi(y), \quad (4.5)$$

are satisfied for every $x, y \in \mathbb{R}$.

Any weight function with growth rate ν also satisfies

$$\phi(x + y) \geq C^{-1} e^{-\nu|x|} \phi(y)$$

for all $x, y \in \mathbb{R}$. Important examples needed here, of weight functions with growth rate ν , are

$$\phi_\mu(x) = \frac{1}{(1 + \frac{1}{2}|\mu x|^2)^{\frac{1}{2}}} \quad \text{and} \quad \theta_\mu(x) = e^{-\mu|x|}, \quad (4.6)$$

where $\mu < \nu$ in the second example. Crucial for what follows is the fact that these functions satisfy (4.5) uniformly with respect to $\mu \rightarrow 0$. Moreover, if $\phi(x)$ satisfies (4.5), then the shifted weight function $\phi(x - x_0)$, $x_0 \in \mathbb{R}$, also satisfies (4.5) with the same constants C and ν .

For any weight function $\phi(x)$ with exponential rate of growth and any $1 \leq p \leq \infty$, the weighted Lebesgue spaces are defined as follows:

$$L_\phi^p(\mathbb{R}) = \left\{ u \in L_{loc}^p(\mathbb{R}) : \|u\|_{L_\phi^p} := \left(\int_{\mathbb{R}} \phi(x)^p |u(x)|^p dx \right)^{1/p} < \infty \right\}$$

Analogously, the weighted Sobolev spaces $W_\phi^{l,p}(\mathbb{R})$ are defined as subspaces of $\mathcal{D}'(\mathbb{R})$ of distributions whose derivatives up to order l inclusively belong to $L_\phi^p(\mathbb{R})$ (this works for natural l only, the weighted Sobolev spaces with fractional/negative number of derivatives can be also defined in a standard way using interpolation/duality).

The following proposition gives the technical tool for estimating the uniformly local norms of solutions using the energy estimates in weighted Sobolev spaces.

Proposition 4.1 *Let ϕ be a weight function with exponential growth such that $\|\phi\|_{L^p(\mathbb{R})} < \infty$ and let $u \in L_b^p(\mathbb{R})$ for some $1 \leq p < \infty$. Then, $u \in L_\phi^p(\mathbb{R})$ and*

$$\|u\|_{L_\phi^p} \leq C_1 \|\phi\|_{L_\phi^p} \|u\|_{L_b^p}, \quad (4.7)$$

where the constant C_1 depends only on p and the constants C and ν in (4.5) (and is independent of the concrete choice of the functions ϕ and u). Moreover,

$$\|u\|_{L_b^2} \leq C_1 \sup_{s \in \mathbb{R}} \|u\|_{L_{\phi(-s)}^2}, \quad (4.8)$$

where C_1 is also independent of the concrete choice of u and ϕ .

For the proof of this proposition, see [30].

We will mainly use estimate (4.7) in the situation where $\phi = \phi_\mu$ is one of the special weight functions of (4.6) and $\mu > 0$ is a small parameter. In this case, for $p > 1$, $\|\phi_\mu\|_{L^p} \sim \mu^{-1/p}$ and (4.7) reads

$$\|v\|_{L_{\phi_\mu}^p} \leq C\mu^{-1/p}\|v\|_{L_b^p} \quad (4.9)$$

where the constant $C > 0$ is independent of $\mu \rightarrow 0$.

The next standard proposition gives a useful equivalent form of the weighted norms.

Proposition 4.2 *Let ϕ be a weight function with exponential rate of growth and let $u \in L_\phi^p(\mathbb{R})$. Then*

$$C_1\|u\|_{L_\phi^p}^p \leq \int_{\mathbb{R}} \phi^p(x_0)\|u\|_{L^p(x_0, x_0+1)}^p dx_0 \leq C_2\|u\|_{L_\phi^p}^p \quad (4.10)$$

where the constants C_1 and C_2 depend only on the constants involving (4.5) and are independent of the concrete choice of p , ϕ and u .

Finally, we will need Gagliardo-Nirenberg type interpolation inequalities in weighted Sobolev spaces. We state below only 2 of them which will be used in the paper and refer to [30] or [31] for more detailed exposition.

Proposition 4.3 *Let ϕ be a weight function with sufficiently small exponential rate of growth ν and let $u \in W_\phi^{1,2}(\mathbb{R})$. Then, for every $2 \leq p \leq \infty$,*

$$\|u\|_{L_\phi^p} \leq C\|u\|_{L_\phi^2}^{\frac{1}{2} + \frac{1}{p}} \|u\|_{W_\phi^{1,2}}^{\frac{1}{2} - \frac{1}{p}}, \quad (4.11)$$

where the constant C is independent of p and of the concrete choice of ϕ and u . In addition, if $u \in W_\phi^{2,2}(\mathbb{R})$, then

$$\|u\|_{W_\phi^{1,2}} \leq C\|u\|_{L_\phi^2}^{\frac{1}{2}} \|u\|_{W_\phi^{2,2}}^{\frac{1}{2}} \quad (4.12)$$

and C is also independent of ϕ and u .

Remark 4.1 In this section the parameter μ in the weight function can be chosen independently. This choice is to be contrasted with §5 where the choice of weight function parameter will be linked to the parameter ε .

4.2 Uniqueness and further regularity of weak solutions

In this subsection, we show that the weak solution of equation (4.1), if it exists, is unique and is smooth if the initial data and \mathbf{Q} are regular enough. We start with uniqueness.

Proposition 4.4 *Let the diffusion matrix \mathbf{E} and the nonlinearity \mathbf{Q} satisfy (4.2) and (4.3). Then the weak solution u of (4.1) which satisfies (4.4) is unique.*

Proof Without loss of generality, we may assume that

$$\mathbf{E} + \mathbf{E}^* > 0. \quad (4.13)$$

Indeed, if (4.2) holds, one always can fix an inner product in \mathbb{R}^k in such way that (4.13) will be also satisfied.

Let \mathbf{u} and \mathbf{v} be two solutions of (4.1) belonging to the class (4.4) and let $\mathbf{w} = \mathbf{u} - \mathbf{v}$. Then, by taking the difference of two equations (4.1) (with solutions \mathbf{u} and \mathbf{v} respectively), we have

$$\partial_t \mathbf{w} - \mathbf{E} \mathbf{w}_{xx} = \partial_x \mathbf{Q}(\mathbf{u}) - \partial_x \mathbf{Q}(\mathbf{v}),$$

where $\mathbf{w} = \mathbf{u} - \mathbf{v}$. Take the dot product of this equation with $\theta_\mu^2 \mathbf{w}$ (see (4.6), where $\mu > 0$ is small enough) and integrate the result over \mathbb{R} to obtain

$$\frac{d}{dt} \|\mathbf{w}\|_{L_{\theta_\mu}^2}^2 + ((\mathbf{E} + \mathbf{E}^*)(\theta_\mu \mathbf{w}_x), (\theta_\mu \mathbf{w}_x))_{L^2} + 2(\mathbf{E} \mathbf{w}_x, \partial_x(\theta_\mu^2 \mathbf{w}))_{L^2} + (\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v}), \partial_x(\theta_\mu^2 \mathbf{w}))_{L^2} = 0, \quad (4.14)$$

where $(u, v)_{L^2}$ denotes the usual inner product in $L^2(\mathbb{R})$ (all the integrals make sense since \mathbf{u} and \mathbf{v} belong to the class (4.4)). Note that, due to (4.13), there exists a positive constant α such that $\mathbf{E} + \mathbf{E}^* \geq 2\alpha$, and the weight function θ_μ satisfies

$$|\theta'_\mu(x)| \leq \mu\theta_\mu(x), \quad x \in \mathbb{R}. \quad (4.15)$$

Using these facts together with the Cauchy-Schwarz inequality, we transform (4.14) to

$$\frac{d}{dt} \|\mathbf{w}\|_{L^2_{\theta_\mu}}^2 + \alpha \|\mathbf{w}_x\|_{L^2_{\theta_\mu}}^2 \leq C \|\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v})\|_{L^2_{\theta_\mu}}^2 + C\mu^2 \|\mathbf{w}\|_{L^2_{\theta_\mu}}^2. \quad (4.16)$$

As usual, we estimate the first term in the right-hand side using the integral mean value theorem in the form

$$\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v}) = \int_0^1 \mathbf{Q}'(s\mathbf{u} + (1-s)\mathbf{v}) ds (\mathbf{u} - \mathbf{v}),$$

so, using assumption (4.3), we obtain

$$\|\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v})\|_{L^2_{\theta_\mu}} \leq C \|(1 + |\mathbf{u}| + |\mathbf{v}|)\mathbf{w}\|_{L^2_{\theta_\mu}}.$$

To estimate the right-hand side of this inequality, we use Proposition 4.2 together with the fact that $\|\mathbf{u}\|_{L^2_b}$ and $\|\mathbf{v}\|_{L^2_b}$ are under control. Then,

$$\begin{aligned} \|\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v})\|_{L^2_{\theta_\mu}}^2 &\leq C \int_{\mathbb{R}} \theta_\mu^2(s) \|(1 + |\mathbf{u}| + |\mathbf{v}|)\mathbf{w}\|_{L^2(s, s+1)}^2 ds \\ &\leq C_1 \int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^\infty(s, s+1)}^2 (1 + \|\mathbf{u}\|_{L^2(s, s+1)}^2 + \|\mathbf{v}\|_{L^2(s, s+1)}^2) ds \\ &\leq C_2 (1 + \|\mathbf{u}\|_{L^2_b}^2 + \|\mathbf{v}\|_{L^2_b}^2) \int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^\infty(s, s+1)}^2 ds \\ &\leq C_3 \int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^\infty(s, s+1)}^2 ds. \end{aligned} \quad (4.17)$$

The right-hand side of (4.17) is estimated using the standard (non-weighted) interpolation inequality

$$\|\mathbf{w}\|_{L^\infty(s, s+1)}^2 \leq C \|\mathbf{w}\|_{L^2(s, s+1)} \|\mathbf{w}\|_{W^{1,2}(s, s+1)},$$

the Cauchy-Schwarz inequality and again Proposition 4.2:

$$\begin{aligned} \|\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v})\|_{L^2_{\theta_\mu}}^2 &\leq C_3 \int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^\infty(s, s+1)}^2 ds \leq C \int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^2(s, s+1)} \|\mathbf{w}\|_{W^{1,2}(s, s+1)} ds \\ &\leq C_4 \left(\int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{L^2(s, s+1)}^2 ds \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} \theta_\mu^2(s) \|\mathbf{w}\|_{W^{1,2}(s, s+1)}^2 ds \right)^{\frac{1}{2}} \\ &\leq C_5 \|\mathbf{w}\|_{L^2_{\theta_\mu}} \|\mathbf{w}\|_{W^{1,2}_{\theta_\mu}} \\ &\leq C \|\mathbf{w}\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{2} \|\mathbf{w}_x\|_{L^2_{\theta_\mu}}^2. \end{aligned} \quad (4.18)$$

Inserting this estimate into the right-hand side of (4.16), we end up with

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{w}\|_{L^2_{\theta_\mu}}^2 \leq C \|\mathbf{w}\|_{L^2_{\theta_\mu}}^2.$$

Integrating this inequality and using that $\mathbf{u}(0) = \mathbf{v}(0)$, we prove that $\|\mathbf{w}(t)\|_{L^2_{\theta_\mu}} \equiv 0$. Thus, $\mathbf{u}(t) \equiv \mathbf{v}(t)$ and the proposition is proved. \square

We now turn to proving the extra regularity of a weak solution. We start with the uniformly local analogue of the second energy estimate.

Proposition 4.5 *Let the assumptions of Proposition 4.2 hold and let \mathbf{u} be a weak solution of (4.1) on the time interval $t \in [0, T]$ such that*

$$\|\mathbf{u}\|_{L^\infty([0, T], L^2_b(\mathbb{R}))} \leq K, \quad \mathbf{u}_0 \in W_b^{1,2}(\mathbb{R}). \quad (4.19)$$

Then, $\mathbf{u} \in L^\infty([0, T], W_b^{1,2}(\mathbb{R}))$, $\mathbf{u}_{xx} \in L^2_b([0, T] \times \mathbb{R})$ and

$$\|\mathbf{u}\|_{L^\infty([0, T], W_b^{1,2}(\mathbb{R}))} + \|\mathbf{u}_{xx}\|_{L^2_b([0, T] \times \mathbb{R})} \leq C(\|\mathbf{u}_0\|_{W_b^{1,2}(\mathbb{R})} + K^5 + 1), \quad (4.20)$$

where the constant C is independent of \mathbf{u} , T and \mathbf{u}_0 .

Proof Take the dot product of equation (4.1) with $-\partial_x(\theta_\mu^2 \mathbf{u}_x)$, where θ_μ is defined by (4.6) and integrate in space. This gives

$$\frac{d}{dt} \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + ((\mathbf{E} + \mathbf{E}^*)(\theta_\mu \mathbf{u}_{xx}), (\theta_\mu \mathbf{u}_{xx}))_{L^2} + 2(\mathbf{E} \partial_x(\theta_\mu^2) \mathbf{u}_x, \mathbf{u}_{xx})_{L^2} = -2(\partial_x \mathbf{Q}(\mathbf{u}), \partial_x(\theta_\mu^2 \mathbf{u}_x))_{L^2}. \quad (4.21)$$

As in the proof of Proposition 4.4, we may assume that (4.13) holds. Then, using (4.15), we obtain

$$|(\mathbf{E} \partial_x(\theta_\mu^2) \mathbf{u}_x, \mathbf{u}_{xx})_{L^2}| \leq \frac{\alpha}{4} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2 + \frac{\mu^2}{\alpha} \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2, \quad (4.22)$$

and again using, in addition, the Cauchy-Schwarz inequality, we estimate the right-hand side of (4.21) as follows

$$|(\partial_x \mathbf{Q}(\mathbf{u}), \partial_x(\theta_\mu^2 \mathbf{u}_x))_{L^2}| \leq C \|\partial_x \mathbf{Q}(\mathbf{u})\|_{L^2_{\theta_\mu}}^2 + C \mu^2 \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{4} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2.$$

Inserting these estimates into (4.21) and using that μ is small enough, we arrive at

$$\frac{d}{dt} \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{2} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2 \leq C \|\partial_x \mathbf{Q}(\mathbf{u})\|_{L^2_{\theta_\mu}}^2 + C \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2. \quad (4.23)$$

Thus, it now remains to estimate the $\partial_x \mathbf{Q}(\mathbf{u})$ term. Using the growth restriction (4.3) and assumption (4.19) and arguing exactly as in (4.17) and (4.18) (where \mathbf{w} should be replaced by \mathbf{u}_x), we get

$$\begin{aligned} C \|\partial_x \mathbf{Q}(\mathbf{u})\|_{L^2_{\theta_\mu}}^2 &\leq C_1 \|(1 + \mathbf{u}) \cdot \mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 \\ &\leq C_2 \|u\|_{L^2_b}^2 \|\mathbf{u}_x\|_{L^2_{\theta_\mu}} \|\mathbf{u}_x\|_{W_b^{1,2}} \\ &\leq C_3 (K^2 + 1)^2 \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{4} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2, \end{aligned}$$

and, therefore, (4.23) now reads

$$\frac{d}{dt} \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{4} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2 \leq C(K^2 + 1)^2 \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2. \quad (4.24)$$

Finally, using the weighted interpolation inequality

$$\|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 \leq C \|\mathbf{u}\|_{L^2_{\theta_\mu}}^2 + C \|\mathbf{u}\|_{L^2_{\theta_\mu}} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}},$$

and assumption (4.19) again, we end up with

$$\frac{d}{dt} \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \|\mathbf{u}_x\|_{L^2_{\theta_\mu}}^2 + \frac{\alpha}{8} \|\mathbf{u}_{xx}\|_{L^2_{\theta_\mu}}^2 \leq C(K^4 + 1)^2 \|\mathbf{u}\|_{L^2_{\theta_\mu}}^2 \leq C(K^5 + 1)^2. \quad (4.25)$$

Applying the Gronwall inequality, we arrive at the desired weighted analogue of the second energy estimate:

$$\|\mathbf{u}_x(t)\|_{L^2_{\theta_\mu}}^2 + \int_t^{\max\{t+1, T\}} \|\mathbf{u}_{xx}(s)\|_{L^2_{\theta_\mu}}^2 ds \leq C(\|\mathbf{u}_0\|_{W_b^{1,2}} + 1 + K^5)^2. \quad (4.26)$$

Moreover, since the shifted weight function $\theta_{\mu,s}(x) := \theta_\mu(x - s)$ satisfies estimates (4.5) and (4.15) uniformly with respect to $s \in \mathbb{R}$, arguing as before, we see that estimate (4.26) remains true if we replace θ_μ by $\theta_{\mu,s}$ and the constant C is independent of $s \in \mathbb{R}$. Taking now the supremum over $s \in \mathbb{R}$ from both parts of the obtained inequality and using estimates (4.7) and (4.8), we end up with (4.20) and finish the proof of Proposition 4.5. \square

The regularity of the solution \mathbf{u} can be further improved in a straightforward way if additional smoothness of \mathbf{Q} and \mathbf{u}_0 is known. In particular, the following result holds.

Corollary 4.1 *Let the assumptions of Proposition 4.5 hold and let, in addition, $\mathbf{Q} \in C^2(\mathbb{R}^2, \mathbb{R}^2)$ and $\mathbf{u}_0 \in W_b^{2,2}(\mathbb{R})$. Then, the solution \mathbf{u} of the \mathbf{Q} -Burgers system (4.1) satisfies:*

$$\mathbf{u} \in L^\infty([0, T], W_b^{2,2}(\mathbb{R})), \quad \mathbf{u}_{tx}, \mathbf{u}_{xxx} \in L_b^2([0, T] \times \mathbb{R}),$$

and the following estimate holds:

$$\|\mathbf{u}\|_{L^\infty([0, T], W_b^{2,2}(\mathbb{R}))} + \|\mathbf{u}_{tx}\|_{L_b^2([0, T] \times \mathbb{R})} + \|\mathbf{u}_{xxx}\|_{L_b^2([0, T] \times \mathbb{R})} \leq C \left(\|\mathbf{u}_0\|_{W_b^{2,2}(\mathbb{R})} + C_K \right), \quad (4.27)$$

where the constant C_K depends on K (from (4.19)), but is independent of T and \mathbf{u} .

The proof of this corollary is analogous to Proposition 4.5: we just need to differentiate (4.1) with respect to x (denoting $\mathbf{v} = \mathbf{u}_x$) and then multiply the obtained equation by $(\theta_\mu \mathbf{v}_x)_x$, so we leave it to the reader.

4.3 A priori estimates and global existence

In this subsection, we present some results on the global existence of weak solutions of the \mathbf{Q} -Burgers system. As we have seen before, the control of the $L_b^2(\mathbb{R})$ -norm of a solution is enough to estimate the higher norms, so finding an appropriate *a priori* estimate of this norm is sufficient for the global existence of weak/strong solutions (which can be proved in a straightforward way, e.g., by approximating the problem in unbounded domains by the analogous problems in large bounded domains, say, with periodic boundary conditions and by passing to the limit, see EFENDIEV & ZELIK [12, 13]). Thus, we will concentrate on deriving proper *a priori* estimates for the L_b^2 -norm of the solution only.

Note that, in contrast to the scalar case $k = 1$, where such an estimate is immediate due to the maximum principle, the vector case looks much more delicate since the maximum principle fails and one should use principally different methods to bound the L_b^2 -norm of the solution. In particular, based on some techniques, initially developed for infinite energy solutions for the Navier-Stokes equations in a strip [31], we succeed in finding such an estimate under the additional assumption that the nonlinearity \mathbf{Q} is gradient

$$\mathbf{Q}(\mathbf{u}) = \nabla_u \Psi(\mathbf{u}), \quad (4.28)$$

for some potential $\Psi \in C^2(\mathbb{R}^2)$ and the diffusion matrix \mathbf{E} satisfying (4.13) (which is now essential since the trick with changing the inner product in \mathbb{R}^k used before may destroy the gradient structure of the problem). Note also that, again in contrast to the scalar case, our estimate allows the norm $\|\mathbf{u}(t)\|_{L_b^2}$ to grow at most linearly in time although we do not know of any examples of growing solutions in that class. To be more precise, the following result holds.

Theorem 4.1 *Let the assumptions of Proposition 4.4 hold and let, in addition, (4.13) and (4.28) hold. Then, for every initial data $\mathbf{u}_0 \in L_b^2(\mathbb{R})$, there exists a unique global weak solution \mathbf{u} of the \mathbf{Q} -Burgers system (4.1) and the following estimate is valid:*

$$\|\mathbf{u}(t)\|_{L_b^2(\mathbb{R})} \leq R(\|\mathbf{u}_0\|_{L_b^2(\mathbb{R})})(t + 1) \quad (4.29)$$

for some monotone increasing function R which is independent of t and \mathbf{u}_0 .

Proof As was already mentioned, we only need to derive an *a priori* estimate (4.29). To this end, we will use as before the weighted energy estimates, in contrast to the previous subsection, the exponential weight function $\theta_\mu(x)$ is no longer appropriate for our purposes and is replaced by the polynomial one $\phi_\mu(x)$ (see (4.6)). The reason is that the polynomial weight function satisfies a stronger inequality than (4.15), namely,

$$|\phi'_\mu(x)| \leq \frac{1}{2} \mu \phi_\mu^2(x) \quad \text{and} \quad \sup_{x \in \mathbb{R}} \phi_\mu(x) \leq 1, \quad (4.30)$$

which is crucial to handle the nonlinearity.

Multiplying equation (4.1) by $\phi_\mu^2 \mathbf{u}$, where $\mu > 0$ is a small parameter which will be fixed below, and integrating in space, we have

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{u}\|_{L_{\phi_\mu}^2}^2 + \frac{1}{2} ((\mathbf{E} + \mathbf{E}^*) \mathbf{u}_x, \phi_\mu^2 \mathbf{u}_x) + (\mathbf{E} \mathbf{u}_x, \mathbf{u} \partial_x (\phi_\mu^2))_{L^2} = (\partial_x (\mathbf{Q}(\mathbf{u})), \phi_\mu^2 \mathbf{u})_{L^2}. \quad (4.31)$$

Note that condition (4.13) implies that there exists a positive constant α such that $\mathbf{E} + \mathbf{E}^* \geq 2\alpha$ and since the weight ϕ_μ satisfies (4.15), we may estimate the last term in the left-hand side of (4.31) analogously to the third term in (4.14). Moreover, using that \mathbf{Q} is gradient, we simplify the right-hand side by integrating by parts:

$$\begin{aligned} (\partial_x(\mathbf{Q}(\mathbf{u})), \phi_\mu^2 \mathbf{u})_{L^2} &= -(\mathbf{Q}(\mathbf{u}), \mathbf{u}_x \phi_\mu^2)_{L^2} - (\mathbf{Q}(\mathbf{u}), \mathbf{u} \partial_x(\phi_\mu^2))_{L^2} = \\ &= -((\Psi(\mathbf{u}))_x, \phi_\mu^2)_{L^2} - (\mathbf{Q}(\mathbf{u}), \mathbf{u} \partial_x(\phi_\mu^2))_{L^2} = (\Psi(\mathbf{u}) - \mathbf{Q}(\mathbf{u}) \cdot \mathbf{u}, \partial_x(\phi_\mu^2))_{L^2}. \end{aligned}$$

Thus, (4.31) reads

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 + \alpha \|\mathbf{u}_x\|_{L^2_{\phi_\mu}}^2 \leq C\mu^2 \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 + 2(\Psi(\mathbf{u}) - \mathbf{u} \cdot \mathbf{Q}(\mathbf{u}), \partial_x(\phi_\mu^2))_{L^2}, \quad (4.32)$$

where the constant C is independent of $\mu \rightarrow 0$.

Now using that $\Psi(\mathbf{u})$ and $\mathbf{Q}(\mathbf{u}) \cdot \mathbf{u}$ are at most cubic, namely,

$$|D\mathbf{Q}(\mathbf{u})| \leq C(1 + |\mathbf{u}|) \quad \text{implies} \quad |\mathbf{Q}(\mathbf{u})| \leq C(1 + |\mathbf{u}|^2) \quad \text{and} \quad |\Psi(\mathbf{u})| \leq C(1 + |\mathbf{u}|^3),$$

together with (4.30) and the fact that $\int_{\mathbb{R}} \phi_\mu(x) dx = C\mu^{-1}$, we estimate the last term in the right-hand side as follows:

$$2(\Psi(\mathbf{u}) - \mathbf{u} \cdot \mathbf{Q}(\mathbf{u}), \partial_x(\phi_\mu^2))_{L^2} \leq C\mu(|\mathbf{u}|^3 + 1, \phi_\mu^3)_{L^2} = C(\mu \|\mathbf{u}\|_{L^3_{\phi_\mu}}^3 + 1),$$

where C is again independent of $\mu \rightarrow 0$. Using now the weighted interpolation inequality (4.11) with $p = 3$, we get

$$\begin{aligned} C\mu \|\mathbf{u}\|_{L^3_{\phi_\mu}}^3 &\leq C_1 \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{5/2} \|\mathbf{u}\|_{W_\mu^{1,2}}^{1/2} \\ &\leq C_1 \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^3 + C_1 \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{5/2} \|\mathbf{u}_x\|_{L^2_{\phi_\mu}}^{1/2} \\ &\leq C_1 \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^3 + C\mu^{4/3} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{10/3} + \alpha \|\mathbf{u}_x\|_{L^2_{\phi_\mu}}^2. \end{aligned} \quad (4.33)$$

Inserting these estimates into the right-hand side of (4.32), we arrive at

$$\frac{d}{dt} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 \leq C \left(\mu^2 \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 + \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^3 + \mu^{4/3} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{10/3} + 1 \right). \quad (4.34)$$

We may simplify (4.34) furthermore using again Young's inequality

$$\mu^2 \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 \leq \frac{2}{3} \mu^3 \|\mathbf{u}\|_{L^2_{\phi_\mu}}^3 + C \quad \text{and} \quad \mu \|\mathbf{u}\|_{L^2_{\phi_\mu}}^3 \leq \mu^{7/6} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{10/3} + C\mu^{-1/2},$$

and using that μ is small. This gives, finally,

$$\frac{d}{dt} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^2 \leq C \left(\mu^{7/6} \|\mathbf{u}\|_{L^2_{\phi_\mu}}^{10/3} + \mu^{-1/2} \right). \quad (4.35)$$

In addition, due to (4.9), we also have

$$\|\mathbf{u}_0\|_{L^2_{\phi_\mu}}^2 \leq C\mu^{-1} \|\mathbf{u}_0\|_{L^2_b}^2. \quad (4.36)$$

Crucial for us is that the constant C in (4.35) and (4.36) is independent of $\mu \rightarrow 0$. Indeed, for every fixed μ , the solutions of the differential inequality (4.35) may blow up in finite time. Nevertheless, choosing properly the parameter μ , dependence of the initial data and the time interval, we gain the global existence as well as the linear growth estimate (4.29). In order to see that, we introduce the function $Y_\mu(t)$ as a solution of the following ODE:

$$Y'_\mu = C \left(\mu^{7/6} Y_\mu^{5/3} + \mu^{-1/2} \right) \quad \text{with} \quad Y_\mu(0) = C\mu^{-1} \|\mathbf{u}_0\|_{L^2_b}^2.$$

Then, due to the comparison principle,

$$\|\mathbf{u}(t)\|_{L^2_{\phi_\mu}}^2 \leq Y_\mu(t), \quad t \geq 0. \quad (4.37)$$

After scaling the variables $Z_\mu(t) = \mu Y_\mu(t\mu^{-1/2})$, we see that the new function $Z_\mu(t) = Z(t)$ is independent of μ and solves

$$Z' \leq C(Z^{5/3} + 1), \quad Z(0) = C\|\mathbf{u}_0\|_{L_b^2}^2.$$

This ODE is locally solvable in time and we have

$$Z(t) \leq 2C(\|u_0\|_{L_b^2}^2 + 1), \quad t \leq \beta(\|\mathbf{u}_0\|_{L_b^2}^2) > 0,$$

for some monotone decreasing positive function β . Thus, due to (4.37),

$$\|\mathbf{u}(t)\|_{L_{\phi_\mu}^2}^2 \leq 2C\mu^{-1}(\|\mathbf{u}_0\|_{L_b^2}^2 + 1), \quad t \leq \mu^{-1/2}\beta(\|\mathbf{u}_0\|_{L_b^2}^2). \quad (4.38)$$

Moreover, as in the previous subsection, this estimate remains valid if we replace $\phi_\mu(x)$ on the shifted one $\phi_{\mu,s}(x) := \phi_\mu(x-s)$ uniformly with respect to the shift $s \in \mathbb{R}$. Taking after that the supremum over $s \in \mathbb{R}$ and using (4.8), we end up with

$$\|\mathbf{u}(t)\|_{L_b^2}^2 \leq 2C\mu^{-1}(\|\mathbf{u}_0\|_{L_b^2}^2 + 1), \quad t \leq \mu^{-1/2}\beta(\|\mathbf{u}_0\|_{L_b^2}^2). \quad (4.39)$$

We are now ready to optimize this estimate with respect to μ . Namely, for every large enough time moment T , we fix $\mu = T^{-2}\beta^2(\|\mathbf{u}_0\|_{L_b^2}^2)$. Then, (4.39) holds for $t = T$ and this gives

$$\|\mathbf{u}(T)\|_{L_b^2}^2 \leq 2CT^2\beta^2(\|\mathbf{u}_0\|)(\|\mathbf{u}_0\|_{L_b^2}^2 + 1)$$

which coincides with (4.29) and finishes the proof of the theorem. \square

5 The analysis of the scaled linearized system

As we have seen in Section 3, in order to justify the validity of the \mathbf{Q} -Burgers reduction, we need to study equation (3.9) for small $\varepsilon > 0$. Note that the case $\varepsilon = 1$ of that equation corresponds to the linearization of the initial problem (1.1) about a constant solution $\mathbf{U} = \mathbf{p}$. The aim of this section is to state the conditions which guarantee the spectral stability of that linearized problem and derive several uniform with respect to $\varepsilon \rightarrow 0$ estimates which are crucial for the validity proof.

5.1 Spectral properties of the linearized system

Of interest is the following ε -family of parabolic PDEs,

$$\mathbf{u}_t + \frac{1}{\varepsilon}\mathbf{A}\mathbf{u}_x - \mathbf{D}\mathbf{u}_{xx} = \mathbf{h}(t), \quad \mathbf{u}|_{t=0} = \mathbf{u}_0, \quad (5.1)$$

for the vector-valued function $\mathbf{u}(x, t)$. Here $\varepsilon > 0$ is any given small parameter, $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{h}(t)$ are given external forces, and $\mathbf{A} = \mathbf{D}\mathbf{F}(\mathbf{p})$ and \mathbf{D} are given constant matrices with the following properties:

$$(1) : \mathbf{A} \text{ has real eigenvalues and is diagonalizable} \quad (5.2)$$

$$(2) : \operatorname{Re} \sigma(\mathbf{D}) > 0 \quad (5.3)$$

$$(3) : \operatorname{Re} \sigma(\mathbf{D}k^2 + ik\mathbf{A}) > 0, \quad \forall k \in \mathbb{R} \setminus \{0\} \quad (5.4)$$

$$(4) : \operatorname{Re} \sigma(\mathbf{D}_\mathbf{A}) > 0. \quad (5.5)$$

In the hypothesis (5.5) $\mathbf{D}_\mathbf{A}$ is a block diagonal matrix built up by the following rule: let $\omega \in \sigma(\mathbf{A})$ be the eigenvalue of \mathbf{A} of multiplicity l and let $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_l\}$ and $\{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_l\}$ be the normalized right and left eigenvectors respectively associated with the eigenvalue ω . Then, the entries of the block in $\mathbf{D}_\mathbf{A}$ associated with ω are $\mathbf{D}_\omega := (\langle \mathbf{D}\boldsymbol{\xi}_i, \boldsymbol{\eta}_j \rangle)_{i,j=1}^l$. Thus,

$$\mathbf{D}_\mathbf{A} = \operatorname{diag}(\mathbf{D}_{\omega_1}, \dots, \mathbf{D}_{\omega_m})$$

where $\omega_1, \dots, \omega_m$ are all different eigenvalues of \mathbf{A} .

We first clarify the above conditions. To this end, we note that the spectral stability requirement for the linearized system (5.1) is *necessary* for the validity of the reduction and reads

$$\operatorname{Re} \sigma(\mathcal{L}_\varepsilon, L^2(\mathbb{R})) \geq 0, \quad \mathcal{L}_\varepsilon \mathbf{u} := \frac{1}{\varepsilon} \mathbf{A} \mathbf{u}_x - \mathbf{D} \mathbf{u}_{xx}. \quad (5.6)$$

The spectrum of this differential operator is independent of ε up to the scaling factor ε^{-2} and can be easily found using the Fourier transform:

$$\sigma(\mathcal{L}_\varepsilon) = \varepsilon^{-2} \sigma(\mathcal{L}_1) = \varepsilon^{-2} \{ \lambda \in \sigma(\mathbf{D}k^2 + ik\mathbf{A}), k \in \mathbb{R} \}. \quad (5.7)$$

Thus, assumption (5.4) is close to being necessary (the necessary non-strict inequality is replaced by the strict one) and guarantees that the spectrum of the operator \mathcal{L}_ε touches the imaginary axis only at wavenumber $k = 0$ and $\lambda = 0$. Furthermore, looking at large wavenumbers k , we see that the spectrum of \mathbf{D} should not contain eigenvalues with negative real part. The slightly stricter assumption (5.3) guarantees in addition the uniform parabolicity of (5.1) and is posed for technical reasons. Although the case of zero eigenvalues of the dissipation matrix \mathbf{D} is interesting from the applications point of view (see Section 9), its rigorous treatment requires essentially more complicated technique and will be not considered here.

Analogously, considering the case of small wavenumbers $k \rightarrow 0$, we see that (5.4) is possible only if the spectrum of \mathbf{A} is real. The rest of hypothesis (5.2) that \mathbf{A} is diagonalizable is also a simplifying assumption, some discussion on what happens in the non-semi-simple case is given in Section 11. Finally, for clarifying hypothesis (5.5), we need the following lemma.

Lemma 5.1 . *Let $\omega \in \sigma(\mathbf{A})$ be a semi-simple eigenvalue of the matrix \mathbf{A} of multiplicity l . Then, it generates l -branches $\lambda_j(k)$, $j = 1, \dots, l$, of the spectrum of \mathcal{L}_ε near $k = 0$ and the following asymptotic expansions hold:*

$$\lambda_j(k) = i\omega k + \mu_j k^2 + o(k^2), \quad \mu_j \in \sigma(\mathbf{D}_\omega) \quad (5.8)$$

for $j = 1, \dots, l$.

The proof of the expansions (5.8) is standard and straightforward and is therefore omitted.

Thus, hypothesis (5.5) is also "almost" necessary and prevents the branches of spectrum of \mathcal{L}_ε from crossing the imaginary axis for small wavenumbers k . Note also that $\hat{\mathbf{D}} = \mathbf{D}_0$, see (1.11), and hypothesis (5.5) also guarantees the uniform parabolicity (and, therefore, the local well-posedness) of the reduced \mathbf{Q} -Burgers equations (1.10).

5.2 Uniform estimates: the unweighted case

We now turn to the study of the linear problem (5.1). We start with the homogeneous system in non-weighted Sobolev spaces. Set $\mathbf{h} = 0$ in (5.1). The following standard Lemma shows that equation (5.1) generates a uniformly (with respect to $\varepsilon \rightarrow 0$) stable semigroup in the space $H = L^2(\mathbb{R})$.

Lemma 5.2 *Let the above assumptions (5.2)-(5.5) hold. Then, the solution $\mathbf{u}(t)$, $t \geq 0$, of equation (5.1) with $\mathbf{h} = 0$ satisfies*

$$\|\mathbf{u}(t)\|_{H^s(\mathbb{R})} \leq C \|\mathbf{u}(0)\|_{H^s(\mathbb{R})}, \quad \text{for any } s \geq 0, \quad (5.9)$$

where the constant C is independent of s , $\varepsilon > 0$ and $t \geq 0$.

Proof It is sufficient to verify the estimate for $s = 0$ only, since the equation is spatially-homogeneous. Applying the x -Fourier transform to equation (5.1) gives

$$\hat{\mathbf{u}}_t = - \left(\mathbf{D}k^2 + \frac{1}{\varepsilon} \mathbf{A}k \right) \hat{\mathbf{u}}(t). \quad (5.10)$$

Thus, due to Parseval's equality, we only need to check that

$$\left\| e^{-(ik\varepsilon^{-1}\mathbf{A} + \mathbf{D}k^2)t} \right\| \leq C,$$

where the constant C is independent of ε , $t \geq 0$ and k . Or, scaling $t \rightarrow \varepsilon^2 t$, $k \rightarrow \varepsilon^{-1} k$, this requirement is transformed to the case of $\varepsilon = 1$:

$$\left\| e^{-t(ik\mathbf{A} + \mathbf{D}k^2)} \right\| \leq C. \quad (5.11)$$

Due to assumption (5.3) on the spectrum of the matrix \mathbf{D} , (5.11) is satisfied uniformly for large k , say, for $|k| > k_+ \gg 1$. On the other hand, due to assumption (5.4), (5.11) is satisfied for every fixed $k \neq 0$. Consequently, using a compactness argument, (5.11) holds uniformly for all $|k| \leq k_- \ll 1$ (where k_- can be chosen arbitrarily small). Finally, the validity of (5.11) for small k is assured by the hypothesis that the matrix \mathbf{A} is diagonalizable and the following expansion: let λ_j^0 , $j = 1, \dots, k$ be the eigenvalues of \mathbf{A} . Then, due to Lemma 5.1, the eigenvalues of $\mathbf{D}k^2 + ik\mathbf{A}$ near $k = 0$ have the form

$$\lambda_j(ik\mathbf{A} + k^2\mathbf{D}) = ik\lambda_j^0 + \mu k^2 + o(k^2), \quad (5.12)$$

for some $\mu \in \sigma(\mathbf{D}\mathbf{A})$. The assumption (5.4) assures that $\operatorname{Re} \mu > 0$ in the expansion (5.12). This completes the proof of estimate (5.11) and finishes the proof of the lemma. \square

Now suppose that $\mathbf{h} \neq 0$.

Lemma 5.3 *Let the above assumptions (5.2)-(5.5) hold, and suppose that $\mathbf{h} \in L^2(\mathbb{R}_t \times \mathbb{R}_x)$. Then, there exists a unique solution $\mathbf{u}(t)$ of (5.1) satisfying the following inequality*

$$\left\| \frac{\partial^2 \mathbf{u}}{\partial x^2} \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \|\mathbf{h}\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)}, \quad (5.13)$$

where the constant C is independent of ε and \mathbf{h} .

Proof Applying the (t, x) -Fourier transform to equation (5.1),

$$[i\omega + \varepsilon^{-1}ik\mathbf{A} + \mathbf{D}k^2]\widehat{\mathbf{u}}(\omega, k) = \widehat{\mathbf{h}}(\omega, k), \quad (\omega, k) \in \mathbb{R}^2.$$

Using Parseval's equality, the estimate (5.13) is satisfied if (and only if) the following inequality holds

$$k^2 \left\| [i\omega + \varepsilon^{-1}ik\mathbf{A} + \mathbf{D}k^2]^{-1} \right\| \leq C,$$

where C is independent of ε , k and ω . Scaling $\omega \rightarrow \varepsilon^2\omega$ and $k \rightarrow k\varepsilon^{-1}$, the problem is reduced to the case $\varepsilon = 1$:

$$k^2 \left\| [i\omega + ik\mathbf{A} + \mathbf{D}k^2]^{-1} \right\| \leq C. \quad (5.14)$$

Estimate (5.14) holds uniformly for all $\omega \in \mathbb{R}$ and large k ($|k| > k_+ \gg 1$) due to the assumption (5.3) on the matrix \mathbf{D} , so only the case of $|k|$ bounded need be considered. In that case, (5.14) is trivially satisfied if $|\omega| \rightarrow \infty$, so only the case of bounded (ω, k) need be considered.

Now, due to (5.4), the matrix on the left-hand side of (5.14) is invertible for $k \neq 0$. Thus, using a compactness argument, equation (5.14) needs to be verified only for the case $k \rightarrow 0$. To this end, scale $\tilde{\omega} = \omega k^{-1}$ and transform (5.14) to

$$k \left\| [i\tilde{\omega} + i\mathbf{A} + \mathbf{D}k]^{-1} \right\| \leq C, \quad (5.15)$$

which should be satisfied for all $\tilde{\omega} \in \mathbb{R}$ and $|k| \ll 1$.

Let $\{\lambda_1^0, \dots, \lambda_m^0\}$ be the (real) eigenvalues of the matrix \mathbf{A} . Then, (5.15) may be violated only in the case where $|\tilde{\omega} - \lambda_i^0| \ll 1$, $i = 1, \dots, m$ and the necessary and sufficient condition for (5.15) to be satisfied in that region of parameters is that

$$|\lambda_i(i\tilde{\omega} + i\mathbf{A} + \mathbf{D}k)| \geq C|k|. \quad (5.16)$$

But this condition is guaranteed by assumption (5.4) and Lemma 5.1. This completes the proof of Lemma 5.3. \square

The next lemma allows for the reduction of the general case of equation (5.1) to the particular case $\mathbf{u}(0) = 0$.

Lemma 5.4 *Let the above assumptions (5.2)-(5.5) hold. Then, for every $\mathbf{u}_0 \in H^1(\mathbb{R})$ and every $\varepsilon > 0$, there exists a function $\mathbf{w} \in C(\mathbb{R}, H^1) \cap L^2(\mathbb{R}, H^2)$ such that $\mathbf{w}(t) \equiv 0$ for $t \notin [-1, 1]$, $\mathbf{w}(0) = \mathbf{u}_0$ and*

$$\|\mathbf{w}\|_{C([-1,1], H^1) \cap L^2([-1,1], H^2)} \leq C \|\mathbf{u}_0\|_{H^1(\mathbb{R})}, \quad (5.17)$$

where C is independent of $\varepsilon \neq 0$ and \mathbf{u}_0 . Moreover,

$$\mathbf{w}_t + \frac{1}{\varepsilon} \mathbf{A} \mathbf{w}_x - \mathbf{D} \mathbf{w}_{xx} = \mathbf{f},$$

and

$$\|\mathbf{f}\|_{L^2([-1,1] \times \mathbb{R})} \leq C_1 \|\mathbf{u}_0\|_{H^1},$$

where C_1 is also independent of ε and \mathbf{u}_0 .

Proof It is sufficient to prove the lemma for $\mathbf{D} = \mathbf{I}$, where \mathbf{I} is the identity matrix. Fix $\mathbf{D} = \mathbf{I}$. Without loss of generality, \mathbf{A} can be assumed to be symmetric. Construct the function $\tilde{\mathbf{w}}(t)$ as a solution of the following equations:

$$\left. \begin{aligned} \tilde{\mathbf{w}}_t + \frac{1}{\varepsilon} \mathbf{A} \tilde{\mathbf{w}}_x - \tilde{\mathbf{w}}_{xx} &= 0, & t \geq 0, & \quad \tilde{\mathbf{w}}(0) = \mathbf{u}_0, \\ \tilde{\mathbf{w}}_t + \frac{1}{\varepsilon} \mathbf{A} \tilde{\mathbf{w}}_x + \tilde{\mathbf{w}}_{xx} &= 0, & t \leq 0, & \quad \tilde{\mathbf{w}}(0) = \mathbf{u}_0. \end{aligned} \right\} \quad (5.18)$$

Then, since \mathbf{A} is symmetric, multiplication of the first of (5.18) by $\tilde{\mathbf{w}} - \tilde{\mathbf{w}}_{xx}$, and integrating over time and space,

$$\frac{1}{2} \|\tilde{\mathbf{w}}\|_{C(0,1;H^1(\mathbb{R}))}^2 + \int_0^1 \|\tilde{\mathbf{w}}_x(t)\|_{H^1(\mathbb{R})}^2 dt \leq \|\tilde{\mathbf{w}}(0)\|_{H^1(\mathbb{R})}^2.$$

For negative times, multiply the second equation of (5.18) by $\tilde{\mathbf{w}} - \tilde{\mathbf{w}}_{xx}$ giving

$$\frac{1}{2} \|\tilde{\mathbf{w}}\|_{C(-1,0;H^1(\mathbb{R}))}^2 + \int_{-1}^0 \|\tilde{\mathbf{w}}_x(t)\|_{H^1(\mathbb{R})}^2 dt \leq \|\tilde{\mathbf{w}}(0)\|_{H^1(\mathbb{R})}^2.$$

These estimates combine to give the desired estimate (5.17). Finally, setting

$$\mathbf{w}(t, x) := \tilde{\mathbf{w}}(t, x) \chi(t),$$

where $\chi(t)$ is a cut-off function which equals one at $t = 0$ and is zero outside of the interval $[-1, 1]$, gives the desired extension \mathbf{w} . \square

The main result of this section can now be formulated and proved.

Theorem 5.1 *Let the assumptions of Lemma 5.2 and 5.3 hold. Then, the solution $\mathbf{u}(t)$ of problem (5.1) satisfies*

$$\|\mathbf{u}(t)\|_{H^1(\mathbb{R})}^2 + \int_0^t \|\mathbf{u}_{xx}\|_{H(\mathbb{R})}^2 ds \leq C \left(\|\mathbf{u}(0)\|_{H^1(\mathbb{R})}^2 + \int_0^t \|\mathbf{h}(s)\|_{H(\mathbb{R})}^2 ds \right), \quad (5.19)$$

where the constant C is independent of ε , \mathbf{h} and $t \geq 0$.

Proof Due to Lemma 5.4, it can be assumed, without loss of generality that $\mathbf{u}(0) = 0$. The general case is reduced to that particular one just by replacing \mathbf{u} by $\mathbf{u} - \mathbf{w}$ where \mathbf{w} is constructed in Lemma 5.4. Now apply Lemma 5.3 with $\tilde{\mathbf{h}}(s) = \mathbf{h}(s)$, for $s \in [0, t]$ and $\mathbf{h}(s) = 0$ for $s \notin [0, t]$. Then, estimate (5.13) reads

$$\int_0^t \|\mathbf{u}_{xx}\|_H^2 ds \leq C \int_0^t \|\mathbf{h}(s)\|_H^2 ds. \quad (5.20)$$

Thus, it is only necessary to verify the estimate of the H^1 -norm of the solution. To this end, due to the estimate (5.20), the term $\mathbf{D}\mathbf{u}_{xx}$ may be interpreted as the right-hand side and write problem (5.1) in the form

$$\mathbf{u}_t + \frac{1}{\varepsilon} \mathbf{A} \mathbf{u}_x = \tilde{\mathbf{h}} := \mathbf{h} + \mathbf{D}\mathbf{u}_{xx}. \quad (5.21)$$

Without loss of generality, assume that \mathbf{A} is symmetric here. Multiply this equation by \mathbf{u}_{xx} and integrate over t and x to obtain the desired estimate for the H^1 -norm of \mathbf{u} , thus finishing the proof. \square

5.3 Uniform estimates: the weighted and uniformly local case

Our aim now is to obtain the analogue of (5.19) in weighted and uniformly local spaces.

A novelty in this section is that ε , which is a measure of the distance from the criticality surface, will also be used for the parameter in the definition of the weighted space: $\mu = \varepsilon$. Indeed, it will be shown that this choice is essential. On the other hand, this use of ε is only an intermediate step as the final results are in terms of the uniformly local spaces which are independent of the choice of weight function.

To this end, we consider the case of weighted spaces with the weight function

$$\theta_\varepsilon(x) := e^{-\varepsilon|x|},$$

where the parameter ε in the weight is the same as in (5.1).

Corollary 5.1 *Let the assumptions of Theorem 5.1 hold and let $\theta_\varepsilon(x) := e^{-\varepsilon|x|}$ be an exponential weight function with the same exponent ε as in (5.1). Then the solution \mathbf{u} of equation (5.1) satisfies the following estimate:*

$$\|\mathbf{u}(t)\|_{H_{\theta_\varepsilon}^1}^2 + \int_0^t e^{C(t-s)} \|\mathbf{u}(s)\|_{H_{\theta_\varepsilon}^2}^2 ds \leq C e^{Ct} \|\mathbf{u}(0)\|_{H_{\theta_\varepsilon}^1}^2 + C \int_0^t e^{C(t-s)} \|\mathbf{h}(s)\|_{L_{\theta_\varepsilon}^2}^2 ds, \quad (5.22)$$

where the constant C is independent of ε .

Proof Let $\bar{\theta}_\varepsilon(x) = e^{-\varepsilon\sqrt{x^2+1}}$ be the smooth analogue of the weight $\theta_\varepsilon(x)$ and let $\mathbf{w}(t) = \bar{\theta}_\varepsilon \mathbf{u}(t)$. Then, this function solves the following equation

$$\mathbf{w}_t + \frac{1}{\varepsilon} \mathbf{A} \mathbf{w}_x - \mathbf{w}_{xx} = \tilde{\mathbf{h}} := \bar{\theta}_\varepsilon \mathbf{h} + \frac{1}{\varepsilon} \bar{\theta}'_\varepsilon (\bar{\theta}_\varepsilon)^{-1} \mathbf{w} - 2(\bar{\theta}_\varepsilon)^{-1} \bar{\theta}'_\varepsilon \mathbf{w}_x - (\bar{\theta}_\varepsilon)^{-1} \bar{\theta}''_\varepsilon \mathbf{w}. \quad (5.23)$$

Using the obvious formula

$$|\bar{\theta}'_\varepsilon(x)| + |\bar{\theta}''_\varepsilon(x)| \leq C \varepsilon \bar{\theta}_\varepsilon(x), \quad (5.24)$$

where the constant C is independent of ε , we see that

$$\|\tilde{\mathbf{h}}(t)\|_{L^2(\mathbb{R})}^2 \leq C \|\bar{\theta}_\varepsilon \mathbf{h}(t)\|_{L^2(\mathbb{R})}^2 + C \|\mathbf{w}(t)\|_{H^1(\mathbb{R})}^2. \quad (5.25)$$

Applying estimate (5.19) to equation (5.23) gives

$$\|\mathbf{w}(t)\|_{H^1}^2 + \int_0^t \|\mathbf{w}_{xx}(s)\|_{L^2}^2 ds \leq C \|\mathbf{w}(0)\|_{H^1}^2 + C \int_0^t \|\bar{\theta}_\varepsilon \mathbf{h}(s)\|_{L^2}^2 ds + C \int_0^t \|\mathbf{w}(s)\|_{H^1}^2 ds. \quad (5.26)$$

This estimate together with Gronwall's inequality and the equivalence

$$\|\mathbf{u}\|_{H_{\theta_\varepsilon}^s} \sim \|\bar{\theta}_\varepsilon \mathbf{u}\|_{H^s},$$

give

$$\|\mathbf{u}(t)\|_{H_{\theta_\varepsilon}^1}^2 + \int_0^t e^{C(t-s)} \|\mathbf{u}(s)\|_{H_{\theta_\varepsilon}^2}^2 ds \leq C \|\mathbf{u}(0)\|_{H_{\theta_\varepsilon}^1}^2 e^{Ct} + C \int_0^t e^{C(t-s)} \|\mathbf{h}(s)\|_{L_{\theta_\varepsilon}^2}^2 ds, \quad (5.27)$$

which is equivalent to (5.22). \square

Corollary 5.2 *Let the assumptions of Theorem 5.1 hold and let $\mathbf{h} \in L_b^2(\mathbb{R}_+ \times \mathbb{R})$ and $\mathbf{u}(0) \in H_b^1(\mathbb{R})$. Then, the following estimate holds for the solution $\mathbf{u}(t)$ of equation (5.1):*

$$\|\mathbf{u}(t)\|_{H_b^1} + \|\mathbf{u}_{xx}\|_{L_b^2([0,t] \times \mathbb{R})} \leq C \varepsilon^{-1/2} e^{Ct} \left(\|\mathbf{u}(0)\|_{H_b^1(\mathbb{R})} + \|\mathbf{h}\|_{L_b^2([0,t] \times \mathbb{R})} \right) \quad (5.28)$$

where the constant t is independent of ε , v and $t \geq 0$.

Proof Indeed writing out the space-shifted version of estimate (5.22) (with the weight function $\theta_\varepsilon(x)$ replaced by $\theta_\varepsilon(x-s)$), taking the supremum with respect to $s \in \mathbb{R}$ from both sides of the obtained inequality and using estimate (4.9) with $p=2$, gives the desired inequality. \square

6 Validity of the reduction to Q -Burgers equation

We now have the technical tools required to study the the remainder equation (3.9) and to verify the long-time validity of the reduced Q -Burgers equations. The strategy is to apply Corollary 5.1 to equation (3.9). To this end, the following assumptions on the linear part of this equation are required:

$$\text{The matrices } \mathbf{A} := \mathbf{D}\mathbf{F}(\mathbf{p}) \text{ and } \mathbf{D} \text{ satisfy the assumptions of Theorem 5.1.} \quad (6.1)$$

In addition, we assume that there exists a solution $\mathbf{w}_0 = (u_1, \dots, u_k)$ of the reduced Q -Burgers equations which does not blow up on the interval $T \in [0, \mathcal{T}]$ and satisfies

$$\|\mathbf{w}_0(T)\|_{L_b^2(\mathbb{R})} := \sum_{j=1}^k \|u_j(T)\|_{L_b^2(\mathbb{R})} \leq K, \quad T \in [0, \mathcal{T}], \quad (6.2)$$

for some constant $K > 0$. Then, according to Proposition 4.5 and Corollary 4.1, the higher norms of \mathbf{w}_0 do not blow up as well and, in particular, if the initial data satisfies

$$\|\mathbf{w}_0|_{t=0}\|_{W_b^{2,2}(\mathbb{R})} \leq K, \quad (6.3)$$

we have

$$\|\mathbf{w}_0\|_{L^\infty([0,\mathcal{T}],W_b^{2,2})} + \|\partial_T \partial_X \mathbf{w}_0\|_{L_b^2([0,\mathcal{T}] \times \mathbb{R})} + \|\partial_X^3 \mathbf{w}_0\|_{L_b^2([0,\mathcal{T}] \times \mathbb{R})} \leq C, \quad (6.4)$$

where the constant C depends only on K and is independent of \mathcal{T} . Therefore, the corrector \mathbf{h} in equation (3.7) satisfies the estimate

$$\|\mathbf{h}\|_{L_b^2([0,\mathcal{T}] \times \mathbb{R})} \leq C, \quad (6.5)$$

where the constant C depends only on K and is independent of \mathcal{T} .

Finally, we assume that the initial data $U(X, 0)$ for equation (1.1) is ε^2 close to the initial data for the Q -Burgers approximation, namely,

$$\|\mathbf{R}|_{t=0}\|_{W_b^{1,2}(\mathbb{R})} \leq K\varepsilon, \quad (6.6)$$

where \mathbf{R} is defined via (3.3).

We are now ready to formulate the main result of the section.

Theorem 6.1 *Suppose assumptions (6.1), (6.2), (6.5) and (6.6) hold. Then, for every $K > 0$ and $\mathcal{T} > 0$, there exists $\varepsilon_0 = \varepsilon_0(K, \mathcal{T}) > 0$ such that, for all $\varepsilon < \varepsilon_0$, the trajectory $U(X, T, \varepsilon)$ of (1.1) remains $\varepsilon^{3/2}$ close to the corresponding trajectory of the Q -Burgers approximation on the time interval $T \in [0, \mathcal{T}]$, namely, the following estimate holds*

$$\|\mathbf{z}(T) - \mathbf{w}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{1/2}, \quad T \in [0, \mathcal{T}], \quad (6.7)$$

for all $\varepsilon \leq \varepsilon_0$. We remind that $\mathbf{z}(X, T) := \varepsilon^{-1}(U(X, T, \varepsilon) - \mathbf{p})$ and the approximation $\mathbf{w}(X, T)$ is built up based on the solution $\mathbf{w}_0 = (u_1, \dots, u_k)$ of the Q -Burgers equation via (3.4) and (2.5).

Remark 6.1 As we have seen in (6.4), the solution $\mathbf{w}_0(T)$ is uniformly bounded in the $W_b^{2,2}$ -norm as $T \in [0, \mathcal{T}]$. By this reason, the component $\mathbf{V}_0(X, T)$ is also uniformly bounded in the $W_b^{1,2}$ -norm and, consequently, $\varepsilon \mathbf{V}_0(X, T)$ is of order ε uniformly with respect to T . Thus, estimate (6.9) can be rewritten as follows:

$$\|\mathbf{z}(T) - \mathbf{w}_0(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{1/2}. \quad (6.8)$$

However, it is more convenient to include the term $\varepsilon \mathbf{V}_0(T)$ into the approximation $\mathbf{z}(X, T, \varepsilon)$.

Proof It is not difficult to see that, in order to prove the theorem, it is sufficient to verify the following conditional result: the estimate

$$\|\mathbf{z}(T) - \mathbf{w}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{1/2}e^{\kappa T}, \quad T \in [0, \mathcal{T}], \quad (6.9)$$

holds if

$$\|\mathbf{z}(T) - \mathbf{w}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq 1 \quad \text{for } T \leq \mathcal{T}. \quad (6.10)$$

Here C and κ are some constants depending on K , but they are independent of $\varepsilon > 0$. Thus, we only need to prove (6.9) under the assumption that (6.10) is valid.

Note that, due to (6.10) and (6.4)

$$\|\mathbf{z}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C, \quad (6.11)$$

and therefore, since $W^{1,2}$ is an algebra (in one space dimension)

$$\|\partial_X \Phi_\varepsilon(\mathbf{z}(T))\|_{L_b^2(\mathbb{R})} \leq \chi\left(\|\mathbf{z}(T)\|_{W_b^{1,2}(\mathbb{R})}\right) \leq C, \quad (6.12)$$

for some function $\chi(\cdot)$. Analogously,

$$\|\mathbf{D}^2 \mathbf{F}(\mathbf{p})[\mathbf{R}(T), 2\mathbf{w}(\mathbf{T}) + \mathbf{R}(T)]\|_{W_{\theta_{\varepsilon,s}}^{1,2}(\mathbb{R})} \leq C\|\mathbf{R}(T)\|_{W_{\theta_{\varepsilon,s}}^{1,2}(\mathbb{R})}, \quad (6.13)$$

where $\theta_{\varepsilon,s}(X) := e^{-\varepsilon|X-s|}$. Therefore, using (4.9) with $p = 2$,

$$\|\mathbf{H}(t)\|_{L_{\theta_{\varepsilon,s}}^2(\mathbb{R})} \leq C\varepsilon^{1/2} + C\|\mathbf{R}(T)\|_{W_{\theta_{\varepsilon,s}}^{1,2}(\mathbb{R})}, \quad (6.14)$$

where $\mathbf{H}(t)$ is defined in (3.10). Applying now estimate (5.22) to equation (3.9), and using estimate (6.14), we have

$$\|\mathbf{R}(T)\|_{W_{\theta\varepsilon,s}^{1,2}(\mathbb{R})}^2 \leq C\varepsilon e^{\kappa T} + C \int_0^T e^{\kappa(T-\tau)} \|\mathbf{R}(\tau)\|_{W_{\theta\varepsilon,s}^{1,2}(\mathbb{R})}^2, \quad (6.15)$$

where the constants C and κ are independent of T and ε . Applying the Gronwall inequality, we finally arrive at

$$\|\mathbf{R}(T)\|_{W_{\theta\varepsilon,s}^{1,2}(\mathbb{R})}^2 \leq C_1 \varepsilon e^{\kappa_1 T},$$

for some new uniform constants C_1 and κ_1 . Taking the supremum over $s \in \mathbb{R}$ from both sides of the last inequality, we deduce (6.9) and finish the proof of the theorem. \square

Note that Theorem 6.1 gives the long-time validity of the Q -Burgers approximation for the trajectories $U(X, T, \varepsilon)$ starting from the small neighborhood of a single smooth solution \mathbf{w}_0 of the Q -Burgers equation only. This result can be essentially improved if the global solvability of the Q -Burgers equations is known.

Corollary 6.1 *Let the assumption (6.1) hold and suppose the associated Q -Burgers system is globally solvable on any time interval $T \in [0, \mathcal{T}]$ and the following estimate holds:*

$$\|\mathbf{w}_0(T)\|_{L_b^2(\mathbb{R})} \leq \chi(T, \|\mathbf{w}_0(0)\|_{L_b^2(\mathbb{R})}), \quad (6.16)$$

for some monotone increasing (in both arguments) function χ . Then, for every $K > 0$ and $\mathcal{T} > 0$, there exists $\varepsilon_0 = \varepsilon_0(K, \mathcal{T})$ such that, for every $\varepsilon < \varepsilon_0$ and every initial data $U(X, 0, \varepsilon)$ satisfying

$$\|\mathbf{P}(U(\cdot, 0, \varepsilon) - \mathbf{p})\|_{W_b^{2,2}(\mathbb{R})} \leq K\varepsilon, \quad \|\mathbf{Q}(U(\cdot, 0, \varepsilon) - \mathbf{p})\|_{W_b^{1,2}(\mathbb{R})} \leq K\varepsilon^2, \quad (6.17)$$

the trajectory $U(X, T, \varepsilon)$ is $\varepsilon^{3/2}$ -close to the Q -Burgers approximation with the initial data

$$\mathbf{w}_0(X, 0, \varepsilon) = \varepsilon^{-1} \mathbf{P}(U(X, 0, \varepsilon) - \mathbf{p}),$$

namely,

$$\|\mathbf{z}(T) - \mathbf{w}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{1/2}, \quad (6.18)$$

for all $T \leq \mathcal{T}$.

Proof Indeed, due to the first assumption of (6.17), $\|\mathbf{w}_0(0)\|_{W_b^{2,2}} \leq K$. Then assumption (6.16) guarantees that the corresponding solution $\mathbf{w}_0(T)$ of the Q -Burgers system exists on the time interval $T \in [0, \mathcal{T}]$ and estimate (6.2) holds for some K which is independent on ε . Finally, due to the second inequality of (6.17), assumption (6.6) is also satisfied. Thus, applying Theorem 6.1, we get (6.18) and prove the corollary. \square

Remark 6.2 For the case of scalar Burgers equation ($k = 1$), assumption (6.16) is automatically satisfied due to the maximum principle. However, there is no maximum principle for the Q -Burgers equations in the vector case ($k > 1$) and we do not know whether or not the global solvability holds in general (some particular cases where (6.16) holds are considered in Sections 4 and 7). For this reason, we have to postulate the existence of a smooth solution \mathbf{w}_0 in Theorem 6.1.

7 Q -Burgers and gradient structure

When $k = 1$ the reduced equation is the classical scalar Burgers equation

$$u_T + \kappa u u_X = \nu u_{XX}, \quad (7.1)$$

about which much is known (e.g. BURGERS [9], BECK & WAYNE [3], DOELMAN ET AL. [11] and references therein). The canonical solution is the traveling front

$$u(X, T) = \frac{c}{\kappa} - \frac{c}{\kappa} \tanh\left(\frac{c}{2\nu}(X - cT)\right).$$

This front connects two constant states as $X \rightarrow \pm\infty$:

$$\lim_{X \rightarrow -\infty} u(X, T) = \frac{2c}{\kappa} \quad \text{and} \quad \lim_{X \rightarrow +\infty} u(X, T) = 0 \quad (T \text{ fixed}).$$

The maximum principle can be used to prove global existence of solutions of scalar Burgers in $L_b^2(\mathbb{R})$, and that result is a by-product of the more general proof in §4. In the context of shallow water hydrodynamics (e.g. [2, 15, 19, 28]), the travelling front is called a *hydraulic jump*.

When $k \geq 2$ the reduced equation has non-trivial coupling and so the analysis is more difficult. For general non-gradient quadratic nonlinearity a global existence proof for the reduced equations is elusive and we have to assume the non-linearity to be gradient in order to be able to establish such a global existence, see subsection §4.3.

However, in the case $k = 2$ there is a theorem of SCHAEFFER & SHEARER [26] which relates the gradient structure with the more natural (from the point of view of conservation laws) concept of hyperbolicity of \mathbf{Q} . We recall that a map $\mathbf{Q}(\mathbf{u})$ is *hyperbolic* if $D\mathbf{Q}(\mathbf{u})$ has real eigenvalues and is diagonalizable for all $\mathbf{u} \in \mathbb{R}^2$.

Lemma 7.1 (Schaeffer & Shearer [26]) *Suppose that $\mathbf{Q}(\mathbf{u})$ is a homogeneous quadratic and is hyperbolic. Then there exists a constant invertible linear transformation \mathbf{M} such that*

$$\mathbf{M}^{-1}\mathbf{Q}(\mathbf{M}\mathbf{v}) = \nabla\Psi(\mathbf{v}), \quad (7.2)$$

for some scalar potential Ψ .

This result is Lemma 3.2 in [26] and a direct proof can be given by noting that $D\mathbf{Q}(\mathbf{u})$ is linear in \mathbf{u} and then using linear algebra.

This transformation is useful since substitution of $\mathbf{u} = \mathbf{M}\mathbf{v}$ in the \mathbf{Q} -Burgers equation (1.10) and use of (7.2) gives

$$\frac{\partial\mathbf{v}}{\partial T} + \frac{\partial}{\partial X}(\nabla\Psi(\mathbf{v})) = \mathbf{E}\frac{\partial^2\mathbf{v}}{\partial X^2}, \quad (7.3)$$

where

$$\mathbf{E} = \mathbf{M}^{-1}\widehat{\mathbf{D}}\mathbf{M}. \quad (7.4)$$

With a gradient nonlinearity, we may apply Theorem 4.1 to the transformed equation (7.3) and to prove global existence of weak solutions if the new diffusion matrix (7.4) has a positive symmetric part. In particular, it will be always true if the initial diffusion matrix \mathbf{D} is a scalar multiple of the identity, $\mathbf{D} = \nu\mathbf{I}$.

For the case $k > 1$ in Q -Burgers (1.10) we may still look for traveling wave solutions $\mathbf{u}(X, T) = \widehat{\mathbf{u}}(X - cT)$. They satisfy

$$\widehat{\mathbf{D}}\widehat{\mathbf{u}}_X = \mathbf{Q}(\widehat{\mathbf{u}}) - c\widehat{\mathbf{u}} + \text{constant}. \quad (7.5)$$

In the case $k = 2$ it is possible to analyze solutions using phase plane techniques, and when the nonlinearity is gradient, it will be a planar gradient ODE. In the latter case, the canonical solution will then be a traveling front, but there is the potential for multiple fronts. For $k > 2$ the phase space will be of dimension 3 or greater and so existence of solutions of (7.5) will be more difficult, and may even be chaotic.

Note also that, to the best of our knowledge, the Schaeffer-Shearer transformation does not generalize to the case $k > 2$: there does not exist in general a constant matrix \mathbf{M} with the property (7.2). This is to be expected since the solutions of (7.5) for $k > 2$ can be much more complicated and are not gradient-like in general.

8 Geometry of the flux vector

To clarify the geometry of the flux vector, consider it as a mapping, $\mathbf{F} : V \rightarrow W$, between two vector spaces V and W of dimension n . Once the geometry is clear, we will revert to a mapping on \mathbb{R}^n .

Degeneracy of the form $\det(D\mathbf{F}(\mathbf{p})) = 0$ of a smooth mapping \mathbf{F} between manifolds is a well-studied problem in singularity theory (cf. Chapter 2 of ARNOLD, GUSEIN-ZADE & VARCHENKO [1], Chapter VI of GOLUBITSKY & GUILLEMIN [17]). The subsets of V where the Jacobian is singular are called the Thom-Boardman singularities of the mapping and are denoted in singularity theory by

$$\Sigma^k(\mathbf{F}) = \{ \mathbf{p} \in V : D\mathbf{F}(\mathbf{p}) \text{ has rank } n - k \}.$$

In this paper, attention will be restricted to the cases $k = 1$ and $k = 2$. There can however be higher order singularities on subsets of $\Sigma^1(\mathbf{F})$. The simplest such set is denoted by

$$\Sigma^{11}(\mathbf{F}) = \Sigma^1 \left(\mathbf{F} \Big|_{\Sigma^1(\mathbf{F})} \right).$$

When $\Sigma^1(\mathbf{F})$ is a manifold, then $\mathbf{F}\Big|_{\Sigma^1(\mathbf{F})}$ is a mapping between two manifolds, and $\Sigma^{11}(\mathbf{F})$ is the subset of $\Sigma^1(\mathbf{F})$ where the differential has a one-dimension kernel (an example is given in §9). Geometrically $\Sigma^{11}(\mathbf{F})$ is the set where the kernel of $D\mathbf{F}(\mathbf{p})$ lies in the tangent space of $\Sigma^1(\mathbf{F})$. This classification, $\Sigma^{ijk\dots}(\mathbf{F})$ can be continued until the dimension is exhausted [1,17]. In this paper, attention will be predominantly restricted to a study of the implications for $\mathbf{F}(\mathbf{p})$ of the hypersurfaces $\Sigma^1(\mathbf{F})$, $\Sigma^2(\mathbf{F})$ and the subset $\Sigma^{11}(\mathbf{F})$.

8.1 The geometry of $\Sigma^1(\mathbf{F})$

When $\mathbf{p} \in \Sigma^1(\mathbf{F})$ the kernel and cokernel of $D\mathbf{F}(\mathbf{p})$ are one dimensional. Define $\mathfrak{h} = \text{Ker}(D\mathbf{F}(\mathbf{p}))$ and $\mathfrak{h}^* = \text{Ker}(D\mathbf{F}(\mathbf{p})^T)$. Then with the assumption $p \notin \Sigma^{11}(\mathbf{F})$, decompose the vector spaces V and W as follows

$$V \cong T_p V = \mathfrak{h} \oplus T_p \Sigma^1(\mathbf{F}) \quad \text{and} \quad W \cong T_{F(p)} W = T_{F(p)} \mathbf{F}(\Sigma^1(\mathbf{F})) \oplus \mathfrak{h}^*,$$

where $\mathbf{F}(\Sigma^1(\mathbf{F}))$ is the image of $\Sigma^1(\mathbf{F})$ in W , which is locally a hypersurface in W . Introduce the mapping on $\mathfrak{h}^* \times \mathfrak{h}$,

$$s \mapsto \mathcal{K}(\mathbf{p}, s) := \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + s\boldsymbol{\xi}) \rangle_W, \quad \mathbf{p} \in \Sigma^1(\mathbf{F}), \quad \boldsymbol{\xi} \in \mathfrak{h}, \quad \boldsymbol{\eta} \in \mathfrak{h}^*, \quad (8.1)$$

where $\langle \cdot, \cdot \rangle_W$ is a pairing on W . With the definition (8.1),

$$\mathcal{K}_s(\mathbf{p}, s) = \langle \boldsymbol{\eta}, D\mathbf{F}(\mathbf{p} + s\boldsymbol{\xi})\boldsymbol{\xi} \rangle_W,$$

using the identification $W \cong T_{F(p)} W$, and so $\mathcal{K}_s(\mathbf{p}, 0) = 0$ when $\mathbf{p} \in \Sigma^1(\mathbf{F})$.

The graph of $\mathcal{K}(\mathbf{p}, s)$ in $\mathfrak{h}^* \times \mathfrak{h}$ is a planar curve. It is the curvature of the graph $(s, \mathcal{K}(\mathbf{p}, s))$ that appears as the coefficient of the nonlinearity in the reduced Burgers equation. The curvature of a planar curve at any s takes the standard form

$$\frac{\mathcal{K}_{ss}}{(1 + \mathcal{K}_s^2)^{3/2}}.$$

However at points with $\mathcal{K}_s = 0$, the denominator reduces to unity, making the second derivative itself invariant under coordinate change. This observation is a special case of the *intrinsic second derivative* of the mapping \mathbf{F} . The intrinsic second derivative is the second derivative evaluated on the kernel of the first derivative (cf. PORTEOUS [25], ARNOLD ET AL. [1], page 149 of GOLUBITSKY & GUILLEMIN [17]). It arises naturally in our construction of the reduced equation. The function $\mathcal{K}(\mathbf{p}, s)$ is quadratic in s for s small when $\mathbf{p} \in \Sigma^1(\mathbf{F})$,

$$\mathcal{K}(\mathbf{p}, s) = \mathcal{K}(\mathbf{p}, 0) + \frac{1}{2}s^2 \langle \boldsymbol{\eta}, D^2\mathbf{D}(\mathbf{p})(\boldsymbol{\xi}, \boldsymbol{\xi}) \rangle_W + \dots \quad (8.2)$$

The quadratic term is precisely the intrinsic second derivative of \mathbf{F} . An additional interesting fact is that $\mathbf{p} \in \Sigma^1(\mathbf{F}) \setminus \Sigma^{11}(\mathbf{F})$ if and only if the second intrinsic derivative is non-vanishing (cf. §3 of [7]),

$$\mathbf{p} \notin \Sigma^{11}(\mathbf{F}) \quad \Leftrightarrow \quad \left. \frac{d^2}{ds^2} \right|_{t=0} \mathcal{K}(\mathbf{p}, s) \neq 0, \quad \mathbf{p} \in \Sigma^1(\mathbf{F}). \quad (8.3)$$

8.2 An example of $\Sigma^1(\mathbf{F})$ in the case $n = 2$

Let V, W be two-dimensional vector spaces. Consider the following example of a flux vector $\mathbf{F} : V \rightarrow W$,

$$F_1(\mathbf{U}) = U_1 + \frac{1}{2}U_2^2 \quad \text{and} \quad F_2(\mathbf{U}) = U_1U_2, \quad (8.4)$$

with Jacobian

$$D\mathbf{F}(\mathbf{U}) = \begin{pmatrix} 1 & U_2 \\ U_2 & U_1 \end{pmatrix}.$$

This example is in fact a simplified version of the flux vector for the shallow water equations (see §9).

The Jacobian is degenerate when $U_2^2 = U_1$ and so

$$\Sigma^1(\mathbf{F}) = \{ \mathbf{p} \in V : p_2^2 - p_1 = 0 \}.$$

The Jacobian has rank 1 as long as $1 + p_1 \neq 0$. But $\mathbf{p} \in \Sigma^1(\mathbf{F})$ implies $p_1 \geq 0$ and so the rank can not drop to zero. Hence $\Sigma^2(\mathbf{F})$ is empty.

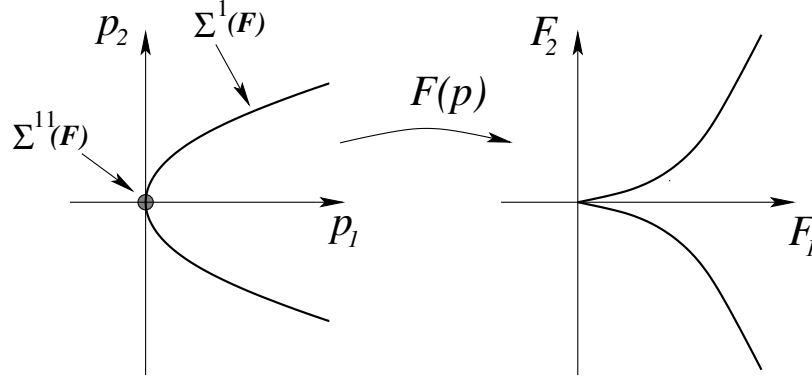


Fig. 8.1 The set $\Sigma^1(\mathbf{F})$ in V and its image under \mathbf{F} in W .

On the other hand, the set $\Sigma^{11}(\mathbf{F})$ is not empty. The tangent space of $\Sigma^1(\mathbf{F})$ is

$$T_p \Sigma^1(\mathbf{F}) = \left\{ \text{span} \begin{pmatrix} 2p_2 \\ 1 \end{pmatrix} \right\} \subset T_p V,$$

whereas,

$$\mathfrak{h} = \text{span}\{\boldsymbol{\xi}\}, \quad \boldsymbol{\xi} = \begin{pmatrix} -p_2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathfrak{h}^* = \text{span}\{\boldsymbol{\eta}\}, \quad \boldsymbol{\eta} = \frac{1}{1+p_1} \begin{pmatrix} -p_2 \\ 1 \end{pmatrix}.$$

The geometry of $\Sigma^1(\mathbf{F})$ and its mapping is illustrated in Figure 8.1. The left figure shows $\Sigma^1(\mathbf{F})$ as a subset of V , and the right figure shows the image of $\Sigma^1(\mathbf{F})$ under the mapping \mathbf{F} in the W -plane. The image is

$$\mathbf{F}(\Sigma^1(\mathbf{F})) = \left(p_1 + \frac{1}{2}p_2^2 \right) \Big|_{p \in \Sigma^1(\mathbf{F})} = \begin{pmatrix} \frac{3}{2}p_2^2 \\ p_2^3 \end{pmatrix}, \quad p_2 \in \mathbb{R},$$

which is a cusp in the W -plane.

When $p_2 \neq 0$ then \mathfrak{h} and $T_p \Sigma^1(\mathbf{F})$ together span V , and it fails when $p_2 = 0$. Hence

$$\Sigma^{11}(\mathbf{F}) = \{(p_1, 0)\} \cap \Sigma^1(\mathbf{F}),$$

which consists of just the origin in V .

The intrinsic second derivative is obtained as follows

$$\begin{aligned} \mathcal{H}(\mathbf{p}, s) &= \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + s\boldsymbol{\xi}) \rangle \\ &= \eta_1((p_1 + s\xi_1) + \frac{1}{2}(p_2 + s\xi_2)^2) + \eta_2(p_1 + s\xi_1)(p_2 + s\xi_2) \\ &= \kappa_0 + \kappa_1 s + \frac{1}{2}\kappa_2 s^2, \end{aligned} \tag{8.5}$$

with

$$\kappa_0 = -\frac{p_2^3}{2(1+p_1)}, \quad \kappa_1 = 0, \quad \kappa_2 = -\frac{3p_2}{1+p_1} \quad \text{when } \mathbf{p} \in \Sigma^1(\mathbf{F}).$$

Hence

$$\kappa = \frac{d^2}{ds^2} \Big|_{s=0} \mathcal{H}(\mathbf{p}, s) = \kappa_2 = -\frac{3p_2}{1+p_1}.$$

κ is clearly the intrinsic curvature of the plane curve in $\mathcal{H}(\mathbf{p}, s)$ in $\mathfrak{h}^* \times \mathfrak{h}$. As expected, the intrinsic second derivative κ vanishes if and only if $\mathbf{p} \in \Sigma^{11}(\mathbf{F})$.

8.3 Summary

Reverting back to $V = W = \mathbb{R}^n$, with inner product $\langle \cdot, \cdot \rangle$, the key geometric features of the flux vector that will be needed are (a) the kernel of $D\mathbf{F}(\mathbf{p})$, denoted by $\mathfrak{h} = \text{span}\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_k\}$, (b) the co-kernel of $D\mathbf{F}(\mathbf{p})$, denoted by $\mathfrak{h}^* = \text{span}\{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_k\}$, and (c) the intrinsic second derivative which is a tensor with entries

$$\Gamma_{ij}^l = \langle \boldsymbol{\eta}_l, D^2\mathbf{F}(\mathbf{p})(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) \rangle, \quad i, j, l = 1, \dots, k. \tag{8.6}$$

9 Example: the shallow water equations near criticality

An example is the shallow water equations with dissipative mollifier

$$\begin{aligned} h_t + (hv)_x &= \mu h_{xx} \\ v_t + \left(gh + \frac{1}{2}v^2\right)_x &= \mu v_{xx}. \end{aligned} \quad (9.1)$$

In these equations, $h(x, t)$ is the depth and $v(x, t)$ the horizontal velocity for a shallow water flow as shown schematically in Figure 9.1. The parameter $g > 0$ is the gravitational constant, and $\mu > 0$ is a positive dissipation parameter. In the shallow water equations (9.1) the state vector is $\mathbf{U} = (h, v)$ and

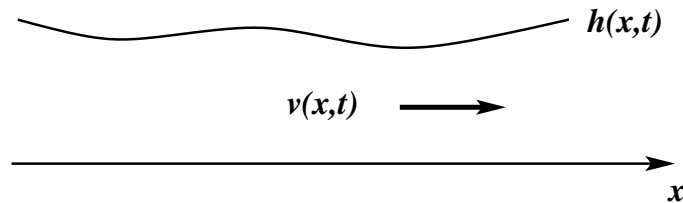


Fig. 9.1 Schematic of shallow water hydrodynamics.

the flux vector is

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} hv \\ gh + \frac{1}{2}v^2 \end{pmatrix}.$$

Constant states (h_0, v_0) are called uniform flows. The derivative of the flux vector evaluated on a uniform flow is

$$D\mathbf{F}(\mathbf{U}_0) = \begin{pmatrix} v_0 & h_0 \\ g & v_0 \end{pmatrix},$$

and criticality occurs when $\mathbf{p} = (h_0, v_0)$ satisfies

$$0 = \det(D\mathbf{F}(\mathbf{p})) = \det \begin{pmatrix} v_0 & h_0 \\ g & v_0 \end{pmatrix} = v_0^2 - gh_0 = -gh_0(1 - F^2),$$

where

$$F = \frac{v_0}{\sqrt{gh_0}},$$

is the *Froude number*. The singularity in this case is in $\Sigma^1(\mathbf{F})$. When $v_0 \neq 0$ this is the highest order singularity since both $\Sigma^{11}(\mathbf{F})$ and $\Sigma^2(\mathbf{F})$ are empty. Hence assume $v_0 \neq 0$. The singular set is a parabola in the (h, v) plane and the image of the singular set in the \mathbf{F} plane is a cusp, similar to Figure 8.1 (in Figure 8.1 the components of the flux vector are permuted and $g = 1$).

According to the theory in this paper, the reduced equation for parameter values near the criticality surface is a single Burgers equation

$$u_T + \kappa uu_X = \nu u_{XX},$$

with

$$\nu = \langle \boldsymbol{\eta}, D\boldsymbol{\xi} \rangle = \langle \boldsymbol{\eta}, \mu I \boldsymbol{\xi} \rangle = \mu \langle \boldsymbol{\eta}, \boldsymbol{\xi} \rangle = \mu.$$

To compute the curvature coefficient, explicit expressions for the eigenvectors are needed. The eigenvectors satisfy

$$D\mathbf{F}(\mathbf{p})\boldsymbol{\xi} = 0 \quad \text{and} \quad D\mathbf{F}(\mathbf{p})^T \boldsymbol{\eta} = 0 \quad \text{with} \quad \langle \boldsymbol{\eta}, \boldsymbol{\xi} \rangle = 1,$$

giving

$$\boldsymbol{\xi} = \begin{pmatrix} -v_0 \\ g \end{pmatrix} \quad \text{and} \quad \boldsymbol{\eta} = \frac{1}{2gv_0} \begin{pmatrix} -g \\ v_0 \end{pmatrix}.$$

Use (1.7) to compute κ ,

$$\kappa = \left. \frac{d^2}{ds^2} \right|_{s=0} \langle \boldsymbol{\eta}, \mathbf{F}(\mathbf{p} + s\boldsymbol{\xi}) \rangle = \frac{3}{2}g.$$

In this case the reduced equation captures the classical result of the generation of a travelling hydraulic jump due to criticality.

9.1 Singular dissipation matrix

The typical physical model for dissipation in the shallow water equations (e.g. [15,21]) is of the form

$$\begin{aligned} h_t + (hv)_x &= 0 \\ v_t + \left(gh + \frac{1}{2}v^2\right)_x &= \mu v_{xx}, \end{aligned} \quad (9.2)$$

that is, the dissipation acts only on the second equation. The dissipation matrix,

$$\mathbf{D} = \begin{pmatrix} 0 & 0 \\ 0 & \mu \end{pmatrix},$$

is singular. Formally, the derivation of Burgers equation goes through as before, and the reduced equation is still Burgers equation since

$$\nu = \langle \boldsymbol{\eta}, \mathbf{D}\boldsymbol{\xi} \rangle = \mu \left\langle \boldsymbol{\eta}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\xi} \right\rangle = \frac{1}{2}\mu.$$

Hence, the reduced Burgers equation is only slightly modified

$$u_T + \kappa uu_X = \frac{1}{2}\mu u_{XX}, \quad (9.3)$$

with the same κ .

9.2 Validity for singular dissipation matrix

Validity of the reduction from (9.2) to (9.3) does not follow from the general Theorem 6.1 of this paper, since a non-singular dissipation matrix is a hypothesis in (5.5) of §5. A validity result would depend on the dimension of the kernel of \mathbf{D} , and extra assumptions would be required in general.

On the other hand, to give some idea of the impact of singular dissipation, a validity proof for the reduction of (9.2) to (9.3) is given in this section. Surprisingly, the only additional assumption in this case is that the initial data for (9.3) requires one additional derivative. A key to the validity proof is that there is a natural energy for (9.2).

Seek a solution of (9.2) in the form

$$\begin{pmatrix} h(T, X) \\ v(T, X) \end{pmatrix} = \begin{pmatrix} h_0 \\ v_0 \end{pmatrix} + \varepsilon \begin{pmatrix} -v_0 \\ g \end{pmatrix} u(T, X) + \varepsilon^2 \begin{pmatrix} v_0 \\ g \end{pmatrix} \bar{v}(T, X) + \varepsilon \begin{pmatrix} H(T, X, \varepsilon) \\ V(T, X, \varepsilon) \end{pmatrix}, \quad (9.4)$$

where $v_0^2 = gh_0$, $\begin{pmatrix} v_0 \\ g \end{pmatrix}$ is the second eigenvector of the matrix $\mathbf{D}\mathbf{F}(\mathbf{p})$, u solves the Burgers equation (9.3), the corrector \bar{v} is found according to (2.5), namely,

$$\bar{v}(T, X) := \frac{g}{8v_0} u^2(T, X) + \frac{\mu}{4v_0} u_X(T, X), \quad (9.5)$$

and $\mathbf{R} := \begin{pmatrix} H \\ V \end{pmatrix}$ is the remainder to be estimated. A straightforward calculation shows that this remainder solves

$$\begin{aligned} \begin{pmatrix} H \\ V \end{pmatrix}_T + \frac{1}{\varepsilon} \begin{pmatrix} v_0 & h_0 \\ g & v_0 \end{pmatrix} \begin{pmatrix} H \\ V \end{pmatrix}_X + u \begin{pmatrix} g & -v_0 \\ 0 & g \end{pmatrix} \begin{pmatrix} H \\ V \end{pmatrix}_X + \varepsilon \bar{v} \begin{pmatrix} g & v_0 \\ 0 & g \end{pmatrix} \begin{pmatrix} H \\ V \end{pmatrix}_X \\ + \begin{pmatrix} HV \\ \frac{1}{2}V^2 \end{pmatrix}_X = \mu \begin{pmatrix} 0 \\ V \end{pmatrix}_{XX} + \varepsilon \begin{pmatrix} -v_0 \bar{v}_T - 2v_0 \varepsilon g \bar{v} \bar{v}_X \\ -g \bar{v}_T + \mu g \bar{v}_{XX} + g^2 (u \bar{v})_X - \varepsilon g^2 \bar{v} \bar{v}_X \end{pmatrix}. \end{aligned} \quad (9.6)$$

Assume now that the initial data $u(0)$ and $\mathbf{R}(0)$ for the Burgers equation (9.3) and the remainder equation (9.6) are chosen in such way that

$$\|u(0)\|_{W_b^{3,2}(\mathbb{R})} \leq K, \quad \|\mathbf{R}(0)\|_{W_b^{1,2}(\mathbb{R})} \leq K\varepsilon \quad (9.7)$$

for some K which is independent of ε . Then, due to the maximum principle for the scalar Burgers equation, it follows that

$$\|u(T)\|_{W_b^{3,2}(\mathbb{R})} + \|\bar{v}(T)\|_{W_b^{2,2}(\mathbb{R})} + \|\bar{v}_{TX}\|_{L_b^2([T, T+1] \times \mathbb{R})} + \|\bar{v}_{XX}\|_{L_b^2([T, T+1] \times \mathbb{R})} \leq K_1 \quad (9.8)$$

for some constant K_1 which depends only on K (and is independent of T).

The following result is the analogue of Theorem 6.1 for equations (9.6).

Theorem 9.1 *Let the above assumptions hold. Then, for every $K > 0$ and $\mathcal{T} > 0$ there exists $\varepsilon_0 = \varepsilon_0(K, \mathcal{T})$ such that, for every initial data satisfying (9.7) and all $\varepsilon \leq \varepsilon_0$, the corresponding solution of (9.6) exists on the time interval $T \in [0, \mathcal{T}]$ and satisfies the estimate*

$$\|\mathbf{R}(T)\|_{W_b^{1,2}(\mathbb{R})} \leq C\varepsilon^{1/2}, \quad T \in [0, \mathcal{T}], \quad (9.9)$$

where the constant C depends only on K and \mathcal{T} (and is independent of $\varepsilon \leq \varepsilon_0$).

Sketch of the proof Take the weight function $\theta_\varepsilon(X) := e^{-\varepsilon|X-X_0|}$, where $X_0 \in \mathbb{R}$ is a parameter and multiply the first and the second equations of (9.6) by $g\theta_\varepsilon^2 H - g(\theta_\varepsilon^2 H_X)_X$ and $h_0\theta_\varepsilon^2 V - h_0(\theta_\varepsilon^2 V_X)_X$ respectively. Then, using integration by parts as well as the estimate (4.15), we see that all terms of order ε^{-1} cancel out. Thus, using also (9.8) and arguing in a standard way, we end up with the following estimate:

$$\begin{aligned} \frac{1}{2} \frac{d}{dT} \left(g\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + h_0\|V\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right) + \mu h_0\|V\|_{W_{\theta_\varepsilon}^{2,2}(\mathbb{R})}^2 &\leq C \left(\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + \|V\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right) + \\ + C\varepsilon^2 \left(\|\bar{v}_T\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + \|\bar{v}_{XX}\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right) &- g((HV)_X, \theta_\varepsilon^2 H - (\theta_\varepsilon^2 H_X)_X) + h_0(VV_X, \theta_\varepsilon^2 V - (\theta_\varepsilon^2 V_X)_X), \end{aligned}$$

where the constant C depends on K , but is independent of ε . Thus, we only need to estimate the last two non-linear terms in the right-hand side of this inequality. We consider in more detail only the most complicated term containing H_{XX} , the other terms can be estimated in a similar way. Indeed, integrating by parts and using the embedding $W_{\theta_\varepsilon}^{1,2} \subset L_{\theta_\varepsilon}^\infty$, we get

$$\begin{aligned} g|((HV))_X, (\theta_\varepsilon^2 H_X)_X| &\leq \frac{3}{2}g|(H_X^2 \theta_\varepsilon^2, V_X)| + g|(HH_X, \theta_\varepsilon^2 V_{XX})| + \frac{1}{2}g|(H_X^2, (\theta_\varepsilon^2)_X V)| \leq \\ &\leq C\|V_X\|_{L_{\theta_\varepsilon}^\infty(\mathbb{R})}\|H_X\|_{L_b^2(\mathbb{R})}\|H_X\|_{L_{\theta_\varepsilon}^2(\mathbb{R})} + C\|H\|_{L^\infty(\mathbb{R})}\|H_X\|_{L_{\theta_\varepsilon}^2(\mathbb{R})}\|V_{XX}\|_{L_{\theta_\varepsilon}^2(\mathbb{R})} + \\ &+ C\|V\|_{L_{\theta_\varepsilon}^\infty(\mathbb{R})}\|H_X\|_{L_b^2(\mathbb{R})}\|H_X\|_{L_{\theta_\varepsilon}^2(\mathbb{R})} \leq C\|H\|_{W_b^{1,2}(\mathbb{R})}\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}\|V\|_{W_{\theta_\varepsilon}^{2,2}(\mathbb{R})} \leq \\ &\leq C_1\|H\|_{W_b^{1,2}(\mathbb{R})}^2\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + \mu h_0/4\|V\|_{W_{\theta_\varepsilon}^{2,2}(\mathbb{R})}^2. \quad (9.10) \end{aligned}$$

Inserting this estimate into the previous one, we end up with the following inequality:

$$\begin{aligned} \frac{1}{2} \frac{d}{dT} \left(g\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + h_0\|V\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right) &\leq \\ &\leq C \left(1 + \|H\|_{L_b^2(\mathbb{R})}^2 + \|V\|_{L_b^2(\mathbb{R})}^2 \right) \left(g\|H\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + h_0\|V\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right) + \\ &+ C\varepsilon^2 \left(\|\bar{v}_T\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + \|\bar{v}_{XX}\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \right), \quad (9.11) \end{aligned}$$

where the constant C depends on K , but is independent of ε . Estimate (9.11) implies the desired estimate (9.9) in a standard way. Indeed, if we assume for a while that

$$\|H(T)\|_{W_b^{1,2}(\mathbb{R})}^2 + \|V(T)\|_{W_b^{1,2}(\mathbb{R})}^2 \leq 1, \quad T \in [0, \mathcal{T}], \quad (9.12)$$

Then applying the Gronwall inequality to (9.11) and using (9.7) and (9.8) together with (4.9), we conclude that

$$\|H(T)\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 + \|V(T)\|_{W_{\theta_\varepsilon}^{1,2}(\mathbb{R})}^2 \leq C_2\varepsilon, \quad T \in [0, \mathcal{T}],$$

where C_2 depends on K and \mathcal{T} , but not on ε and the parameter $X_0 \in \mathbb{R}$ in the definition of the weight. Taking the supremum with respect to $X_0 \in \mathbb{R}$ and using (4.8), we see that

$$\|H(T)\|_{W_b^{1,2}(\mathbb{R})}^2 + \|V(T)\|_{W_b^{1,2}(\mathbb{R})}^2 \leq C_2\varepsilon, \quad T \in [0, \mathcal{T}].$$

Thus, assumption (9.12) is indeed satisfied if ε is small enough and the theorem is proved. \square

As we see, Theorem 9.1 gives almost a complete analogue of our general result of Theorem 6.1 for the case of partially dissipative system (9.2). The only difference is that we need a bit more smoothness of the solution u of the Q -Burgers equation ($u \in W_b^{3,2}(\mathbb{R})$ instead of $u \in W_b^{2,2}(\mathbb{R})$, see (6.3)) in order to compensate for the lack of dissipation in the h -equation.

Another feature of the proof is that there is a well-defined energy. Let $\mathbf{U} = \begin{pmatrix} h \\ v \end{pmatrix}$. Then (9.2) can be written in the form

$$\mathbf{U}_t + (\mathbf{M}\nabla E)_x = 0, \quad \text{with} \quad \mathbf{M} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

with energy

$$E(h, v) = \frac{1}{2}hv^2 + \frac{1}{2}gh^2,$$

and this energy plays a central role in the proof. General systems of this type, where the flux vector is the product of a symmetric matrix and the gradient of an energy, have been considered in general by BENJAMIN & BOWMAN [5]. For example, the case of three layers considered in the next section falls into this class of systems. When the energy is positive, systems of the type

$$\mathbf{U}_t + (\mathbf{M}\nabla E)_x = \mathbf{D}\mathbf{U}_{xx},$$

with \mathbf{D} singular, are candidates for extending the theory to more general systems with singular dissipation matrix. In addition the dimension of the kernel of \mathbf{D} will play a role.

10 Example: stratified three-layer shallow water equations

The shallow water equations with multiple layers of differing density are a rich source of singularities [2, 8, 28, 19]. Here we will consider the case of three layers since there are parameter values at which the Jacobian of the flux vector has a double zero eigenvalue with linearly independent eigenvectors. Let ρ_j , $h_j(x, t)$ and $v_j(x, t)$ be the density, depth and horizontal velocity in each layer. For stable stratification we require

$$\rho_1 > \rho_2 > \rho_3 > 0. \quad (10.1)$$

A schematic of the flow field is shown in Figure 10.1. The governing equations can be written in form of

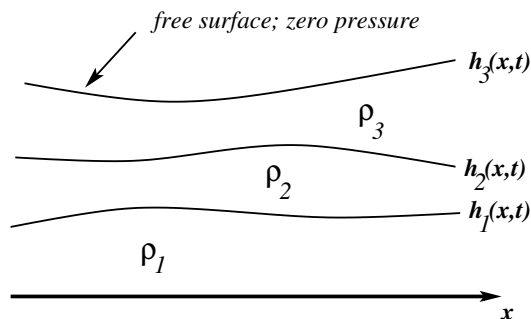


Fig. 10.1 A schematic of the flow configuration for three-layer shallow water theory.

a dissipative conservation law (1.1)

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{D}\mathbf{U}_{xx}, \quad (10.2)$$

with \mathbf{D} a constant matrix satisfying (1.2).

The state and flux vectors are

$$\mathbf{U} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{and} \quad \mathbf{F}(\mathbf{U}) = \begin{pmatrix} h_1 v_1 \\ h_2 v_2 \\ h_3 v_3 \\ \frac{1}{2}v_1^2 + gh_1 + r_2 gh_2 + r_3 gh_3 \\ \frac{1}{2}v_2^2 + gh_1 + gh_2 + \frac{r_3}{r_2} gh_3 \\ \frac{1}{2}v_3^2 + gh_1 + gh_2 + gh_3 \end{pmatrix}$$

where $r_2 = \rho_2/\rho_1$ and $r_3 = \rho_3/\rho_1$, and g is the gravitational constant. These equations are derived (for any number of layers) in §2 of BAINES [2]. These equations are for the case where the upper surface is free (zero pressure) and the pressure is hydrostatic in the other layers.

10.1 Jacobian of the flux vector

Let $\mathbf{U}_0 = (h_1, h_2, h_3, v_1, v_2, v_3)$ be a constant uniform state, and look at the derivative of the flux vector evaluated on this state,

$$\mathbf{DF}(\mathbf{U}_0) = \begin{bmatrix} v_1 & 0 & 0 & h_1 & 0 & 0 \\ 0 & v_2 & 0 & 0 & h_2 & 0 \\ 0 & 0 & v_3 & 0 & 0 & h_3 \\ g & gr_2 & gr_3 & v_1 & 0 & 0 \\ g & g & gr_3/r_2 & 0 & v_2 & 0 \\ g & g & g & 0 & 0 & v_3 \end{bmatrix}.$$

Criticality corresponds to the subset of uniform states, denoted by

$$\mathbf{p} = (h_1^0, h_2^0, h_3^0, v_1^0, v_2^0, v_3^0),$$

where

$$\det(\mathbf{DF}(\mathbf{p})) = 0.$$

Generically, the set of $\mathbf{p} \in \mathbb{R}^6$ satisfying this condition is a codimension one surface in \mathbb{R}^6 . Writing out this determinant gives

$$\det(\mathbf{DF}(\mathbf{p})) = F_2^2(r_3 - r_2(1 - F_3^2)) - (1 - r_2 - F_1^2) \left(\frac{r_3}{r_2} - (1 - F_2^2)(1 - F_3^2) \right), \quad (10.3)$$

where F_j corresponds to the local Froude number in each layer, e.g.,

$$F_j^2 = \frac{v_j^2}{gh_j}.$$

Setting the expression (10.3) to zero gives the criticality surface. However, it is not easy to illustrate this surface or to find the higher-order singularities as it is a hypersurface in \mathbb{R}^6 . Moreover we have to find the case where there are two independent eigenvectors. Hence it is simpler to reduce the eigenvalue and eigenvector problem together.

Suppose $\det(\mathbf{DF}(\mathbf{p})) = 0$ and let $\boldsymbol{\xi} \in \mathbb{R}^6$ be an eigenvector of $\mathbf{DF}(\mathbf{p})$

$$\mathbf{DF}(\mathbf{p})\boldsymbol{\xi} = 0. \quad (10.4)$$

For physical reasons, h_1^0 , h_2^0 and h_3^0 are never zero. Hence the first three equations of (10.4) give

$$\xi_{j+3} = -\frac{v_j^0}{h_j^0} \xi_j, \quad j = 1, 2, 3.$$

Substitution into the second three equations then gives a 3×3 system

$$\begin{bmatrix} 1 - F_1^2 & r_2 & r_3 \\ 1 & 1 - F_2^2 & \frac{r_3}{r_2} \\ 1 & 1 & 1 - F_3^2 \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0.$$

Now, ξ_2 can be eliminated from the third equation, giving the 2×2 system

$$\begin{bmatrix} F_2^2 & \frac{r_3}{r_2} - (1 - F_2^2)(1 - F_3^2) \\ 1 - r_2 - F_1^2 & r_3 - r_2(1 - F_3^2) \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_3 \end{pmatrix} = 0, \quad (10.5)$$

It is apparent that the matrix in (10.5) has a double zero eigenvalue if

$$F_2 = 0 \quad \text{and} \quad F_3^2 = 1 - \frac{r_3}{r_2}. \quad (10.6)$$

However, the 2×2 matrix in (10.5) has a non-trivial Jordan block and so the double zero eigenvalue is nonsemisimple. For it to be semisimple we require the additional condition

$$F_1^2 = 1 - r_2.$$

Hence when the following three conditions are satisfied the Jacobian has a double zero eigenvalue with two linearly independent eigenvectors

$$v_2 = 0, \quad \frac{v_1^2}{gh_1} = 1 - \frac{\rho_2}{\rho_1}, \quad \frac{v_3^2}{gh_3} = 1 - \frac{\rho_3}{\rho_2}. \quad (10.7)$$

Physically at these parameter values the middle layer is passive, and the upper and lower layers decouple and each is independently critical.

It is now straightforward to calculate the eigenvectors

$$\boldsymbol{\xi}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ -v_1^0/h_1^0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\xi}_2 = \begin{pmatrix} 0 \\ -\rho_3/\rho_2 \\ 1 \\ 0 \\ 0 \\ -v_3^0/h_3^0 \end{pmatrix},$$

and the adjoint eigenvectors, which satisfy

$$\mathbf{DF}(\mathbf{p})^T \boldsymbol{\eta} = 0,$$

are

$$\boldsymbol{\eta}_1 = \frac{h_1^0}{2v_1^0} \begin{pmatrix} v_1^0/h_1^0 \\ 0 \\ 0 \\ -1 \\ \rho_2/\rho_1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\eta}_2 = \frac{h_3^0}{2v_3^0} \begin{pmatrix} 0 \\ 0 \\ v_3^0/h_3^0 \\ 0 \\ 1 \\ -1 \end{pmatrix},$$

They have been normalised so that

$$\begin{pmatrix} \langle \boldsymbol{\eta}_1, \boldsymbol{\xi}_1 \rangle & \langle \boldsymbol{\eta}_1, \boldsymbol{\xi}_2 \rangle \\ \langle \boldsymbol{\eta}_2, \boldsymbol{\xi}_1 \rangle & \langle \boldsymbol{\eta}_2, \boldsymbol{\xi}_2 \rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (10.8)$$

Note that the assumptions $v_1^0 \neq 0$ and $v_2^0 \neq 0$ are not necessary since they follow from the criticality conditions (10.7) and the stratification condition (10.1).

10.2 Derivatives of the flux vector

The nonlinearity in the reduced system is

$$\begin{aligned} & \frac{\partial}{\partial x} \left(\frac{1}{2} \Gamma_{11}^1 u_1^2 + \Gamma_{12}^1 u_1 u_2 + \frac{1}{2} \Gamma_{22}^1 u_2^2 \right) \\ & \frac{\partial}{\partial x} \left(\frac{1}{2} \Gamma_{11}^2 u_1^2 + \Gamma_{12}^2 u_1 u_2 + \frac{1}{2} \Gamma_{22}^2 u_2^2 \right) \end{aligned} \quad (10.9)$$

where

$$\Gamma_{ij}^k = \left\langle \boldsymbol{\eta}_k, \frac{\partial^2}{\partial s_i \partial s_j} \mathbf{F}(\mathbf{p} + s_1 \boldsymbol{\xi}_1 + s_2 \boldsymbol{\xi}_2) \Big|_{s_1=s_2=0} \right\rangle.$$

Now, \mathbf{p} is the uniform state satisfying (10.7), and the eigenvectors are calculated above. Hence this is a straightforward computation. Only 2 out of 6 coefficients are nonzero in this example. Computing we find

$$\begin{aligned} \Gamma_{11}^1 &= -\frac{3}{2} \frac{v_1^0}{h_1^0}, & \Gamma_{12}^1 &= 0, & \Gamma_{22}^1 &= 0 \\ \Gamma_{11}^2 &= 0, & \Gamma_{12}^2 &= 0, & \Gamma_{22}^2 &= -\frac{3}{2} \frac{v_3^0}{h_3^0}, \end{aligned}$$

where v_1^0, v_3^0 and h_1^0, h_3^0 have to satisfy (10.7).

Hence the reduced system is

$$\begin{aligned}\frac{\partial u_1}{\partial T} - \frac{3}{2} \frac{v_1^0}{h_1^0} u_1 \frac{\partial u_1}{\partial X} &= \nu_{11} \frac{\partial^2 u_1}{\partial X^2} + \nu_{12} \frac{\partial^2 u_2}{\partial X^2} \\ \frac{\partial u_2}{\partial T} - \frac{3}{2} \frac{v_3^0}{h_3^0} u_2 \frac{\partial u_2}{\partial X} &= \nu_{21} \frac{\partial^2 u_1}{\partial X^2} + \nu_{22} \frac{\partial^2 u_2}{\partial X^2}.\end{aligned}\tag{10.10}$$

Steady solutions of the reduced system, relative to a frame moving at speed c , $u_j(X, T) = \hat{u}_j(X - cT)$, satisfy

$$\begin{aligned}\nu_{11} \hat{u}'_1 + \nu_{12} \hat{u}'_2 &= -3 \frac{v_1^0}{h_1^0} \hat{u}_1^2 - c \hat{u}_1 + A_1 \\ \nu_{21} \hat{u}'_1 + \nu_{22} \hat{u}'_2 &= -3 \frac{v_3^0}{h_3^0} \hat{u}_2^2 - c \hat{u}_2 + A_2,\end{aligned}$$

where A_1 and A_2 are arbitrary constants of integration. There are between 0 and 4 equilibrium points, and connecting orbits between the equilibria represent travelling fronts (or travelling hydraulic jumps) in the original system. In the context of three-layer fluids, these solutions model the birth of hydraulic jumps (see §4 of BAINES [2] for a discussion of hydraulic jumps in three-layer fluids).

11 Concluding remarks

The results in this paper are based on the uniformly local spaces $L_b^2(\mathbb{R})$. These results could be improved, at the expense of greater technicality, by going to the spaces $L_b^p(\mathbb{R})$ with p arbitrary.

11.1 Non-semisimple zero eigenvalues of \mathbf{DF}

We have assumed that the algebraic and geometric multiplicity of zero as an eigenvalue of \mathbf{DF} are equal. What happens when they are not equal? The simplest case of this type is when zero is an eigenvalue of \mathbf{DF} with algebraic multiplicity two but geometric multiplicity one. Indeed this case already arises in the example in §10 (see discussion just below equation (10.6)).

In the non-semisimple case the reduced modulation equation may no longer be Burgers or coupled Burgers. A new scaling is required which may change the nature of the reduced equation. Suppose $\mathbf{DF}(\mathbf{p})$ has a non-semisimple double zero eigenvalue with eigenvectors

$$\mathbf{DF}(\mathbf{p})\xi_1 = 0 \quad \text{and} \quad \mathbf{DF}(\mathbf{p})\xi_2 = \xi_1,$$

and left eigenvectors η_1 and η_2 and usual normalization. Introduce scaled variables

$$X = \varepsilon x \quad \text{and} \quad T = \varepsilon^{3/2} t,$$

and express \mathbf{U} in (1.1) in the form

$$\mathbf{U} = \mathbf{p} + \varepsilon u \xi_1 + \varepsilon^{3/2} v \xi_2 + \varepsilon^2 \mathbf{V}, \quad \text{with} \quad \langle \eta_j, \mathbf{V} \rangle = 0, \quad j = 1, 2.\tag{11.1}$$

Substitution into (1.1) and projection then leads to the following pair of equations at leading order

$$u_T + v_X = 0 \quad \text{and} \quad v_T + \Gamma_{11}^2 u u_X = \nu_{21} u_{XX},\tag{11.2}$$

where $\nu_{ij} = \langle \eta_i, \mathbf{D}\xi_j \rangle$ and Γ_{ij}^k is defined as in (8.6). This reduced equation is formally asymptotically correct, but it is ill posed even locally unless $\Gamma_{11}^2 = \nu_{21} = 0$.

Try a different scaling

$$X = \varepsilon^{1/2} x \quad \text{and} \quad T = \varepsilon t,$$

but express \mathbf{U} in the same form (11.1). Substitution into (1.1) and projection then leads to the following pair of equations at leading order

$$u_T + v_X = \nu_{11} u_{XX} \quad \text{and} \quad v_T + \Gamma_{11}^2 u u_X = \frac{1}{\varepsilon^{1/2}} \nu_{21} u_{XX} + \nu_{22} v_{XX}.\tag{11.3}$$

This equation looks worse since there is a singular term in the limit $\varepsilon \rightarrow 0$. However, a calculation shows that if $\nu_{21} \neq 0$ then the equilibrium \mathbf{U}_0 of (1.1) is exponentially unstable and there is no reason to expect

that a solution started from an ε -neighborhood of \mathbf{U}_0 will be close to it for a long time ($T \sim \frac{1}{\varepsilon}$). This explains, in particular, why both reductions described above lead to "bad" equations if $\nu_{21} \neq 0$.

Hence the assumption $\nu_{21} = 0$ is necessary for the validity of the modulation equations derived above. With this assumption the reduced modulation equation is *locally* well posed when $\nu_{11} > 0$ and $\nu_{22} > 0$ and is in the form of \mathbf{Q} -Burgers equations (see (1.10)) with

$$\mathbf{Q} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ \frac{1}{2} \Gamma_{11}^2 u^2 \end{pmatrix} \text{ and } \widehat{\mathbf{D}} = \begin{pmatrix} \nu_{11} & 0 \\ 0 & \nu_{22} \end{pmatrix}.$$

However, this nonlinearity is not gradient and we cannot guarantee the *global* well posedness using Theorem 4.1. In fact, this global well posedness *fails* and there are solutions which blow up in finite time. Indeed, system (11.3) (with $\nu_{21} = 0$) is equivalent to the so-called strongly damped Boussinesq equation

$$u_{TT} - (\nu_{11} + \nu_{22})u_{TXX} + (\nu_{11}\nu_{22})u_{XXX} = \left(\frac{1}{2}\Gamma_{11}^2 u^2\right)_{XX}, \quad (11.4)$$

for which the finite-time blow up is known (see Theorem 4 in LEVINE [20]). In summary, although we believe that the strongly damped Boussinesq equation (11.4) is a correct modulation equation for the non-semisimple case, its validity does not follow directly from the theory developed in this paper and therefore remains an open question.

11.2 Criticality relative to a moving frame

The results in the paper have been developed relative to a *stationary frame of reference* (a laboratory frame). If we allow for moving frames of reference then criticality and the appearance of Burgers equation becomes even more pervasive. Introduce the shifted coordinate $x \mapsto x - ct$ for some constant c (which may be positive or negative) into (1.1),

$$\mathbf{U}_t + (\mathbf{F}(\mathbf{U}) - c\mathbf{U})_x = \mathbf{D}\mathbf{U}_{xx}, \quad x \in \mathbb{R}, \quad \mathbf{U} \in \mathbb{R}^n. \quad (11.5)$$

The flux vector is just shifted. Constant states $\mathbf{U}_0 \in \mathbb{R}^n$ are still solutions but the derivative of the flux vector is modified to $\mathbf{D}\mathbf{F}(\mathbf{U}_0) - c\mathbf{I}$ and so the criticality condition is

$$\det[\mathbf{D}\mathbf{F}(\mathbf{U}_0) - c\mathbf{I}] = 0. \quad (11.6)$$

The following result is an immediate consequence of the theory in this paper: given *any* constant state $\mathbf{U}_0 \in \mathbb{R}^n$ there exists at least one frame of reference relative to which it is critical. Indeed there may exist as many as n frames of reference at which the basic state \mathbf{U}_0 is critical, since the critical speeds are just the eigenvalues of $\mathbf{D}\mathbf{F}(\mathbf{U}_0)$. Fixing c at one of the semisimple eigenvalues of $\mathbf{D}\mathbf{F}(\mathbf{U}_0)$, the theory of this paper goes through *mutatis mutandis*, showing that the dynamics is modeled by a Burgers equation relative to the selected moving frame. Mathematically, this result is clear, but the physical implications are not, and are an intriguing direction for further study.

Acknowledgements The authors are grateful to Michael Shearer (North Carolina State) for pointing out the Schaeffer-Shearer transformation, and to Roger Grimshaw (Loughborough) for helpful discussions on the three-layer example in §10. The work of JP was supported by a grant from EPSRC for PhD studies, and the work of TB was partially supported by a Leverhulme Trust Fellowship.

References

1. Arnol'd, V.I., Gusein-Zade, S.M., & Varchenko, A.N. *Singularities of Differentiable Maps, Volume I*, Birkhäuser: Boston (1985).
2. Baines, P.G. *A general method for determining upstream effects in stratified flow of finite depth over long two-dimensional obstacles*, J. Fluid Mech. **188** 1–22 (1988).
3. Beck, M. & Wayne, C.E. *Invariant manifolds and the stability of traveling waves in scalar viscous conservation laws*, J. Diff. Eqns. **244** 87–116 (2008).
4. Beck, M. & Wayne, C.E. *Using global invariant manifolds to understand metastability in the Burgers equation with small viscosity*, SIAM J. Appl. Dyn. Syst. **8** 1043–1065 (2009).
5. Benjamin, T.B. & Bowman, S. *Discontinuous Solutions of One-Dimensional Hamiltonian Systems*, Proc. Roy. Soc. London A **413** 263–295 (1987).
6. Bernoff, A.J. *Slowly varying fully nonlinear wavetrains in the Ginzburg-Landau equation*, Physica D **30** 363–381 (1988).
7. Bridges, T.J. *Degenerate relative equilibria, curvature of the momentum map and homoclinic bifurcation*, J. Diff. Eqns. **244** 1629–1674 (2008).

8. Bridges, T.J. & Donaldson, N.M. *Criticality manifolds and their role in the generation of solitary waves for two-layer flow with a free surface*, Euro. J. Mech. B/Fluids **28** 117–126 (2009).
9. Burgers, J.M. *The Nonlinear Diffusion Equation*, D. Reidel Publishing Company: Dordrecht (1974)
10. Crighton, D. *Model equation of nonlinear acoustics*, Ann. Rev. Fluid Mech. **11** 11–33 (1979)
11. Doelman, A., Sandstede, B., Scheel, A. & Schneider, G. *The Dynamics of Modulated Wave Trains*, Memoirs AMS **199**, AMS: Providence (2009).
12. Efendiev, M.A. & Zelik, S.V. *The attractor for a nonlinear reaction-diffusion system in an unbounded domain*, Comm. Pure Appl. Math. **54** 625–688 (2001).
13. Efendiev, M.A. & Zelik, S.V. *Upper and lower bounds for the Kolmogorov entropy of the attractor for an RDE in an unbounded domain*, J. Dyn. Diff. Eqns **14** 369–403 (2002).
14. Esipov, S.E. *Coupled Burgers equations: a model of polydispersive sedimentation*, Phys. Rev. E **52** 3711–3718 (1995).
15. Frankcombe, L.M. & Hogg, A. *Tidal modulation of two-layer hydraulic exchange flows*, Ocean Sci. **3** 179–188 (2007).
16. H. FREISTÜHLER & P. SZMOLYAN. *Spectral stability of small shock waves*, Arch. Rat. Mech. Anal. **164** 287–309 (2002).
17. Golubitsky, M. & Guillemin, V. *Stable Mappings and Their Singularities*, Springer-Verlag: New York (1973).
18. Hewitt, R.E. & Hall, P. *The evolution of finite-amplitude wavetrains in plane channel flow*, Phil. Trans. Royal Soc. London A **356** 2413–2446 (1998).
19. Lane-Serff, G.F., Smeed, D.A. & Postlethwaite, C.R. *Multi-layer hydraulic exchange flows*, J. Fluid Mech. **416** 269–296 (2000).
20. Levine, H.A. *Some additional remarks on the non-existence of global solutions to nonlinear wave equations*, SIAM J. Math. Anal. **5** 138–146 (1974).
21. Marche, F. *Derivation of a new two-dimensional viscous shallow-water model with varying topography, bottom friction and capillary effects*, Euro. J. Mech. B/Fluids **26** 49–63 (2007).
22. Mielke, A. & Schneider, G. *Attractors for modulation equations on unbounded domains – existence and comparison*, Nonlinearity **8** 743–768 (1995).
23. Miranville, A. & Zelik, S. *Attractors for dissipative partial differential equations in bounded and unbounded domains*, Handbook of Differential Equations **4** 103–200 (2008).
24. Plaza, R. & Zumbrun, K. *An Evans function approach to spectral stability of small-amplitude shock profiles*, Discrete Contin. Dyn. Syst. **10** 885–924 (2004).
25. Porteous, I.R. *Simple singularities of maps*, in Proc. Liverpool Singularities Symposium, pp. 286–307. Lecture Notes in Math. **192** Springer-Verlag: Berlin (1971).
26. Schaeffer, D.G. & Shearer, M. *The classification of 2×2 systems of non-strictly hyperbolic conservation laws, with applications to oil recovery*, Comm. Pure Appl. Math. **40** 141–178 (1987).
27. Szepessy, A. & Zumbrun, K. *Stability of rarefaction waves in viscous media*, Arch. Rat. Mech. Anal. **133** 249–298 (1996).
28. Smeed, D.A. *Hydraulic control of three-layer exchange flows: application to the Bab al Mandab*, J. Phys. Ocean. **30** 2574–2588 (2000).
29. Whitham, G. *Linear and Nonlinear Waves*, Wiley Interscience: New York (1974).
30. Zelik, S.V. *Attractors of reaction-diffusion systems in unbounded domains and their spatial complexity*, Comm. Pure Appl. Math. **56** 584–637 (2003).
31. Zelik, S.V. *Spatially nondecaying solutions of 2D Navier-Stokes equations in a strip*, Glasgow Math Jour., **49**, no. 3, 525–588 (2007)