# CONVERGENCE OF LINEARISED AND ADJOINT APPROXIMATIONS FOR DISCONTINUOUS SOLUTIONS OF CONSERVATION LAWS. PART 1: LINEARISED APPROXIMATIONS AND LINEARISED OUTPUT FUNCTIONALS

MIKE GILES* AND STEFAN ULBRICH†

**Abstract.** This paper analyses the convergence of discrete approximations to the linearised equations arising from an unsteady one-dimensional hyperbolic equation with a convex flux function. A simple modified Lax-Friedrichs discretisation is used on a uniform grid, and a key point is that the numerical smoothing increases the number of points across the nonlinear discontinuity as the grid is refined. It is proved that this gives pointwise convergence almost everywhere for the solution to the linearised discrete equations with smooth initial data, and also convergence in the discrete approximation of linearised output functionals. In Part 2 [GU09] we will extend the results to Dirac initial data for the linearised equation and will prove the pointwise convergence almost everywhere for the solution of the adjoint discrete equations. Moreover, we will present numerical results.

## 1. Introduction.

**1.1. Background.** In recent years there has been considerable research in the computational fluid dynamics community into the development and use of adjoint equations for design optimisation (e.g. [Jam95, AV99]), data assimilation (e.g. [CT87, TC87]), and error analysis (e.g. [BR01, BD02]). In almost every case, the adjoint equations have been formulated under the assumption that the original nonlinear flow solution is smooth. Since most applications have been for incompressible or subsonic flow, this has been valid, however there is also considerable interest in transonic design applications for which there are shocks. The correct formulation and discretisation of adjoint equations in the presence of shocks is therefore important, and that is the main motivation for the analysis in this paper.

The reason that shocks present a problem is that adjoint equations are defined to be adjoint to the equations obtained by linearising the original nonlinear equations. Therefore, this raises the issue of linearised perturbations to the shock. A correct treatment of the inviscid analytic equations must linearise the shock jump equations which arise from conservation properties at the shock. However, for numerical approximations which rely on shock capturing, as opposed to shock fitting, it is not clear whether linearised shock capturing will yield the correct results.

The validity of linearised shock capturing for the particular application of shocks oscillating harmonically in flutter analysis was investigated by Lindquist and Giles [LG94]. Their numerical results demonstrated that shock capturing produces the correct prediction of integral quantities such as unsteady lift and moment provided the shock is spread over a number of grid points. It was argued, but not proved, that this is because the "viscous" shock profile remains invariant, to leading order, as the shock oscillates, and therefore the integral effect of the linearised shock motion is correct. As a result, linearised shock capturing is now the standard method of turbomachinery aeroelastic analysis [HCL94, SW98], even though there has been no proof of convergence of the numerical discretisation.

There has been very little prior research into adjoint equations for flows with shocks. Giles and Pierce [GP01] have shown that the analytic derivation of the adjoint

*Professor, Oxford University Mathematical Institute, mike.giles@maths.ox.ac.uk
†Professor, Dept. of Mathematics, TU Darmstadt, ulbrich@mathematik.tu-darmstadt.de

equations for the steady quasi-one-dimensional Euler equations requires the specification of an internal adjoint boundary condition at the shock. However, the numerical evidence [GP98] suggests that for this steady one-dimensional application the correct adjoint solution is obtained using either the "discrete" approach (in which one linearises the discrete equations and then forms their adjoint) or the "continuous" approach (in which one discretises the analytic adjoint equations). In the case of the discrete approach, this is due to the second order accuracy of conservative quasi-one-dimensional shock capturing [Gil96], whereas with the continuous approach it is thought to be because the use of numerical smoothing automatically selects the correct numerical solution which is smooth at the shock [GP98]. Homescu and Navon [HN03] and Bardos and Pironneau [BP03] have also addressed the correct formulation and approximation of adjoint equations in flows with shocks.

Ulbrich has derived the adjoint equations for 1D conservation laws with source terms, and using the method of generalised characteristics he has analysed the differentiability of objective functionals with respect to controls [Ulb02, Ulb03]. In these papers he proved that the correct formulation leads to an interior boundary condition for the adjoint equations along discontinuities in the original nonlinear solutions, which is automatically satisfied for so-called reversible solutions of the adjoint equation in the sense of [BJ98].

Looking at nonlinear hyperbolic equations with scalar conservation laws, there has been considerable prior research into the convergence of numerical approximations to the nonlinear equations. Mentioning just a few key papers, Crandall and Majda [CM80] proved convergence to the unique entropy solution for monotone difference approximations of scalar conservation laws, and Nessyahu and Tadmor [NT92] proved an optimal bound on the order of convergence for a certain class of numerical methods, using the *Lip'* norm introduced by Tadmor [Tad91].

Linearised conservation laws with discontinuous coefficients have been analysed in [LeF90, BJ98]. The appropriate definition of measure valued solutions, as well as their existence and uniqueness, have been considered in [LeF90]. The definition of the measure valued solutions in [LeF90] is based on the averaged superposition of Volpert for functions of bounded variation to define the non-conservative product in the flux term. In this context, Volpert's definition is appropriate; a detailed study of more general definitions of non-conservative products is carried out in [DMLM95]. The solutions in [LeF90] coincide with the duality solutions considered in [BJ98] for general linear conservation laws with discontinuous coefficients satisfying a one-sided Lipschitz condition. Moreover, [BJ98] also introduces reversible solutions for the backward problem of the corresponding non-conservative transport equation with discontinuous coefficients.

Giles has obtained numerical results for Burgers equation [Gil03] showing that numerical discretisations of the adjoint equations can converge to incorrect solutions as the grid is refined uniformly, unless the number of grid points across the discontinuity increases as the grid is refined. Using the discrete linearised scheme and the discrete adjoint scheme ensures automatically that the conservative linearised equations are discretised by a conservative scheme and the non-conservative adjoint equations by a non-conservative scheme. This avoids approximation errors which are studied in detail by [HL94].

It is the desire for a theoretical understanding of these numerical results which has motivated the research in this present paper and in Part 2 [GU09]. We consider an unsteady one-dimensional hyperbolic equation with a convex scalar flux, such as Burgers equation. We use a specific explicit discretisation with a modified Lax-Friedrichs flux with a smoothing coefficient which varies with grid resolution $h$ to increase the number of grid points across any discontinuity as $h \to 0$. Because the numerical discretisation is monotone, for sufficiently small timesteps, the classical results of Crandall and Majda [CM80], as well as the more recent results of Nessyahu and Tadmor [NT92], prove convergence of the nonlinear discretisation to the unique entropy solution. In this paper we prove that for initial data which is smooth apart from one or more discontinuities the corresponding linearised discretisation yields solutions which converge pointwise to the analytic solution everywhere except along the discontinuities. Furthermore, it is proved that the discrete approximation of the linearised perturbation to output integrals converges to the analytic value. From this we will deduce in Part 2 [GU09] that the discrete adjoint approximation must converge to the analytic adjoint solution as $h \to 0$, everywhere except along two characteristics across which it is discontinuous. Furthermore, by an appropriate choice of the numerical smoothing, the order of convergence is $O(h^\alpha)$, for any $\alpha < 1$.

**1.2. The model problem.** The model problem is the equation

$$N(u) \equiv \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \qquad -\infty < x < \infty, \quad 0 < t < T, \tag{1.1}$$

in which $f(u)$ is a $C^\infty$ convex flux function with derivative $a(u)$.

Numerical results will be presented in Part 2 for initial data $u_0(x)$ which is continuous and leads to the formation of a single shock after a finite time. However, the numerical analysis in this paper will be performed for initial data with a single initial discontinuity at $x = 0$, and with all derivatives in $(-\infty, 0)$ and $(0, \infty)$ having a finite $L_1$ norm. This condition implies that $u_0(x)$ has bounded variation, and hence $u(x, t)$ is bounded. An extension to much more general initial data will be done in Part 2 [GU09].

The shock moves with a velocity given by the Rankine-Hugoniot jump relation

$$\dot{x}_s[u] - [f] = 0, \tag{1.2}$$

where $[u] \equiv u(x_s^+, t) - u(x_s^-, t)$ denotes the jump in $u$ across the shock.

Linear perturbations to the inviscid solution due to perturbed initial conditions are governed by the linear p.d.e.

$$L(u)\, \widetilde{u} \equiv \frac{\partial \widetilde{u}}{\partial t} + \frac{\partial}{\partial x}\left(a\, \widetilde{u}\right) = 0, \tag{1.3}$$

The corresponding linearised perturbation to the shock position satisfies the o.d.e.

$$\dot{\widetilde{x}}_s \left[u\right] + \dot{x}_s \left[\widetilde{u} + \widetilde{x}_s \frac{\partial u}{\partial x}\right] - \left[a\, \widetilde{u} + \widetilde{x}_s \frac{\partial f}{\partial x}\right] = 0.$$

Eliminating $\partial f / \partial x$ using Equation (1.1), and noting that

$$\frac{\mathrm{d}}{\mathrm{d}t}[u] = \left[\frac{\partial u}{\partial t} + \dot{x}_s \frac{\partial u}{\partial x}\right],$$

one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\widetilde{x}_s[u]\right) = \left[(a - \dot{x}_s)\, \widetilde{u}\right], \tag{1.4}$$

subject to initial data $\widetilde{x}_s(0)=0$.

By integrating Equations (1.3) and (1.4) it can be verified that if the initial data $\widetilde{u}_0(x)$ has compact support then the quantity

$$\int_{-\infty}^{x_s(t)} \widetilde{u}(x,t)\ \mathrm{d}x\ +\ \int_{x_s(t)}^{\infty} \widetilde{u}(x,t)\ \mathrm{d}x\ -\ \widetilde{x}_s(t)\,[u]_t$$

is invariant in time. Here $[u]_t$ represents the jump in $u(x,t)$ across the shock at time $t$. At the final time $T$, if all of the characteristics from the compact support of $\widetilde{u}_0(x)$ have entered the shock, and so $\widetilde{u}(x,T)$ is zero on either side of the shock, it follows that

$$-\ \widetilde{x}_s(T)\ \ [u]_T = \int_{-\infty}^{\infty} \widetilde{u}_0(x)\ \mathrm{d}x. \tag{1.5}$$

Thus the final shock perturbation is proportional to the integral of the initial solution perturbation.

The linearised perturbations $\widetilde{u}$ and $\widetilde{x}_s$ can be used to analyse the linearised perturbations of output functionals. The output functional of interest is a tracking-type functional of the form

$$J = \int_{-\infty}^{\infty} \gamma(x)\ G(u(x,T))\,\mathrm{d}x = \int_{-\infty}^{x_s(T)} \gamma(x)\ G(u(x,T))\,\mathrm{d}x + \int_{x_s(T)}^{\infty} \gamma(x)\ G(u(x,T))\,\mathrm{d}x,$$

where $\gamma$ is a weighting function with compact support and bounded variation. The corresponding linear perturbation is

$$\begin{aligned}\widetilde{J} = &\int_{-\infty}^{x_s(T)} \gamma(x)\ G'(u(x,T))\ \widetilde{u}(x,T)\,\mathrm{d}x\ +\ \int_{x_s(T)}^{\infty} \gamma(x)\ G'(u(x,T))\ \widetilde{u}(x,T)\,\mathrm{d}x \\ &-\ \widetilde{x}_s(T)\,\gamma(x_s(T))\ [G]_T\end{aligned}$$

where $[G]_T$ represents the jump in $G(u(x,T))$ across the shock at the final time.

It can be shown [Ulb02, Ulb03, Gil03] that the adjoint formulation of the linearised functional perturbation is

$$\widetilde{J} = \int_0^1 w(x,0)\ \widetilde{u}(x,0)\,\mathrm{d}x,$$

where $w(x,t)$ satisfies the adjoint p.d.e.

$$\frac{\partial w}{\partial t} + a\frac{\partial w}{\partial x} = 0,$$

in the smooth regions on either side of the shock, with 'initial' conditions at the final time,

$$w(x,T) = \gamma(x)\ G'(x,T)$$

and along the shock the interior boundary condition

$$w(x_s(t),t) = \gamma(x_s(T))\ [G]_T/[u]_T.$$

Note that it follows from this that the adjoint solution has the uniform constant value $[G]_T/[u]_T$ on all characteristics leading into the shock.

The central objective in this paper is to prove that under certain conditions a linearised discretisation of a regularised p.d.e. yields an approximation to the linear inviscid solution which is convergent pointwise everywhere except at the shock, and also gives a convergent approximation to the linearised functional $\widetilde{J}$.

In Part 2 [GU09] we will show that from this it will then follow that the solution to the corresponding adjoint discretisation converges almost everywhere to the inviscid adjoint solution.

**1.3. Numerical discretisations.** The nonlinear equation is approximated on a mesh with uniform spacing $h$ and timestep $k$, by the finite difference equation

$$k\ N_j(U_j^n) \equiv U_j^{n+1} - U_j^n + \tfrac{1}{2}\, r \left( f(U_{j+1}^n) - f(U_{j-1}^n) \right) - \varepsilon\, d \left( U_{j+1}^n - 2U_j^n + U_{j-1}^n \right) = 0, \tag{1.6}$$

with initial data $U_j^0 = u_0(x_j)$, where

$$r \equiv \frac{k}{h}, \quad d \equiv \frac{k}{h^2}, \quad F_j^n \equiv f(U_j^n)$$

and $\varepsilon = h^\alpha$ for some constant $0 < \alpha < 1$. The inequality $\alpha > 0$ ensures that the discretisation is consistent when the analytic solution $u(x,t)$ is smooth. The inequality $\alpha < 1$ ensures that the shock is spread over an increasing number of grid points as $h \to 0$. It will be proved that stability and monotonicity is achieved for sufficiently small $h$ by choosing $k$ such that $\varepsilon\, d = c$, for some positive constant $c < \frac{1}{2}$.

Linearising the nonlinear discretisation yields the following approximation of the linear p.d.e.,

$$\begin{aligned} &k\ L_j(U_j^n)\ \widetilde{U}_j^n \\ &\equiv \widetilde{U}_j^{n+1} - \widetilde{U}_j^n + \tfrac{1}{2} r \left( a(U_{j+1}^n)\, \widetilde{U}_{j+1}^n - a(U_{j-1}^n)\, \widetilde{U}_{j-1}^n \right) - \varepsilon\, d \left( \widetilde{U}_{j+1}^n - 2\widetilde{U}_j^n + \widetilde{U}_{j-1}^n \right) \\ &= 0, \end{aligned} \tag{1.7}$$

to be solved subject to initial data $\widetilde{U}_j^0 = \widetilde{u}_0(x_j)$. Again, this is a consistent approximation if both $u(x,t)$ and $\widetilde{u}(x,t)$ are smooth. Note also that if $\widetilde{u}_0(x)$ has compact support then summing over $j$ yields the result that

$$\sum_{j=-\infty}^{\infty} \widetilde{U}_j^n = \text{const.} \tag{1.8}$$

If the nonlinear tracking-type functional is approximated by trapezoidal integration, then the corresponding discretisation of the linearised functional is

$$\widetilde{J}_h = h \sum_j \gamma(x_j)\ G'(U_j^N)\ \widetilde{U}_j^N,$$

where $Nk = T$. The exactly equivalent adjoint formulation for this is

$$\widetilde{J}_h = h \sum_j W_j^0\ \widetilde{U}_j^0,$$

where $W_j^n$ satisfies the discrete adjoint equation [Gil03]

$$W_j^{n-1} = W_j^n + \tfrac{1}{2}\, r a(U_j^{n-1}) \left( W_{j+1}^n - W_{j-1}^n \right) + \varepsilon\, d \left( W_{j+1}^n - 2W_j^n + W_{j-1}^n \right),$$

with 'initial' conditions

$$W_j^N = \gamma(x_j)\; G'(U_j^N).$$

This adjoint formulation follows immediately from the identity

$$\sum_j W_j^n\; \widetilde{U}_j^n = \sum_j W_j^{n+1}\; \widetilde{U}_j^{n+1}$$

which is easily verified. In this paper we will prove that $\widetilde{J}_h \to \widetilde{J}$ as $h \to 0$ for smooth linear initial perturbations $\widetilde{u}_0$.

In the particular case of Dirac initial data for the linear discretisation,

$$\widetilde{U}_j^0 = \begin{cases} h^{-1}, & j = J \\ 0, & \text{otherwise} \end{cases}$$

then $\widetilde{J}_h = W_J^0$. Thus, the adjoint solution at a particular point is equal to the linear functional arising from Dirac initial data for the linearised equations at that same point. By proving in Part 2 [GU09] that the linearised functional converges to the correct value also for Dirac initial data, we will also be proving that the adjoint approximation converges to the analytic solution.

**1.4. Outline of paper.** This first paper is devoted to the proof of the convergence of the discrete linear functional $\widetilde{J}_h$ to the analytic value $\widetilde{J}$ as $h \to 0$. The convergence is analysed by using the technique of matched inner and outer asymptotic expansions [BO78, KC81] to construct approximations to both $U_j^N$ and $\widetilde{U}_j^n$. Discrete stability estimates are used to bound the errors in the asymptotic approximations.

- Section 2 derives the stability estimates which are used later in Section 4 to bound the errors in the asymptotic approximations.
- Section 3 derives the asymptotic form of the discrete approximation of a viscous travelling wave on a uniform grid, and then re-scales this to obtain the asymptotic form of the discrete approximation to a moving shock with uniform conditions on either side.
- Section 4 uses the moving shock approximation to form blended inner/outer asymptotic approximations of both the nonlinear and linear discrete solutions for a particular choice of discrete initial data for the nonlinear and linearised equations. Together with the stability estimates, this proves the convergence of the linear solution away from the shock.
- Section 5 completes the main analysis by bounding the error in the linearised discrete functional approximation.

Part 2 [GU09] extends the analysis to linear problems with Dirac initial data, which implies the convergence of the discrete adjoint solution. Moreover, the results of this paper are extended to more general nonlinear initial data.

## 2. Discrete stability estimates.

**2.1. Nonlinear equations.** THEOREM 2.1. *Suppose that $U_j^n$ is a solution of the equation*

$$U_j^{n+1} = U_j^n - \tfrac{1}{2}\, r\left(f(U_{j+1}^n) - f(U_{j-1}^n)\right) + \varepsilon\, d\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right),$$

*where $f(u)$ is a $C^\infty$ function with derivative $a(u) = f'(u)$, and $r = k/h$, $d = k/h^2$, $\varepsilon = h^\alpha$, $0 < \alpha < 1$, and subject to specified initial data $U_j^0$ with $L_\infty$ bound $U_\infty$.*

*Furthermore, let $V_j^n$ be an approximation to $U_j^n$ which satisfies the equation*

$$V_j^{n+1} = V_j^n - \tfrac{1}{2}\, r \left( f(V_{j+1}^n) - f(V_{j-1}^n) \right) + \varepsilon\, d \left( V_{j+1}^n - 2V_j^n + V_{j-1}^n \right) + k\, \tau_j^n,$$

*and the same initial data $U_j^0$, and let $U_\infty$ also be an upper bound on $\|V_j^n\|_\infty$.*

*Then, provided that $h < (2/A_\infty)^{1/(1-\alpha)}$ where $A_\infty = \sup_{|u|<U_\infty} |a(u)|$, and $k$ is chosen so that $\varepsilon d = c$ for some constant $c < \frac{1}{2}$, then $E_j^n = V_j^n - U_j^n$ satisfies the bound*

$$\|E^n\|_1 \;\leq\; \|\tau\|_{1,n},$$

*where the $l_1$ norm is defined as $\|E^n\|_1 \;\equiv h\; \sum_j |E_j^n|$ and $\|\tau\|_{1,n} = k \sum_{m=1}^n \|\tau^m\|_1$.*

*Proof.* The conditions on $h$ and $k$ ensure that $\frac{1}{2} r A_\infty < c$. Hence, the equations for $U_j^n$ are monotone and therefore $|U_j^n| \leq U_\infty$ and $|a(U_j^n)| \leq A_\infty$, for all $j, n$.

Defining

$$A_j^n = \begin{cases} \dfrac{f(V_j^n) - f(U_j^n)}{V_j^n - U_j^n}, & V_j^n \neq U_j^n \\[2ex] a(U_j^n), & V_j^n = U_j^n \end{cases}$$

the difference $E_j^n = V_j^n - U_j^n$ satisfies the equation

$$E_j^{n+1} = E_j^n - \tfrac{1}{2}\, r \left( A_{j+1}^n E_{j+1}^n - A_{j-1}^n E_{j-1}^n \right) + c \left( E_{j+1}^n - 2E_j^n + E_{j-1}^n \right) + k\, \tau_j^n,$$

with homogeneous initial data. $A_j^n$ satisfies the bound $|A_j^n| \leq A_\infty$, and hence $\frac{1}{2}\, r |A_j^n| < c, \;\; \forall j, n$.

If $\|\tau^n\|_1$ is finite, then taking the absolute magnitude and summing over the entire interval gives

$$\sum_j h\, |E_j^{n+1}| \leq \sum_j h \left( |E_j^n| + k |\tau_j^n| \right),$$

and hence $\|E^n\|_1 \leq \|\tau\|_{1,n}$. $\square$

**2.2. Linearised equations.** THEOREM 2.2. *Given $U_j^n$ and $V_j^n$ as defined in Theorem 2.1, let $\widetilde{U}_j^n$ be a solution of the linearised difference equation*

$$\widetilde{U}_j^{n+1} = \widetilde{U}_j^n - \tfrac{1}{2}\, r \left( a(U_{j+1}^n)\; \widetilde{U}_{j+1}^n - a(U_{j-1}^n)\; \widetilde{U}_{j-1}^n \right) + \varepsilon\, d \left( \widetilde{U}_{j+1}^n - 2\widetilde{U}_j^n + \widetilde{U}_{j-1}^n \right),$$

*and let $\widetilde{V}_j^n$ be an approximation to it which satisfies the equation*

$$\widetilde{V}_j^{n+1} = \widetilde{V}_j^n - \tfrac{1}{2}\, r \left( a(V_{j+1}^n)\; \widetilde{V}_{j+1}^n - a(V_{j-1}^n)\; \widetilde{V}_{j-1}^n \right) + \varepsilon\, d \left( \widetilde{V}_{j+1}^n - 2\widetilde{V}_j^n + \widetilde{V}_{j-1}^n \right) + k\, \widetilde{\tau}_j^n,$$

*with initial data $\widetilde{V}_j^0$ which may differ from $\widetilde{U}_j^0$.*

*Then, provided $h$ and $k$ satisfy the same conditions as in Theorem 2.1, the difference $\widetilde{E}_j^n = \widetilde{V}_j^n - \widetilde{U}_j^n$ satisfies the bound*

$$\|\widetilde{E}^n\|_1 \;\leq\; \|\widetilde{E}^0\|_1 + h^{-1} t^n\; B_\infty \widetilde{V}_\infty \|\tau\|_{1,n} + \|\widetilde{\tau}\|_{1,n},$$

*where $B_\infty = \sup_{|u|<U_\infty} |a'(u)|$, and $\widetilde{V}_\infty$ is an upper bound for $|\widetilde{V}_j^n|$.*

*Proof.* Defining

$$B_j^n = \begin{cases} \dfrac{a(V_j^n) - a(U_j^n)}{V_j^n - U_j^n}, & V_j^n \neq U_j^n \\[2ex] a'(U_j^n), & V_j^n = U_j^n \end{cases}$$

with bound $|B_j^n| < B_\infty$, the difference $\widetilde{E}_j^n = \widetilde{V}_j^n - \widetilde{U}_j^n$ satisfies the equation

$$\begin{aligned} \widetilde{E}_j^{n+1} = \widetilde{E}_j^n - \tfrac{1}{2}\, r \left( a(U_{j+1}^n)\ \widetilde{E}_{j+1}^n - a(U_{j-1}^n)\ \widetilde{E}_{j-1}^n \right) + \varepsilon\, d \left( \widetilde{E}_{j+1}^n - 2\widetilde{E}_j^n + \widetilde{E}_{j-1}^n \right) \\ - \tfrac{1}{2}\, r \left( B_{j+1}^n E_{j+1}^n \widetilde{V}_{j+1}^n - B_{j-1}^n E_{j-1}^n \widetilde{V}_{j-1}^n \right) + k\, \widetilde{\tau}_j^n. \end{aligned}$$

Taking the absolute magnitude, using the triangle inequality and then summing over the entire interval gives

$$\sum_j h\, |\widetilde{E}_j^{n+1}| \leq \sum_j h \left( |\widetilde{E}_j^n| + r B_\infty |\widetilde{V}_j^n|\ |E_j^n| + k |\widetilde{\tau}_j^n| \right).$$

Hence,

$$\|\widetilde{E}^{n+1}\|_1 \leq \|\widetilde{E}^n\|_1 + r\ \ B_\infty \widetilde{V}_\infty \|\tau\|_{1,n} + k\, \|\widetilde{\tau}^n\|_1,$$

and therefore

$$\|\widetilde{E}^n\|_1 \ \leq\ \|\widetilde{E}^0\|_1 + h^{-1} t^n\ B_\infty \widetilde{V}_\infty \|\tau\|_{1,n} + \|\widetilde{\tau}\|_{1,n}.$$

□

## 3. Analytic and discrete travelling wave solutions.

**3.1. Viscous travelling wave.** We begin with a theorem which establishes the existence of a unique travelling wave solution to the viscous convection/diffusion equation with unit viscosity and a convex flux function.

THEOREM 3.1. *Given the viscous equation*

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = \frac{\partial^2 u}{\partial x^2}, \tag{3.1}$$

*with a $C^\infty$ convex flux function $f(u)$ with derivative $a(u)$, then for any values of the constants $\dot{x}_s$ and $\Delta u$ satisfying the inequalities $\sup_u a(u) > \dot{x}_s > \inf_u a(u),\ \ \Delta u < 0$, there exists a unique travelling wave solution of the form*

$$u(x,t) = s(\dot{x}_s, \Delta u;\, x - \dot{x}_s t),$$

*with the properties that as $x^* \to \infty$, $s(-x^*) \to s_{-\infty}$, $s(x^*) \to s_\infty = s_{-\infty} + \Delta u$ for some constant $s_{-\infty}$, and*

$$\int_{-x^*}^{x^*} s(x)\ dx \to x^*(s_{-\infty} + s_\infty).$$

*Furthermore, all derivatives of $s(x)$ decay exponentially as $|x| \to \infty$.*

*Proof.* Substituting the travelling wave ansatz into the partial differential equation gives the equation

$$-\dot{x}_s \frac{\mathrm{d}s}{\mathrm{d}x} + \frac{\mathrm{d}f(s)}{\mathrm{d}x} = \frac{\mathrm{d}^2 s}{\mathrm{d}x^2}. \tag{3.2}$$

Integrating this yields

$$f(s) - \dot{x}_s s - \frac{\mathrm{d}s}{\mathrm{d}x} = C$$

for some constant $C$. Given particular values for $\dot{x}_s$ and $\Delta u$ satisfying the specified inequalities, because of the convexity of $f(u)$ there exists a unique value $s_{-\infty}$ such that

$$\dot{x}_s = \frac{f(s_{-\infty} + \Delta u) - f(s_{-\infty})}{\Delta u}. \tag{3.3}$$

Defining $s_\infty = s_{-\infty} + \Delta u$, and setting $C = f(s_\infty) - \dot{x}_s s_\infty$, gives $s(x)$ defined implicitly by

$$\int_{s_{-\infty}+\frac{1}{2}\Delta u}^{s} \frac{\mathrm{d}u}{f(u) - \dot{x}_s u - C} = x - x_0.$$

The quantity $f(u) - \dot{x}_s u - C$ is strictly negative for $s_\infty < u < s_{-\infty}$, because of the convexity of $f(u)$, and approaches zero linearly as $u \to s_\infty$ or $u \to s_{-\infty}$. Hence, all derivatives of $s(x)$ decay exponentially as $|x| \to \infty$.

Finally, the unique value of $x_0$ is determined by the requirement that as $x^* \to \infty$,

$$\int_{-x^*}^{x^*} s(x)\,\mathrm{d}x \to x^*(s_{-\infty} + s_\infty).$$

□

Next, we consider linear perturbations to Equation (3.2), under the influence of a source term $g(x)$, giving the equation

$$-\dot{x}_s \frac{\mathrm{d}\widetilde{s}}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(a(s)\,\widetilde{s}\right) = \frac{\mathrm{d}^2 \widetilde{s}}{\mathrm{d}x^2} + g(x). \tag{3.4}$$

Theorem 3.2. *When $g(x) = 0$, all solutions of Equation (3.4), subject to the boundary conditions $\widetilde{s}(x) \to 0$ as $|x| \to \infty$, are of the form*

$$\widetilde{s} = c\,\frac{ds}{dx},$$

*for some constant $c$.*

*When $g(x)$ is not identically zero, but $g(x)$ and its derivatives all decay exponentially as $|x| \to \infty$, there exist non-unique solutions of Equation (3.4), subject to the boundary conditions $\widetilde{s}(x) \to \widetilde{s}_\infty$ and $\widetilde{s}(-x) \to \widetilde{s}_{-\infty}$ as $x \to \infty$, iff $g(x)$ satisfies the solvability condition*

$$\int_{-\infty}^{\infty} g(\xi)\,d\xi = (a(s_\infty) - \dot{x}_s)\,\widetilde{s}_\infty - (a(s_{-\infty}) - \dot{x}_s)\,\widetilde{s}_{-\infty}.$$

*Furthermore, the derivatives of $\widetilde{s}(x)$ all decay exponentially as $|x| \to \infty$.*

*Proof.* Integrating Equation (3.4) with $g(x) = 0$ gives

$$(a(s) - \dot{x}_s)\ \widetilde{s} = \frac{\mathrm{d}\widetilde{s}}{\mathrm{d}x} + \text{const.}$$

Because $a(s_\infty) - \dot{x}_s < 0$ and $a(s_{-\infty}) - \dot{x}_s > 0$, the boundary conditions are satisfied provided the integration constant is zero.

This equation can then be integrated subject to an arbitrary value of $\widetilde{s}(0)$ to give a one-parameter family of homogeneous solutions which satisfy the boundary conditions as $|x| \to \infty$.

Differentiating Equation (3.2) establishes that $\mathrm{d}s/\mathrm{d}x$ satisfies the homogeneous version of Equation (3.4) and the specified boundary conditions. Therefore, $\mathrm{d}s/\mathrm{d}x$ is a basis for the one-parameter family of homogeneous solutions with the specified boundary conditions.

When $g(x)$ is not identically zero, the solvability condition arises immediately from integrating Equation (3.4) and applying the boundary conditions.

The exponential decay in the derivatives of $\widetilde{s}(x)$ follows from the exponential decay in $g(x)$ and its derivatives, and also the exponential decay in the derivatives of $s(x)$. □

**3.2. Discrete travelling wave with unit viscosity.** The objective in this section is to derive an approximation to the discrete travelling wave solution which arises when approximating Equation (3.1) using the nonlinear discretisation

$$k\ N_j(U_j^n) \equiv U_j^{n+1} - U_j^n + \tfrac{1}{2}\, r \left(f(U_{j+1}^n) - f(U_{j-1}^n)\right) - d \left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right) = 0,$$

where $r = k/h$, and $d = k/h^2$ is held fixed as $h \to 0$.

THEOREM 3.3. *For any values of the constants $\dot{x}_s$ and $\Delta u$ satisfying the conditions of Theorem 3.1, there exists a sequence of functions $c_n(x)$, with $c_n(x) \to 0$, as $|x| \to \infty$, such that for all integer $M \geq 0$ the function $S_M(x)$ defined by*

$$S_M(x) = s(x) + \sum_{n=1}^{M} h^{2n} c_n(x)$$

*has the properties that*

$$N_j\left(S_M(x_j - \dot{x}_s t^n)\right) = o(h^{2M}),$$

*and as $x^* \to \infty$, $S_M(-x^*) \to s_{-\infty}$, $S_M(x^*) \to s_\infty = s_{-\infty} + \Delta u$, and*

$$\int_{-x^*}^{x^*} S_M(x)\ dx \to x^*(s_{-\infty} + s_\infty).$$

*Furthermore, all derivatives of $S_M(x)$ decay exponentially as $|x| \to \infty$.*

*Proof.* The proof is by induction. Suppose that for a given $M \geq 0$ it is true that there exist functions $c_n(x)$, for $n \leq M$, such that $S_M(x)$ has the specified properties.

A truncated Taylor series expansion of $f(u)$ gives

$$f(u) = f(u_0) + \sum_{n=1}^{M+1} \frac{(u - u_0)^n}{n!}\ \left.\frac{\mathrm{d}^n f}{\mathrm{d}u^n}\right|_{u_0} + \frac{(u - u_0)^{M+2}}{(M+2)!}\ f_{M+2}(u_0, u),$$

where

$$f_{M+2}(u_0,u) = \left(\int_{u_0}^{u} (u-\xi)^{M+1}\ \mathrm{d}\xi\right)^{-1} \quad \int_{u_0}^{u} (u-\xi)^{M+1}\ \left.\frac{\mathrm{d}^{M+2} f}{\mathrm{d}u^{M+2}}\right|_{\xi}\ \mathrm{d}\xi,$$

is a weighted average value of $\mathrm{d}^{M+2}f/\mathrm{d}u^{M+2}$ on the interval $[u_0, u]$.

Using this expansion with $u_0 = U_j^n = S_M(x_j - \dot{x}_s t^n)$ and $u = U_{j\pm1}^n$, and then making similar truncated Taylor series expansions for $U_{j\pm1}^n$ and $U_j^{n+1}$ (with $k = dh^2$), one finds that the residual error has an expansion in even powers of $h^2$ of the form

$$N_j\left(S_M(x_j - \dot{x}_s t^n)\right) = \sum_{n=1}^{M+1} h^{2n}\ \left.\frac{\mathrm{d}r_n}{\mathrm{d}x}\right|_{x_j - \dot{x}_s t^n} + o(h^{2M+2}),$$

where $r_n(x) \to 0$, as $|x| \to \infty$. However, by the inductive hypothesis, this residual error is $o(h^{2M})$. Therefore,

$$N_j\left(S_M(x_j - \dot{x}_s t^n)\right) = h^{2M+2}\ \left.\frac{\mathrm{d}r_{M+1}}{\mathrm{d}x}\right|_{x_j - \dot{x}_s t^n} + o(h^{2M+2}).$$

Now, let $c_{M+1}(x - \dot{x}_s t)$ be defined by the linear differential equation

$$L(s(x - \dot{x}_s t))\ c_{M+1}(x - \dot{x}_s t) = -\left.\frac{\mathrm{d}r_{M+1}}{\mathrm{d}x}\right|_{x - \dot{x}_s t}, \tag{3.5}$$

giving the ODE

$$-\dot{x}_s \frac{\mathrm{d}c_{M+1}}{\mathrm{d}x} + \frac{\mathrm{d}}{\mathrm{d}x}\Big(a(s(x))\ c_{M+1}\Big) - \frac{\mathrm{d}^2 c_{M+1}}{\mathrm{d}x^2} = -\frac{\mathrm{d}r_{M+1}}{\mathrm{d}x},$$

subject to the boundary conditions $c_{M+1}(x) \to 0$, as $|x| \to \infty$ Because the right-hand-side of this equation satisfies the necessary solvability condition of Theorem 3.2, there exists a solution $c_{M+1}(x)$, and it is unique if we impose the additional constraint that

$$\int_{-\infty}^{\infty} c_{M+1}(x)\ \mathrm{d}x = 0.$$

Also, due to Theorem 3.2, all derivatives of $c_{M+1}$ decay exponentially as $|x| \to \infty$.

Finally, we obtain

$$\begin{aligned}
&N_j\left(S_M(x_j - \dot{x}_s t^n) + h^{2M+2} c_{M+1}(x_j - \dot{x}_s t^n)\right) \\
&\quad = N_j\left(S_M(x_j - \dot{x}_s t^n)\right) + h^{2M+2} L_j(U_j^n)\ c_{M+1}(x_j - \dot{x}_s t^n) + o(h^{2M+2}) \\
&\quad = N_j\left(S_M(x_j - \dot{x}_s t^n)\right) + h^{2M+2}\ L_{s(x - \dot{x}_s t)}\ c_{M+1}\big|_{x_j - \dot{x}_s t^n} + o(h^{2M+2}) \\
&\quad = o(h^{2M+2}),
\end{aligned}$$

and

$$\int_{-x^*}^{x^*} S_M(x) + h^{2M+2} c_{M+1}(x)\ \mathrm{d}x \longrightarrow \int_{-x^*}^{x^*} S_M(x)\ \mathrm{d}x, \quad \text{as } x^* \to \infty,$$

so the inductive hypothesis is true for $M+1$. The hypothesis is trivially true for the initial value $M=0$, concluding the proof. □

Although the above proof has used the shorthand $S_M(x)$, we should more properly express it as $S_M(d,h,\dot{x}_s,\Delta u;x)$ to make clear its dependence on the parameters $d$, $h$, $\dot{x}_s$ and $\Delta u$.

The next theorem establishes that $S'_M(d,h,\dot{x}_s,\Delta u;x)$ is an approximate solution of the linearised discrete equations,

$$\begin{aligned} &k\ L_j(U_j^n)\ \widetilde{U}_j^n \\ &\equiv \widetilde{U}_j^{n+1}-\widetilde{U}_j^n + \tfrac{1}{2}r\left(a(U_{j+1}^n)\,\widetilde{U}_{j+1}^n - a(U_{j-1}^n)\,\widetilde{U}_{j-1}^n\right) - d\left(\widetilde{U}_{j+1}^n-2\widetilde{U}_j^n+\widetilde{U}_{j-1}^n\right) = 0, \end{aligned}$$

which are an approximation of the linear differential equation

$$L_u\,\widetilde{u} \equiv \frac{\partial \widetilde{u}}{\partial t} + \frac{\partial}{\partial x}\left(a(u)\,\widetilde{u}\right) - \frac{\partial^2 \widetilde{u}}{\partial x^2} = 0.$$

THEOREM 3.4. *If $S_M(x)$ is as defined in Theorem 3.3, then*

$$L_j(S_M(x_j-\dot{x}_s t^n+X))\ S'_M(x_j-\dot{x}_s t^n+X) = o(h^{2M}).$$

*Proof.* Using truncated Taylor series expansions, the residual $N_j\left(S_M(x_j-\dot{x}_s t^n+X)\right)$, can be expressed as a sum of a finite number of terms, each one of the form $e_n h^{2n}$ for $n\geq M+1$, with the coefficients $e_n$ being products of derivatives of $s(x)$ either evaluated at $x_j-\dot{x}_s t^n+X$ or averaged over a small interval in the neighbourhood of this point, or derivatives of $f(u)$ evaluated at $s(x_j-\dot{x}_s t^n+X)$ or averaged over a small interval in its neighbourhood.

Differentiation with respect to $X$ does not introduce any new powers of $h$, and therefore

$$\frac{\partial}{\partial X}\ N_j\Big(S_M(x_j-\dot{x}_s t^n+X)\Big) \ = \ L_j(S_M(x_j-\dot{x}_s t^n+X))\ S'_M(x_j-\dot{x}_s t^n+X)) \ = \ o(h^{2M}).$$

∎

**3.3. Mesh-dependent viscosity.** Switching to the numerical discretisation with $\varepsilon=h^\alpha$ and $\varepsilon d$ fixed as $h\to 0$, we come to the key result which will be used in the general asymptotic analysis.

THEOREM 3.5. *For constants $\dot{x}_s$ and $\Delta u$ satisfying the inequalities in Theorem 3.1, and with the discrete operators $N_j$ and $L_j$ as defined in Equations (1.6) and (1.7), there exists a function $S(x)$ (which also depends on the parameters $\varepsilon\,d, \varepsilon^{-1}h, \dot{x}_s, \Delta u$) such that*

$$N_j\left(S(\varepsilon^{-1}(x_j-\dot{x}_s t^n))\right) = o(\varepsilon), \tag{3.6}$$

*and*

$$L_j\left(S(\varepsilon^{-1}(x_j-\dot{x}_s t^n))\right)\ S'(\varepsilon^{-1}(x_j-\dot{x}_s t^n))) = o(\varepsilon), \tag{3.7}$$

*and as $x^*\to\infty$, $S(-x^*)\to s_{-\infty}$, $S(x^*)\to s_\infty = s_{-\infty}+\Delta u$, and*

$$\int_{-x^*}^{x^*} S(x)\ dx \to x^*(s_{-\infty}+s_\infty).$$

*Proof.* The key is to note that the finite difference equation with $\varepsilon = h^\alpha$ is identical to that for unit viscosity if we make the substitutions $h = \varepsilon\, h_{unit},\ k = \varepsilon\, k_{unit}$.

Hence, $S_M(\varepsilon\, d, \varepsilon^{-1}h, \dot{x}_s, \Delta u; \varepsilon^{-1}(x - \dot{x}_s t))$, is an approximation to the discrete solution with residual error $o(\varepsilon^{-2M-1}h^{2M})$. This residual error can be made $o(\varepsilon)$ as $h \to 0$ by choosing $M$ such that $(-2M-1)\,\alpha + 2M > \alpha \iff M > \alpha/(1-\alpha)$.

Also, for the same value of $M$, $S'_M(\varepsilon\, d, \varepsilon^{-1}h, \dot{x}_s, \Delta u; \varepsilon^{-1}(x - \dot{x}_s t))$ is an approximate solution of the linearised discrete equation with residual error $o(\varepsilon)$. □

## 4. Asymptotic approximation of discrete solutions.

**4.1. Nonlinear analysis.** The numerical examples in Part 2 [GU09] will consider initial data which lead to the formation of a shock after a finite time. Unfortunately, the numerical analysis of such a problem would require a detailed analysis of the neighbourhood of the shock formation. To simplify the analysis, and keep the focus on the key aspect which is the profile of the numerical solution across the shock, we choose instead to consider first initial data $u_0(x)$ with a finite strength shock at $x_s(0) = 0$. Based on this, we will be able in Part 2 [GU09] to consider more general initial data.

To be specific, the initial data is assumed to satisfy the following conditions (this will be relaxed in Part 2 [GU09]):

**(A1)** apart from a discontinuity at $x_s(0)$, $u_0(x)$ is $C^\infty$ with all derivatives having a finite $L_1$ norm over $(-\infty, x_s(0))$ and $(x_s(0), \infty)$.

**(A2)** the discontinuity has finite strength for the entire time interval $[0, T]$, and no other discontinuity is formed during this time interval;

The initial discontinuity raises a new question, how to specify the initial data for the numerical discretisation. The previous section has shown that if the initial data is piecewise constant, then we can expect the numerical solution to approach the travelling wave solution $S(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s, [u]; \varepsilon^{-1}(x - \dot{x}_s t))$. To avoid the requirement of performing an additional asymptotic analysis of this relaxation process, we will assume a very particular form of the initial data in the neighbourhood of the initial shock. The details will be given later, but on either side of the shock the initial data is simply $U^0_j = u_0(x_j)$.

The objective now is to use a matched asymptotic analysis to construct a smooth function $V(x,t)$ such that $V^n_j = V(x_j, t^n)$ is a very close approximation to the discrete solution $U^n_j$. The stability estimate in Theorem 2.1 will be used to bound the difference $V^n_j - U^n_j$ based on the magnitude of the residual error $N_j(V^n_j)$.

The matched asymptotic analysis [BO78, KC81] breaks the domain into three overlapping regions:

A: $x_s - x > \varepsilon^\beta$

B: $|x - x_s| < 2\varepsilon^\beta$

C: $x - x_s > \varepsilon^\beta$

Here $\beta$ is a constant just slightly less than unity so that the overlap regions contain less and less of the exponential tails of the travelling wave profile as $h \to 0$. A lower bound on $\beta$ will be determined in Theorem 4.2.

We begin with a result concerning an approximate solution in the outer region.

THEOREM 4.1. *In outer regions A and C, there exists a function $V_o(x,t)$, with a parametric dependence on $h$ and satisfying initial data $V_o(x,0) = u_0(x)$, such that $N_j\,(V_o(x_j, t^n)) = o(h^{2+\alpha})$.*

*Proof.* In the outer regions, the leading order term in $V_o(x,t)$ is the inviscid solution $u(x,t)$ which gives a residual error whose leading order term is

$$-\varepsilon \frac{\partial^2 u}{\partial x^2}.$$

Following the same procedure as in the proof of Theorem 3.3, this residual error can be eliminated to leading order through the addition of a correction term $\varepsilon V_{o,1}(x,t)$ which satisfies the equation

$$\frac{\partial V_{o,1}}{\partial t} + \frac{\partial}{\partial x}\Big(a(u)\, V_{o,1}\Big) = \frac{\partial^2 u}{\partial x^2},$$

subject to homogeneous initial conditions. Continuing with this process, it can be proved inductively that each correction term is $O(h^{m\alpha+2n})$ for integers $m,n$ with $m\geq 0, n\geq 0, m+n>0$, with the powers of $h^2$ arising from the second order accuracy of the central finite difference approximations. Therefore after a finite number of steps this process gives an approximate solution $V_o(x,t)$ of the form

$$V_o(x,t) = u(x,t) + \varepsilon\, V_{o,1}(x,t) + \sum_{\alpha < m\alpha+2n \leq \alpha+2} h^{m\alpha+2n} V_{o,m,n}(x,t) \tag{4.1}$$

which satisfies the initial data $V_o(x,0) = u_0(x)$ and has a residual error which is $o(h^{2+\alpha})$. ∎

The next theorem considers the more difficult construction of an approximate solution in the inner region, using an inner coordinate $X \equiv \varepsilon^{-1}(x - x_s(t))$. The finite difference discretisation can be viewed as approximating the viscous differential equation

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = \varepsilon \frac{\partial^2 u}{\partial x^2},$$

with $\varepsilon = h^\alpha$. Changing to the new inner coordinate, this equation becomes

$$\varepsilon \frac{\partial u}{\partial t} - \dot{x}_s(t) \frac{\partial u}{\partial X} + \frac{\partial f}{\partial X} = \frac{\partial^2 u}{\partial X^2},$$

and small linearised perturbations of this equation due to the introduction of an inhomogeneous source term $s(x,t)$ satisfy the equation

$$\varepsilon \frac{\partial \widetilde{u}}{\partial t} + \frac{\partial}{\partial X}\Big( (a(u) - \dot{x}_s(t))\, \widetilde{u} \Big) - \frac{\partial^2 \widetilde{u}}{\partial X^2} = s(x,t).$$

Neglecting the $O(\varepsilon)$ unsteady term, this gives the equation

$$\frac{\partial}{\partial X}\Big( (a(u) - \dot{x}_s(t))\, \widetilde{u} \Big) - \frac{\partial^2 \widetilde{u}}{\partial X^2} = s(x,t), \tag{4.2}$$

which plays a central role in the proof of the following theorem.

THEOREM 4.2. *In the inner region B, there exists a function $V_i(x,t)$, with a parametric dependence on $h$, such that $N_j\left(V_i(x_j,t^n)\right) = o(h^{2+\alpha})$, and $V_i(x,t) - V_o(x,t) = o(h^{2+\alpha})$ in the region of overlap with the outer regions A and C, provided $\alpha > \frac{2}{3}$ and $\beta < 1$ is sufficiently large.*

*Proof.* In the overlap between regions A and B, a Taylor series expansion of $V_o(x,t)$ gives

$$V_o(x,t) = V_o(x_s^-,t) + \sum_{p=1}^{3} \frac{h^{p\alpha} X^p}{p!} \left.\frac{\partial^p V_o}{\partial x^p}\right|_{(x_s^-(t),t)} + O(h^{4\alpha} X^4).$$

Note that $V_o(x,t)$ itself has an asymptotic expansion in $h$ given by (4.1), and inserting this in the above will give an expansion in powers of $h^{m\alpha+2n}$ for $m\geq 0, n\geq 0$.

A similar expression holds for the overlap between regions B and C, with $x_s^-$ replaced by $x_s^+$. These form the boundary conditions for the asymptotic expansion in the inner region. To leading order, they give

$$V_{i,0}(X,t) \to \begin{cases} u(x_s^-(t),t), & X\to -\infty, \\ u(x_s^+(t),t), & X\to \infty, \end{cases}$$

and hence the leading term in the asymptotic expansion in inner region B is

$$V_{i,0}(X,t) = S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t), [u(t)]; X - X_s(t)\right),$$

where $X_s(t)$ is initially an arbitrary function of time. $V_{i,0}(X,t)$ satisfies the boundary conditions since $\dot{x}_s(t), u(x_s^-(t),t), u(x_s^+(t),t)$ satisfy the jump condition (1.2) which matches the jump condition (3.3) in Theorem 3.1 (with $s_{-\infty}\equiv u(x_s^-(t),t)$ and $s_\infty \equiv u(x_s^+(t),t)$), which in turn is the basis for the construction of the discrete travelling wave solution in Theorem 3.5.

The non-uniqueness due to $X_s(t)$ in the leading order inner solution is typical of interior boundary layers (see [HY01] and page 160 in [KC81]) and is resolved through a solvability condition for a later term in the asymptotic expansion.

From Equation (3.6) in Theorem 3.5, we have

$$N_j\left(S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s, \Delta u; \varepsilon^{-1}(x_j - \dot{x}_s t^n) - X_s\right)\right) = o(\varepsilon),$$

when $\dot{x}_s, \Delta u, X_s$ are constant, so writing

$$V_j^n = S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^n), [u(t^n)]; \varepsilon^{-1}(x_j - x_s(t^n)) - X_s(t^n)\right),$$

one obtains

$$\begin{aligned} & V_j^n - \tfrac{1}{2}r\left(f(V_{j+1}^n) - f(V_{j-1}^n)\right) + \varepsilon d\left(V_{j+1}^n - 2V_j^n + V_{j-1}^n\right) \\ &= S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^n), [u(t^n)]; \varepsilon^{-1}(x_j - x_s(t^n) - k\dot{x}_s(t^n)) - X_s(t^n)\right) + o(k\varepsilon) \\ &= S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^n), [u(t^n)]; \varepsilon^{-1}(x_j - x_s(t^{n+1})) - X_s(t^n)\right) + o(k\varepsilon), \end{aligned}$$

since $k = o(\varepsilon^2)$. Hence, it follows that

$$\begin{aligned} & N_j\left(S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^n), [u(t^n)]; \varepsilon^{-1}(x - x_s(t^n)) - X_s(t^n)\right)\right) \\ &= k^{-1}\Big(\; S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^{n+1}), [u(t^{n+1})]; \varepsilon^{-1}(x_j - x_s(t^{n+1})) - X_s(t^{n+1})\right) \\ &\qquad - S\left(\varepsilon d, \varepsilon^{-1}h, \dot{x}_s(t^n), [u(t^n)]; \varepsilon^{-1}(x_j - x_s(t^{n+1})) - X_s(t^n)\right)\Big) \; + \; o(\varepsilon) \\ &= \left.\frac{\partial V_{i,0}}{\partial t}\right|_{X_j, t^n} \quad + \quad o(\varepsilon). \end{aligned}$$

Following the methodology of the proof of Theorem 3.3, this $O(1)$ residual error can be corrected for by the addition of a term $\varepsilon V_{i,1}(x,t)$ which must satisfy the equation

$$\frac{\partial}{\partial X}\left(\left(a(V_{i,0}) - \dot{x}_s\right)V_{i,1}\right) - \frac{\partial^2 V_{i,1}}{\partial X^2} = -\frac{\partial V_{i,0}}{\partial t}, \tag{4.3}$$

subject to the boundary conditions

$$V_{i,1}(X,t) \longrightarrow \begin{cases} X \left.\dfrac{\partial u}{\partial x}\right|_{(x_s^-(t),t)} + V_{o,1}(x_s^-(t),t) & X \to -\infty, \\ X \left.\dfrac{\partial u}{\partial x}\right|_{(x_s^+(t),t)} + V_{o,1}(x_s^+(t),t) & X \to \infty. \end{cases}$$

As with Theorem 3.2, there is a solvability condition which must be satisfied, and this is found by integrating Equation (4.3) over the interval $[-X^*, X^*]$. Using the integral property in Theorem 3.5, as $X^* \to \infty$,

$$\int_{-X^*}^{X^*} V_{i,0}\ \mathrm{d}X \to X^* \left(u(x_s^-(t),t) + u(x_s^+(t),t)\right) - X_s(t)\ [u],$$

and hence the integral of the right-hand-side of Equation (4.3) asymptotically approaches

$$-X^* \frac{\mathrm{d}}{\mathrm{d}t}\Big(u(x_s^-(t),t) + u(x_s^+(t),t)\Big) + \frac{\mathrm{d}}{\mathrm{d}t}\Big(X_s(t)\ [u]\Big).$$

Using the boundary conditions as $X^* \to \infty$, the integral of the left-hand-side of Equation (4.3) becomes

$$\begin{aligned}\left[\Big(a(V_{i,0}) - \dot{x}_s\Big) V_{i,1} - \frac{\partial V_{i,1}}{\partial X}\right]_{-X^*}^{X^*} \longrightarrow & \left[\Big(a(u) - \dot{x}_s\Big) V_{o,1} - \frac{\partial u}{\partial x}\right] \\ & + X^* \Big(a(u(x_s^+,t)) - \dot{x}_s\Big) \frac{\partial u}{\partial x}(x_s^+,t) \\ & + X^* \Big(a(u(x_s^-,t)) - \dot{x}_s\Big) \frac{\partial u}{\partial x}(x_s^-,t).\end{aligned}$$

Now, noting that

$$\frac{\mathrm{d}}{\mathrm{d}t} u(x_s^\pm(t),t) = -\Big(a(u(x_s^\pm,t)) - \dot{x}_s(t)\Big) \frac{\partial u}{\partial x}(x_s^\pm,t),$$

it follows that the $X^*$ components on the two sides of the integrated equation are equal. Equating the other components gives the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(X_s\,[u]\Big) = \left[\Big(a(u) - \dot{x}_s\Big) V_{o,1} - \frac{\partial u}{\partial x}\right],$$

governing the evolution of $X_s(t)$ from the initial value $X_s(0) = 0$. Equation (4.3) can now be integrated to obtain $V_{i,1}$, with uniqueness being determined through the solvability condition for the $O(\varepsilon^2)$ correction term.

By continuing the asymptotic expansion and analysis, we eventually obtain an inner solution $V_i(x,t)$ of the form

$$V_i(x,t) = V_{i,0}(X,t) + \varepsilon\, V_{i,1}(X,t) + \sum_{\alpha < m\alpha + 2n \leq \alpha + 2} h^{m\alpha+2n} V_{i,m,n}(X,t)$$

which matches all of the terms in the asymptotic expansion of the outer solution $V_o$ up to and including terms proportional to $h^{2+\alpha}$, and also has a residual error which is $o(h^{2+\alpha})$.

In the overlap region, the exponential tails of the travelling wave solution behave like $\exp(-c|X|)$ for some constant $c$. Since $|X| = O(\varepsilon^{\beta-1}) = O(h^{-\alpha(1-\beta)})$, these exponential tails are $o(h^q)$ for all $q>0$. Furthermore, noting that $4\alpha > 2+\alpha$ because of the lower bound on $\alpha$, let $\gamma$, satisfying the inequalities $4\alpha \geq \gamma > 2+\alpha$, be the lowest power of $h$ for which the corresponding term in the asymptotic expansion of the outer solution given at the beginning of this proof does not have a matching counterpart in the inner expansion. In this case, at worst the mis-match between the inner and outer solutions in the overlap region is $O(h^\gamma X^4) = O(h^{\gamma-4\alpha(1-\beta)})$. Provided $\beta$ satisfies the lower bound given by the condition

$$\gamma - 4\alpha(1-\beta) > 2+\alpha \quad \Longrightarrow \quad \beta > 1 - \frac{\gamma-2-\alpha}{4\alpha}$$

it follows that the inner and outer approximate solutions match to within $O(h^{2+\alpha})$. $\square$

We can now combine these approximate solutions to form a patched solution which gives an accurate approximation of a discrete solution.

Theorem 4.3. *There exists a function $V(x,t)$ with parametric dependence on $h$, and initial data $U_j^0 \equiv V(x_j,0)$ producing the numerical solution $U_j^n$, such that $\|U_j^0 - u_0(x_j)\|_1 = O(h^\alpha)$ and $\|U_j^n - V(x_j,t^n)\|_1 = o(h^{2+\alpha})$.*

*Proof.* The patched solution $V(x,t)$ is defined using a $C^\infty$ blending function $P(x)$ which has constant value $P(x)=1$ for $|x|<1$, and $P(x)=0$ for $|x|>2$. Using this, we define $V(x,t)$ through

$$\begin{aligned} V(x,t) &= V_i(x,t) + \left(1 - P\left(\varepsilon^{-\beta}(x-x_s(t))\right)\right)\Big(V_o(x,t) - V_i(x,t)\Big) \\ &= V_o(x,t) + P\left(\varepsilon^{-\beta}(x-x_s(t))\right)\Big(V_i(x,t) - V_o(x,t)\Big). \end{aligned}$$

Setting $V_j^n = V(x_j,t^n)$, the residual error $N_j(V_j^n)$ is $o(h^{2+\alpha})$ outside the overlap regions. Within the overlap region, defining $P_j^n = P\left(\varepsilon^{-\beta}(x_j-x_s(t^n))\right)$ and $\Delta V_j^n = V_i(x_j,t^n) - V_o(x_j,t^n)$, it can be verified that

$$\begin{aligned} N_j(V_j^n) = {} & P_j^n\, N_j(V_{i,j}^n) + (1-P_j^n)\, N_j(V_{o,j}^n) \\ & + \frac{1}{k}\left(P_j^{n+1} - P_j^n\right)\Delta V_j^{n+1} \\ & + \frac{1}{2h}\left(P_{j+1}^n - P_j^n\right)\Big[f(V_{i,j+1}^n) - f(V_{o,j+1}^n)\Big] \\ & + \frac{1}{2h}\left(P_j^n - P_{j-1}^n\right)\Big[f(V_{i,j-1}^n) - f(V_{o,j-1}^n)\Big] \\ & + \frac{1}{2h}\Big[f(V_{j+1}^n) - P_{j+1}^n\, f(V_{i,j+1}^n) - (1-P_{j+1}^n)\, f(V_{o,j+1}^n)\Big] \\ & - \frac{1}{2h}\Big[f(V_{j-1}^n) - P_{j-1}^n\, f(V_{i,j-1}^n) - (1-P_{j-1}^n)\, f(V_{o,j-1}^n)\Big] \\ & - \frac{\varepsilon}{h^2}\left(P_{j+1}^n - 2P_j^n + P_{j-1}^n\right)\Delta V_j^n \\ & - \frac{\varepsilon}{h^2}\Big[\left(P_{j+1}^n - P_j^n\right)\left(\Delta V_{j+1}^n - \Delta V_j^n\right) + \left(P_j^n - P_{j-1}^n\right)\left(\Delta V_j^n - \Delta V_{j-1}^n\right)\Big] \end{aligned}$$

On line 1 of this equation, the linear interpolation of the residual errors is $o(h^{2+\alpha})$. Since $k^{-1}(P_j^{n+1} - P_j^n) = O(\varepsilon^{-\beta})$ and $\Delta V = o(h^{2+\alpha})$, the term on line 2 is $o(h^{-\alpha\beta+2+\alpha})$. Similarly, the terms on lines 3 and 4 are also $o(h^{-\alpha\beta+2+\alpha})$. Due to standard quadratic error bounds for linear interpolation, the terms on lines 5 and 6 are $o(h^{-1+2(2+\alpha)})$.

The terms on lines 7 and 8 are $o(h^{\alpha-2\alpha\beta+2+\alpha})$, since each derivative of $P$ and $\Delta V$ introduces a factor $h^{-\alpha\beta}$. Since the overlap regions have measure $O(h^{\alpha\beta})$, it follows that the $l_1$ norm of the residual on the overlap regions is $o(h^{2+\alpha})$.

From the above, plus assumption (A1) which gives a finite bound for the $L_1$ norm of the derivatives in the residual, we conclude that $\|N_j(V_j^n)\|_1 = o(h^{2+\alpha})$. Hence, since $U_j^n$ is the numerical solution corresponding to initial data $U_j^0 = V_j^0$, then by the stability estimate in Theorem 2.1 it follows that

$$\|U_j^n - V_j^n\|_1 = o(h^{2+\alpha}).$$

□

**4.2. Linear analysis.** We now construct a matched asymptotic approximation to the numerical solution of the linear discrete equations subject to the initial data $\widetilde{U}_j^0 = \widetilde{u}_0(x_j)$, where throughout this subsection $\widetilde{u}_0$ is $C^\infty$.

THEOREM 4.4. *In outer regions A and C, there exists a function $\widetilde{V}_o(x,t)$, with parametric dependence on h, satisfying initial data $\widetilde{V}_o(x,0) = \widetilde{u}_0(x)$, such that*

$$L_j(V_o(x_j,t^n))\ \widetilde{V}_o(x_j,t^n) = o(\varepsilon)$$

*Proof.* As with the nonlinear analysis, the leading order term in the asymptotic expansion for the linear solution is the analytic solution $\widetilde{u}(x,t)$ which gives a residual error which is $O(\varepsilon)$. By matching all residual error terms which are $O(\varepsilon)$, including those coming from the asymptotic expansion of the nonlinear discrete solution, this can be compensated for through the addition of a correction $\varepsilon\, \widetilde{V}_{o,1}(x,t)$ satisfying the equation

$$\frac{\partial \widetilde{V}_{o,1}}{\partial t} + \frac{\partial}{\partial x}\left(a(u)\, \widetilde{V}_{o,1}\right) = \frac{\partial^2 \widetilde{u}}{\partial x^2} - \frac{\partial}{\partial x}\left(a'(u)\, V_{o,1}\, \widetilde{V}_{o,0}\right),$$

subject to homogeneous initial data. Setting $\widetilde{V}_o(x,t) = \widetilde{u}(x,t) + \varepsilon\, \widetilde{V}_{o,1}(x,t)$ then gives an outer solution with the required properties. □

THEOREM 4.5. *In the inner region B, there exists a function $\widetilde{V}_i(x,t)$, with parametric dependence on h, such that $L_j(V_i(x_j,t^n))\ \widetilde{V}_i(x_j,t^n) = O(\varepsilon)$ and $\widetilde{V}_i(x,t) - \widetilde{V}_o(x,t) = o(\varepsilon)$ in the region of overlap with the outer regions A and C.*

*Proof.* Because the shock width is $O(\varepsilon)$, a linearised displacement of the shock of unit magnitude corresponds to an $O(\varepsilon^{-1})$ linear solution. Accordingly, in region B, the leading term in the asymptotic expansion is

$$\begin{aligned}\varepsilon^{-1}\widetilde{V}_{i,-1}(x,t) &= -\widetilde{X}_s(t)\ \frac{\partial V_{i,0}}{\partial x}\\ &= -\varepsilon^{-1}\, \widetilde{X}_s(t)\ S'(\dot{x}_s(t), [u(t)]; \varepsilon^{-1}(x - x_s(t)) - X_s(t))\end{aligned}$$

for some as yet undetermined function $\widetilde{X}_s(t)$, and with the parametric dependence of $S'$ on the constants $\varepsilon d, \varepsilon^{-1}h$ omitted for brevity.

Since, from Equation (3.7) in Theorem 3.5,

$$L_j\left(S(\dot{x}_s, \Delta u; \varepsilon^{-1}(x_j - \dot{x}_s t^n) - X_s)\right)\ S'\left(\dot{x}_s, \Delta u; \varepsilon^{-1}(x_j - \dot{x}_s t^n) - X_s\right) = o(\varepsilon),$$

when $\dot{x}_s, \Delta u, X_s$ are constant, it follows that

$$
\begin{aligned}
& L_j\left(V_{i,0}(x_j,t^n)+\varepsilon V_{i,1}(x_j,t^n)\right)\ \widetilde{V}_{i,-1}(x_j,t^n) \\
&= -k^{-1}\Big(\widetilde{X}_s(t^{n+1})\, S'\left(\dot{x}_s(t^{n+1}),[u(t^{n+1})];\varepsilon^{-1}(x_j-x_s(t^{n+1}))-X_s(t^{n+1})\right) \\
&\quad -tX(t^n)\ \ S'\left(\dot{x}_s(t^n),\ [u(t^n)];\ \varepsilon^{-1}(x_j-x_s(t^{n+1}))-X_s(t^n)\right)\Big) \\
&\quad +\varepsilon\ \frac{\partial}{\partial x}\left(a'(V_{i,0}(x_j,t^n))\ V_{i,1}(x_j,t^n)\ \widetilde{V}_{i,-1}(x_j,t^n)\right) \quad + \quad o(\varepsilon) \\
&= \frac{\partial}{\partial t}\ \widetilde{V}_{i,-1}(X_j,t^n)+\frac{\partial}{\partial X}\left(a'(V_{i,0}(x_j,t^n))\ V_{i,1}(x_j,t^n)\ \widetilde{V}_{i,-1}(x_j,t^n)\right)\ +\ o(\varepsilon)
\end{aligned}
$$

This $O(1)$ residual error can be corrected for by the addition of a term $\widetilde{V}_{i,0}(x,t)$ which must satisfy (see Equation (4.2) )

$$
\frac{\partial}{\partial X}\left(\left(a(V_{i,0})-\dot{x}_s\right)\widetilde{V}_{i,0}\right)-\frac{\partial^2\widetilde{V}_{i,0}}{\partial X^2}=-\frac{\partial\widetilde{V}_{i,-1}}{\partial t}-\frac{\partial}{\partial X}\left(a'(V_{i,0})\ V_{i,1}\ \widetilde{V}_{i,-1}\right).
$$

subject to the boundary conditions

$$
\widetilde{V}_{i,0}(X,t)\to\begin{cases}\widetilde{u}(x_s^-(t),t), & X\to-\infty,\\ \widetilde{u}(x_s^+(t),t), & X\to\infty.\end{cases}
$$

Again, as in Theorem 3.2, there is a solvability condition which is found by integrating over an interval $[-X^*, X^*]$, and then taking the limit $X^*\to\infty$ to obtain

$$
\frac{\mathrm{d}}{\mathrm{d}t}\left(\widetilde{X}_s\ [u]\right)=\left[(a(u)-\dot{x}_s)\,\widetilde{u}_{o,0}\right],
$$

subject to initial condition $\widetilde{X}_s(0)=0$. This shows that $\widetilde{X}_s(t)$ satisfies the same equation as the inviscid linearised shock displacement $\widetilde{x}_s(t)$, Equation (1.4), and so is identical to it.

Continuing in this way with a further level of additional correction, we obtain an inner solution $\widetilde{V}_i(x,t)$ of the form

$$
\widetilde{V}_i(x,t)=\varepsilon^{-1}\widetilde{V}_{i,-1}(X,t)+\widetilde{V}_{i,0}(X,t)+\varepsilon\,\widetilde{V}_{i,1}(X,t)
$$

which has a residual error which is $O(\varepsilon)$, and whose value and derivatives match the outer solution in the overlap regions to within $o(\varepsilon)$. ◻

We can now construct a patched asymptotic approximation $\widetilde{V}(x,t)$ over the whole interval.

THEOREM 4.6. *Given the nonlinear initial conditions $U_j^0$ as defined in Theorem 4.3, there exists a function $\widetilde{V}(x,t)$ such that $\|\widetilde{U}_j^n-\widetilde{V}(x_j,t^n)\|_1=o(\varepsilon)$.*

*Proof.* $\widetilde{V}(x,t)$ is defined using the same blending function as in the proof of Theorem 4.3 to give

$$
\begin{aligned}
\widetilde{V}(x,t) &= \widetilde{V}_i(X,t)+\left(1-P\left(\varepsilon^{-\beta}(x-x_s(t))\right)\right)\left(\widetilde{V}_o(x,t)-\widetilde{V}_i(X,t)\right) \\
&= \widetilde{V}_o(x,t)+P\left(\varepsilon^{-\beta}(x-x_s(t))\right)\left(\widetilde{V}_i(x,t)-\widetilde{V}_o(X,t)\right).
\end{aligned}
$$

Setting $\widetilde{V}_j^n=\widetilde{V}(x_j,t^n)$, following the same approach as in the proof of Theorem 4.3 it can be proved that the $l_1$ norm of the residual error $L_j(V_j^n)\ \widetilde{V}_j^n$ is $o(\varepsilon)$. Hence, combining the stability estimate in Theorem 2.2 and the result in Theorem 4.3 we obtain

$$\|\widetilde{U}_j^n - \widetilde{V}_j^n\|_1 = o(\varepsilon).$$

□

**5. Functional errors.** We now come to the main theorem of the paper, proving the convergence of the numerical approximation to the linear functional.

THEOREM 5.1. *If $U_j^n$ satisfies the nonlinear discrete equations subject to the initial data $U_j^0$ specified in Theorem 4.3, and $\widetilde{U}_j^n$ satisfies the linear discrete equations with initial data $\widetilde{U}_j^0 = \widetilde{u}_0(x_j)$, and if $\widetilde{J}$ and $\widetilde{J}_h$ are as defined in the Introduction, then $\left|\widetilde{J}_h - \widetilde{J}\right| = O(\varepsilon)$.*

*Proof.* Taking $V_j^n$ and $\widetilde{V}_j^n$ to be as defined in Theorems 4.3 and 4.6, the difference between the true value of the linear functional

$$\begin{aligned} \widetilde{J} = & \int_{-\infty}^{x_s(T)} \gamma(x)\ G'(u(x,T))\ \widetilde{u}(x,T)\ \mathrm{d}x \ -\ \widetilde{x}_s(T)\,\gamma(x_s(T))\,[G]_T \\ & + \int_{x_s(T)}^{\infty} \gamma(x)\ G'(u(x,T))\ \widetilde{u}(x,T)\ \mathrm{d}x \end{aligned}$$

and the discrete approximation

$$\widetilde{J}_h = h \sum_j \gamma(x_j)\ G'(U_j^N)\ \widetilde{U}_j^N,$$

can be bounded by the sum of three terms:

$$\begin{aligned} \left|\widetilde{J}_h - \widetilde{J}\right| \le & \left| h \sum_j \gamma(x_j) \left( G'(U_j^N)\ \widetilde{U}_j^N - G'(V_j^N)\ \widetilde{V}_j^N \right) \right| \\ & + \left| \ h \sum_j \gamma(x_j)\ G'(V_j^N)\ \widetilde{V}_j^N - \int_{-\infty}^{\infty} \gamma(x)\ G'(V(x,T))\ \widetilde{V}(x,T)\ \mathrm{d}x\ \right| \\ & + \left| \ \int_{-\infty}^{\infty} \gamma(x)\ G'(V(x,T))\ \widetilde{V}(x,T)\ \mathrm{d}x \ -\ \widetilde{J}\ \right| \end{aligned}$$

i) Considering the first term,

$$G'(U_j^N)\ \widetilde{U}_j^N - G'(V_j^N)\ \widetilde{V}_j^N = G'(U_j^N) \left( \widetilde{U}_j^N - \widetilde{V}_j^N \right) + \left( G'(U_j^N) - G'(V_j^N) \right) \widetilde{V}_j^N,$$

and hence

$$\begin{aligned} & \left| h \sum_j \gamma(x_j) \left( G'(U_j^N)\ \widetilde{U}_j^N - G'(V_j^N)\ \widetilde{V}_j^N \right) \right| \\ & \qquad \le \|\gamma\|_\infty \left( g_1\ \|\widetilde{U}^N - \widetilde{V}^N\|_1 + g_2\ \|U^N - V^N\|_1\ \|\widetilde{V}^N\|_\infty \right), \end{aligned}$$

where

$$g_1 = \sup_{|v| \le \|u\|_\infty} G'(v), \qquad g_2 = \sup_{|v| \le \|u\|_\infty} G''(v).$$

Theorems 4.3, 4.5 and 4.6, give $\|\widetilde{U}^N - \widetilde{V}^N\|_1 = o(\varepsilon)$, $\|U^N - V^N\|_1 = o(h^{2+\alpha}) = o(h^2\varepsilon)$ and $\|\widetilde{V}^N\|_\infty = O(\varepsilon^{-1})$, and hence, this first term is $o(\varepsilon)$.

ii) Since $V_j^N = V(x_j, T)$ and $\widetilde{V}_j^N = \widetilde{V}(x_j, T)$, the second term corresponds to the error in using trapezoidal integration to approximate the integral of

$$f(x) \equiv \gamma(x)\ G'(V(x,T))\ \widetilde{V}(x,T).$$

$V(x,T)$ and $\widetilde{V}(x,T)$ are both smooth, so the error is bounded by the Euler-Maclaurin error formula [SB80]

$$\left| h \sum_j f(x_j) - \int_{-\infty}^{\infty} f(x)\ \mathrm{d}x \right| \leq \frac{h^{2M}}{(2M)!} \int_{-\infty}^{\infty} \left| \overline{B}_{2m}\!\left(\frac{x}{h}\right)\ f^{(2M)} \right|\ \mathrm{d}x,$$

for any integer $M$, with $\overline{B}_{2m}(x)$ being the periodic extension of Bernoulli polynomials on $[0,1]$ [SB80]. Since $\|f^{(2M)}\|_\infty = O(h^{-(2M+1)\alpha})$ it is possible to choose $M$ sufficiently large so that $2M - (2M+1)\,\alpha > \alpha$ and hence this second term is $o(\varepsilon)$.

iii) For the third term, we start by evaluating the leading order behaviour of the integral in the inner region.

$$\begin{aligned}
&\int_B \gamma(x)\ G'(V(x,T))\ \widetilde{V}(x,T)\ \mathrm{d}x \\
&= \int_B \gamma(x)\ G'\Big(V_{i,0}(x,T) + \varepsilon V_{i,1}(x,T)\Big)\ \Big(-\widetilde{x}_s\, \frac{\partial V_{i,0}}{\partial x}(x,T) + \widetilde{V}_{i,0}(x,T)\Big)\ \mathrm{d}x + o(\varepsilon) \\
&= -\ \widetilde{x}_s\, \gamma(x_s(T)) \int_B G'(V_{i,0}(x,T))\ \frac{\partial V_{i,0}}{\partial x}(x,T)\ \mathrm{d}x \\
&\quad -\ \varepsilon\, \widetilde{x}_s \int_B \gamma(x)\ G''(V_{i,0}(x,T))\ V_{i,1}(x,T)\ \frac{\partial V_{i,0}}{\partial x}(x,T)\ \mathrm{d}x \\
&\quad -\ \widetilde{x}_s \int_B \Big(\gamma(x) - \gamma(x_s(T))\Big)\, G'(V_{i,0}(x,T))\ \frac{\partial V_{i,0}}{\partial x}(x,T)\ \mathrm{d}x \\
&\quad +\ \int_B \gamma(x)\ G'(V_{i,0}(x,T))\ \widetilde{V}_{i,0}(x,T)\ \mathrm{d}x\ +\ o(\varepsilon)
\end{aligned}$$

In this final expression, the first integral gives $[G(u)]_T$. The second integral is $O(1)$ since $\frac{\partial V_{i,0}}{\partial x}(x,T) = O(\varepsilon^{-1})$ in the innermost region of size $O(\varepsilon)$ and tails off exponentially outside this, and hence $\|\frac{\partial V_{i,0}}{\partial x}(x,T)\|_1 = O(1)$. The third integral is $O(\varepsilon)$ since the product $(\gamma(x) - \gamma(x_s(T)))\,\frac{\partial V_{i,0}}{\partial x}(x,T)$ is $O(1)$ in the innermost region and tails off exponentially. Finally, noting that $G'(V_{i,0}(x,T)) - G'(u(x,T)) = O(1)$ and $\widetilde{V}_{i,0}(x,T) - \widetilde{u}_{o,0}(x,T) = O(1)$ uniformly in the innermost region and tail off exponentially to an $O(\varepsilon^\beta)$ value outside this, the fourth integral is equal to

$$\int_B \gamma(x)\ G'(u(x,T))\ \widetilde{u}_{o,0}\ \mathrm{d}x +\ O(\varepsilon).$$

Adding in the contributions from the outer regions then gives

$$\begin{aligned}
&\int_{-\infty}^{\infty} \gamma(x)\ G'(V(x,T))\ \widetilde{V}(x,T)\ \mathrm{d}x \\
&= -\widetilde{x}_s\, \gamma(x_s(T))\ [G(u)]_T + \int_{-\infty}^{\infty} \gamma(x)\ G'(u(x,T))\ \widetilde{u}_{o,0}(x,T)\ \mathrm{d}x\ +\ O(\varepsilon) \\
&= \widetilde{J}\ +\ O(\varepsilon).
\end{aligned}$$

Hence the third term is $O(\varepsilon)$, giving the dominant contribution of the three terms.

This concludes the proof that the error in the discrete approximation to the linearised functional is $O(\varepsilon)$, as it would be for a model problem without a shock. $\square$

**6. Conclusions.** This paper has analysed the convergence of approximate linear solutions for a class of convex flux functions using a particular modified Lax-Friedrichs discretisation. It has been proved that in the case of a single shock, the linear solution converges pointwise everywhere except at the shock, and the shock itself is treated correctly in the sense that the value of integral output functionals also converges.

The proofs rely upon the facts that 1) the linear discretisation is a linearisation of the nonlinear discretisation; and 2) the number of mesh points across the smeared shock increases as $h \to 0$.

In Part 2 [GU09] we will continue our analysis. We will present numerical results that confirm our analytical findings. Moreover, we will show that the results of the present paper hold also for Dirac initial data for the linearised equation. From this we deduce that the adjoint approximation also converges pointwise everywhere except along the two characteristics at which it is discontinuous. Finally, the convergence of the adjoint solution will be extended to cases with multiple shocks.

The modified Lax-Friedrichs discretisation which is analysed in this paper is not a great choice as a practical numerical method, since it provides only $O(h^\alpha)$ convergence for $0<\alpha<1$. A better numerical method would use adaptive smoothing, reducing the magnitude of $\varepsilon$ or switching to a fourth difference smoothing in the smooth regions on either side of the shocks, together with adaptive grid resolution to reduce the magnitude of the grid spacing $h$ in the vicinity of the shock. Thus, the contribution of this paper and Part 2 [GU09] is to prove convergence of a simplified discretisation, in order to provide insight and guidance to those trying to construct more accurate, practical methods.

REFERENCES


[AV99] W.K. Anderson and V. Venkatakrishnan. Aerodynamic design optimization on unstructured grids with a continuous adjoint formulation. *Computers & Fluids*, 28(4-5):443–480, 1999.

[BD02] T. Barth and H. Deconinck, editors. *Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics.* Number 25 in Lecture Notes in Computational Science and Engineering. Springer-Verlag, 2002.

[BJ98] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal.*, 32(7):891–933, 1998.

[BO78] C.M. Bender and S.A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers.* McGraw-Hill Inc, 1978.

[BP03] C. Bardos and O. Pironneau. Derivatives and control in the presence of shocks. *Computational Fluid Dynamics Journal*, 11(4):383–392, 2003.

[BR01] R. Becker and R. Rannacher. An optimal control approach to error control and mesh adaptation. In A. Iserles, editor, *Acta Numerica 2001.* Cambridge University Press, 2001.

[CM80] M. Crandall and A. Majda. Monotone difference approximations for scalar conservation laws. *Mathematics of Computation*, 34:1–21, 1980.

[CT87] P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation, II, Numerical results. *Q. J. R. Meteorol. Soc.*, 113:1329–1347, 1987.

[DMLM95] Gianni Dal Maso, Philippe G. Lefloch, and François Murat. Definition and weak stability of nonconservative products. *J. Math. Pures Appl. (9)*, 74(6):483–548, 1995.

[Gil96] M.B. Giles. Analysis of the accuracy of shock-capturing in the steady quasi-1D Euler equations. *International Journal of Computational Fluid Dynamics*, 5(2):247–258, 1996.

[Gil03] M.B. Giles. Discrete adjoint approximations with shocks. In T. Hou and E. Tadmor, editors, *Hyperbolic Problems: Theory, Numerics, Applications*. Springer-Verlag, 2003.

[GP98] M.B. Giles and N.A. Pierce. On the properties of solutions of the adjoint Euler equations. In M. Baines, editor, *Numerical Methods for Fluid Dynamics VI*. ICFD, June 1998.

[GP01] M.B. Giles and N.A. Pierce. Analytic adjoint solutions for the quasi-one-dimensional Euler equations. *Journal of Fluid Mechanics*, 426:327–345, 2001.

[GU09] M.B. Giles and S. Ulbrich. Convergence of linearised and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions. Technical Report, Mathematical Institute, University of Oxford, 2009.

[HCL94] K.C. Hall, W.S. Clark, and C.B. Lorence. A linearized Euler analysis of unsteady transonic flows in turbomachinery. *Journal of Turbomachinery*, 116:477–488, 1994.

[HL94] Thomas Y. Hou and Philippe G. LeFloch. Why nonconservative schemes converge to wrong solutions: error analysis. *Math. Comp.*, 62(206):497–530, 1994.

[HN03] C. Homescu and I.M. Navon. Optimal control of flow with discontinuities. *Journal of Computational Physics*, 187:660–682, 2003.

[HY01] S.I. Hariharan and G.W. Young. Comparison of asymptotic solutions of a phase-field model to a sharp-interface model. *SIAM Journal of Applied Mathematics*, 62(1), 2001.

[Jam95] A. Jameson. Optimum aerodynamic design using control theory. In M. Hafez and K. Oshima, editors, *Computational Fluid Dynamics Review 1995*, pages 495–528. John Wiley & Sons, 1995.

[KC81] J. Kevorkian and J.D. Cole. *Perturbation Methods in Applied Maths*. Number 34 in Applied Math Series. Springer-Verlag, 1981.

[LeF90] Philippe LeFloch. An existence and uniqueness result for two nonstrictly hyperbolic systems. In *Nonlinear evolution equations that change type*, volume 27 of *IMA Vol. Math. Appl.*, pages 126–138. Springer, New York, 1990.

[LG94] D.R. Lindquist and M.B. Giles. Validity of linearized unsteady Euler equations with shock capturing. *AIAA Journal*, 32(1):46, 1994.

[NT92] H. Nessyahu and E. Tadmor. The convergence rate of approximate solutions for nonlinear scalar conservation laws. *SIAM Journal of Numerical Analysis*, 29:1505–1519, 1992.

[SB80] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer-Verlag, 1980.

[SW98] K. Sreenivas and D.L. Whitfield. Time- and frequency-domain numerical simulation of linearized Euler equations. *AIAA Journal*, 36(6):968–975, 1998.

[Tad91] E. Tadmor. Local error estimates for discontinuous solutions of nonlinear hyperbolic equations. *SIAM Journal of Numerical Analysis*, 28:891–906, 1991.

[TC87] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation, I, Theory. *Q. J. R. Meteorol. Soc.*, 113:1311–1328, 1987.

[Ulb02] S. Ulbrich. A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM Journal of Control and Optimization*, 41(3):740–797, 2002.

[Ulb03] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems & Control Letters*, 48(3-4):313–328, 2003.