Figure 1: Diagram demonstrating a free surface flow (▢) over a bed (▢) where $w(x,t)$ is the absolute location of the free surface, $h(x,t)$ is the height of a column of fluid, $u(x,t)$ is the horizontal velocity of a column of fluid and $b(x)$ is the stationary bed profile.

Numerical solution of the fully non-linear weakly dispersive Serre equations for flows over dry beds.

J.P.A. Pitt[a,*], C. Zoppou[a], S.G. Roberts[a]

[a]*Mathematical Sciences Institute, Australian National University, Canberra, ACT 0200, Australia*

Abstract

Keywords: Serre equations, dry bed

1. Introduction

2. Serre Equations

2.1. Conservation Law Form

$$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} = 0, \tag{1a}$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x}\left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3\left[\frac{\partial u}{\partial x}\right]^2 + h^2 u\frac{\partial u}{\partial x}\frac{\partial b}{\partial x}\right)$$

$$\underbrace{+ \frac{1}{2}h^2 u\frac{\partial u}{\partial x}\frac{\partial^2 b}{\partial x^2} - hu^2\frac{\partial b}{\partial x}\frac{\partial^2 b}{\partial x^2} + gh\frac{\partial b}{\partial x}}_{\text{source term}} = 0 \tag{1b}$$

with

$$G = uh\left(1 + \frac{\partial h}{\partial x}\frac{\partial b}{\partial x} + \frac{1}{2}h\frac{\partial^2 b}{\partial x^2} + \left[\frac{\partial b}{\partial x}\right]^2\right) - \frac{\partial}{\partial x}\left(\frac{1}{3}h^3\frac{\partial u}{\partial x}\right) \tag{2}$$

2.2. Conservation Properties

The total amount of a quantity q in a system occurring on the interval $[a, b]$ at time t is

$$C_q(t) = \int_a^b q(x, t)\, dx.$$

A system is conservative for a given quantity q if $C_q(0) = C_q(t)$ for all t. The Serre equations conserve mass (h), momentum (uh), G and the energy

$$\mathcal{H}(x, t) = \frac{1}{2}\left(gh\,(h + 2b) + hu^2 + \frac{h^3}{3}\left[\frac{\partial u}{\partial x}\right]^2 + u^2 h\left[\frac{\partial b}{\partial x}\right]^2 - uh^2\frac{\partial u}{\partial x}\frac{\partial b}{\partial x}\right).$$

The conservation of h, uh, G is a result of integrating the Serre equations in conservative law form, and in their non-conservative form []. While the conservation of the energy $\mathcal{H}(x, t)$ is given by the derivation of the the Green-Naghdi equations [?] which are equivalent to the Serre equations for one-dimensional flows. Indeed $\mathcal{H}$ is just a sum of the gravitational and kinetic energy throughout the depth of water.

*Corresponding author

Email addresses: jordan.pitt@anu.edu.au (J.P.A. Pitt), christopher.zoppou@anu.edu.au (C. Zoppou), stephen.roberts@anu.edu.au (S.G. Roberts)

3. Method

- Reconstruction

- Finite Element Method

- Finite Volume Update

- Source Term

3.1. Reconstruction

3.1.1. h, w and G

We reconstruct h, w and G with piecewise linear functions over a cell from neighbouring cell averages. Since h, w and G use the same reconstruction operators we demonstrate them for a general quantity q. For the jth cell we reconstruct the values of q at $x_{j-1/2}$, x_j and $x_{j+1/2}$ in the following way

$$q_{j-1/2}^{+} = \mathcal{R}_{j-1/2}^{+}(\overline{q}) = \overline{q}_j - \frac{\Delta x}{2} d_j, \tag{3a}$$

$$q_j = \mathcal{R}_j(\overline{q}) = \overline{q}_j, \tag{3b}$$

$$q_{j+1/2}^{-} = \mathcal{R}_{j+1/2}^{-}(\overline{q}) = \overline{q}_j + \frac{\Delta x}{2} d_j \tag{3c}$$

where

$$d_j = \text{minmod}\left(\theta \frac{\overline{q}_j - \overline{q}_{j-1}}{\Delta x}, \frac{\overline{q}_{j+1} - \overline{q}_{j-1}}{2\Delta x}, \theta \frac{\overline{q}_{j+1} - \overline{q}_j}{\Delta x}\right) \tag{4}$$

with $\theta \in [1, 2]$.

3.1.2. Bed Profile

For the bed profile we require a reconstruction that is at least second-order accurate for b, $\partial b/\partial x$ and $\partial^2 b/\partial x^2$. To accomplish this b is reconstructed with a cubic polynomial $C_j(x)$ centred around x_j

$$C_j(x) = c_0 \left(x - x_j\right)^3 + c_1 \left(x - x_j\right)^2 + c_2 \left(x - x_j\right) + c_3.$$

This cubic has the coefficients

$$c_0 = \frac{-b_{j-2} + 2b_{j-1} - 2b_{j+1} + b_{j+2}}{12\Delta x^3}, \qquad c_1 = \frac{b_{j-2} - b_{j-1} - b_{j+1} + b_{j+2}}{6\Delta x^2},$$

$$c_2 = \frac{b_{j-2} - 8b_{j-1} + 8b_{j+1} - b_{j+2}}{12\Delta x}, \qquad c_3 = \frac{-b_{j-2} + 4b_{j-1} + 4b_{j+1} - b_{j+2}}{6}.$$

We require a continuous bed profile and so we average the two reconstructions at the cell edge from the adjacent cells. Therefore, our reconstruction of the bed profile in the

j^{th} cell is the cubic which takes these values

$$b_{j-1/2} = \mathcal{B}_{j-1/2}(\boldsymbol{b}) = \frac{1}{2}\left(C_j(x_{j-1/2}) + C_{j-1}(x_{j-1/2})\right), \tag{5a}$$

$$b_{j-1/6} = \mathcal{B}_{j-1/6}(\boldsymbol{b}) = C_j(x_{j-1/6}), \tag{5b}$$

$$b_{j+1/6} = \mathcal{B}_{j+1/6}(\boldsymbol{b}) = C_j(x_{j+1/6}), \tag{5c}$$

$$b_{j+1/2} = \mathcal{B}_{j+1/2}(\boldsymbol{b}) = \frac{1}{2}\left(C_j(x_{j+1/2}) + C_{j+1}(x_{j+1/2})\right). \tag{5d}$$

3.2. Velocity Solve using a Finite Element Method

In the FEVM we solve for the primitive variable u given h, G and b using a finite element approximation to (2). For the FEM we begin with the weak form of (2) using test function v over the spatial domain Ω which is

$$\int_\Omega Gv\,dx = \int_\Omega uh\left(1 + \frac{\partial h}{\partial x}\frac{\partial b}{\partial x} + \frac{1}{2}h\frac{\partial^2 b}{\partial x^2} + \left[\frac{\partial b}{\partial x}\right]^2\right)v - \frac{\partial}{\partial x}\left(\frac{1}{3}h^3\frac{\partial u}{\partial x}\right)v\,dx.$$

Integrating by parts with zero Dirichlet boundary conditions we get

$$\int_\Omega Gv\,dx = \int_\Omega uh\left(1 + \left[\frac{\partial b}{\partial x}\right]^2\right)v\,dx + \int_\Omega \frac{1}{3}h^3\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\,dx$$
$$- \int_\Omega \frac{1}{2}uh^2\frac{\partial b}{\partial x}\frac{\partial v}{\partial x}\,dx - \int_\Omega \frac{1}{2}h^2\frac{\partial b}{\partial x}\frac{\partial u}{\partial x}v\,dx. \tag{6}$$

By assuming that time is fixed so that all the functions only vary in space, this formulation implies that by ensuring that G, h, b and $\partial b/\partial x$ have finite integrals over Ω, then u and $\partial u/\partial x$ must have finite integrals as well. Since we require $\partial u/\partial x$ to be well defined to approximate the flux and the source terms (1) and thus have finite integrals we will assume that for each time t that $h, G \in \mathbb{L}^2(\Omega)$ and $b \in \mathbb{W}^{1,2}(\Omega)$ so that $u \in \mathbb{W}^{1,2}(\Omega)$. See Appendix **??** for a precise definition of $\mathbb{L}^2(\Omega)$ and $\mathbb{W}^{1,2}(\Omega)$.

We simplify (6) by performing the integration over the cells and then summing the integrals together to get the equation for the entire domain

$$\sum_{j=0}^{m}\left(\int_{x_{j-1/2}}^{x_{j+1/2}}\left[\left(uh\left(1 + \left[\frac{\partial b}{\partial x}\right]^2\right) - \frac{1}{2}h^2\frac{\partial b}{\partial x}\frac{\partial u}{\partial x} - G\right)v\right.\right.$$
$$\left.\left. + \left(\frac{1}{3}h^3\frac{\partial u}{\partial x} - \frac{1}{2}uh^2\frac{\partial b}{\partial x}\right)\frac{\partial v}{\partial x}\right]dx\right) = 0 \tag{7}$$

which holds for all test functions v. The next step is to replace the functions for h, G, b, v and u with their corresponding basis function approximations.

For h and G we use the basis functions ψ (**??**) which are linear inside a cell and zero elsewhere and so are not continuous as shown in Figure 2. This is consistent
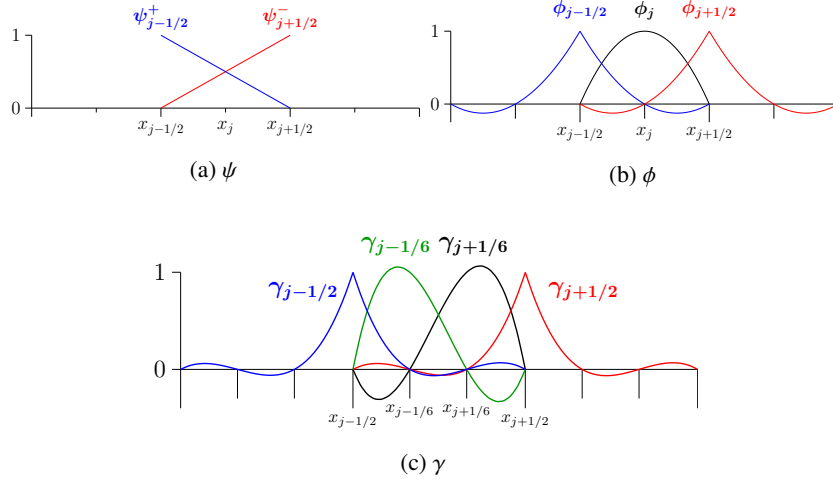
4

(a) ψ

(b) ϕ

(c) γ

Figure 2: Support of the basis functions ψ, ϕ and γ which are non-zero over the jth cell.

with our reconstruction which is second-order accurate inside the cell and possesses discontinuities at the cell edges. Since these basis functions are in $\mathbb{L}^2(\Omega)$ our basis function approximations to h and G are in the appropriate function space.

From the basis functions ψ we have the following representation for h and G in our FEM written for the generic quantity q

$$q = \sum_{j=0}^{m} \left(q_{j-1/2}^{+} \psi_{j-1/2}^{+} + q_{j+1/2}^{-} \psi_{j+1/2}^{-}, \right). \tag{8}$$

For u we use

$$u = u_{-1/2} \phi_{-1/2} + \sum_{j=0}^{m} \left(u_j \phi_j + u_{j+1/2} \phi_{j+1/2} \right). \tag{9}$$

While for b we use

$$b = b_{-1/2} \gamma_{-1/2} + \sum_{j=0}^{m} \left(b_{j-1/6} \gamma_{j-1/6} + b_{j+1/6} \gamma_{j+1/6} + b_{j+1/2} \gamma_{j+1/2} \right). \tag{10}$$

By combining all the matrices generated by the integral of each of the u terms we get the contribution of the jth cell to the stiffness matrix $\mathbf{A}_j$. Likewise all the integrals of the remaining term Gv in (7) generate the element wise vector $\boldsymbol{g}_j$. These element wise matrices and vectors are then assembled into the global stiffness matrix $\mathbf{A}$ and the global right hand-side term $\boldsymbol{g}$ thus (7) is rewritten as

$$\mathbf{A}\hat{\boldsymbol{u}} = \boldsymbol{g}. \tag{11}$$

This is a penta-diagonal matrix equation which can be solved by direct banded matrix solution techniques such as those of **?**] to obtain

$$\hat{\boldsymbol{u}} = \mathcal{G}\left(\hat{\boldsymbol{h}}, \hat{\boldsymbol{G}}, \hat{\boldsymbol{b}} \right) = \mathbf{A}^{-1} \boldsymbol{g} \tag{12}$$

5

as desired.

3.3. Flux Approximation

We use the method of **?**] to calculate the flux across a cell interface. This method was employed because it can handle discontinuities across the cell boundary and only requires an estimate of the maximum and minimum wave speeds. This is precisely the situation for the Serre equations which do not have a known expression for the characteristics but do possess estimates on the maximum and minimum wave speeds (**??**).

Only the calculation of the flux term $F_{j+1/2}$ is demonstrated as the process to calculate the flux term $F_{j-1/2}$ is identical but with different cells. For a general quantity q the approximation of the flux term given by **?**] is

$$F_{j+\frac{1}{2}} = \frac{a^+_{j+\frac{1}{2}} f\left(q^-_{j+\frac{1}{2}}\right) - a^-_{j+\frac{1}{2}} f\left(q^+_{j+\frac{1}{2}}\right)}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} + \frac{a^+_{j+\frac{1}{2}} a^-_{j+\frac{1}{2}}}{a^+_{j+\frac{1}{2}} - a^-_{j+\frac{1}{2}}} \left(q^+_{j+\frac{1}{2}} - q^-_{j+\frac{1}{2}}\right) \tag{13}$$

where $a^+_{j+\frac{1}{2}}$ and $a^-_{j+\frac{1}{2}}$ are given by bounds on the wave speed. Applying the wave speed bounds (**??**) we obtain

$$a^-_{j+\frac{1}{2}} = \min\left\{0, \ u^-_{j+1/2} - \sqrt{gh^-_{j+1/2}}, \ u^+_{j+1/2} - \sqrt{gh^+_{j+1/2}}\right\}, \tag{14}$$

$$a^+_{j+\frac{1}{2}} = \max\left\{0, \ u^-_{j+1/2} + \sqrt{gh^-_{j+1/2}}, \ u^+_{j+1/2} + \sqrt{gh^+_{j+1/2}}\right\}. \tag{15}$$

To recover the lake at rest steady state solution we follow the work of **?**], who accomplished this for the SWWE. Previously, we demonstrated that this process can be extended to the Serre equations [**?**]. To enforce well-balancing the reconstruction of h is modified at the cell edges in the following way.

First calculate

$$b^-_{j+1/2} = w^-_{j+1/2} - h^-_{j+1/2}, \qquad\qquad b^+_{j+1/2} = w^+_{j+1/2} - h^+_{j+1/2}. \tag{16}$$

Find the maximum

$$\ddot{b}_{j+1/2} = \max\left\{b^-_{j+1/2}, b^+_{j+1/2}\right\}$$

then define

$$\ddot{h}^-_{j+1/2} = \max\left\{0, w^-_{j+1/2} - \ddot{b}_{j+1/2}\right\}, \tag{17a}$$

$$\ddot{h}^+_{j+1/2} = \max\left\{0, w^+_{j+1/2} - \ddot{b}_{j+1/2}\right\}. \tag{17b}$$

This generates the vector $\ddot{\boldsymbol{h}}$

$$\ddot{\boldsymbol{h}} = \begin{bmatrix} \ddot{h}^+_{-1/2} \\ h_0 \\ \ddot{h}^-_{1/2} \\ \vdots \\ \ddot{h}^-_{m+1/2} \end{bmatrix}$$

6

which we use to calculate the flux term $F_{j+1/2}$ in (13) for h and G instead of $\hat{h}$. Applying the same process but with different cells we obtain $F_{j-1/2}$ and we have

$$F^n_{j\pm1/2} = \mathcal{F}_{j\pm1/2}\left(\ddot{h}, \hat{G}, \hat{b}, \hat{u}\right).$$

65   for the evolution of h and G equations as desired.

    The flux functions $f(q^-_{j+\frac{1}{2}})$ and $f(q^+_{j+\frac{1}{2}})$ across the cell edge $x_{j+1/2}$ are evaluated using the reconstructed values $q^-_{j+\frac{1}{2}}$ from the jth cell and $q^+_{j+\frac{1}{2}}$ from the $(j+1)^{th}$ cell. From the continuity equation (1a) we have

$$f\left(h^\pm_{j+\frac{1}{2}}\right) = u^\pm_{j+1/2}h^\pm_{j+1/2}.$$

For the evolution of G equation (1b) we have

$$f\left(G^\pm_{j+\frac{1}{2}}\right) = u^\pm_{j+1/2}G^\pm_{j+1/2} + \frac{g}{2}\left(h^\pm_{j+1/2}\right)^2 - \frac{2}{3}\left(h^\pm_{j+1/2}\right)^3\left[\left(\frac{\partial u}{\partial x}\right)^\pm_{j+1/2}\right]^2$$

$$+ \left(h^\pm_{j+1/2}\right)^2 u^\pm_{j+1/2}\left(\frac{\partial u}{\partial x}\right)^\pm_{j+1/2}\left(\frac{\partial b}{\partial x}\right)^\pm_{j+1/2}. \tag{18}$$

66   The quantities $h^+_{j-1/2}$, $h^-_{j+1/2}$, $G^+_{j-1/2}$ and $G^-_{j+1/2}$ were calculated during the recon-
67 struction and the FEM provided $u^\pm_{j+1/2} = u_{j+1/2}$ as u is continuous across the cell bound-
68 aries.

    For $P^u_j(x)$ we obtain the coefficients

$$p^u_0 = \frac{u_{j-1/2} - 2u_j + u_{j+1/2}}{2\Delta x^2},$$
$$p^u_1 = \frac{-u_{j-1/2} + u_{j+1/2}}{\Delta x},$$
$$p^u_2 = u_j.$$

While for $P^b_j(x)$ the coefficients are

$$p^b_0 = \frac{-9b_{j-1/2} + 27b_{j-1/6} - 27b_{j+1/6} + 9b_{j+1/2}}{2\Delta x^3},$$

$$p^b_0 = \frac{9b_{j-1/2} - 9b_{j-1/6} - 9b_{j+1/6} + 9b_{j+1/2}}{4\Delta x^2},$$

$$p^b_0 = \frac{b_{j-1/2} - 27b_{j-1/6} + 27b_{j+1/6} - b_{j+1/2}}{8\Delta x},$$

$$p^b_0 = \frac{-b_{j-1/2} + 9b_{j-1/6} + 9b_{j+1/6} - b_{j+1/2}}{16}.$$

7

Taking the derivative of the polynomials (**??**) and (**??**) we get

$$\frac{\partial}{\partial x} P_j^u(x) = 2p_0^u \left(x - x_j\right) + p_1^u,$$

$$\frac{\partial}{\partial x} P_j^b(x) = 3p_0^b \left(x - x_j\right)^2 + 2p_1^b \left(x - x_j\right) + p_2^b.$$

This gives a second-order approximation to the derivative of u and b at $x_{j+1/2}$ for the jth cell. The process for the $(j + 1)^{th}$ cell is the same and we get

$$\left(\frac{\partial u}{\partial x}\right)_{j+1/2}^{-} = \frac{\partial}{\partial x} P_j^u(x_{j+1/2}),$$

$$\left(\frac{\partial u}{\partial x}\right)_{j+1/2}^{+} = \frac{\partial}{\partial x} P_{j+1}^u(x_{j+1/2}),$$

$$\left(\frac{\partial b}{\partial x}\right)_{j+1/2}^{-} = \frac{\partial}{\partial x} P_j^b(x_{j+1/2}),$$

$$\left(\frac{\partial b}{\partial x}\right)_{j+1/2}^{+} = \frac{\partial}{\partial x} P_{j+1}^b(x_{j+1/2}).$$

3.4. Source Term Approximation

To evolve the Serre equations (1), we require an approximation to the source term at the cell centre x_j from time t^n to t^{n+1} which we denote as S_j^n. Equation (1a) has no source term, therefore we only present the calculation of the source term for equation (1b).

Following the work of **?**] to produce a well-balanced method, we split our approximation to S_j^n into the centred source term S_{ci} and the corrective interface source terms $S_{j+\frac{1}{2}}^{-}$ and $S_{j+\frac{1}{2}}^{+}$

$$S_j^n = S_{j+\frac{1}{2}}^{-} + \Delta x S_{ci} + S_{j-\frac{1}{2}}^{+}.$$

Where S_{ci} is the naive source term approximation and $S_{j+\frac{1}{2}}^{-}$ and $S_{j+\frac{1}{2}}^{+}$ are correction terms that ensure that the flux and source term cancel for the lake at rest solution.

We calculate the centred source term using

$$S_{ci} = -\frac{1}{2} \left(h_j\right)^2 u_j \left(\frac{\partial u}{\partial x}\right)_j \left(\frac{\partial^2 b}{\partial x^2}\right)_j + h_j \left(u_j\right)^2 \left(\frac{\partial b}{\partial x}\right)_j \left(\frac{\partial^2 b}{\partial x^2}\right)_j - gh_j \left(\frac{\partial b}{\partial x}\right)_j.$$

Where we use h_j from the reconstruction process (3) and u_j from the solution of (12). To calculate the derivatives we employ our polynomial representations of u (**??**) and b (**??**) inside a cell. However, to ensure that the terms cancel properly for a lake at rest we modify our approximation to $\frac{\partial b}{\partial x}$ to use $\dot{b}_{j+1/2}^{-}$ and $\dot{b}_{j+1/2}^{+}$ from (16). Therefore, the

8

following approximations are used to calculate S_{ci}

$$\left(\frac{\partial u}{\partial x}\right)_j = \frac{\partial}{\partial x}P^u_j(x_j),$$

$$\left(\frac{\partial b}{\partial x}\right)_j = \frac{\dot{b}^-_{j+1/2} - \dot{b}^+_{j-1/2}}{\Delta x},$$

$$\left(\frac{\partial^2 b}{\partial x^2}\right)_j = \frac{\partial^2}{\partial x^2}P^b_j(x_j).$$

To ensure well-balancing the corrective interface source terms

$$S^-_{j+\frac{1}{2}} = \frac{g}{2}\left(\ddot{h}^-_{j+\frac{1}{2}}\right)^2 - \frac{g}{2}\left(h^-_{j+\frac{1}{2}}\right)^2,$$

$$S^+_{j-\frac{1}{2}} = \frac{g}{2}\left(h^+_{j-\frac{1}{2}}\right)^2 - \frac{g}{2}\left(\ddot{h}^+_{j-\frac{1}{2}}\right)^2$$

are also added. These corrective terms make use of $h^-_{j+\frac{1}{2}}$ and $h^+_{j+\frac{1}{2}}$ obtained from the reconstruction (3) and the modified values $\ddot{h}^-_{j+\frac{1}{2}}$ and $\ddot{h}^+_{j+\frac{1}{2}}$ from (17). Combining the centred and interface source terms our approximation to the source term for G is

$$S^n_j = S_j\left(\hat{h}, \ddot{h}, \hat{w}, \hat{b}, \hat{u}\right) = S^-_{j+\frac{1}{2}} + \Delta x S_{ci} + S^+_{j-\frac{1}{2}}.$$

76 This operator S_j is slightly different to the example given in the overview of the evo-
77 lution step as it takes $\hat{h}$ and $\ddot{h}$ as inputs. This change was made to make clear that $\ddot{h}$ is
78 required to calculate S^n_j. This would be obscured by a more consistent notation that is
79 possible since $\ddot{h}$ only depends on $\hat{h}$, $\hat{w}$ and $\hat{b}$.

80 *3.5. Time-Stepping*

To increase the order of accuracy in time we employ the Strong Stability Preserving
(SSP) Runge-Kutta method [?] which is a convex combination of the first-order time
steps (??) in the following way

$$\bar{q}^{(1)}_j = \bar{q}^n_j + \frac{\Delta t}{\Delta x}\left(F^n_{j+\frac{1}{2}} - F^n_{j-\frac{1}{2}} + S^n_j\right), \tag{19a}$$

$$\bar{q}^{(2)}_j = \bar{q}^{(1)}_j + \frac{\Delta t}{\Delta x}\left(F^{(1)}_{j+\frac{1}{2}} - F^{(1)}_{j-\frac{1}{2}} + S^{(1)}_j\right), \tag{19b}$$

$$\bar{q}^{n+1}_j = \frac{1}{2}\left(\bar{q}^n_j + \bar{q}^{(2)}_j\right). \tag{19c}$$

81 This results in a time stepping method that preserves the stability of the first-order
82 method (??) and is second-order accurate in time. Since all the spatial approximations
83 are second-order accurate, the steps (i)-(vi) should result in a second-order accurate
84 FEVM for the Serre equations, as desired.

9

$$h_{j-1/2}^{+} = h_{j-1/2}^{+} \left(\frac{h_{j-1/2}^{+} + h_{base}}{h_{j-1/2}^{+} + h_{tol}} \right), \tag{20a}$$

$$h_{j+1/2}^{-} = h_{j+1/2}^{-} \left(\frac{h_{j+1/2}^{-} + h_{base}}{h_{j+1/2}^{-} + h_{tol}} \right) \tag{20b}$$

86

4. Lake at Rest Solution Validation

87

$$h(x,t) = \max\{a_0 - b(x), 0\}, \qquad b(x) = a_1 \sin(a_2 x), \tag{21a}$$
$$u(x,t) = 0, \qquad G(x,t) = 0. \tag{21b}$$

88   was chosen for this validation to ensure that all terms with derivatives of the bed were
89 tested. To demonstrate the capability of the methods in the presence of dry and wet
90 beds the parameter values $a_0 = 0m$, $a_1 = 1m$ and $a_2 = 2\pi/50m^{-1}$ were chosen. These
91 parameter values result in wet regions with a horizontal free surface where the stage
92 $w(x,t) = h(x,t) + b(x) = a_0 = 0$ (21). Therefore, we have a periodic bed where water
93 submerges the troughs of the bed while the peaks of the bed are dry.
94   For the numerical solutions the spatial domain was $x \in [-112.5m, 87.5m]$ and the
95 final time was $t = 10s$, with the standard gravitational acceleration $g = 9.81m/s^2$. The
96 spatial resolution of the method was varied so that $\Delta x = 100/2^k m$ with $k \in [8, \ldots, 17]$
97 and the CFL condition (**??**) was satisfied by having $\Delta t = Cr\Delta x/\sqrt{g}$ with condition
98 number $Cr = 0.5$. The standard limiting parameter $\theta = 1.2$ was used in the generalised
99 minmod limiter, (4) for both FEVM$_2$ and FDVM$_2$. Dirichlet boundary conditions were
100 used at both ends as the analytic solution is stationary.
101   The numerical methods are assessed by using the specified lake at rest solution
102 as initial conditions and comparing the numerical solutions of FEVM$_2$ and FDVM$_2$
103 at $t = 10s$ to the analytic solution, which are the initial conditions. To demonstrate
104 the utility of the well-balancing method the results from two versions of FEVM$_2$ and
105 FDVM$_2$ are presented, where the well-balancing method described in Chapter **??** is
106 and is not employed.
107   Example numerical solutions with $\Delta x = 100/2^{10} m \approx 0.0977m$ at $t = 10s$ for all
108 versions of FEVM$_2$ and FDVM$_2$ are given in Figure **??**. The numerical solutions in
109 these figures are indistinguishable from the analytic solutions at this scale and so the
110 analytic solutions have been omitted from the plots.

5. Forced Solution Validation

111

5.2. Forced Solution

112 *5.1. Forced Solution*

113   The known analytic solutions of the Serre equations provide a stringent test when
114 the bed is horizontal, as all terms in the equations are non-zero and vary in space and
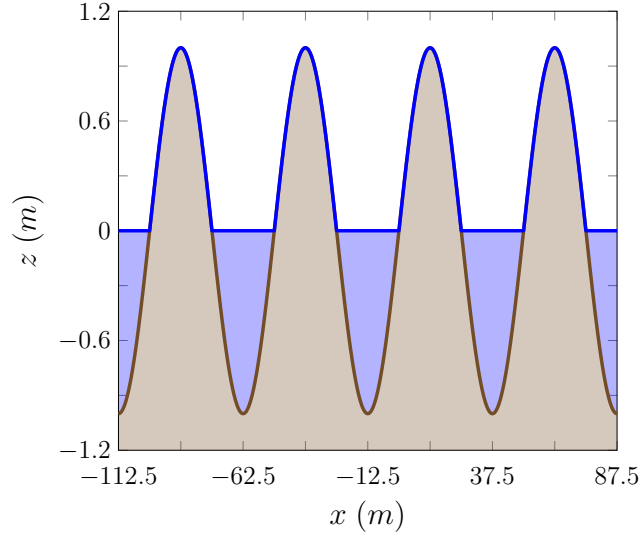
Figure 3: Numerical solution for w (□) and b (□) with $\Delta x = 100/2^{10}m$ for the lake at rest problem at $t = 10s$.
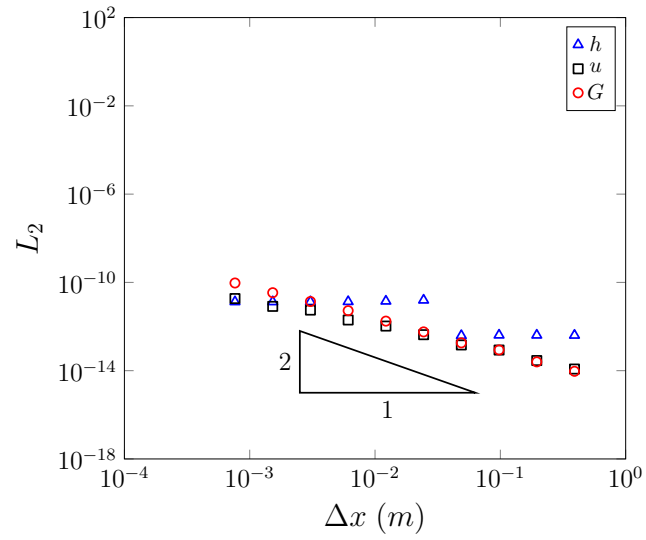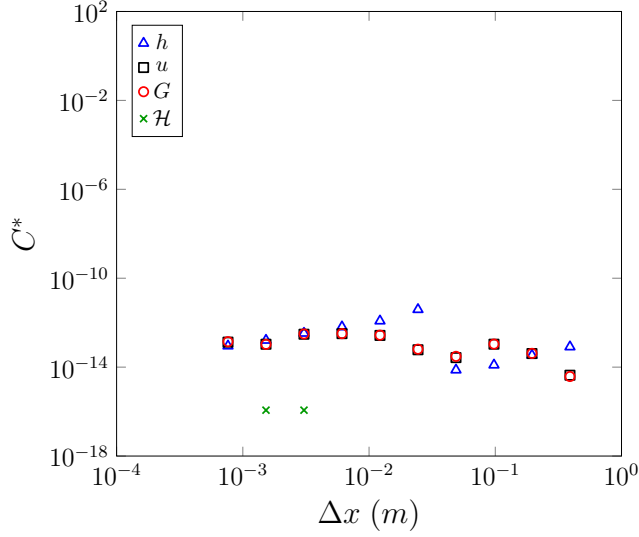

Figure 4: Convergence as measured by the L_2 norm against Δx for h , u and G for the lake at rest problem at $t = 10s$.

(a) numerical conservation error

(b) Conservation error against Δx for h, u, G and $\mathcal{H}$ for the lake at rest problem at $t = 10s$.

time. For varying bathymetry there is only the lake at rest solution where all terms are constant in time and some vanish. Therefore, the accuracy of the approximations of all terms of the Serre equations in the numerical method is not adequately assessed using only the currently available analytic solutions.

Currently the verification of the order of accuracy of the numerical methods for transient solutions with varying bathymetry requires the use of forced solutions. To do this we select some particular functions for all of the primitive quantities; h, u and b which we denote h^*, u^* and b^* respectively. To force these functions h^*, u^* and b^* to be solutions of the Serre equations (1) we add the terms S_h and S_G to obtain the forced Serre equations

$$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} + S_h = 0, \tag{22a}$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 + h^2 u \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right)$$

$$+ \frac{1}{2}h^2 u \frac{\partial u}{\partial x} \frac{\partial^2 b}{\partial x^2} - hu^2 \frac{\partial b}{\partial x} \frac{\partial^2 b}{\partial x^2} + gh \frac{\partial b}{\partial x} + S_G = 0 \tag{22b}$$

12

where

$$S_h = -\frac{\partial h^*}{\partial t} - \frac{\partial (u^* h^*)}{\partial x},$$

$$S_G = -\frac{\partial G^*}{\partial t} - \frac{\partial}{\partial x}\left(u^* G^* + \frac{g\,[h^*]^2}{2} - \frac{2}{3}\,[h^*]^3\left[\frac{\partial u^*}{\partial x}\right]^2 + [h^*]^2\,u^*\frac{\partial u^*}{\partial x}\frac{\partial b^*}{\partial x}\right)$$

$$-\frac{1}{2}\,[h^*]^2\,u^*\frac{\partial u^*}{\partial x}\frac{\partial^2 b^*}{\partial x^2} + h^*[u^*]^2\frac{\partial b^*}{\partial x}\frac{\partial^2 b^*}{\partial x^2} - gh^*\frac{\partial b^*}{\partial x}.$$

These forced Serre equations are then numerically solved by solving the Serre equations (1) with the analytic values of S_h and S_G given h^*, u^* and b^*. So that, the only error present in the numerical solutions of the forced Serre equations is the error produced by the numerical methods used to solve the Serre equations.

Note that since the choice of the forced solutions h^*, u^* and b^* is arbitrary the solutions of the forced Serre equations need not be conservative or retain any properties of the underlying Serre equations.

5.2. Dry Bed Forced Solution Problem

$$h^*(x,t) = a_0 + a_1 \exp\left(-\frac{[(x - a_2 t) - a_3]^2}{2a_4}\right), \tag{23a}$$

$$u^*(x,t) = a_5 \exp\left(-\frac{[(x - a_2 t) - a_3]^2}{2a_4}\right), \tag{23b}$$

$$b^*(x) = a_6 \sin(a_7 x) \tag{23c}$$

for the primitive variables were chosen. These functions produce an a_1 high Gaussian bump for h and u that travels at a fixed speed a_2 over a periodic bed. Thus, h and u will have constant shape and travel to the right over time. However, this is not the case for G as u and h have constant shape but the bed is periodic. With the bed terms in G (2) changing the shape of G as the Gaussian bump in h and u encounters different bed slopes.

For non-trivial choices of the parameters a_i all terms in the Serre equations vary in space and time and so all terms must be accurately approximated by the numerical method to adequately reproduce the forced solution.

Both validation studies used the values $a_1 = 0.5m$, $a_2 = 2\pi/(10a_7)\,m/s$, $a_3 = -3\pi/(2a_7)\,m$, $a_4 = \pi/(16a_7)m^2$, $a_5 = 0.5m/s$, $a_6 = 1.0m$ and $a_7 = \pi/25m^{-1}$ with $a_0 = 1m$ for the finite water depth forced solution and $a_0 = 0m$ for the dry bed forced solution. These parameter values result in a Gaussian bump in h and u that has a width much smaller than the wavelength of the bed profile and travels precisely one wavelength in $10s$.

The domain of the numerical solutions was $x \in [-112.5m, 87.5m]$ with $t \in [0s, 10s]$. The standard gravitational acceleration $g = 9.81m/s^2$ was used. The spatial resolution
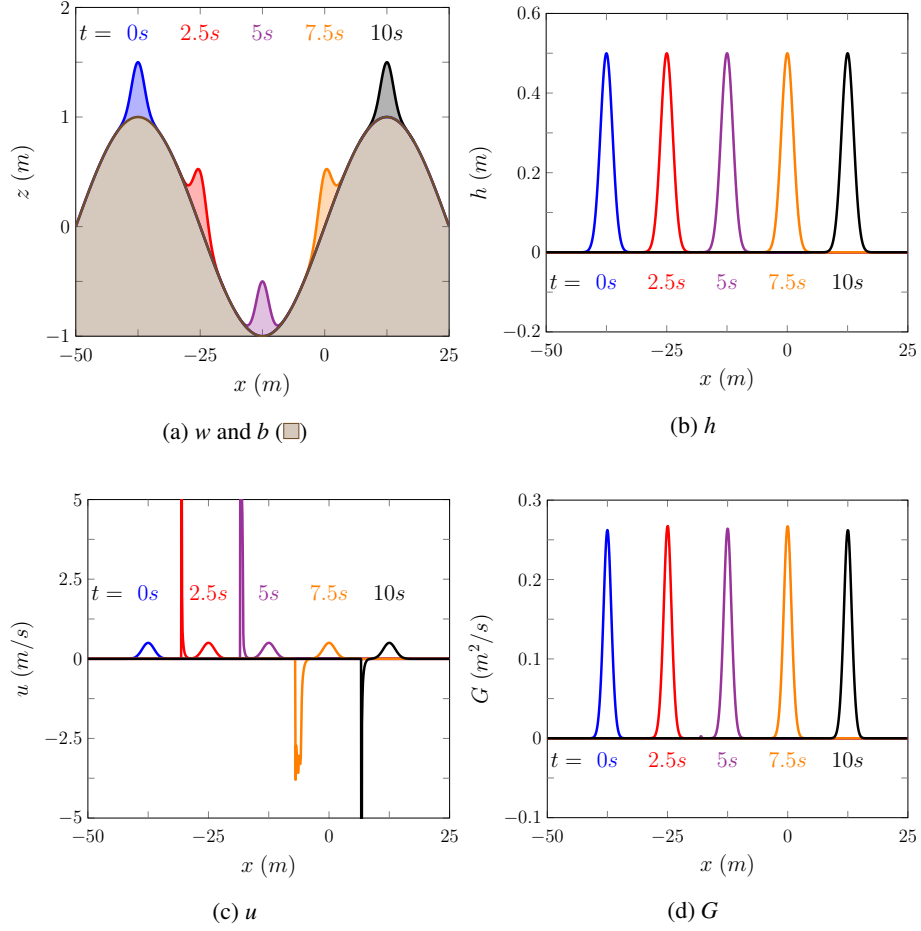
13

(a) w and b (▨)

(b) h

(c) u

(d) G

Figure 6: Example numerical solutions for w, b, h, G with $\Delta x = 100/2^{10}m$ at various times to the dry bed forced solution problem.

of numerical methods was varied like so $\Delta x = 100/2^{k}m$ with $k \in [8, \ldots, 17]$. To satisfy the CFL condition, (**??**) the temporal resolution $\Delta t = Cr\Delta x/\left(a_2 + a_5 + \sqrt{g(a_0 + a_1)}\right)$ was chosen with condition number $Cr = 0.5$. The value $\theta = 1.2$ was used in the generalised minmod limiter (4) for both FEVM$_2$ and FDVM$_2$ and Dirichlet boundary conditions were applied at the boundaries of the domain.

6. Run-up of a Solitary Wave
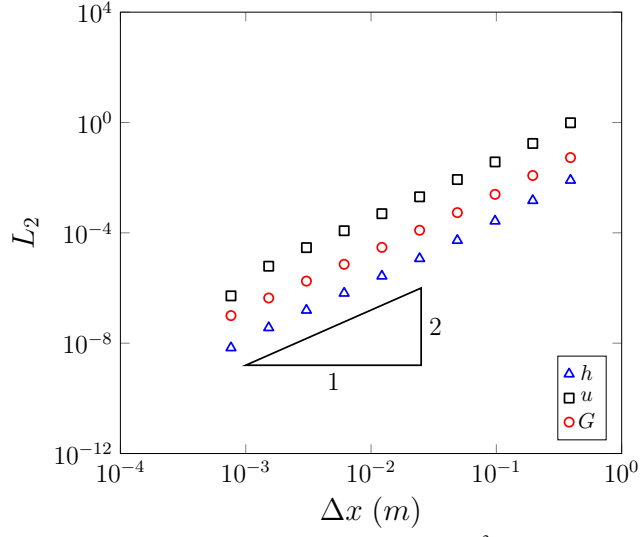
7. Comparison To Ritter's Dry-Bed Dam-Break Solution

Figure 7: Convergence as measured by the L_2 in regions where $h > 10^{-3}m$ norm against Δx for h, u, and G for the dry bed forced solution problem at $t = 10s$.
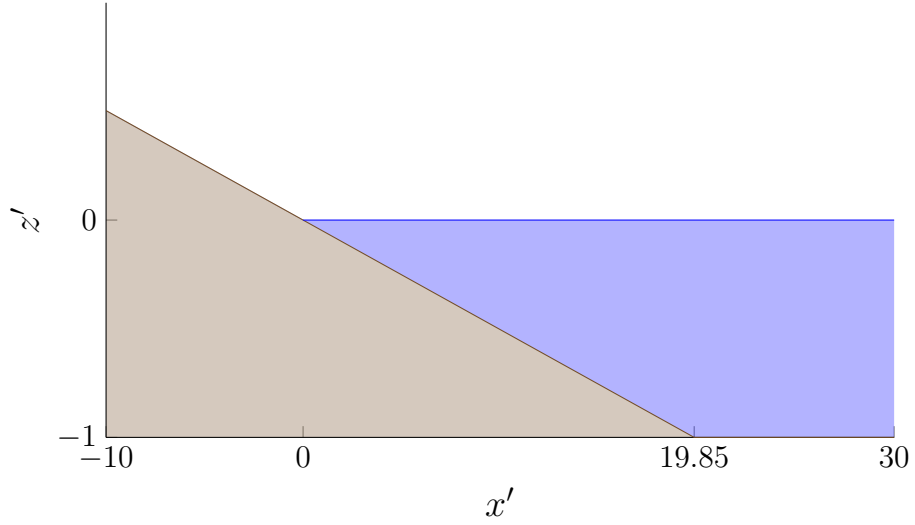

Figure 8: Diagram showing a longitudinal section of the wave tank for the run-up experiment with the water (■) and the bed (■) where the coordinates have been non-dimensionalised [?].
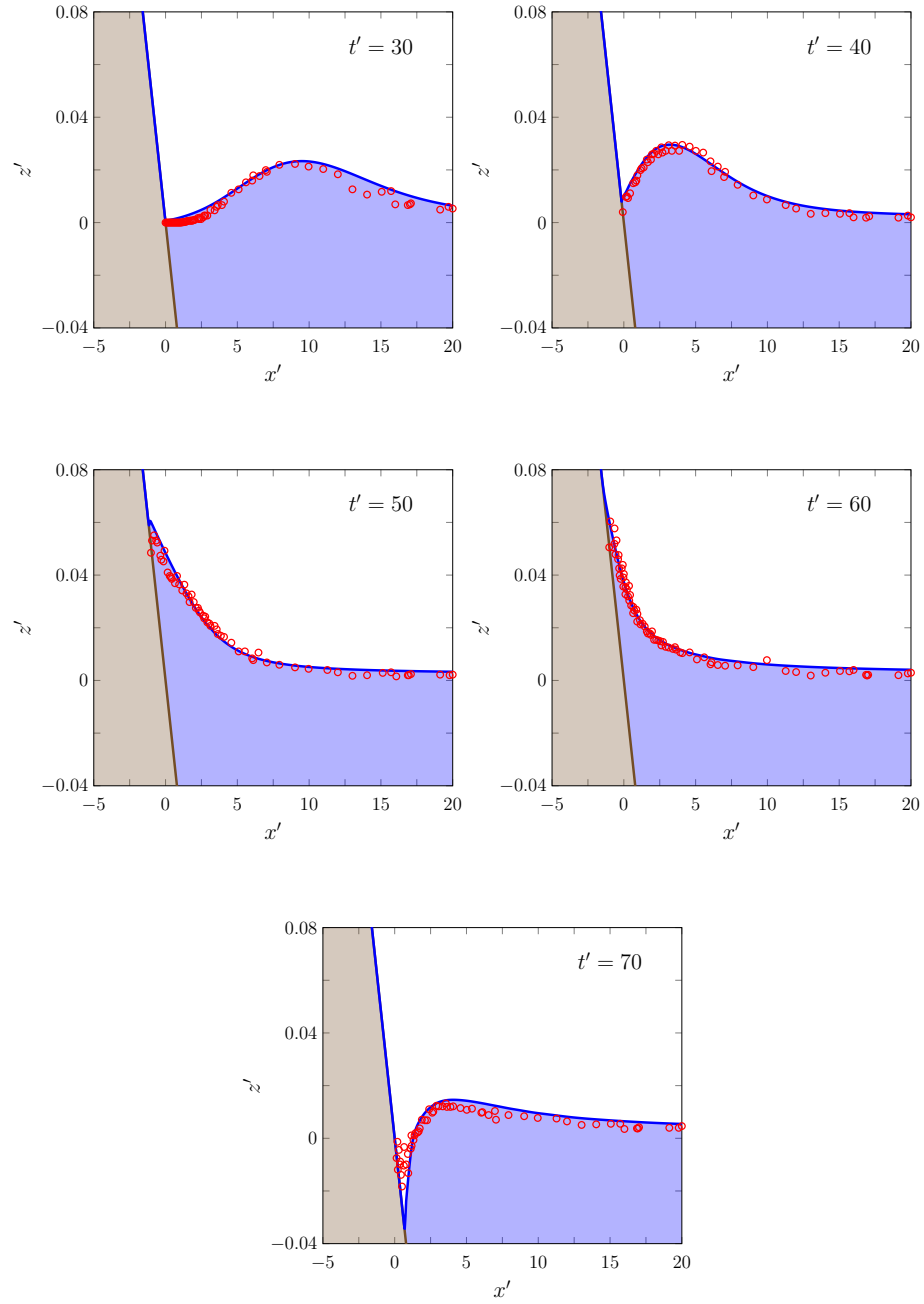
15

Figure 9: A comparison of the water surface profiles $w'(x', t')$ for the experiment (○) and the numerical solution (▢) over the bed (▢) at various times.

16

Quantity	$C^* \left(\boldsymbol{q}^0 \right)$	$C^* \left(\boldsymbol{q}^* \right)$	$C^* \left(\boldsymbol{q}^0, \boldsymbol{q}^* \right)$
h'	240.416965344	240.416965376	1.33×10^{-10}
$u'h'$	-0.319050138516	0.318891991793	4.96×10^{-4}
G'	-0.319073723126	0.318886191223	5.88×10^{-4}
$\mathcal{H}'$	-118.389958187	-118.3900028	3.77×10^{-7}

Table 1: Initial and final ($t' = 200$) total amounts and the conservation error for the conserved quantities in the numerical solution of the run-up experiment. Here the absolute value of the total amount of uh and G are taken in the error as the wave is reflected off the beach.