A CFL-Free Explicit Scheme with Compression for Linear Hyperbolic Equations

Ronald A. DeVore   Hong Wang   Jiang-Guo Liu
Hong Xu

Abstract

We develop an unconditionally stable, explicit numerical scheme for linear hyperbolic equations, which arises as an advection-reaction equation in porous medium flows. The derived scheme generates accurate numerical solutions even if large time steps are used, and conserves mass. Furthermore, this scheme has the capability of performing adaptive compression while maintaining the accuracy of the compressed solution and mass conservation. Numerical results show the strong potential of the method.

KEYWORDS: characteristic methods, hyperbolic equations, multiresolution analysis, wavelet decomposition

1   Introduction

Hyperbolic partial differential equations describe the displacement of oil by injected fluid in petroleum recovery, the subsurface contaminant transport and remediation, aerodynamics, and many other applications. Because of the moving steep fronts present in their solutions, the numerical treatment of these equations often presents severe difficulties. In late 1920's, Courant, Friedrichs, and Lewy proved the famous result that there are no explicit, unconditionally stable, consistent finite difference schemes for hyperbolic partial differential equations [2]. Consequently, most numerical methods developed for hyperbolic equations are explicit but subject to the well-known Courant-Friedrichs-Lewy (CFL) condition. Very small time steps often have to be used in numerical simulations to meet the stability requirement of the numerical methods. Implicit methods could be unconditionally stable, but an algebraic system must be solved at each time step in order to obtain numerical solutions.

In this paper we develop an unconditionally stable, explicit numerical scheme for linear hyperbolic equations. The derived scheme generates accurate numerical solutions even if large time steps are used, and conserves mass. Furthermore, this scheme has the capability of carrying out adaptive compression while maintaining the accuracy of the numerical solution and mass conservation. Computational results are presented to show the strong potential of the method developed.

2   A Reference Equation

We consider the initial-value problem for the linear hyperbolic equation

$$
\begin{aligned}
u_t + \nabla \cdot (\mathbf{v}u) + Ku &= f(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \mathbb{R}^d \times (0, T], \\
u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d,
\end{aligned} \tag{2.1}
$$

where $\mathbf{v}(\mathbf{x}, t) := (V_1(\mathbf{x}, t), V_2(\mathbf{x}, t), \ldots, V_d(\mathbf{x}, t))$ is a fluid velocity field, $K(\mathbf{x}, t)$ is a first-order reaction coefficient, $f(\mathbf{x}, t)$ is a given source function, $u(\mathbf{x}, t)$ is the unknown function, and $u_0(\mathbf{x})$ is a given initial condition. $\mathbf{x} := (x_1, \ldots, x_d)$, $u_t := \frac{\partial u}{\partial t}$, $\nabla := (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_d})$. We assume that $u_0(\mathbf{x})$ and $f(\mathbf{x}, t)$ have compact support, so the exact solution $u(\mathbf{x}, t)$ also has compact support for any finite time $t > 0$.

We define a uniform partition on the temporal interval $[0, T]$ by $t_n := n\Delta t$ with $\Delta t = T/N$. If we choose the test functions $w(\mathbf{x}, t)$ to be of compact support in space, to vanish outside the interval $(t_{n-1}, t_n]$, and to be discontinuous in time at time t_{n-1}, the weak formulation for Eq. (2.1) is

$$
\begin{aligned}
\int_{\mathbb{R}^d} u(\mathbf{x}, t_n) w(\mathbf{x}, t_n) d\mathbf{x} &- \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} u(w_t + \mathbf{v} \cdot \nabla w - Kw)(\mathbf{x}, t) d\mathbf{x} dt \\
&= \int_{\mathbb{R}^d} u(\mathbf{x}, t_{n-1}) w(\mathbf{x}, t_{n-1}^+) d\mathbf{x} + \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{x}, t) w(\mathbf{x}, t) d\mathbf{x} dt,
\end{aligned} \tag{2.2}
$$

where $w(\mathbf{x}, t_{n-1}^+) := \lim_{t \to t_{n-1}^+} w(\mathbf{x}, t)$ takes into account the fact that $w(\mathbf{x}, t)$ is discontinuous in time at time t_{n-1}.

Motivated by the ELLAM framework of Celia *et al* [1], the test functions w in Eq. (2.2) are chosen from the solution space of the adjoint equation of Eq. (2.1). Let $\mathbf{y} = \mathbf{r}(\theta; \bar{\mathbf{x}}, \bar{t})$, with $\bar{t} \in [t_{n-1}, t_n]$, be the characteristic determined by

$$
\frac{d\mathbf{y}}{d\theta} = \mathbf{v}(\mathbf{y}, \theta), \quad \text{with} \quad \mathbf{y}|_{\theta=\bar{t}} = \bar{\mathbf{x}}. \tag{2.3}
$$

Then the adjoint equation of Eq. (2.1) is rewritten as

$$-\frac{d}{d\theta}w(\mathbf{r}(\theta;\bar{\mathbf{x}},\bar{t}),\theta) + K(\mathbf{r}(\theta;\bar{\mathbf{x}},\bar{t}),\theta)w(\mathbf{r}(\theta;\bar{\mathbf{x}},\bar{t}),\theta) = 0,$$
$$w(\mathbf{r}(\theta;\bar{\mathbf{x}},\bar{t}),\theta)|_{\theta=\bar{t}} = w(\bar{\mathbf{x}},\bar{t}),$$
(2.4)

leading to the following expression for w

$$w(\mathbf{r}(\theta;\bar{\mathbf{x}},\bar{t}),\theta) = w(\bar{\mathbf{x}},\bar{t})e^{-\int_\theta^{\bar{t}} K(\mathbf{r}(\gamma;\bar{\mathbf{x}},\bar{t}),\gamma)d\gamma}.$$
(2.5)

Therefore, once the test functions $w(\mathbf{x},t_n)$ are specified at time t_n, they are determined completely and in fact vary exponentially along the characteristic $\mathbf{r}(\theta;\mathbf{x},t_n)$ for $\theta \in [t_{n-1},t_n]$.

To avoid confusion in the derivation, we replace the dummy variables $\mathbf{x}$ and t in the second term on the right-hand side of Eq. (2.2) by $\mathbf{y}$ and θ. For any $\mathbf{y} \in \mathbb{R}^d$, there exists an $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{y} = \mathbf{r}(\theta;\mathbf{x},t_n)$. We obtain

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} f(\mathbf{y},\theta)w(\mathbf{y},\theta)d\mathbf{y}d\theta \\
&= \int_{\mathbb{R}^d} \int_{t_{n-1}}^{t_n} f(\mathbf{r}(\theta;\mathbf{x},t_n),\theta)w(\mathbf{r}(\theta;\mathbf{x},t_n),\theta)\left|\frac{\partial\mathbf{r}(\theta;\mathbf{x},t_n)}{\partial\mathbf{x}}\right|d\theta d\mathbf{x} \\
&= \int_{\mathbb{R}^d} F(\mathbf{x},t_n)w(\mathbf{x},t_n)\left[\int_{t_{n-1}}^{t_n} e^{-K(\mathbf{x},t_n)(t_n-\theta)}d\theta\right]d\mathbf{x} + E_1(f,w) \\
&= \int_{\mathbb{R}^d} \Lambda(\mathbf{x},t_n)f(\mathbf{x},t_n)w(\mathbf{x},t_n)d\mathbf{x} + E_1(f,w).
\end{aligned}
$$
(2.6)

Here $\Lambda(x,t_n) = (1 - e^{-K(\mathbf{x},t_n)\Delta t})/K(\mathbf{x},t_n)$ if $K(\mathbf{x},t_n) \neq 0$, or Δt otherwise. $E_1(f,w)$ is the local truncation error.

Incorporating Eq. (2.6) into Eq. (2.2), we obtain a reference equation

$$
\begin{aligned}
&\int_{\mathbb{R}^d} u(\mathbf{x},t_n)w(\mathbf{x},t_n)d\mathbf{x} \\
&= \int_{\mathbb{R}^d} u(\mathbf{x},t_{n-1})w(\mathbf{x},t_{n-1}^+)d\mathbf{x} + \int_{\mathbb{R}^d} \Lambda(\mathbf{x},t_n)(fw)(\mathbf{x},t_n)d\mathbf{x} + E(w),
\end{aligned}
$$
(2.7)

where $E(w) = \int_{t_{n-1}}^{t_n} \int_{\mathbb{R}^d} u[w_t + \mathbf{v}\cdot\nabla w - Kw](\mathbf{x},t)d\mathbf{x}dt + E_1(f,w)$.

3   Multiresolution and Wavelet Decomposition

To develop a numerical scheme based on the reference equation (2.7), we need to define the trial and test functions at time t_n. To do so, we recall the notions of multiresolution analysis and wavelet decompositions.

In the standard Fourier analysis, L^2-functions are represented by linear combinations of sines and cosines. In 1910, Haar studied the representation of L^2-functions by step functions taking values ± 1 [5]. In the 1980's, these ideas were explored and developed further into the theory of wavelets. The first wavelets were introduced in early 1980's by Stromberg [10] and Morlet *et al* [8]. One of the best ways of constructing wavelets is multiresolution analysis. This approach began in image-processing [9, 11] and was introduced into mathematics by Mallat [6]. Multiresolution analysis was used by Daubechies [3] to construct compactly supported orthogonal wavelets with arbitrary smoothness, which include the Haar wavelets as the simplest case.

Definition A sequence of closed subspaces $\{\mathcal{V}_j\}_{j\in\mathbf{Z}}$ ($\mathbf{Z}$-the set of all integers) of $L^2(\mathbb{R})$ is a *Multiresolution Analysis* if

(a) these spaces are nested: $\mathcal{V}_j \subset \mathcal{V}_{j+1}$ for $\forall j \in \mathbf{Z}$;

(b) these spaces are dense in $L^2(\mathbb{R})$: $\overline{\cup}_{j\in\mathbf{Z}}\mathcal{V}_j = L^2(\mathbb{R})$ and $\cap_{j\in\mathbf{Z}}\mathcal{V}_j = \emptyset$;

(c) $\mathcal{V}_0$ is invariant under integer shifts: $f(\cdot) \in \mathcal{V}_0 \Longrightarrow f(\cdot - k) \in \mathcal{V}_0$, $\forall k \in \mathbf{Z}$;

(d) $\mathcal{V}_j$ is obtained from $\mathcal{V}_0$ by dilation: $f(\cdot) \in \mathcal{V}_j \Longleftrightarrow f(2^{-j}\cdot) \in \mathcal{V}_0$, $\forall j \in \mathbf{Z}$;

(e) $\mathcal{V}_0$ is generated by a single (the so-called scaling) ϕ and its translates: $\{\phi_{0,k} : k \in \mathbf{Z}\}$ is an orthonormal basis in $\mathcal{V}_0$, where

$$\phi_{j,k}(x) := 2^{j/2}\phi(2^j x - k), \quad j, k \in \mathbf{Z}. \tag{3.1}$$

The conditions (d) and (e) in the definition implies that the family $\{\phi_{j,k} : k \in \mathbf{Z}\}$ forms an orthonormal basis for $\mathcal{V}_j$. Let $\mathcal{P}_j : L^2(\mathbb{R}) \longrightarrow \mathcal{V}_j$ be the orthogonal projection operator, then the conditions (a) and (b) in the definition concludes that for any $f \in L^2(\mathbb{R})$

$$\lim_{j\to+\infty} \|\mathcal{P}_j f - f\|_{L^2(\mathbb{R})} = 0, \tag{3.2}$$

i.e.,

$$\mathcal{P}_j f = \sum_{k\in\mathbf{Z}}(f, \phi_{j,k})\phi_{j,k} \longrightarrow f, \qquad \text{as} \quad j \to +\infty. \tag{3.3}$$

Let $\mathcal{W}_{j-1}$ be the orthogonal complement of $\mathcal{V}_{j-1}$ in $\mathcal{V}_j$

$$\begin{aligned} \mathcal{V}_j &= \mathcal{V}_{j-1} \oplus \mathcal{W}_{j-1} = \ldots \\ &= \mathcal{V}_J \oplus \mathcal{W}_J \oplus \mathcal{W}_{J+1} \oplus \ldots \oplus \mathcal{W}_{j-1}, \quad \text{for } j > J. \end{aligned} \tag{3.4}$$

Then, if $f \in L^2(\mathbb{R})$, the telescoping sum

$$\mathcal{P}_j f = \mathcal{P}_J f + \sum_{i=J}^{j-1} (\mathcal{P}_{i+1} f - \mathcal{P}_i f) \tag{3.5}$$

represents $\mathcal{P}_j f \in \mathcal{V}_j$ as an element of $\mathcal{V}_J \oplus \mathcal{W}_J \oplus \mathcal{W}_{J+1} \oplus \ldots \oplus \mathcal{W}_{j-1}$. By (3.2), Eq. (3.5) provides an approach of approximating arbitrary L^2-functions by sequences of functions from $\mathcal{V}_J$ and $\mathcal{W}_i$ for $J \leq j < +\infty$. More importantly, the spaces $\mathcal{W}_j$ can be generated by a single (the so-called wavelet) function ψ. In other words, ψ and its integer translates $\psi_{0,k}$, with $\psi_{j,k}$ being defined by

$$\psi_{j,k}(x) := 2^{j/2} \psi(2^j x - k), \quad j, k \in \mathbf{Z}, \tag{3.6}$$

constitute an orthonormal basis for $\mathcal{W}_0$. For each fixed j, the $\psi_{j,k}$ ($k \in \mathbf{Z}$) form an orthogonal basis for $\mathcal{W}_j$.

4   A CFL-Free Explicit Scheme with Conservative Compression

In this section we develop a CFL-free, explicit scheme for the initial-value problem (2.1). We define the finite-dimensional space $\mathcal{S}_j(\mathbb{R}^d)$ to be the tensor product of the one-dimensional space $\mathcal{V}_j$

$$\mathcal{S}_j(\mathbb{R}^d) := (\mathcal{V}_j(\mathbb{R}))^d = \text{span} \{\Phi_{j,\mathbf{k}}(\mathbf{x})\}_{\mathbf{k}=(k_1,k_2,\ldots,k_d) \in \mathbf{Z}^d}, \tag{4.1}$$

where $\Phi_{j,\mathbf{k}}(\mathbf{x}) := \phi_{j,k_1}(x_1) \phi_{j,k_2}(x_2) \ldots \phi_{j,k_d}(x_d)$ and $\phi_{j,k}(x) \in \mathcal{V}_j(\mathbb{R})$ is defined by Eq. (3.1).

With these preparations, we replace the exact solution u in Eq. (2.7) by the trial functions $U(\mathbf{x}, t_n) \in \mathcal{S}_j(\mathbb{R}^d)$ and drop the error term $E(w)$ in Eq. (2.7), leading to the following numerical scheme: find $U(\mathbf{x}, t_n) \in \mathcal{S}(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} U(\mathbf{x}, t_n) w(\mathbf{x}, t_n) d\mathbf{x}$$
$$= \int_{\mathbb{R}^d} U(\mathbf{x}, t_{n-1}) w(\mathbf{x}, t_{n-1}^+) d\mathbf{x} + \int_{\mathbb{R}^d} \Lambda(\mathbf{x}, t_n) f(\mathbf{x}, t_n) w(\mathbf{x}, t_n) d\mathbf{x}, \tag{4.2}$$
$$\forall w(\mathbf{x}, t_n) = \Phi_{j,\mathbf{k}}(\mathbf{x}) \in \mathcal{S}_j(\mathbb{R}^d).$$

Although the integrals in Eq. (4.2) are formally defined on the space $\mathbb{R}^d$, they are in fact supported locally since the test functions $w(\mathbf{x}, t_n) = \Phi_{j,\mathbf{k}}(\mathbf{x})$ have compact support. Moreover, because the solution $U(\mathbf{x}, t_{n-1})$ and $f(\mathbf{x}, t_n)$

have compact support, Eq. (4.2) only needs a finite number of operations at each time level. Thirdly, because the scaling functions $\Phi_{j,\mathbf{k}}(\mathbf{x})$, $\mathbf{k} \in \mathbf{Z}^d$, form an orthonormal system for $\mathcal{S}_j(\mathbb{R}^d)$, the scheme (4.2) is explicit. It can be proven that the scheme (4.2) is unconditionally stable.

It is well known that in practice the solutions of linear transport equations are fairly smooth outside some very small (dynamic) regions, where the solutions can develop steep fronts or even shock discontinuities. Therefore, some kind of local refinement or adaptive techniques should be used. Extensive research has been conducted in this regard in the finite difference, finite element, and finite volume methods. In the current context, this can be realized in terms of compression in a very natural way. For example, we can apply a thresholding to the scheme (4.2) to minimize the number of equations that need to be formulated and solved, and so significantly improve the computational efficiency. However, a direct application to the scaling functions $\Phi_{j,\mathbf{k}}$ will introduce mass balance error. In contrast, because of their vanishing moments, applying such a technique to the wavelets does not affect mass balance. Hence, we use the wavelet decomposition (3.4) to choose another basis for the space $\mathcal{V}_j(\mathbb{R})$

$$\mathcal{V}_j(\mathbb{R}) = \mathrm{span}\left\{\phi_{l,k}^{(\alpha)}(x)\right\}_{k \in \mathbf{Z}, \ J \leq l \leq j, \ \alpha = 0,1}, \tag{4.3}$$

where $\phi_{l,k}^{(0)}(x) := \phi_{l,k}(x)$ and $\phi_{l,k}^{(1)}(x) := \psi_{l,k}(x)$. This leads to another expression for the space $\mathcal{S}_j(\mathbb{R}^d)$

$$\mathcal{S}_j(\mathbb{R}^d) = \mathrm{span}\left\{\Phi_{l,\mathbf{k}}^{(\boldsymbol{\alpha})}(\mathbf{x})\right\}_{\mathbf{k} \in \mathbf{Z}^d, \ J \leq l \leq j, \ \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \{0,1\}^d}. \tag{4.4}$$

In this case, the numerical scheme is still defined by Eq. (4.2), but using a different basis for $\mathcal{S}_j(\mathbb{R}^d)$. We begin by projecting the initial condition $u_0(\mathbf{x})$ onto the space $\mathcal{S}_j(\mathbb{R}^d)$ to obtain an approximant $\tilde{U}_0(\mathbf{x}, t_0)$. For $n = 1, 2, \ldots, N$, we apply a thresholding technique to compress $\tilde{U}_0(\mathbf{x}, t_{n-1})$ to obtain $U(\mathbf{x}, t_{n-1})$. Then, using its wavelet expansion as an error indicator, we solve the reduced scheme (4.2) at time step t_n with the minimal number of basis functions used to obtain $\tilde{U}(\mathbf{x}, t_n)$. Next, we compress $\tilde{U}(\mathbf{x}, t_n)$ to obtain $U(\mathbf{x}, t_n)$. We continue this process until we reach the final time step $t_N = T$.

5 Numerical Experiments

In this section we present numerical results to show the potential of the numerical scheme developed in this paper. The spatial domain is $\Omega = (0, 5)$, the time interval is $[0, T] = [0, 2]$, a velocity field $V = 1 + 0.1x$ is chosen, $K = f = 0$. The initial condition $u_0(x)$ is
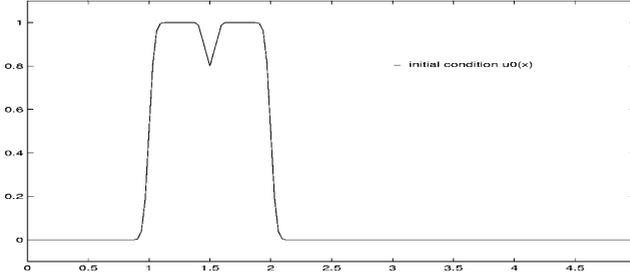

Figure 1:The initial condition $u_0(x)$.

$$u_0(x) := \text{Imp}(x; 1, 2, 0.05) - 0.2\, \text{Tri}(x; 1.5, 0.1), \tag{5.1}$$

with the hat function $\text{Tri}(x; a, b)$ and the impulse function $\text{Imp}(x; a, b, c)$ being defined by

$$\text{Tri}(x; a, b) := \begin{cases} 1 + \dfrac{x - a}{b} & \text{if } a - b \leq \quad x \quad \leq a, \\ 1 - \dfrac{x - a}{b} & \text{if } \quad\;\; a \leq \quad x \quad \leq a + b, \\ 0 & \text{otherwise,} \end{cases} \tag{5.2}$$

and

$$\text{Imp}(x; a, b, c) := \frac{1}{2} \left[\text{erf}\left(\frac{x - 1}{0.05} \right) - \text{erf}\left(\frac{x - 2}{0.05} \right) \right], \tag{5.3}$$

where the error function $\text{erf}(x) := (2/\sqrt{\pi}) \int_0^x e^{-z^2} dz$. The initial condition $u_0(x)$ is plotted in Figure 1.

In the numerical experiments, we use a time step of $\Delta t = 0.4$. The coarse and fine spatial levels are $J = 3$ and $j = 7$, respectively. This leads to a maximal Courant number of 192. At the final time $T = 2$, the L^1 error of the uncompressed solution is 0.00477 while the L^1 error of the compressed solution is 0.00515. The compression ratio is 30.19. We plot the uncompressed and compressed solutions against the analytical solution in Figure 2.
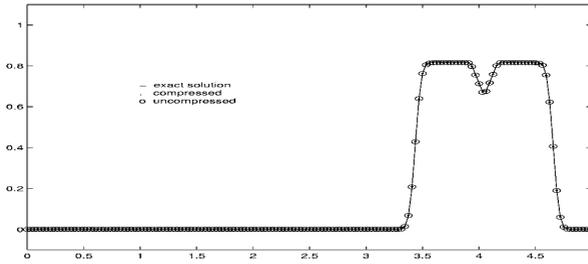
FIGURE 2:The analytical, uncompressed, and compressed numerical solutions.

References

[1] Celia, M. A., Russell, T. F., Herrera, I., Ewing R. E., An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation, *Advances in Water Resources* **13** (1990), 187–206.

[2] Courant, R., Friedrichs, K. O., and Lewy, H., Uber die partiellen differenzen-gleichungen der mathematisches physik, *Math. Annalen* **100** (1928), 32–74.

[3] Daubechies, I., Orthogonal bases of compactly supported wavelets, *Commun. Pure and Appl. Math.* **41** (1988), 909–996.

[4] I. Daubechies: *Ten Lectures on Wavelets*, **61** of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[5] Haar, A., Zur Theorie der Orthogonalen Funktionen-Systeme,*Math. Ann.* **69** (1910), 331–371.

[6] Mallet, S. G., Multiresolution and wavelet orthonormal bases in $L^2(\mathbb{R})$, *Trans. Amer. Math. Soc.* **315** (1989), 69–87.

[7] Meyer, Y., *Ondelettes et Opérateurs*, **1** and **2**, Hermann, Paris, 1992.

[8] Morlet, J., Arens, G., Fourgeau, I., and Giard, D., Wave propagation and sampling theory, *Geophysics* **47** (1982), 203–236.

[9] Smith, M. J. and Barnwell, D. P., Exact reconstruction for tree-structured subband coders, *IEEE Trans. ASSP* **34** (1986), 434–441.

[10] Stromberg, J. O., A modified Franklin system and higher order spline on $\mathbb{R}^n$ as unconditional bases for Hardy spaces, Becker W. *et al* (eds.), Wadsworth Math. Series, Belmont CA, 1982, 475–493.

[11] Vetterli, M., Filter banks allowing perfect reconstruction,*Signal Processing* **10** (1986), 219–244.