INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids (in press)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1070

MPDATA: Gauge transformations, limiters and monotonicity§

L. G. Margolin[1,*,†] and M. Shashkov[2,‡]

[1] *MS F699, Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.*
[2] *MS B284, Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.*

SUMMARY

MPDATA is a flexible and computationally efficient methodology that has been applied to advection, remapping, and full fluid solvers. In this paper, we extend the fundamental concept, iterated upwind compensation of error, to incorporate a new degree of freedom—that of gauge transformations—with the goal of constructing a monotonicity preserving option for MPDATA. We further augment this scheme by adapting the idea of summing the recursive relations to improve the overall accuracy. This process leads to a theoretical connection of this MPDATA scheme to flux-limited algorithms. Published in 2005 by John Wiley & Sons, Ltd.

KEY WORDS: advection; remapping; nonoscillatory methods

1. INTRODUCTION

The multidimensional positive definite advection transport algorithm (MPDATA) was introduced by Smolarkiewicz [1] as a flexible and computationally efficient algorithm for modelling advection. Through more than twenty years, its use has been extended to nonoscillatory interpolation [2], as a full fluid solver [3], and recently as a remapper for arbitrary Lagrange–Eulerian codes [4, 5]. Most recently, MPDATA has been successfully used as a basis for implicit turbulence modelling [6, 7]. A review of MPDATA and its many options can be found in Reference [8].

The basic idea underlying all MPDATA applications is the compensation of error by iterated upwind approximation. Diffusion-like truncation errors are rewritten in the form of an

*Correspondence to: L. G. Margolin, MS F699, Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.
†E-mail: len@lanl.gov
‡E-mail: shashkov@lanl.gov

Contract/grant sponsor: U.S. Department of Energy; contract/grant number: W-7405-ENG-36

Received 30 March 2005
Revised 17 June 2005
Published in 2005 by John Wiley & Sons, Ltd. *Accepted 25 June 2005*

advection term, and compensated in a second upwind step. The velocity of this second step results from a modified equation analysis of the first-order donor-cell algorithm; it is termed an antidiffusive velocity, or a pseudo velocity as it has no physical interpretation. The form of the pseudo velocity is constrained by the requirements of computational stability, but is not unique, and many MPDATA options are created through modification of this term.

The goal of this paper is to describe a new degree of freedom for the form of the pseudo velocity, through the introduction of a gauge transformation in the advected field. In Section 2, we will briefly review the basic ideas of MPDATA. In Section 3, we introduce the gauge transformation. We describe the modified 'gauge' MPDATA algorithm in Section 4 and show that it preserves monotonicity, but is more diffusive than the basic algorithm. In Section 5, we adapt the idea of analytic recursion to the gauge algorithm. This modification improves the accuracy of the overall algorithm. At the same time, it demonstrates a theoretical connection between MPDATA and flux-limited schemes, which is further discussed in Section 6. We conclude the paper in Section 7 with some ideas for exploiting our theoretical framework to treat problems of advection and remapping on multidimensional, unstructured grids.

Throughout the paper, we illustrate our results with examples of constant velocity transport. However, all statements regarding donor cell and basic MPDATA have been shown to apply for the case of variable velocity [8]. Further, all new proofs in this paper concerning the preservation of monotonicity, both for the gauge algorithm and for the flux-limited algorithm, are derived for the case of spatially variable velocity.

2. A NONLINEAR APPROXIMATION OF UNITY

MPDATA belongs to the general class of Lax–Wendroff methods, where the second-order error of donor cell advection is estimated by means of modified equation analysis (MEA; see Reference [9]) and compensated. Consider the one-dimensional linear transport equation

$$\psi_t = -(u\psi)_x \tag{1}$$

The donor cell approximation to (1) is

$$\psi_j^{n+1} - \psi_j^n = -F(\psi_j^n, \psi_{j+1}^n, u_{j+(1/2)}) + F(\psi_{j-1}^n, \psi_j^n, u_{j-(1/2)}) \tag{2}$$

Here, superscripts indicate the discretized time level, and integer subscripts indicate the centres of the computational cells. Δt is the computational time step and Δx is the cell size. The flux function is defined as

$$F(\psi_{j-1}, \psi_j, u) \equiv \psi_{j-1}\left(\frac{C+|C|}{2}\right) + \psi_j\left(\frac{C-|C|}{2}\right) \tag{3}$$

where $C \equiv u\Delta t/\Delta x$ is the dimensionless Courant number.

Let us now assume (for the moment) that u is constant in space. A straightforward MEA shows that the donor cell algorithm more accurately approximates the advection–diffusion equation

$$\psi_t = -(u\psi)_x + K\psi_{xx} \tag{4}$$

where the diffusion coefficient $K \equiv (\Delta x^2/2\Delta t)(|C| - C^2)$. Here, the model PDE has been used to rewrite $\psi_{tt} = u^2\psi_{xx}$. More generally, we use the modified equation to eliminate the time derivatives for the sake of deriving higher-order algorithms, see, e.g. Reference [10]. To ensure a computationally stable algorithm, we require $K \geqslant 0$; then it is necessary that $|C| \leqslant 1$, which is the well known Courant–Friedrichs–Lewy (CFL) condition on the time step.

From this brief discussion, we see that donor cell is a first-order accurate algorithm—i.e. truncation terms first appear linear in the spatial scale Δx or the computational time step Δt. Further, donor cell is sign preserving. For example, if $u > 0$, we can rewrite (2) in the form

$$\psi_j^{n+1} = C\psi_{j-1}^n + (1 - C)\psi_j^n \tag{5}$$

i.e. $\psi_j^n > 0 \,\forall\, j$ and $C \in [0, 1] \rightarrow \psi_j^{n+1} > 0 \,\forall\, j$. We remark that in the case of constant velocity, donor cell has the stronger property of preserving monotonicity—e.g. if

$$\psi_{j-1}^n < \psi_j^n < \psi_{j+1}^n, \rightarrow \psi_{j-1}^{n+1} < \psi_j^{n+1} < \psi_{j+1}^{n+1}$$

However, when u varies in space arbitrarily, donor cell is only sign-preserving.

To improve the accuracy of the donor cell algorithm, one must compensate the second-order error. A straightforward linear approximation

$$K\,\psi_{xx} \approx \frac{K}{\Delta x^2}(\psi_{j+1} - 2\psi_j + \psi_{j-1}) \tag{6}$$

leads to the classic Lax–Wendroff algorithm, which is oscillatory and does not preserve sign of the advected field. Smolarkiewicz' idea [1] is to use the properties of donor cell in the approximation by writing the diffusive error in the form of an advective flux

$$K\psi_{xx} = \frac{\partial}{\partial x}\left(\frac{K}{\psi}\frac{\partial \psi}{\partial x}\psi\right) = \frac{\partial}{\partial x}(v\psi) \tag{7}$$

where the pseudo velocity $v \equiv (K/\psi)\psi_x$. Numerically, we write

$$v_{j+(1/2)} = \frac{K}{\Delta x}\,\frac{\psi_{j+1} - \psi_j}{\psi_{j+1} + \psi_j} \tag{8}$$

Then one can compensate the truncation in a second donor cell pass

$$\psi_j^{n+1} - \tilde{\psi}_j^n = -F(\psi_j^n, \psi_{j+1}^n, v_{j+(1/2)}) + F\left(\psi_{j-1}^n, \psi_j^n, v_{j-(1/2)}\right) \tag{9}$$

Here, $\tilde{\psi}$ is the result of the first, donor cell pass as in (5)

$$\tilde{\psi}_j = C\psi_{j-1}^n + (1 - C)\psi_j^n$$

Writing out (9) in detail, one sees that we have approximated the analytic value of unity at the node $j - \frac{1}{2}$, e.g. for $u > 0$

$$1 = \frac{\psi}{\psi} \approx \frac{2\psi_{j-1}}{\psi_j + \psi_{j-1}} \tag{10}$$

This nonlinear approximation is the key to the nonoscillatory behaviour of MPDATA; see, for example, the discussion of Godunov's theorem in Reference [11]. We have chosen to write the flux functions in terms of ψ^n rather than $\tilde{\psi}$. The two choices are equivalent to the second-order of accuracy, but may require different time step restrictions to ensure positivity.

Note that when the advected field ψ is everywhere positive, or negative, $|v| \leqslant |u|$. Thus the stability of the donor cell pass guarantees the stability of the second pass.

Let us now make the following observations. First, the MPDATA algorithm (2) and (9) is applicable for the more general case that the physical velocity u is not constant in space; the constant K will then become variable and a more restrictive time step criterion may be required. Second, we note that the pseudo velocity v is not constant in space, even when the physical velocity is constant. Thus the MPDATA algorithm (2) and (9) is second-order accurate, sign preserving, but not monotonicity preserving. Third, we note that if the field ψ is not uniformly of one sign, then the stability of the algorithm is lost.

The last two observations are critical if one intends to apply the algorithm to the momentum equation. Several modifications are proposed in Reference [8] to ensure the necessary boundedness of the pseudo velocity for variably signed fields. However, it has been found that while sign preservation is sufficient for many applications, the stronger property of monotonicity preservation is sometimes essential. This is particularly true in the case of the momentum equation. At present, our strategy has been to augment MPDATA with a specially tailored flux corrected transport procedure [12]. This is an effective strategy to ensure preservation of monotonicity, but does significantly complicate the computational procedure and also changes somewhat the character of MPDATA. In the next section, we present an idea for an alternative strategy that we believe retains more of the philosophical approach of MPDATA.

3. THE GAUGE TRANSFORMATION

There is a close connection between the preservation of sign and the preservation of monotonicity. This can be illustrated by means of the following simple example. Consider the transport of a square pulse by a constant velocity through a nonzero (constant) background. The MPDATA result, shown in the left panel of Figure 1 after 40 time steps at Courant number $C = 0.5$, shows unphysical oscillations, both at the foot and the top of the pulse.

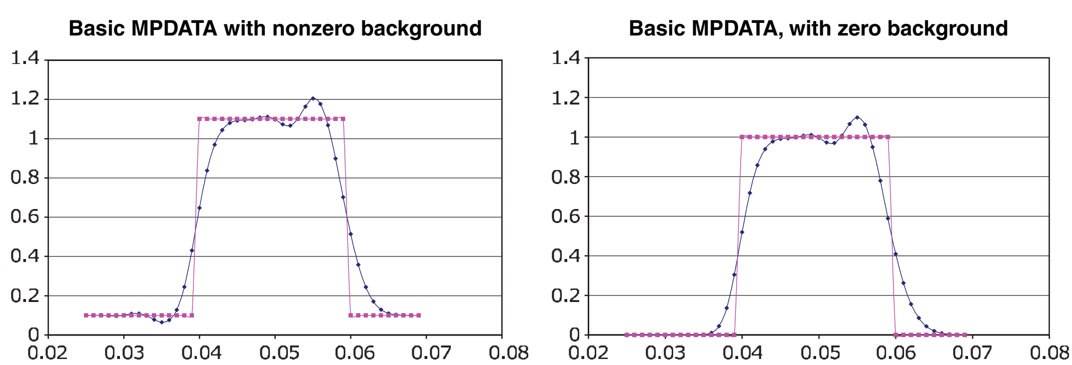

Figure 1. The left panel shows a square pulse, transported with a constant velocity through a nonzero background by basic MPDATA. Note the undershoots at the foot of the pulse and the overshoots at the top of the pulse. The right panel shows the same problem, but with no constant background. The undershoots are gone, but the overshoots remain.

Consider next what happens as the value of the constant background decreases—the magnitude of the oscillations as the foot of the pulse decrease, and ultimately vanish as the background vanishes. This is shown in the right panel of Figure 1.

This behaviour should be seen as somewhat surprising; the problem is linear and homogeneous, and should be insensitive to the background. The actual behaviour is the result of the nonlinearity of MPDATA itself, as described in the previous section, and specifically in the nonlinear approximation (10). One could write this approximation more generally at the node $j - \frac{1}{2}$ in the form

$$1 = \frac{\psi - \psi_G}{\psi - \psi_G} \approx \frac{2(\psi_{j-1} - \psi_G)}{\psi_j + \psi_{j-1} - 2\psi_G} \tag{11}$$

where ψ_G is any constant. Now let us return to the simple problem of the beginning of the section, and set ψ_G equal to the background value. The result (not shown) is identical to that of the right panel of Figure 1. This illustrates that the implicit choice $\psi_G = 0$ in (10) is what makes the value $\psi = 0$ special in the basic MPDATA algorithm. We refer to (11) as the global gauge transformation.

From an implementation point of view, the gauge transformation alters both the form of the pseudo velocity and of the flux function. Now

$$v^G_{j-(1/2)} = \frac{K}{\Delta x} \frac{\psi_j - \psi_{j-1}}{\psi_j + \psi_{j-1} - 2\psi_G} \tag{12}$$

and the new flux function

$$F(\psi_{j-1}, \psi_j, C) \equiv (\psi_{j-1} - \psi_G)\left(\frac{C + |C|}{2}\right) + (\psi_j - \psi_G)\left(\frac{C - |C|}{2}\right) \tag{13}$$

where $C \equiv v^G_{j-(1/2)} \Delta t/\Delta x$.

More generally, we can choose ψ_G as a function of both space and time without altering the consistency of the approximation with the model equation. This would allow us to set the zero value of the field locally; we refer to this as the local gauge transformation. Note that from (12), the gauge value ψ_G lives at the node.

There is one more idea to discuss in this section. Referring back to Figure 1, right panel, we see that changing the background value has eliminated the oscillations at the foot of the pulse, but has not affected the oscillations at the top of the pulse. To eliminate these oscillations, one might first multiply the entire field by -1, use the negative of the maximum pulse height as the gauge, transport the field, and then restore the original field by again multiplying by -1. Of course, this procedure will not mitigate the oscillations at the foot of the pulse.

The point is that, in general, there are two gauge values that must be considered, a minimum value and a maximum value. This corresponds to the two bounds that must be enforced to ensure the preservation of monotonicity. We will refer to these as $\psi^{\min}_{j+(1/2)}$ and $\psi^{\max}_{j+(1/2)}$. We note that the gauge, like the new pseudo velocity which will depend on the gauge, lives at the cell vertices rather than cell centres. In the next section, we will justify how these values are chosen and used to preserve monotonicity.

4. MONOTONICITY

Because the choice of gauge is local, we expect it should depend only on local data. We will define

$$\psi^{\min}_{j+(1/2)} \equiv \min(\psi_{j-1}, \psi_j, \psi_{j+1}, \psi_{j+2}) \tag{14}$$

and

$$\psi^{\max}_{j+(1/2)} \equiv \max(\psi_{j-1}, \psi_j, \psi_{j+1}, \psi_{j+2}) \tag{15}$$

and demonstrate that these choices are effective. Recall that the gauge is associated with the vertex, not the cell centre. We refer to the algorithm defined by (12) and (13) with the above choices of gauges, as the *basic gauge* algorithm.

4.1. Constant velocity case

Let us consider the case of a monotonically increasing distribution

$$\psi_{j-1} \leqslant \psi_j \leqslant \psi_{j+1} \leqslant \psi_{j+2}$$

where all time levels not explicitly stated are assumed to be at level n. To begin, we will also assume that the velocity is positive and constant in space, $u \geqslant 0$. Let us first show that

$$\psi^{n+1}_j \geqslant \psi_{j-1} \tag{16}$$

Explicitly,

$$\psi^{n+1}_j = (1-C)\psi_j + C\psi_{j-1} + A^{\min}_{j-(1/2)} - A^{\min}_{j+(1/2)} \tag{17}$$

where the antidiffusive fluxes associated with the minimum gauge

$$A^{\min}_{j-(1/2)} = \frac{K\Delta t}{\Delta x^2}\left[\frac{\psi_j - \psi_{j-1}}{0.5\,(\psi_j + \psi_{j-1}) - \psi^{\min}_{j-(1/2)}}\right][\tilde{\psi}_{j-1} - \psi^{\min}_{j-(1/2)}] \tag{18}$$

and

$$\begin{aligned} A^{\min}_{j+(1/2)} &= \frac{K\Delta t}{\Delta x^2}\left[\frac{\psi_{j+1} - \psi_j}{0.5\,(\psi_{j+1} + \psi_j) - \psi^{\min}_{j+(1/2)}}\right][\tilde{\psi}_j - \psi^{\min}_{j+(1/2)}] \\ &= \frac{K\Delta t}{\Delta x^2}\left[\frac{\psi_{j+1} - \psi_j}{0.5\,(\psi_{j+1} + \psi_j) - \psi^{\min}_{j+(1/2)}}\right][(1-C)\psi_j + C\psi_{j-1} - \psi^{\min}_{j+(1/2)}] \end{aligned} \tag{19}$$

Note that in this case, all antidiffusive fluxes are positive. Thus, to prove the inequality (16), we can ignore $A^{\min}_{j-(1/2)}$. Further, $\psi^{\min}_{j+(1/2)} = \psi_{j-1}$. Substituting into (17) and rearranging, we derive

$$\psi^{n+1}_j - \psi_{j-1} \geqslant (1-C)(\psi_j - \psi_{j-1})\left[1 - \frac{2K\Delta t}{\Delta x^2}\left(\frac{R_j}{R_j + 2}\right)\right] \tag{20}$$

Here, the cell-centred quantity $R_j \equiv (\psi_{j+1} - \psi_j)/(\psi_j - \psi_{j-1})$. This quantity is typically used in flux-limiting schemes, see Reference [13]. Finally, we note that, in the case of monotonic distributions, $R \geqslant 0$ and always $2K\Delta t/\Delta x^2 \leqslant 1$. Thus, all the factors on the right-hand side of (20) are positive and the inequality is proved.

A similar proof is easily constructed to show that $\psi_{j+1}^{n+1} \leqslant \psi_{j+2}$, if the antidiffusive fluxes are constructed with the maximum gauge. Proofs of the satisfaction of bounds for the additional cases where the (constant) velocity is negative, and/or where the distribution is monotonically decreasing, are also easily constructed and will not be detailed here.

4.2. *Spatially variable velocity*

The situation is a little more subtle when the velocity field is not constant in space. In this case, it is not clear whether the continuous solution, i.e. to (1), will be monotone. From the numerical point of view, the donor cell solution will not *necessarily* preserve monotonicity; then the antidiffusive fluxes may not have the same sign as the donor cell fluxes, which was an essential element of the proof above. To summarize, the numerical solution can only be guaranteed to preserve monotonicity *if* the donor cell scheme does. Since it is usually not possible to determine whether the continuous solution should be monotone in the context of a numerical simulation, we make the practical choice of asserting that monotonicity should be preserved by the gauge scheme whenever it has been preserved by the donor cell scheme. We note that this choice is equivalent to the definition of monotonicity preservation in the flux corrected transport schemes (FCT, [14]), where oscillations in the high-order scheme are only deemed unphysical when they are not found in the low-order reference scheme.

Let us now return to the example above allowing the dimensionless velocity to vary in space, which we will now write as $C_{j+(1/2)}$. We assume that the donor cell solution preserves monotonicity

$$\tilde{\psi}_j - \psi_{j+(1/2)}^{\min} = \psi_j(1 - C_{j-(1/2)}) - \psi_{j-1}(1 - C_{j-(3/2)}) \geqslant 0$$

which implies $A_{j-(1/2)}^{\min} \geqslant 0$. Then we have

$$\psi_{j-1}^{n+1} - \psi_{j-1} \geqslant \psi_j(1 - C_{j-(1/2)}) - \psi_{j-1}(1 - C_{j-(3/2)}) \left[1 - \frac{2K\Delta t}{\Delta x^2}\left(\frac{R_j}{R_j + 2}\right)\right] \geqslant 0 \qquad (21)$$

That is, when the donor cell solution preserves monotonicity, then the gauge MPDATA does also.

It is easy to verify that, when the spatial distribution is not monotone, i.e. there is a local maximum or minimum, then the antidiffusive fluxes vanish and the donor cell solution is used. There are theoretical ways to decide whether the flux associated with the minimum gauge $A_{j+(1/2)}^{\min}$ or that associated with the maximum gauge $A_{j+(1/2)}^{\max}$ should be used, for example based on the second derivative (curvature) of the field. However, as a practical matter, one should always choose the flux with the smaller magnitude.

The result of a square wave propagation (same problem as in the left panel of Figure 1) using the *basic* gauge MPDATA algorithm is shown in the left panel of Figure 2. The pulse is now monotone; however the edges of the pulse are less steep. That is, the gauge MPDATA algorithm is more diffusive than basic MPDATA. This observation is quantified in Table I. In the next section, we discuss how one might improve this result. This will lead to the main theoretical result of this paper, a relation between the gauge transformation and flux limiting.

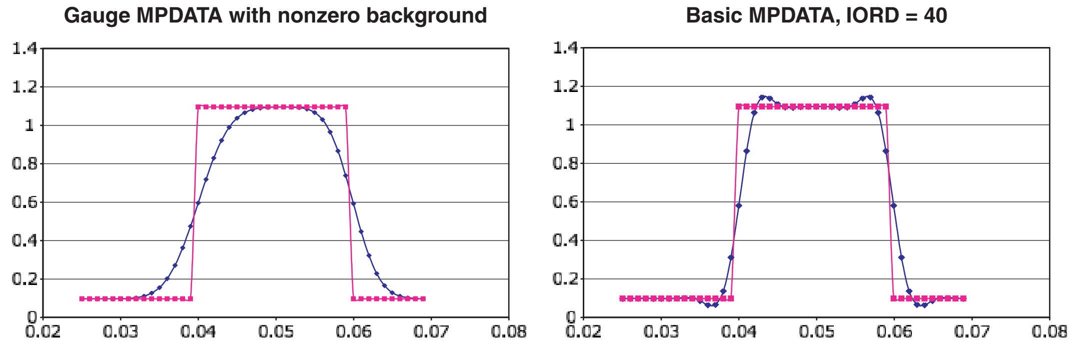

Figure 2. The left panel shows a square pulse, transported with a constant velocity through a constant background by gauge MPDATA. Note the monotone profile. The right panel shows the same problem, but with a large number of iterative corrections defined by the parameter IORD = 40, in basic MPDATA. Note that the solution is now more accurate and symmetric than that shown in the left panel of Figure 1, but still does not preserve monotonicity.

Table I. The errors in $L1$ and $L2$ norms for the various algorithms simulating the problem of Figure 1.

Algorithm	L_1 error	L_2 error
Donor cell	5.14	1.240
Basic MPDATA, IORD = 2	2.96	0.963
Basic MPDATA, IORD = 40	2.43	0.848
Gauge MPDATA	4.73	1.190
Flux-limited	2.34	0.824

5. RELATION OF THE GAUGE TRANSFORMATION TO FLUX LIMITING

5.1. Recursion

It was noted in the original MPDATA Reference [1] that the error after the second pass can also be analysed, and can be compensated in a similar process. Indeed, this process can be repeated indefinitely. Each pass further reduces the error, but does not increase the order of the algorithm beyond second-order. The process is described, e.g. in Reference [8], by the index IORD, where IORD = 1 is donor cell, IORD = 2 is the basic MPDATA algorithm, etc.

In Reference [10], an interesting extension to the basic MPDATA algorithm is described. Effectively, the results of an infinite number of corrective passes are estimated analytically, and applied in a single corrective step. The details of this process are different for the gauge algorithm, but the overall idea carries over readily.

Although one can readily do a MEA, there is an even simpler way to derive the appropriate recursion relation of the gauge algorithm. Consider the case of monotonically increasing field

$$\psi_{j-1} \leqslant \psi_j \leqslant \psi_{j+1}$$

positive velocity $u > 0$, and the minimum gauge. We will focus only on the interface $j+\frac{1}{2}$, and so it is not necessary to assume that u is constant in space. In this case, we have $\psi^{\min}_{j+(1/2)} = \psi_{j-1}$. Employing the gauge MPDATA algorithm (17), the (dimensionless) antidiffusive flux associated with the minimum gauge for the first corrective pass is

$$A^{\min}_{j+(1/2)} = \frac{K_{j+(1/2)}\Delta t}{\Delta x^2}\mathscr{A}(\tilde{\psi}_j - \psi_{j-1}) = V^{(1)}_{j+(1/2)}(\tilde{\psi}_j - \psi_{j-1}) \tag{22}$$

where

$$\mathscr{A} \equiv \frac{\psi_{j+1} - \psi_j}{0.5(\psi_{j+1} + \psi_j) - \psi_{j-1}}, \quad V^{(1)}_{j+(1/2)} \equiv \frac{K_{j+(1/2)}\Delta t}{\Delta x^2}\mathscr{A} \tag{23}$$

Rather than go to the continuous PDE, let us note that a third-order accurate representation of the continuous antidiffusive flux is the Lax–Wendroff corrective flux (i.e. when $\mathscr{A} \to 1$)

$$A^{\mathrm{LW}}_{j+(1/2)} \approx \frac{K_{j+(1/2)}\Delta t}{\Delta x^2}(\psi_{j+1} - \psi_j) \tag{24}$$

Therefore, the error in our approximation after one corrective step IORD = 2 is

$$\mathrm{error}^{(1)}_{j+(1/2)} = A^{\mathrm{LW}}_{j+(1/2)} - A^{\min}_{j+(1/2)} = \frac{K_{j+(1/2)}\Delta t}{\Delta x^2}(\psi_{j+1} - \psi_j)\left[1 - \frac{\mathscr{A}}{\mathscr{R}_j}\right] \tag{25}$$

where

$$\mathscr{R}_j \equiv \frac{\psi_{j+1} - \psi_j}{\tilde{\psi}_j - \psi_{j-1}}$$

Note that $\mathscr{R}$ differs from R defined below (20) because of the different time levels of the various terms. Thus the next corrective step would set

$$V^{(2)}_{j+(1/2)} = \left(1 - \frac{\mathscr{A}}{\mathscr{R}_j}\right) V^{(1)}_{j+(1/2)} \tag{26}$$

Following this chain of reasoning, it is easy to show that the recursive relation we are looking for is

$$V^{(k)}_{j+(1/2)} = \left(1 - \frac{\mathscr{A}}{\mathscr{R}_j}\right)^{k-1} V^{(1)}_{j+(1/2)} \tag{27}$$

where $V^{(1)}_{j+(1/2)}$ is defined in (22).

The idea of the recursion then is to calculate the sum of all the terms

$$V^T \equiv \sum_{\mathrm{IORD}=2}^{\infty} V^{(\mathrm{IORD})} = \frac{K_{j+(1/2)}\mathscr{A}}{1 - \left(1 - \dfrac{\mathscr{A}}{\mathscr{R}_j}\right)} = K_{j+(1/2)}\mathscr{R}_j \tag{28}$$

This leads to the final result

$$A^{\min}_{j+(1/2)} = V^T \frac{\tilde{\psi}_j - \psi_{j-1}}{\Delta x} = \frac{K_{j+(1/2)}\Delta t}{\Delta x^2}(\psi_{j+1} - \psi_j) = A^{\mathrm{LW}}_{j+(1/2)} \tag{29}$$

that is, *the result of analytically summing all the antidiffusive corrections is the Lax–Wendroff corrective flux*. On the one hand, this is the expected result since each correction pass reduces the error. On the other hand, it appears that we have used a succession of monotonicity preserving passes, and ended with a nonmonotonic result.

The resolution of this apparent paradox is readily understood. In performing the sum, we must insure that the test for monotonicity, e.g. Equation (20), remains valid. Allowing for the possibility that we would want to limit the sum V^T by a *velocity* limiter ϕ, the condition that we must enforce is

$$\psi_j^{n+1} - \psi_{j-1} \geqslant (1 - C)(\psi_j - \psi_{j-1}) \left[1 - \frac{CR_j\,\phi}{2}\right] \geqslant 0 \tag{30}$$

This leads to the following form for the velocity limiter:

$$\phi(R, C) = \max\left[0, \min\left(\frac{2}{R_j C}, 1\right)\right] \tag{31}$$

5.2. The connection to flux limiting

A limiter for the pseudo velocity is interesting, and will be given a physical interpretation in the next section. However, it is easy to reformulate our results in more conventional terms. Returning to the example of Section 4—i.e. monotonically increasing field and constant positive velocity, let us exploit our result that the recursion relations lead to the Lax–Wendroff form, to write

$$\psi_j^{n+1} = (1 - C)\psi_j + C\psi_j + L_{j-(1/2)} - L_{j+(1/2)} \tag{32}$$

where

$$L_{j+(1/2)} \equiv A^{\mathrm{LW}}_{j+(1/2)} \Phi_{j+(1/2)} = \frac{K\Delta t}{\Delta x^2}\, \Phi_{j+(1/2)}(\psi_{j+1} - \psi_j) \tag{33}$$

In the equation above, Φ is a flux-limiter, i.e. a function that limits the Lax–Wendroff correction to the donor cell flux to insure the preservation of monotonicity, see Reference [13]. Then the condition that $\psi_j^{n+1} \geqslant \psi_{j-1}$ leads directly to the result

$$\Phi_{j+(1/2)} \leqslant \frac{2}{R_j C}$$

Further, the condition that $\psi_{j+1}^{n+1} \leqslant \psi_{j+2}$ leads directly to the result

$$\Phi_{j+(1/2)} \leqslant \frac{2(R_{j+1} + C)}{C(1 - C)}$$

It is now easy to verify that in the rest of the cases, where one considers monotonically decreasing flow, and/or negative velocities, the same two conditions appear, if one replaces $C \to |C|$. As noted previously, it is necessary to enforce the more restrictive of these conditions. Further, we will limit $\Phi \leqslant 1$. It is possible to allow larger values, see, e.g. Reference [13] where a maximum value of 2 is allowed. However, the Lax–Wendroff flux is the most accurate (to second-order) and there seems to be no reason to allow a larger flux. Setting an upper limit of 1 is also consistent with Sweby's constraint that $R = 1 \Rightarrow \Phi = 1$. Finally, we must require $\Phi \geqslant 0$ to reproduce the result that in nonmonotone regions $R < 0$, we use the donor cell solution. Putting all this together leads to our final result for the flux limiter

$$\Phi_{j+(1/2)} = \max\left[0, \min\left(\frac{2}{R_j\,|C|}, \frac{2(R_{j+1} + |C|)}{|C|(1 - |C|)}, 1\right)\right] \tag{34}$$

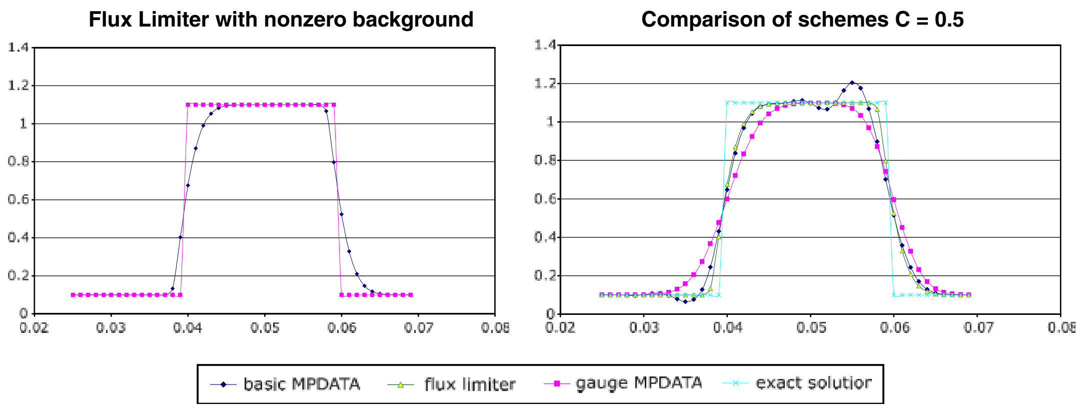

Figure 3. The left panel shows the same calculation as the left panel in Figure 1, but using the flux limiter. The solution remains monotone, but is less diffusive. The right panel shows an overlay to facilitate comparison of the various advection methods.

The result from a simulation of the square wave pulse using the flux-limiter algorithm (same problem as in the left panel of Figure 1) is shown in the left panel of Figure 3. Note the improvement over the gauge algorithm, in terms of steepness of the edges. However, there is now a pronounced lack of symmetry in the pulse. This is attributable to higher-order, dispersive errors, see discussion in Reference [10]. A composite plot of the solutions of transporting a square pulse through a nonzero background using the basic MPDATA, the gauge MPDATA and the flux-limiter algorithms is shown in the right panel of Figure 3.

6. DISCUSSION POINTS

(1) We begin our discussion by noting an interesting interpretation of our results in summing the recursion relation. Using (28) and (31), we have

$$C + V_T\,\Phi \leqslant C + C(1 - C) = 1 \tag{35}$$

That is, monotonicity is the result of limiting the 'total' Courant number $C + V_T\,\Phi$ in the same way that limiting the Courant number itself ensures stability.

(2) We note that the form of the flux-limiter (34) is more general than those reported by Sweby [13] in that it depends on the Courant number C, and that it depends on both the 'upstream' and the 'downstream' ratios R_j and R_{j+1}.

(3) All the simulation results shown in this paper have used a Courant number $C = 0.5$. We have explored all methods over a range of values $C \in [0.1, 0.8]$. In general, the larger the Courant number, the better the results. However, the qualitative results that we have reported are unchanged.

(4) The errors in the discrete L_1 norm

$$E_1 = \sum_{j \in \text{mesh}} |\psi_j - \psi_{\text{exact}}|$$

and in the discrete L_2 norm

$$E_2 = \sqrt{\sum_{j \in \text{mesh}} (\psi_j - \psi_{\text{exact}})^2}$$

are shown in Table I. Here ψ_{exact} is the square pulse transported downstream, but otherwise unchanged. All MPDATA schemes are more accurate than the donor cell. Note that the gauge algorithm, though preserving monotonicity, is less accurate than basic MPDATA.

(5) It is interesting to consider the behaviour of the basic and gauge algorithms if more *explicit* recursions are used—i.e. if IORD $\to \infty$. In the case of the basic MPDATA, the iterative process leads to reduction of the error. However, the asymptotic limit is quickly reached, by IORD equals 3 or 4. A calculation of the square wave propagation is shown in the right panel of Figure 2. The transported pulse is much more symmetric than the result shown in the left panel of this figure, but remains oscillatory.

Increasing IORD is more problematic in the gauge algorithm. For one thing, the reduction of error is much slower than for the basic algorithm. However, it is not possible to recover the flux-limited result of Figure 3 (left panel) by simply increasing IORD. Whereas the stability of the donor step in the basic algorithm guarantees the stability of the corrective steps, the stability of the corrective steps of the gauge algorithm is not guaranteed, but requires limiting of the pseudo velocity. This was effectively demonstrated in Section 5.1. If the velocity is limited, then the corresponding asymptotic result is the flux-limited result. Note in Table I that the flux-limited result is more accurate than the basic MPDATA result with many iterations.

(6) Returning to Equation (11), we can also consider the limit $\psi_G \to -\infty$. In this limit, the gauge algorithm becomes equivalent to a two-step Lax–Wendroff method. We refer to this limit, when combined with FCT [12], as the 'infinite gauge' option.

7. CONCLUSION

We began this paper by summarizing the fundamental idea upon which MPDATA is based—that of upwind compensation of diffusive truncation error. We described a simple generalization of the basic algorithm that transformed the basic, sign-preserving method to a monotonicity-preserving method. We then derived a theoretical connection between our new gauge algorithm and more conventional approaches to monotonicity-preserving algorithms based on flux-limiting.

The gauge transformation represents a potentially valuable addition to the available options for MPDATA. However, more work remains to extend it to multiple dimensional simulations, and to integrate it with other MPDATA options. We believe that the theoretical connection that we have established is also important. There is, of course, a large literature on flux-limiters, which may lead to new improvements in MPDATA. In addition, we see this flow of ideas going in both directions. MPDATA has several advantageous features: it is truly multidimensional, i.e. unsplit spatially; it can be extended to irregular grids and to unstructured grids [15]. The capability to deal with such grids is particularly important for its use in remapping, see Reference [4].

We conclude this paper with two specific examples that illustrate how the connection to MPDATA may impact flux limited schemes. First, we note that the quantities R_j, which play an essential role in the flux-limiter form (34), are a one-dimensional idea. The extension to multiple dimensions on a structured grid requires some form of direction splitting. The further extension to unstructured grids is problematic. However, in the context of the gauge algorithm, it became clear that the ψ terms in R not directly adjacent to the interface arise from the gauge condition and not from estimates of the gradient; this idea is readily extended to multiple dimensions and unstructured grids, since the gauge can simply be defined in terms of the next nearest neighbours of an interface.

A second direction concerns relaxing the strict imposition of monotonicity. The value of nonoscillatory algorithms is well recognized in the computational fluid dynamics (CFD) community. Basic MPDATA is, strictly speaking, nonoscillatory only in the neighbourhood of zero although it does have the important property of ensuring nonlinear stability. Our focus in this paper has been the construction of an MPDATA scheme that is always nonoscillatory. However, there has been an effort in the CFD community to relax these conditions, see References [16, 17]. A physical motivation for this has been suggested in Reference [6], where the connection between the mathematical condition of monotonicity preservation and the physical constraints of the second law of thermodynamics is explored. In that paper, it is shown that strict monotonicity preservation is a sufficient, but not a necessary condition for the second law, and it is suggested that relaxing this constraint would better represent the physical process of backscatter in turbulent flows. As a practical matter, relaxing monotonicity in the gauge algorithm could be easily achieved by widening the neighbourhood over which the min and max gauge values are chosen.

8. HISTORICAL NOTE

The concept of applying gauge transformations to advection algorithms was described in an earlier publication in the context of relaxing monotonicity constraints in flux-limited schemes [17]. However, the idea originated in the framework of MPDATA more than fifteen years ago [18]. This work, like much research done at Livermore and Los Alamos National Laboratories, remains unpublished. The LLNL report is now available on the web; however, those early results are entirely reproduced in this paper.

ACKNOWLEDGEMENTS

The authors acknowledge the early contributions of P.K. Smolarkiewicz and C.W. Beason to this work, as well as discussions with W.J. Rider. Los Alamos National Laboratory is operated by the US Department of Energy under Contract W-7405-ENG-36.

REFERENCES

1. Smolarkiewicz PK. A simple positive definite advection scheme with small implicit diffusion. *Monthly Weather Review* 1983; **111**:479–486.
2. Smolarkiewicz PK, Grell GA. A class of monotone interpolation schemes. *Journal of Computational Physics* 1992; **101**:431–440.
3. Smolarkiewicz PK, Margolin LG. On forward-in-time differencing for fluids. *Monthly Weather Review* 1993; **125**:1847–1859.
4. Margolin LG, Shashkov M. Second-order sign-preserving conservative interpolation (remapping) on general grids. *Journal of Computational Physics* 2003; **184**:266–298.
5. Margolin LG, Shashkov M. Remapping, recovery and repair on a staggered grid. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:4139–4155.

6. Margolin LG, Rider WJ. A rationale for implicit turbulence modeling. *International Journal for Numerical Methods in Fluids* 2002; **39**:821–841.
7. Margolin LG, Smolarkiewicz PK, Sorbjan Z. Large eddy simulations of convective boundary layers using nonoscillatory differencing. *Physica D* 1999; **133**:390–397.
8. Smolarkiewicz PK, Margolin LG. MPDATA: a finite-difference solver for geophysical flows. *Journal of Computational Physics* 1998; **140**:459–480.
9. Hirt C. Heuristic stability theory for finite difference equations. *Journal of Computational Physics* 1968; **2**: 339–355.
10. Margolin LG, Smolarkiewicz PK. Antidiffusive velocities for multipass donor cell advection. *SIAM Journal on Scientific Computing* 1988; **20**:907–929.
11. LeVeque RJ. *Numerical Methods for Conservation Laws* (2nd edn). Birkhauser: Boston, 1999.
12. Smolarkiewicz PK, Grabowski WW. The multidimensional positive definite advection transport algorithm: nonoscillatory option. *Journal of Computational Physics* 1990; **86**:355–375.
13. Sweby PK. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM Journal on Numerical Analysis* 1984; **210**:995–1010.
14. Boris JP, Book DL. Flux-corrected transport: 1. SHASTA. A fluid transport algorithm that works. *Journal of Computational Physics* 1973; **11**:38–69.
15. Smolarkiewicz PK, Szmelter J. MPDATA: an edge-based unstructured-grid formulation. *Journal of Computational Physics* 2005, in press.
16. Harten A, Engquist B, Osher S, Chakravarthy S. Uniformly high order accurate essentially nonoscillatory schemes—III. *Journal of Computational Physics* 1987; **71**:231–301.
17. Rider WJ, Margolin LG. Simple modifications of monotonicity-preserving limiters. *Journal of Computational Physics* 2001; **174**:473–488.
18. Margolin LG, Beason CW. DPDC: a second-order monotone scheme for advection. *Lawrence Livermore National Laboratory Report UCRL-99731*, 1988. Available at: http://www.llnl.gov/tid/lof/documents/pdf/208688.pdf