

Chapter 4

Finite volume schemes

In this chapter the principle of finite volume schemes and their ability to conserve the integral of the transported properties will be introduced (section 4.1). The principles of monotonicity and total variation will be introduced in section 4.2. Monotonicity-preserving schemes (here flux-limiter methods) will be discussed in this given framework in section 4.3.

We have so far discussed the one-dimensional advection equation

$$\partial_t S + u \partial_x S = 0 \quad (4.1)$$

with constant advection velocity u . In the ocean, however, this assumption is generally not correct. Therefore, general three-dimensional discretisation schemes have to consider velocity fields with variable velocity.

4.1 Conservative formulation

Let $S(x(t), y(t), z(t), t)$ denote a tracer in the ocean being advected with the flow fields $(u, v, w) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$. Then the transport equation for the advective transport of S (neglecting diffusive mixing) is given by setting the total derivative of S to zero:

$$D_t S(x(t), y(t), z(t), t) = \partial_t S + u \partial_x S + v \partial_y S + w \partial_z S = 0. \quad (4.2)$$

Together with the incompressibility condition

$$\partial_x u + \partial_y v + \partial_z w = 0 \quad (4.3)$$

this may be transformed into

$$\partial_t S + \partial_x(uS) + \partial_y(vS) + \partial_z(wS) = 0. \quad (4.4)$$

This is the so-called conservative form in contrast to (4.2) which is the advective form. (4.4) may also be written in vector notation:

$$\partial_t S + \vec{\nabla} \cdot (\vec{u}S) = 0, \quad (4.5)$$

such that integration over a finite volume V gives according to the Gauss theorem

$$\partial_t \int_V S dV + \int_V \vec{\nabla} \cdot (\vec{u}S) dV = \partial_t \int_V S dV + \int_A \vec{u}S \cdot d\vec{A} = 0, \quad (4.6)$$

where A is the surface of the finite volume V and $d\vec{A}$ is the vector normal to the surface with length dA where dA is the surface area of the surface element. The meaning of (4.6) is that all changes of S contained in the volume are determined by the advective flow of S through the surface A , i.e., S is conserved inside V .

Let us consider a Cartesian grid with

$$V_{i,j,k} = \Delta x_i \Delta y_j \Delta z_k \quad (4.7)$$

and

$$\begin{aligned} \Delta x_i &= x_{i+1/2} - x_{i-1/2}, \\ \Delta y_j &= y_{j+1/2} - y_{j-1/2}, \end{aligned} \quad (4.8)$$

$$\Delta z_k = z_{k+1/2} - z_{k-1/2},$$

being a discretisation of a volume of water V ,

$$V = \sum_{i,j,k} V_{i,j,k} \quad (4.9)$$

with

$$i_{\min} \leq i \leq i_{\max}, \quad j_{\min} \leq j \leq j_{\max}, \quad k_{\min} \leq k \leq k_{\max}. \quad (4.10)$$

Integration of (4.4) over $V_{i,j,k}$ leads to

$$\begin{aligned} d_t (V_{i,j,k} S_{i,j,k}) + (uS)|_{i+1/2,j,k} - (uS)|_{i-1/2,j,k} + (vS)|_{i,j+1/2,k} - (vS)|_{i,j-1/2,k} \\ + (wS)|_{i,j,k+1/2} - (wS)|_{i,j,k-1/2} = 0 \end{aligned} \quad (4.11)$$

with

$$S_{i,j,k} = V_{i,j,k}^{-1} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \int_{z_{k-1/2}}^{z_{k+1/2}} S dz dy dx. \quad (4.12)$$

Assuming for simplicity, that there are no fluxes of S through the surface of V (closed basin), summation of (4.11) over all discrete finite volumes $V_{i,j,k}$ results

in cancellation of all internal fluxes such that:

$$\begin{aligned}
& d_t \sum_{i,j,k} V_{i,j,k} S_{i,j,k} \\
& + \sum_{j,k} \left((uS)|_{i_{\max}+1/2,j,k} - (uS)|_{i_{\min}-1/2,j,k} \right) \\
& + \sum_{i,k} \left((vS)|_{i,j_{\max}+1/2,k} - (vS)|_{i,j_{\min}-1/2,k} \right) \\
& + \sum_{i,j} \left((wS)|_{i,j,k_{\max}+1/2} - (wS)|_{i,j,k_{\min}-1/2} \right) = 0 \\
\Rightarrow & d_t \sum_{i,j,k} V_{i,j,k} S_{i,j,k} = 0.
\end{aligned} \tag{4.13}$$

With this, the total mass of S is exactly conserved even in the numerical discretisation.

Methods which are based on the principle demonstrated above are called finite volume methods.

Let us now return to one-dimensional problems. The conservative form of a first-order PDE is then

$$\partial_t \psi + \partial_x f(\psi) = 0. \tag{4.14}$$

A finite-volume discretisation of (4.14) is given by

$$\frac{\phi_j^{n+1} - \phi_j^n}{\Delta t} + \frac{F_{j+1/2} - F_{j-1/2}}{\Delta x} \tag{4.15}$$

with $F_{j\pm 1/2}$ being numerical approximations to $f(\psi((j \pm 1/2)\Delta x))$ of the form

$$\begin{aligned}
F_{j+1/2} &= F(\phi_{j-p}^n, \phi_{j-p+1}^n, \dots, \phi_{j+q+1}^n), \\
F_{j-1/2} &= F(\phi_{j-p-1}^n, \phi_{j-p}^n, \dots, \phi_{j+q}^n),
\end{aligned} \tag{4.16}$$

with p and q being integers. If F is sufficiently smooth and

$$F(\psi_0, \psi_0, \dots, \psi_0) = f(\psi_0), \tag{4.17}$$

then (4.15) converges to a weak¹ solution of (4.14), as shown by *Lax and Wendroff* [1960].

For the linear advection equation in conservative form the flux from equation (4.14) is simply of the following form:

$$f(\psi) = c\psi. \tag{4.18}$$

¹weak solutions are a concept for hyperbolic partial differential equations discontinuous solutions, e.g. shock waves, see *LeVeque* [1992].

The first-order first-order upstream discretisation for this flux is simply (for $c > 0$):

$$F_{j+1/2}^{n,up} = c\phi_j^n, \quad (4.19)$$

which gives the standard upstream formulation for the discretisation of (4.14), see e.g. (2.8). The Lax-Wendroff scheme is constructed by subtracting the numerical diffusion caused by the first-order upstream scheme, which has a diffusivity of

$$\nu_{num} = \frac{c}{2}\Delta x(1 - \mu), \quad (4.20)$$

such the numerical anti-diffusive flux must have the form

$$F_{j+1/2}^{n,diff} = -\nu_{num} \frac{\phi_{j+1}^n - \phi_j^n}{\Delta x} = -\frac{c}{2}(1 - \mu)(\phi_{j+1}^n - \phi_j^n). \quad (4.21)$$

Thus the Lax-Wendroff flux is of the form:

$$F_{j+1/2}^{n,lw} = F_{j+1/2}^{n,up} - F_{j+1/2}^{n,diff} = c\phi_j^n + \frac{c}{2}(1 - \mu)(\phi_{j+1}^n - \phi_j^n). \quad (4.22)$$

The complete Lax-Wendroff scheme is thus

$$\frac{\phi_j^{n+1} - \phi_j^n}{\Delta t} + c \frac{\phi_j^n - \phi_{j-1}^n}{\Delta x} + \frac{c}{2}(1 - \mu) \frac{\phi_{j+1}^n - 2\phi_j^n + \phi_{j-1}^n}{\Delta x} = 0, \quad (4.23)$$

which is the well-known result from assignment 3. On the same principles, another numerical scheme for the linear advection equation is constructed, which uses an upstream instead of a central approximation of the numerical anti-diffusion term. The resulting flux is of this so-called *Warming and Beam* [1976] scheme is

$$F_{j+1/2}^{n,wb} = c\phi_j^n + \frac{c}{2}(1 - \mu)(\phi_j^n - \phi_{j-1}^n). \quad (4.24)$$

such that the complete scheme is of the following form:

$$\frac{\phi_j^{n+1} - \phi_j^n}{\Delta t} + c \frac{\phi_j^n - \phi_{j-1}^n}{\Delta x} + \frac{c}{2}(1 - \mu) \frac{\phi_j^n - 2\phi_{j-1}^n + \phi_{j-2}^n}{\Delta x} = 0, \quad (4.25)$$

which is like the Lax Wendroff scheme of second-order accuracy and even more stable with $0 \leq \mu \leq 2$.

In figure 4.1, the upstream and the Lax-Wendroff scheme are applied to the linear advection of a rectangle and a Gaussian bell, and the numerical results are compared to the analytical solution and to the performance of highly accurate monotone schemes, which we are going to develop in the next sections. The advantage of these non-oscillating schemes with low numerical diffusion is obvious.

4.2 Monotonicity and Total Variation

We have already seen (e.g., in figure 4.1) that many numerical discretisations of the one-dimensional linear advection problem develop unphysical oscillations.

In fact, the only scheme we have seen so far which does not have this property is the first order upstream scheme, which is however very diffusive.

The concept of monotonicity helps to find numerical schemes without unphysical oscillations. A function g is generally called *monotone* when $g(x) \leq g(y)$ for $x \leq y$. Formulating a numerical scheme as

$$\phi_j^{n+1} = H(\phi_{j-p}^n, \phi_{j-p+1}^n, \dots, \phi_{j+q+1}^n) \quad (4.26)$$

and transferring the concept of monotonicity to finite difference schemes, the condition of monotonicity would be formulated as

$$\partial_{\phi_i} H(\phi_{j-p}, \phi_{j-p+1}, \dots, \phi_{j+q+1}) \geq 0, \quad (4.27)$$

for $j-p \leq i \leq j+q+1$. For linear schemes this has the consequence that they are monotone only if all coefficients are non-negative. This may be demonstrated for the first-order upstream approximation to the linear advection equation in flux form,

$$\partial_t \psi + \partial_x(x\psi) = 0, \quad (4.28)$$

which is

$$\phi_j^{n+1} = (1 - \mu)\phi_j^n + \mu\phi_{j-1}^n \quad (4.29)$$

with $\mu = c\Delta t/\Delta x$. It can be easily seen that the monotonicity condition (4.27) is fulfilled for $0 \leq \mu \leq 1$ which is identical to the stability condition.

There are however many stable numerical schemes for the linear advection equation (e.g. the Lax-Wendroff scheme) which are not monotone. It can actually be proven (*Godunov* [1959]) that

any monotone scheme is at most first-order accurate.

In the following it will be the goal to construct numerical methods which are monotone (and thus first-order accurate) only at sharp gradients which tend to create oscillations, but higher-order accurate at sufficiently smooth locations.

For doing so, we introduce the concept of total variation (TV), which is defined for any discrete solution as

$$\text{TV}(\phi) = \sum_{j=j_{\min}}^{j_{\max}} |\phi_{j+1} - \phi_j|. \quad (4.30)$$

A numerical method is *total variation non-increasing* (or, less accurate but nonetheless more often used *total variation diminishing* or TVD), when

$$\text{TV}(\phi^{n+1}) \leq \text{TV}(\phi^n). \quad (4.31)$$

Another class of numerical schemes are the *monotonicity preserving* schemes, with

$$\phi_j^n \leq \phi_{j+1}^n \implies \phi_j^{n+1} \leq \phi_{j+1}^{n+1}, \quad (4.32)$$

for all j . Thus, monotonicity preserving schemes generate numerical solutions without unphysical oscillations.

It can be shown that

$$\text{monotone} \implies \text{TVD} \implies \text{monotonicity preserving}. \quad (4.33)$$

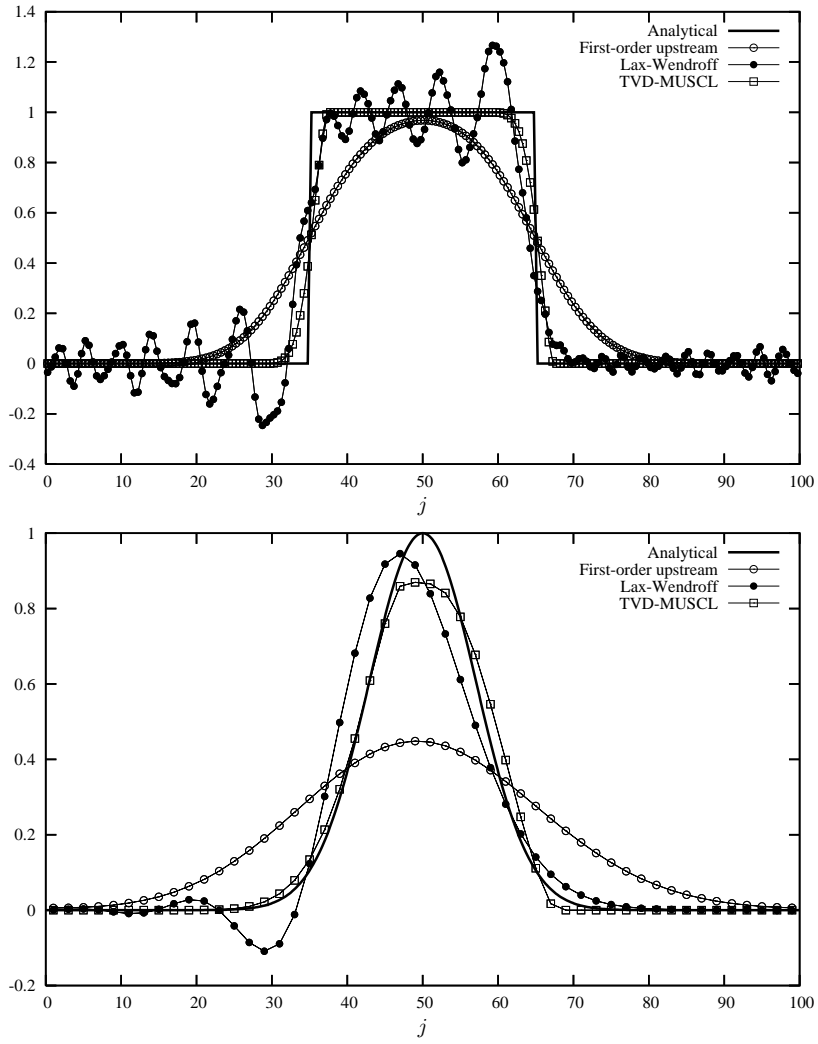


Figure 4.1: Analytical and numerical solutions for the linear advection equation for rectangle and bell shaped initial conditions. The rectangular shape was solved with $j_{\max} = 200$ spatial steps, the bell-shape with $j_{\max} = 50$ spatial steps.

4.3 Monotone schemes

The basic principle for deriving TVD-schemes is to combine a low-order monotone scheme (usually the first-order upstream scheme) with a high-order linear scheme. With such methods, the fluxes $F_{j+1/2}$ are calculated as a weighted average of the low-order flux $F_{j+1/2}^l$ and the high-order flux $F_{j+1/2}^h$:

$$F_{j+1/2} = F_{j+1/2}^l + C_{j+1/2} \left(F_{j+1/2}^h - F_{j+1/2}^l \right) \quad (4.34)$$

with the non-negative multiplicative limiter $C_{j+1/2}$, which may however be larger than unity in some cases. This limiter is calculated as a function of the slope ratio, which is in the case of a positive advection velocity c of the following form:

$$r_{j+1/2} = \frac{\phi_j - \phi_{j-1}}{\phi_{j+1} - \phi_j}. \quad (4.35)$$

For mild slope changes, $r_{j+1/2}$ has a value of about unity, and for local maxima or local minima just upstream of the flux interface under consideration, the slope ratio becomes negative.

In order to obtain TVD-schemes, a numerical scheme is formulated as

$$\phi_j^{n+1} = \phi_j^n - G_{j-1/2} (\phi_j^n - \phi_{j-1}^n) + H_{j+1/2} (\phi_{j+1}^n - \phi_j^n). \quad (4.36)$$

Harten [1983] proved that a sufficient condition for the TVD property of the above formulation is:

$$0 \leq G_{j+1/2}, \quad 0 \leq H_{j+1/2}, \quad G_{j+1/2} + H_{j+1/2} \leq 1. \quad (4.37)$$

Proof:

$$\begin{aligned} |\phi_{j+1}^{n+1} - \phi_j^{n+1}| &= \\ &|\phi_{j+1}^n - \phi_j^n - G_{j+1/2} (\phi_{j+1}^n - \phi_j^n) \\ &+ G_{j-1/2} (\phi_j^n - \phi_{j-1}^n) + H_{j+3/2} (\phi_{j+2}^n - \phi_{j+1}^n) - H_{j+1/2} (\phi_{j+1}^n - \phi_j^n)| \\ &= |(1 - G_{j+1/2} - H_{j+1/2}) (\phi_{j+1}^n - \phi_j^n) \\ &+ G_{j-1/2} (\phi_j^n - \phi_{j-1}^n) + H_{j+3/2} (\phi_{j+2}^n - \phi_{j+1}^n)| \\ &\leq (1 - G_{j+1/2} - H_{j+1/2}) |\phi_{j+1}^n - \phi_j^n| \\ &+ G_{j-1/2} |\phi_j^n - \phi_{j-1}^n| + H_{j+3/2} |\phi_{j+2}^n - \phi_{j+1}^n|, \end{aligned} \quad (4.38)$$

where the last relation holds with (4.37) and the triangle inequality. Building now the sum of (4.38), shifting indices for two terms (assuming periodic boundary conditions) and applying again (4.37), we obtain

$$\sum_j |\phi_{j+1}^{n+1} - \phi_j^{n+1}| \leq |\phi_{j+1}^n - \phi_j^n|, \quad (4.39)$$

such that schemes of form (4.36) with (4.37) are proven to be TVD.

Recalling the flux for the Lax-Wendroff scheme from (4.22), we see that is may be used as a basis for an interpolation between a low-order and a high-order flux according to (4.34):

$$F_{j+1/2}^{n,lw} = c\phi_j^n + \frac{c}{2}(1-\mu)(\phi_{j+1}^n - \phi_j^n)C_{j+1/2}. \quad (4.40)$$

For $C_{j+1/2} = 0$, the low-order upstream scheme is obtained, and for $C_{j+1/2} = 1$, the high-order Lax-Wendroff scheme is retained. It should be noted that in contrast to this, the flux of the Warming and Beam scheme may be formulated as follows:

$$F_{j+1/2}^{n,wb} = c\phi_j^n + \frac{c}{2}(1-\mu)(\phi_{j+1}^n - \phi_j^n)r_{j+1/2}. \quad (4.41)$$

With (4.40), the complete scheme is of the following form:

$$\begin{aligned} \phi_j^{n+1} = & \phi_j^n - \left(\mu - \frac{\mu}{2}(1-\mu)C_{j-1/2}\right)(\phi_j^n - \phi_{j-1}^n) \\ & - \frac{\mu}{2}(1-\mu)C_{j+1/2}(\phi_{j+1}^n - \phi_j^n), \end{aligned} \quad (4.42)$$

such that it would be straight-forward to define in accordance to (4.36):

$$G_{j-1/2} = \mu - \frac{\mu}{2}(1-\mu)C_{j-1/2}, \quad (4.43)$$

$$H_{j+1/2} = -\frac{\mu}{2}(1-\mu)C_{j+1/2}.$$

This approach however fails, since with $C_{j+1/2} \geq 0$, we obtain $H_{j+1/2} \leq 0$ for all stable $0 \leq \mu \leq 1$.

An alternative approach has been suggested by *Sweby* [1984]:

$$G_{j-1/2} = \mu + \frac{\mu}{2}(1-\mu) \left(C_{j+1/2} \left(\frac{\phi_{j+1}^n - \phi_j^n}{\phi_j^n - \phi_{j-1}^n} \right) - C_{j-1/2} \right), \quad (4.44)$$

$$H_{j+1/2} = 0.$$

For this formulation, the TVD criteria from (4.37) will be fulfilled for

$$0 \leq G_{j-1/2} \leq 1, \quad (4.45)$$

or, equivalently,

$$0 \leq \mu \left(1 + \frac{1}{2}(1-\mu) \left(\frac{C_{j+1/2}}{r_{j+1/2}} - C_{j-1/2} \right) \right) \leq 1. \quad (4.46)$$

For stable schemes with $0 \leq \mu \leq 1$, this reduces to

$$\frac{-2}{1-\mu} \leq \frac{C_{j+1/2}}{r_{j+1/2}} - C_{j-1/2} \leq \frac{2}{\mu} \quad (4.47)$$

or

$$\left| \frac{C_{j+1/2}}{r_{j+1/2}} - C_{j-1/2} \right| \leq 2. \quad (4.48)$$

For $r_{j+1/2} > 0$, (4.48) is fulfilled for functions $C(r)$ with

$$0 \leq \frac{C(r)}{r} \leq 2, \quad 0 \leq C(r) \leq 2. \quad (4.49)$$

For negative $r_{j+1/2}$, when poorly resolved local minima and maxima occur just upstream of the flux interface, a monotone scheme must be used in order to avoid unphysical oscillations, such that for $r < 0$ the condition $C(r) = 0$ must hold.

Examples for flux limiters $C(r)$ which fulfill (4.49) are the minmod limiter,

$$C(r) = \max\{0, \min(1, r)\}, \quad (4.50)$$

the superbee limiter,

$$C(r) = \max\{0, \min(1, 2r), \min(2, r)\}, \quad (4.51)$$

the van Leer limiter,

$$C(r) = \frac{r + |r|}{1 + |r|}, \quad (4.52)$$

and the MUSCL limiter,

$$C(r) = \max \left\{ 0, \min \left(2, 2r, \frac{1+r}{2} \right) \right\} \quad (4.53)$$

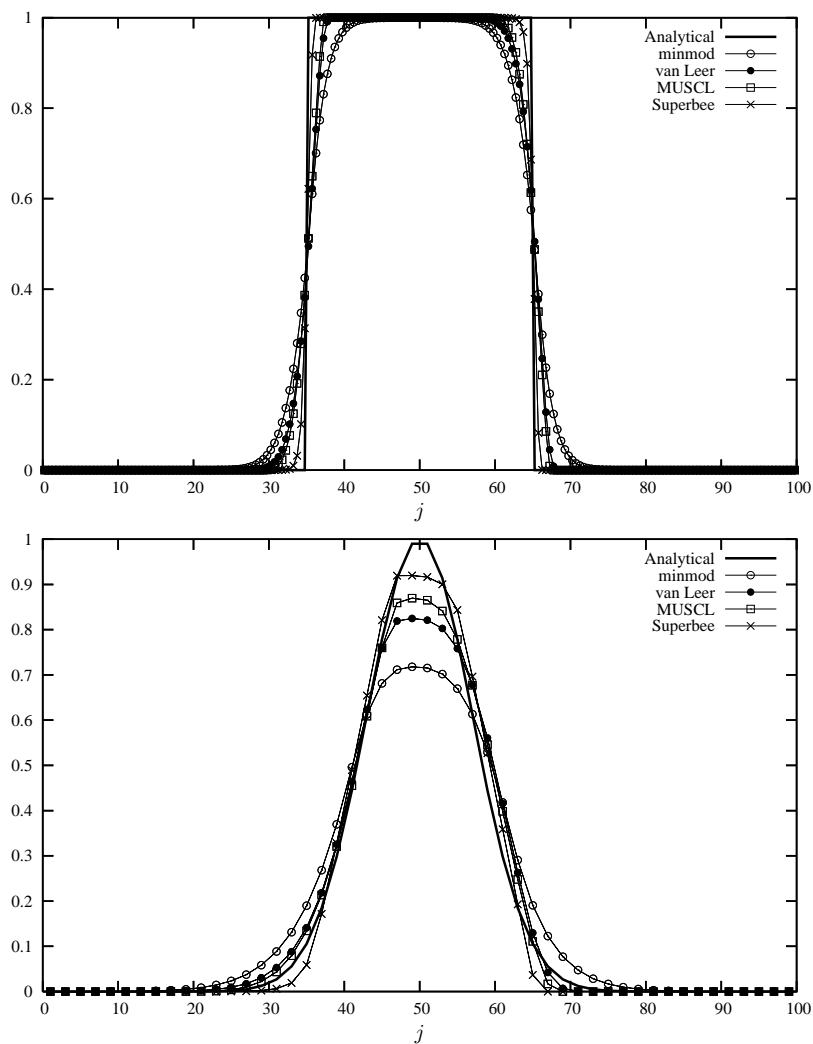


Figure 4.2: Analytical and numerical solutions for the linear advection equation for rectangle and bell shaped initial conditions. The rectangular shape was solved with $j_{\max} = 200$ spatial steps, the bell-shape with $j_{\max} = 50$ spatial steps.