UNIFORM ASYMPTOTIC STABILITY OF STRANG'S EXPLICIT COMPACT SCHEMES FOR LINEAR ADVECTION*

BRUNO DESPRÉS†

Abstract. We consider a family of explicit compact schemes for advection in one dimension. The order is arbitrarily high. These stencils may be called Strang's stencils after the seminal work of Strang [J. Math. Phys., 41 (1962), pp. 147–154]. We prove that odd order schemes are stable in all L^q under CFL one. The strategy of the proof is similar to the one of Thomée [J. Differential Equations, 1 (1965), pp. 273–292] with a careful verification that all sharp estimates on the amplification factor are independent of the CFL number. This is possible based on a general representation formula for the amplification factor. Numerical results in one dimension confirm the analysis.

Key words. stability in L^q , advection equation, compact explicit schemes

AMS subject classifications. 65M06, 65M12

DOI. 10.1137/080734571

1. Introduction. We consider the numerical discretization of the advection equation

$$\partial_t u + a \partial_x u = 0, \quad a > 0,$$

in one dimension by means of some very high order finite difference schemes [15, 6, 5, 18, 22], which can also be interpreted as finite volume schemes. Our interest in this family of schemes stems from the use of such methods in more complicated problems, where they have proved to be very efficient in terms of both stability and accuracy. See, for example, the three-dimensional (3D) computation of an acoustic wave in the atmosphere [7], where we used the 17th scheme in each direction (x, y, z) of a 3D Cartesian mesh. Among many other possibilities for the discretization of advection and the one-dimensional (1D) related wave equation, we single out [2, 1]. In a simplified finite difference form [9, 16] on a Cartesian grid, the family of linear schemes may be written as

(1.2)
$$u_j^{n+1} = \sum_{r=k-p}^k \alpha_r u_{j+r}^n, \qquad \alpha_r = \alpha_r(\nu).$$

The coefficients of the scheme are the α_r 's. These are functions of the CFL number $\nu = a \frac{\Delta t}{\Delta x}$. The analysis is limited to explicit and compact schemes with a stencil of p+1 contiguous cells. We will show that this is compatible with the order in space and time of the scheme which is p. Once p has been chosen, k determines the shift of the scheme, as described in Figure 1.1. We will show in section 2 that there is a one-to-one correspondence between the schemes described in this paper and the pairs (p,k).

Basic examples are the well-known upwind scheme $u_j^{n+1}=(1-\nu)u_j^n+\nu u_{j-1}^n$ such that (p,k)=(1,0), the Lax–Wendroff [14] scheme $u_j^{n+1}=(1-\nu^2)u_j^n+\frac{\nu+\nu^2}{2}u_{j-1}^n+\frac{\nu+\nu^2}{2}u_{j-1}^n$

^{*}Received by the editors September 5, 2008; accepted for publication (in revised form) September 30, 2009; published electronically December 16, 2009.

http://www.siam.org/journals/sinum/47-5/73457.html

[†]Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, 75252 Paris Cedex 05, France (bruno.despres@cea.fr).

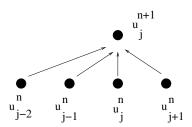


Fig. 1.1. Graphical representation of the O3 scheme (p,k) = (3,1). The number of cells at stage n is p+1=4. The stencil is shifted k=1 cell on the right.

 $\frac{\nu^2-\nu}{2}u_{j+1}^n$ such that (p,k)=(2,1), and the Beam–Warming scheme [24] $u_j^{n+1}=\left(1-\frac{3}{2}\nu+\frac{1}{2}\nu^2\right)u_j^n+(2\nu-\nu^2)u_{j-1}^n+\frac{\nu^2-\nu}{2}u_{j-2}^n$ such that (p,k)=(2,0). These schemes may be rewritten also as incremental finite volume methods

(1.3)
$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + a \frac{u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n}{\Delta x} = 0.$$

A third order in space and time O3 scheme (p,k)=(3,1) is defined by a convex combination [8] of the Lax-Wendroff scheme and the Beam-Warming scheme: $O3=(1-\alpha)LW+\alpha BW$ with $\alpha=\frac{1+\nu}{3}$.

As usual it is convenient to define the amplification factor for Fourier modes $\lambda(\theta) = \sum_{r=k-p}^{k} \alpha_r(\nu) e^{-ri\theta}$, where $\mathbf{i} \in \mathbb{C}$ is the imaginary number such that $\mathbf{i}^2 = -1$. The criterion for stability in L^2 that we retain is [16]

$$(1.4) |\lambda(\theta)| \le 1 \quad \forall \theta \in \mathbb{R}$$

What is important is to have the stability inequality (1.4) for ν in the largest interval. The seminal works of Iserles and Strang [20, 12] answer this question and show that the order in time and space p can be arbitrary large. See also [11, 3, 13]. In our framework the Iserles–Strang theorem can be stated as follows.

THEOREM 1 (Iserles and Strang [20, 12]). The only pairs (p,k) for which the stability inequality (1.4) is true for all $\nu \leq 1$ are p=2k+1, p=2k, and p=2k+2. This is a special case of Iserles and Strang's more general theorem [12].

So stable and compact schemes can be grouped into three families. The first family p=2k+1 begins with the upwind scheme and the O3 scheme; the second family p=2k begins with the Lax-Wendroff scheme; and the third family begins with the Beam-Warming scheme. We will give a new proof of this property. We will also prove in the appendix that for p=2k+2, inequality (1.4) is true for all $\nu \leq 2$.

Actually the "sufficient" condition of the theorem was already given by Strang [20] for the special stencils considered in this work (one may call them Strang's stencils). What is more difficult to prove is the "necessary" part. In [12] it was proved for general stencils with the theory of order stars. In our case the proof is much easier, since it is based on a compact formula that we obtain for the amplification factor

(1.5)
$$\lambda(\theta) = \left[1 - \mathbf{i}^{p+1} \alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} e^{(k - \frac{p}{2} + \nu) \mathbf{i}(\varphi - \theta)} d\varphi \right] e^{\nu \mathbf{i} \theta},$$

where for convenience

(1.6)
$$\alpha_{k,p} = \frac{\prod_{q=0}^{p} (k + \nu - q)}{p!}.$$

Our main result is a new stability inequality in all L^q for the p = 2k + 1 family.

THEOREM 2. Assume moreover that the order is odd, that is, p = 2k + 1. Then the scheme is stable in all L^q : there exists a constant $D_p > 0$ such that

$$(1.7) ||u^n||_{L^q} \le D_p||u^0||_{L^q} \quad \forall n, \ \forall \nu \in]0,1], \ \forall u^0, \ \forall q \in [1,\infty].$$

That is, only the first family of odd order schemes is stable in L^q for $1 \le q \le \infty$. In [8] the property (1.7) is called asymptotic stability because the constant D_p does not depend on n nor on the CFL number ν . The proof follows the strategy in Fourier proposed in [8] with a careful verification that the technical conditions are fulfilled for all members of the family p = 2k + 1. In [8], the proof of the stability inequality (1.7) was restricted to the O3 scheme in the case q = 1 and $q = \infty$. The fundamental reason why only p = 2k + 1 schemes are stable in all L^q is because the leading order of their amplitude coefficient in Fourier is real negative. In section 5.3 we analyze in more detail why the theorem cannot be true for even order schemes.

Actually the property is similar to a result proved by Thomée [21]; see also [4]. What is new with respect to Thomée's work is the following: (a) the proof that the general theory applies for all Strang's stencils provided p is odd, and the proof relies on sharp estimates for $\lambda(\theta)$; (b) the fact that the stability estimates do not depend on the CFL number ν . The satisfaction of (b) is important for practical situations where one can rarely choose the CFL number. On the contrary, even order schemes p = 2k and p = 2k + 2 are not stable in L^1 and L^{∞} even if they are stable in L^2 : moreover their dispersion is highly dependent on the CFL number. Some results provided in the numerical section confirm this theoretical analysis.

Some consequences of these very strong stability properties have been explored in [8]. In particular it is shown that it is possible to derive optimal convergence estimates, and also that such methods have very low dispersion and dissipation. Basic numerical experiments reproduced at the end of this work will illustrate these facts. Let us stress one consequence of (1.7), which is that

(1.8)
$$TV(u^n) \le D_p TV(u^0), \quad p = 2k + 1,$$

where TV(u) is the total variation of the numerical profile u. It is sufficient to apply the inequality (1.7) for q=1 to the numerical gradient of u to obtain this inequality. Therefore these methods are TVB (total variation bounded), which explains why they control the oscillations. In view of these considerations, it is possible to think that the stencils p=2k+1 have potential applications to ENO and WENO schemes (we refer to [10, 19]). This is completely open.

This work is organized as follows. We show that two schemes with the same pair (p,k) are equal. It gives a compact representation formula of the amplification factor. In section 4 we give a new proof of the Iserles–Strang theorem for p=2k+1. The next section is devoted to showing that all odd schemes p=2k+1 are also stable in all L^q , as stated in Theorem 2. In the next section we show that the incremental or finite volume formulation is nondegenerate for $\nu \approx 0$. Then we give the results of basic numerical experiments, which illustrate that odd order advection schemes are indeed stable in all L^q and illustrate also that even order advection schemes oscillate much more. In the appendix, we show how to extend our proof of the Iserles–Strang theorem for even order schemes p=2k and p=2k+2. For p=2k+2 the CFL requirement for stability in L^2 is $\nu \leq 2$. Finally we give a theoretical explanation why even order schemes are not stable in all L^q .

2. Uniqueness property.

LEMMA 3. Two schemes (1.2) with the same order p and the same shift k are equal.

Consider a scheme (1.2) of order $p \ge 1$ with a stencil of p+1 contiguous points. Since the order is p then the scheme is exact over all polynomials of degree less than or equal to p. That is,

(2.1) if
$$u_j^n = P(j\Delta x) \ \forall j$$
, then $u_j^{n+1} = P(j\Delta x - a\Delta t) \ \forall j$, $d(P) \le p$.

Now consider two schemes with the same order p and the same shift k. The first one is $u_j^{n+1} = \sum_{r=k-p}^k \alpha_r u_{j+r}^n$, and the second one is $u_j^{n+1} = \sum_{r=k-p}^k \beta_r u_{j+r}^n$. These two formulas are exact for polynomials of order less than or equal to p. Define $\gamma_r = \alpha_r - \beta_r$. Then by subtraction

(2.2)
$$0 = \sum_{r=k-p}^{k} \gamma_r u_{j+r}^n, \quad u_{j+r}^n = P((j+r)\Delta x), \quad d(P) \le p.$$

Consider the polynomial sequence $P_q(x) = (x - (k - p)\Delta x)^q$, $0 \le q \le p$, and let us write (2.2) at j = 0. So

$$\begin{cases}
0 = \sum_{k-p \le r \le k} \gamma_r, \\
0 = \sum_{k-p \le r \le k} (r - k + p) \gamma_r, \\
0 = \cdots, \\
0 = \sum_{k-p \le r \le k} (r - k + p)^p \gamma_r.
\end{cases}$$

The p+1 unknowns of this linear system of size p+1 are the γ_r 's. The matrix of the system is a Vandermonde matrix with a nonzero determinant. Therefore the matrix is nonsingular and $\gamma_r=0$ for all $k-p\leq r\leq k$. Then $\alpha_r=\beta_r,\ k-p\leq r\leq k$. This completes the proof. \square

A consequence is that all of the schemes derived in [6, 5, 18, 22] are the same, provided they have the same pair (p, k). It also covers the explicit schemes of [12, 20].

We now use this property to obtain a direct derivation of the schemes by means of a Fourier expansion of the symbol of the translation operator. The Fourier symbol of the translation operator (that is, translation by one cell Δx) is $e^{i\theta}$, while $e^{i\nu\theta}$ is the Fourier symbol of the advection operator by a smaller length $\nu\Delta x = a\Delta t$. The discretization process amounts to approximating $e^{i\nu\theta}$ by a truncated expansion of $e^{i\nu\theta}$ into Laurent series with respect to $e^{i\theta}$,

$$(2.3) e^{\mathbf{i}\nu\theta} = \left(1 + (e^{\mathbf{i}\theta} - 1)\right)^{k+\nu} e^{-k\mathbf{i}\theta} = \left(P_p\left(e^{\mathbf{i}\theta} - 1\right) + \mathcal{O}\left(e^{\mathbf{i}\theta} - 1\right)^{p+1}\right) e^{-k\mathbf{i}\theta}.$$

The degree of P_p is p. So the expression $P_p(e^{i\theta}-1)e^{-ki\theta}$ is an approximation at order p+1 of the exact symbol. Therefore this expression is the Fourier symbol of the (p,k) scheme.

For example, take p=2 and k=0. Then

$$P_{2}(e^{i\theta} - 1) = 1 + \nu(e^{i\theta} - 1) + \frac{\nu(\nu - 1)}{2}(e^{i\theta} - 1)^{2}$$
$$= \left(1 - \frac{3}{2}\nu + \frac{1}{2}\nu^{2}\right) + \left(2\nu - \nu^{2}\right)e^{i\theta} + \frac{-\nu + \nu^{2}}{2}e^{2i\theta}.$$

The coefficients coincide with those of the Beam-Warming scheme.

To extend further the analysis, we define $f_k(z) = (1+z)^{k+\nu}$. The Taylor formula states

(2.4)
$$f_k(z) = \sum_{q>0}^p \mu_q z^q + \int_0^1 \frac{(1-t)^p}{p!} z^{p+1} f_k^{(p+1)}(tz) dt,$$

with

(2.5)
$$\int_0^1 \frac{(1-t)^p}{p!} z^{p+1} f_k^{(p+1)}(tz) dt = \frac{\prod_{q=0}^p (k+\nu-q)}{p!} \int_0^1 \frac{(1-t)^p z^{p+1}}{(1+tz)^{p+1-k-\nu}} dt.$$

One has to restrict the formula to $\theta \in [0, \pi)$ because there may be (we have to assume that $p+1-k-\nu \geq 1$) a singularity in the integral for $\theta=\pi$ and $t=\frac{1}{2}$: in this case $1+tz=1+\frac{1}{2}(e^{i\theta}-1)=0$.

LEMMA 4. For a given (p,k), the amplification factor of the scheme (1.2) is

$$(2.6) \quad \lambda(\theta) = \left[f_k \left(e^{\mathbf{i}\theta} - 1 \right) - \int_0^1 \frac{(1-t)^p}{p!} f_k^{(p+1)} \left(t \left(e^{\mathbf{i}\theta} - 1 \right) \right) dt \right] e^{-k\mathbf{i}\theta}, \quad \theta \in [0, \pi).$$

Since by construction the schemes have real coefficients, then

(2.7)
$$\lambda(2\pi - \theta) = \overline{\lambda(\theta)}.$$

Formula (2.3) shows that the Fourier symbol is $\lambda(\theta) = P_p \left(e^{\mathbf{i}\theta} - 1\right) e^{-k\mathbf{i}\theta}$. Now $P_p(z)$ is the leading term in the Taylor expansion of $f_k(z)$ with $z = e^{\mathbf{i}\theta} - 1$. Therefore $P_p(z) = \sum_{q \geq 0}^p \mu_q z^q = f_k(z) - \int_0^1 \frac{(1-t)^p}{p!} z^{p+1} f_k^{(p+1)}(tz) dt$. This completes the proof. \square

3. Study of the amplification factor. The integral (2.5) has a singularity for $\theta = \pi$. So we modify the path of integration in the complex plane, as explained in Figure 3.1, in order to remove this singularity [17, 23] and to obtain a more convenient formula.

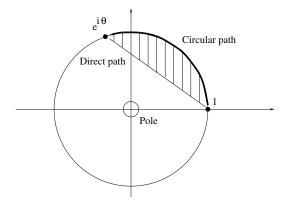


Fig. 3.1. The direct path of integration is replaced by the circular path.

Let us study the integral in formula (2.5):

$$A = \int_0^1 \frac{(1-t)^p \left(e^{i\theta} - 1\right)^{p+1}}{(1+t(e^{i\theta} - 1))^{p+1-k-\nu}} dt.$$

Set $z=1+t\left(e^{\mathbf{i}\theta}-1\right)$ so that A can be seen as an integral in the complex plane along the straight line that joins z=1 to $z=e^{\mathbf{i}\theta}$. Instead of integrating along the straight line, we integrate along the circle |z|=1, which is another possibility to connect z=1 to $z=e^{\mathbf{i}\theta}$. This path change is valid, provided the pole z=0 does not lie in the closure of the domain bounded by the direct path and the circular path. This holds true if $\theta < \pi$. So let us assume for the moment that $0 \le \theta < \pi$. One has $dz=\left(e^{\mathbf{i}\theta}-1\right)dt$ and $\left(e^{\mathbf{i}\theta}-z\right)=(1-t)\left(e^{\mathbf{i}\theta}-1\right)$. So

$$A = \int_0^{e^{i\theta}} \frac{\left(e^{i\theta} - z\right)^p}{z^{p+1-k-\nu}} dz, \quad |z| = 1.$$

Next we define $\varphi \in [0, \theta]$ such that $z = e^{i\varphi}$. So after the change of integral path,

$$A = \int_0^\theta \frac{\left(e^{\mathbf{i}\theta} - e^{\mathbf{i}\varphi}\right)^p}{e^{\mathbf{i}(p+1-k-\nu)\varphi}} \mathbf{i} e^{\mathbf{i}\varphi} d\varphi.$$

Plugging A in (2.6), one gets

$$\lambda(\theta) = \left[e^{(k+\nu)i\theta} - i\alpha_{k,p} \int_0^\theta \left(e^{i\theta} - e^{i\varphi} \right)^p e^{i(-p+k+\nu)\varphi} d\varphi \right] e^{-ki\theta},$$

where $\alpha_{k,p}$ is given in (1.6). After simplifications, one obtains the formula

(3.1)
$$\lambda(\theta) = \left[1 - \mathbf{i}^{p+1} \alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} e^{(k - \frac{p}{2} + \nu) \mathbf{i}(\varphi - \theta)} d\varphi \right] e^{\nu \mathbf{i} \theta}.$$

This formula is valid by continuity for $\theta = \pi^-$. Since $\lambda(\theta)$ is a truncated Laurent series with respect to $e^{i\theta}$, it is also analytical for $\theta \in \mathbb{R}$. Therefore formula (3.1) is valid for all real θ .

To pursue the discussion we assume for convenience that p is odd. In this case the stability inequality in L^2 , that is, $|\lambda(\theta)| \leq 1$ for all real θ , is equivalent to

$$|\lambda(\theta)|^2 = \left(1 - (-1)^{\frac{p+1}{2}} \alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2 + \left(\alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} \sin\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2 \le 1.$$

The less interesting complementary case where p is even is discussed in the appendix.

- 4. Stability in L^2 . In this section we give a new proof of the Iserles–Strang theorem. This is based on the examination of formula (3.2). We think the sufficient part of this new proof is much more elementary than the previous Iserles–Strang proof, which is based on order stars theory. For the sake of convenience we restrict the analysis to odd p's.
- **4.1. Necessary condition.** A first condition for (3.2) to be true for all θ is that the integral

$$(4.1) \int_{0}^{\theta} \sin^{p} \frac{\theta - \varphi}{2} \cos \left(k - \frac{p}{2} + \nu \right) (\varphi - \theta) d\varphi = \int_{0}^{\theta} \sin^{p} \frac{\varphi}{2} \cos \left(k - \frac{p}{2} + \nu \right) \varphi d\varphi$$

does not change sign when θ varies, because if the sign changes, then the first square in (3.2) shall be strictly greater than one.

Let assume for a moment that $\left|k-\frac{p}{2}+\nu\right|\geq 1$. Define $\overline{\theta}=\frac{\pi}{\left|k-\frac{p}{2}+\nu\right|}\in [0,\pi]$. Then

$$\int_{0}^{\overline{\theta}} \sin^{p} \frac{\varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) \varphi \, d\varphi$$

$$= \int_{0}^{\frac{\overline{\theta}}{2}} \sin^{p} \frac{\varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) \varphi \, d\varphi + \int_{\frac{\overline{\theta}}{2}}^{\overline{\theta}} \sin^{p} \frac{\varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) \varphi \, d\varphi.$$

The first right-hand side is positive: $\int_0^{\frac{\overline{\theta}}{2}} \cdots > 0$. The second integral is negative: $\int_{\frac{\overline{\theta}}{2}}^{\frac{\overline{\theta}}{2}} \cdots < 0$. Let us perform the change of variable $\psi = \overline{\theta} - \varphi$ in the second integral. Then

$$\int_0^{\overline{\theta}} \sin^p \frac{\varphi}{2} \cos \left(k - \frac{p}{2} + \nu\right) \varphi \ d\varphi = \int_0^{\frac{\overline{\theta}}{2}} \left[\sin^p \frac{\varphi}{2} - \sin^p \frac{\overline{\theta} - \varphi}{2} \right] \cos \left(k - \frac{p}{2} + \nu\right) \varphi d\varphi < 0.$$

By construction $0 \le \frac{\varphi}{2} \le \frac{\overline{\theta} - \varphi}{2} \le \frac{\pi}{2}$. So the term $[\cdot \cdot \cdot]$ is negative, and the whole sum is also negative. On the other hand the integral (4.1) is positive for a small $\theta > 0$. The sign of the integral changes strictly when θ increases from 0 to π , which means that (3.2) is violated for some θ . Therefore a necessary condition of stability is $|k - \frac{p}{2} + \nu| < 1$ for all $\nu \in [0, 1]$. That is, $-1 < k - \frac{p}{2} \le k - \frac{p}{2} + \nu \le k - \frac{p}{2} + 1 < 1$. Since p is odd by hypothesis, there is only one solution, which is $k - \frac{p}{2} = -\frac{1}{2} \iff p = 2k + 1$.

4.2. Sufficient condition. Now we prove that p = 2k + 1 is sufficient to have (3.2) for all $\theta \in \mathbb{R}$, provided $0 < \nu \le 1$. Set $B(\theta) = |\lambda(\theta)|^2$. We define $\beta = \nu - \frac{1}{2} \in (-\frac{1}{2}, \frac{1}{2}]$. Then

$$B(\theta) = \left(1 - |\alpha_{k,2k+1}| \, 2^{2k+1} \int_0^\theta \sin^{2k+1} \frac{\varphi}{2} \cos \beta \varphi \, d\varphi\right)^2 + \left(|\alpha_{k,2k+1}| \, 2^{2k+1} \int_0^\theta \sin^{2k+1} \frac{\varphi}{2} \sin \beta \varphi \, d\varphi\right)^2.$$

The derivative of B is $B'(\theta) = \left(2 |\alpha_{k,2k+1}| 2^{2k+1} \sin^{2k+1} \frac{\theta}{2}\right) h(\theta)$ with

(4.2)
$$h(\theta) = |\alpha_{k,2k+1}| 2^{2k+1} \int_0^\theta \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\theta - \varphi) d\varphi - \cos \beta \theta.$$

The symmetry of $\lambda(\theta)$ implies that

$$(4.3) B(2\pi - \theta) = B(\theta) \Longrightarrow B'(\pi) = 0 \Longrightarrow h(\pi) = 0,$$

that is, $|\alpha_{k,2k+1}| 2^{2k+1} \int_0^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\pi - \varphi) \ d\varphi - \cos \beta \pi = 0$. One can express the coefficient $|\alpha_{k,2k+1}| 2^{2k+1}$ as

$$|\alpha_{k,2k+1}| \, 2^{2k+1} = \frac{\cos \beta \pi}{\int_0^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\pi - \varphi) \, d\varphi}$$

and replace it in (4.2). We obtain

$$h(\theta) = \frac{\int_0^{\theta} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\theta - \varphi) \, d\varphi}{\int_0^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\pi - \varphi) \, d\varphi} \cos \beta \pi - \cos \beta \theta.$$

One can write

$$h(\theta) = \frac{\int_0^\theta \sin^{2k+1} \frac{\varphi}{2} P(\theta, \varphi) d\varphi}{\int_0^\pi \sin^{2k+1} \frac{\varphi}{2} \cos \beta(\pi - \varphi) d\varphi} - \frac{\int_\theta^\pi \sin^{2k+1} \frac{\varphi}{2} \cos \beta(\pi - \varphi) d\varphi}{\int_0^\pi \sin^{2k+1} \frac{\varphi}{2} \cos \beta(\pi - \varphi) d\varphi} \cos \beta \theta,$$

where

$$P(\theta, \varphi) = \cos \beta \pi \cos \beta (\theta - \varphi) - \cos \beta \theta \cos \beta (\pi - \varphi)$$

$$= \cos \beta \pi (\cos \beta \theta \cos \beta \varphi + \sin \beta \theta \sin \beta \varphi) - \cos \beta \theta (\cos \beta \pi \cos \beta \varphi + \sin \beta \pi \sin \beta \varphi)$$

$$= \sin \beta \varphi (\cos \beta \pi \sin \beta \theta - \cos \beta \theta \sin \beta \pi) = \sin \beta \varphi \sin \beta (\theta - \pi).$$

Since $0 \le \varphi \le \theta \le \pi$ and $\beta \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ then $P(\theta, \varphi) \le 0$. So

(4.4)
$$h(\theta) \le -\frac{\int_{\theta}^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta(\pi - \varphi) \, d\varphi}{\int_{0}^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta(\pi - \varphi) \, d\varphi} \cos \beta\theta,$$

and $h(\theta) \leq 0$ for $\theta \in [0, \pi]$. Therefore $B'(\theta) \leq 0$, $\theta \in [0, \pi] \Longrightarrow B(\theta) \leq B(0) = 1$, $\theta \in [0, \pi]$. Using (2.7), one has proved that $|\lambda(\theta)|^2 = B(\theta) \leq 1$ for all $\theta \in [0, 2\pi]$. It means the scheme is L^2 stable in the case p = 2k + 1.

4.3. More estimates. In the following we will need sharper estimates, which we prove now. From now on we consider that $0 < \nu \le 1$. We also restrict the analysis to $\theta \in [0, \pi]$ since the other case, $\theta \in [\pi, 2\pi[$, is obtained by symmetry; see (2.7).

Proposition 5. One has the estimate

(4.5)
$$|\lambda(\theta)| \le 1 - m_1(k)\nu(1-\nu)\theta^{2k+2}, \quad 0 \le \theta \le \pi,$$

for a suitable constant $m_1(k) > 0$.

The rather technical proof is split into two steps.

First step. Consider the fraction in the right-hand side of (4.4),

$$\mu(\theta) = \frac{\int_{\theta}^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\pi - \varphi) \, d\varphi}{\int_{0}^{\pi} \sin^{2k+1} \frac{\varphi}{2} \cos \beta (\pi - \varphi) \, d\varphi}, \qquad \beta = \frac{1}{2} - \nu \in \left] - \frac{1}{2}, \frac{1}{2} \right[,$$

so that (4.4) may be rewritten as $h(\theta) \leq -\mu(\theta)\cos(\beta\theta)$. It is immediate to see that $\mu(0) = 1, \mu(\pi) = 0$,

$$\mu'(\theta) = \frac{-\sin^{2k+1}\frac{\theta}{2}\cos\beta(\pi-\theta)}{\int_0^{\pi}\sin^{2k+1}\frac{\varphi}{2}\cos\beta(\pi-\varphi)\,d\varphi}$$

and

$$\mu''(\theta) = \frac{-\frac{2k+1}{2}\cos\frac{\theta}{2}\sin^{2k}\frac{\theta}{2}\cos\beta(\pi-\theta) - \sin^{2k+1}\frac{\theta}{2}\beta\sin\beta(\pi-\theta)}{\int_{0}^{\pi}\sin^{2k+1}\frac{\varphi}{2}\cos\beta(\pi-\varphi)\,d\varphi} \le 0.$$

Therefore μ is concave and lies above the chord joining (0,1) to $(\pi,0)$, that is, $\mu(\theta) \ge 1 - \frac{\theta}{\pi}$. So

(4.6)
$$h(\theta) \le -\left(1 - \frac{\theta}{\pi}\right)\cos\beta\theta, \qquad 0 \le \theta \le \pi.$$

Therefore

$$B(\theta) = 1 + \int_0^\theta B'(\varphi) d\varphi \le 1 - 2 |\alpha_{k,2k+1}| 2^{2k+1} \int_0^\theta \sin^{2k+1} \frac{\varphi}{2} \left(1 - \frac{\varphi}{\pi}\right) \cos \beta \varphi \ d\varphi.$$

On the interval $[0,\pi]$ one has $1-\frac{\varphi}{\pi}\geq \frac{2}{\pi}\cos\frac{\varphi}{2}$. Therefore

$$B(\theta) \le 1 - \frac{4}{\pi} |\alpha_{k,2k+1}| \, 2^{2k+1} \cos \beta \theta \int_0^\theta \sin^{2k+1} \frac{\varphi}{2} \cos \frac{\varphi}{2} \, d\varphi$$
$$\le 1 - \frac{2}{\pi (2k+2)} |\alpha_{k,2k+1}| \, 2^{2k+1} \cos \beta \theta \sin^{2k+2} \frac{\theta}{2}.$$

So one gets the intermediate result

(4.7)
$$|\lambda(\theta)| = \sqrt{B(\theta)} \le 1 - \frac{1}{\pi(k+1)} |\alpha_{k,2k+1}| \, 2^{2k} \cos \beta \theta \sin^{2k+2} \frac{\theta}{2}$$

By definition (1.6) one has $\nu(1-\nu)\frac{(k!)^2}{2k+1} \leq |\alpha_{k,2k+1}| \leq \nu(1-\nu)\frac{((k+1)!)^2}{2k+1}$. So (4.7) implies the claim (4.5), provided $\cos\beta\theta$ is lower bounded by a positive constant. Unfortunately this is not the case if $\nu=0$ or $\nu=1$, because in this case $\beta=\pm\frac{1}{2}$ and $\cos\beta\theta=0$ for $\theta=\pi$. Therefore the estimate (4.7) does not imply directly the claim. The second step of the proof is here to remove carefully the term $\cos\beta\theta$.

Second step. To simplify the notation, we set $\delta_k = |\alpha_{k,2k+1}| \, 2^{2k}$. Let us choose an arbitrary $\theta(k) \in]0, \pi[$. The function $x \mapsto \cos x$ is decreasing on the interval $[0, \frac{\pi}{2}]$, so $\cos \beta x \geq \cos \frac{x}{2}$. Then one has

$$|\lambda(\theta)| \le 1 - \frac{\delta_k}{\pi(k+1)} \cos \frac{\theta(k)}{2} \sin^{2k+2} \frac{\theta}{2} \text{ for } 0 \le \theta \le \theta(k).$$

Therefore we can write

$$(4.8) |\lambda(\theta)| \le 1 - \frac{\delta_k}{\pi(k+1)} \cos \frac{\theta(k)}{2} \sin^{2k+2} \frac{\theta(k)}{2} \sin^{2k+2} \frac{\theta}{2}$$

for all $\theta \in [0, \theta(k)]$. On the other hand $\theta \mapsto |\lambda(\theta)|$ is decreasing because $B' \leq 0$. So

$$|\lambda(\theta)| \le 1 - \frac{\delta_k}{\pi(k+1)} \cos \frac{\theta(k)}{2} \sin^{2k+2} \frac{\theta(k)}{2} \text{ for } \theta \in [\theta(k), \pi].$$

It implies that (4.8) is true for all $\theta \in [0, \pi]$. Now we take the simple choice $\theta(k) = \text{constant} = \frac{\pi}{2}$. One also has $\sin \frac{\theta}{2} \ge \frac{\theta}{\pi}$ for $0 \le \theta \le \pi$. So one can replace $\sin \frac{\theta}{2}$ by $\frac{\theta}{\pi}$ in (4.8). Finally all contributions which depend solely on k are aggregated in the definition of $m_1(k)$. It ends the proof of the proposition. \square

Let us now turn to other consequences of formula (3.1). Let us define

$$\sigma(\theta) = \lambda(\theta)e^{-\nu i\theta} = 1 - i^{p+1}\alpha_{k,p}2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} e^{(k - \frac{p}{2} + \nu)i(\varphi - \theta)} d\varphi.$$

Proposition 6. One has the bounds

(4.9)
$$|\sigma'(\theta)| \le m_2(k)\nu(1-\nu)\theta^{2k+1}$$

and

$$(4.10) |\sigma''(\theta)| \le m_3(k)\nu(1-\nu)\theta^{2k},$$

where $m_2(k) > 0$ and $m_3(k) > 0$.

One has

$$|\sigma'(\theta)| \le |\alpha_{k,p}| \, 2^p \left(\frac{p}{2}\theta \sin^{p-1}\frac{\theta}{2} + \left|k - \frac{p}{2} + \nu\right| \theta \sin^p\frac{\theta}{2}\right) \le |\alpha_{k,p}| \, 2^p \left(\frac{p}{2}\pi + \frac{\pi}{2}\right) \sin^p\frac{\theta}{2},$$

that is, (4.9). The second derivative of σ is bounded as

$$|\sigma''(\theta)| \leq |\alpha_{k,p}| \, 2^p \left(\frac{p(p-1)}{4}\theta \sin^{p-2}\frac{\theta}{2} + \left|k - \frac{p}{2} + \nu\right|^2 \theta \sin^p\frac{\theta}{2} + 2\frac{p}{2}\left|k - \frac{p}{2} + \nu\right| \theta \sin^{p-1}\frac{\theta}{2}\right).$$

After evident simplifications one gets (4.10).

5. Stability in L^q . Let us define the translation operator T:

$$(Tu)_j = u_{j-1}, \qquad u = (u_j)_{j \in \mathbb{Z}}.$$

The powers of T are $(T^r u)_j = u_{j-r}$ for all $r \in \mathbb{Z}$ and all $j \in \mathbb{Z}$. The L^q norm $(1 \le q < \infty)$ is $\|u\|_{L^q} = \left(\sum_{j \in \mathbb{Z}} |u_j|^q\right)^{\frac{1}{q}}$. The L^∞ norm is $\|u\|_{L^\infty} = \sup_{j \in \mathbb{Z}} |u_j|$. One has $\|T^r\|_{L^q} = 1$ for all r and all q. The scheme (1.2) can be rewritten as $u^n = \left(\sum_{s=k-p}^k \alpha_s T^s\right)^n u^0$. The expansion of the operator is written as $\left(\sum_{s=k-p}^k \alpha_s T^s\right)^n = \sum_{r \in \mathbb{Z}} b_{r,n} T^r$, which implies

(5.1)
$$\left\| \left(\sum_{s=k-p}^{k} \alpha_s T^s \right)^n \right\|_{L^q} \le \sum_{r \in \mathbb{Z}} |b_{r,n}| \|T^r\|_{L^q} = \sum_{r \in \mathbb{Z}} |b_{r,n}|.$$

The inequality is actually an equality for q = 1 and $q = \infty$. Note that this holds true for all norms $\|\cdot\|$ for which $\|T\| = 1$. Therefore, uniform stability is not restricted to L^q norms alone. The coefficients $b_{r,n}$ are expressed as a function of the amplification factor as follows:

$$\lambda(\theta) = \sum_{r=k-p}^{k} \alpha_r e^{-r\mathbf{i}\theta}, \quad \mathbf{i}^2 = -1.$$

One has the formula

$$b_{r,n} = \frac{1}{2\pi} \int_0^{2\pi} \lambda(\theta)^n e^{\mathbf{i}r\theta} d\theta.$$

The rest of the proof consists in getting various estimates for $b_{r,n}$, which will be enough to bound $\sum_{r\in\mathbb{Z}}|b_{r,n}|$.

5.1. Estimates for $b_{r,n}$ **.** The first estimate is immediate since $|\lambda(\theta)| \leq 1$ for all $\theta \in \mathbb{R}$ under CFL.

Proposition 7. Assume $0 < \nu \le 1$. Then $|b_{r,n}| \le 1$.

For the second estimate we use inequality (4.7).

PROPOSITION 8. Assume $0 < \nu < 1$. Then $|b_{r,n}| \le m_4(k) (n\nu(1-\nu))^{-\frac{1}{2k+2}}$ for some suitable constant $m_4(k) > 0$.

Using the symmetry relation (2.7) one gets $|b_{r,n}| \leq \frac{1}{\pi} \int_0^{\pi} |\lambda(\theta)|^n d\theta$. Therefore

$$|b_{r,n}| \le \frac{1}{\pi} \int_0^{\pi} \left| 1 - m_1(k)\nu(1-\nu)\theta^{2k+2} \right|^n d\theta \le \frac{1}{\pi} \int_0^{\pi} e^{-nm_1(k)\nu(1-\nu)\theta^{2k+2}} d\theta.$$

It is a matter of standard manipulations to get

$$|b_{r,n}| \le \frac{1}{\pi \left(nm_1(k)\nu(1-\nu)\right)^{\frac{1}{2k+2}}} \int_0^\infty e^{-y^{2k+2}} dy.$$

For a general $k \ge 1$ the integral $\int_0^\infty e^{-y^{2k+2}} dy$ is uniformly bounded by 2. It proves the claim.

The last estimate comes from (3.1), which allows one to write

$$b_{r,n} = \frac{1}{2\pi} \int_0^{2\pi} \sigma(\theta)^n e^{\mathbf{i}(r+n\nu)\theta} d\theta,$$

where σ has been defined previously. Integrating by parts twice, one gets another expression:

$$b_{r,n} = -\frac{1}{2\pi (r + n\nu)^2} \int_0^{2\pi} \left(n\sigma''(\theta)\sigma(\theta)^{n-1} + n(n-1)(\sigma'(\theta))^2 \sigma(\theta)^{n-2} \right) e^{\mathbf{i}(r + n\nu)\theta} d\theta.$$

The function $g(\theta) = (n\sigma''\sigma^{n-1} + n(n-1)(\sigma')^2\sigma^{n-2})e^{\mathbf{i}(r+n\nu)\theta}$ which is under the integral, satisfies, like the function λ , the conjugation relation (2.7). So

$$(5.2) |b_{r,n}| \le \frac{1}{\pi |r + n\nu|^2} \int_0^{\pi} \left(n|\sigma''||\sigma| + n|n - 1||\sigma'|^2 \right) |\lambda|^{n-2} d\theta,$$

for which we can use our estimates on the derivatives of σ .

Proposition 9. Assume $0 < \nu \le 1$. Then

$$|b_{r,n}| \le \frac{m_5(k)}{|r+n\nu|^2} (\nu(1-\nu)n)^{\frac{1}{2k+2}}, \quad m_5(k) > 0.$$

The right-hand side of (5.2) depends on two terms. The first one, multiplied by $|r + n\nu|^2$, is

$$A = \frac{1}{\pi} \int_0^{\pi} n|\sigma''(\theta)||\lambda(\theta)|^{n-1}d\theta$$

$$\leq \frac{m_3(k)\nu(1-\nu)n}{\pi} \int_0^{\pi} \left|1-m_1(k)\nu(1-\nu)\theta^{2k+2}\right|^{n-1}\theta^{2k}d\theta$$

$$\leq \frac{m_3(k)\nu(1-\nu)n}{\pi} \frac{1}{(m_1(k)\nu(1-\nu)(n-1))^{\frac{2k+1}{2k+2}}} \int_0^{\infty} e^{-y^{2k+2}}y^{2k}dy.$$

One has

$$\frac{1}{(n-1)^{\frac{2k+1}{2k+2}}} = \left(\frac{n}{n-1}\right)^{\frac{2k+1}{2k+2}} \times \frac{1}{n^{\frac{2k+1}{2k+2}}} \le 2\frac{1}{n^{\frac{2k+1}{2k+2}}}, \quad \text{provided } 2 \le n.$$

Therefore

(5.3)
$$A \le m_6(k)(\nu(1-\nu)n)^{1-\frac{2k+1}{2k+2}} = m_6(k)(\nu(1-\nu)n)^{\frac{1}{2k+2}}$$

with $m_6(k) > 0$. If n = 0, this inequality is trivially true. If n = 1, then

$$A \le \frac{m_3(k)\nu(1-\nu)}{\pi} \int_0^{\pi} \theta^{2k} d\theta,$$

which means that inequality (5.3) is also trivially true.

For $n \geq 3$ the second term in (5.2) implies

$$\begin{split} B &= \frac{1}{\pi} \int_0^\pi n(n-1) |\sigma'|^2 |\lambda|^{n-2} d\theta \\ &\leq \frac{(m_2(k)\nu(1-\nu))^2 n^2}{\pi} \int_0^\pi \left| 1 - m_1(k)\nu(1-\nu) \theta^{2k+2} \right|^{n-2} \theta^{2(2k+1)} d\theta \\ &\leq \frac{(m_2(k)\nu(1-\nu))^2 n^2}{\pi} \frac{1}{(m_1(k)\nu(1-\nu)(n-2))^{\frac{4k+3}{2k+2}}} \int_0^\infty e^{-y^{2k+2}} y^{4k+2} dy. \end{split}$$

One has

$$\frac{1}{(n-2)^{\frac{2k+1}{2k+2}}} = \left(\frac{n}{n-2}\right)^{\frac{2k+1}{2k+2}} \times \frac{1}{n^{\frac{2k+1}{2k+2}}} \le 3\frac{1}{n^{\frac{2k+1}{2k+2}}}, \quad \text{provided } n \ne 0, 1, 2.$$

Therefore

(5.4)
$$B \le m_7(k)(\nu(1-\nu)n)^{2-\frac{4k+3}{2k+2}} = m_7(k)(\nu(1-\nu)n)^{\frac{1}{2k+2}}$$

with $m_7(k) > 0$. If n = 0 or n = 1, this inequality is trivially true. If n = 2, then

$$B \le \frac{(m_2(k)\nu(1-\nu))^2 n^2}{\pi} \int_0^{\pi} \theta^{2(2k+1)} d\theta$$

and inequality (5.4) is also trivially true. Defining $m_5 = m_6 + m_7$ proves the claim.

5.2. Proof of Theorem 2. We use all the previous estimates to get a bound for

$$C = \sum_{r \in \mathbb{Z}} |b_{r,n}|.$$

Let us define an arbitrary integer number N > 0. We split the sum into two parts:

$$\begin{split} C &= \sum_{|r+n\nu| \geq N+1} |b_{r,n}| + \sum_{|r+n\nu| < N+1} |b_{r,n}| \\ &\leq \sum_{|r+n\nu| \geq N+1} \frac{m_5(k)}{|r+n\nu|^2} (\nu(1-\nu)n)^{\frac{1}{2k+2}} + \sum_{|r+n\nu| < N+1} \min\left(1, m_4(k) \left(n\nu(1-\nu)\right)^{-\frac{1}{2k+2}}\right). \end{split}$$

For the first sum, one notices that

$$\sum_{|r+n\nu| \geq N+1} \frac{1}{|r+n\nu|^2} \leq 2 \sum_{j \in \mathbb{N}, j \geq N+1} \frac{1}{j^2} \leq \frac{2}{N}.$$

Therefore

(5.5)
$$C \le 2m_5(k) \frac{(\nu(1-\nu)n)^{\frac{1}{2k+2}}}{N} + 2N \min\left(1, m_4(k) \left(n\nu(1-\nu)\right)^{-\frac{1}{2k+2}}\right).$$

To finish the discussion, we distinguish between two cases.

• Assume that $m_4(k) (n\nu(1-\nu))^{-\frac{1}{2k+2}} > 1$. So $(n\nu(1-\nu))^{\frac{1}{2k+2}} < m_4(k)$. Then we take N=1, which turns into

$$C \le 2m_5(k)m_4(k) + 2.$$

• Assume on the contrary that $m_4(k) \left(n\nu(1-\nu)\right)^{-\frac{1}{2k+2}} < 1$. Then we take $N = \left[\frac{(n\nu(1-\nu))^{\frac{1}{2k+2}}}{m_4(k)}\right] + 1 \ge 1$, so

$$\frac{(n\nu(1-\nu))^{\frac{1}{2k+2}}}{m_4(k)} \le N \le 2\frac{(n\nu(1-\nu))^{\frac{1}{2k+2}}}{m_4(k)}.$$

Inserting (5.5), we get $C \le 2 \frac{m_5(k)}{m_4(k)} + 4$.

Finally taking the maximum of the two bounds for C, it proves the claim of Theorem 2 with

$$D_p = \max\left(2m_5(k)m_4(k) + 2, 2\frac{m_5(k)}{m_4(k)} + 4\right), \qquad p = 2k + 1.$$

5.3. Why Theorem 2 is not true for even order schemes. The proof of Theorem 2 given in this paper is essentially a consequence of estimates (4.5)–(4.10), rewritten as (p = 2k + 1)

(5.6)
$$\begin{cases} |\lambda(\theta)| \leq 1 - m_1(k)\nu(1-\nu)\theta^{p+1}, \\ |\sigma'(\theta)| \leq m_2(k)\nu(1-\nu)\theta^p, \\ |\sigma''(\theta)| \leq m_3(k)\nu(1-\nu)\theta^{p-1}. \end{cases}$$

In the even order these estimates are no longer true, and this is the reason why Theorem 2 does not hold for even order schemes.

Indeed consider the case p=2k. From (A.3) one gets $h(\theta) \approx -\nu\theta$. Using the inequality $\nu(1-\nu)C_k^1 \leq |\alpha_{k,2k}| \leq \nu(1-\nu)C_k^2$, one gets

(5.7)
$$|\lambda(\theta)| \le 1 - \widehat{m_1(k)} \nu^2 (1 - \nu) \theta^{p+2}.$$

Inequalities (4.9)–(4.10) remain the same, that is,

$$|\sigma'(\theta)| \le \widehat{m_2(k)}\nu(1-\nu)\theta^p$$

and

$$|\sigma''(\theta)| \le \widehat{m_3(k)}\nu(1-\nu)\theta^{p-1}.$$

We observe two problems. The first one is that the power of θ is one more in (5.7) than it is in (5.6). The reason is that this inequality measures the damping provided by the amplification factor. For even order schemes, the leading term is necessarily purely imaginary and so cannot contribute to the damping. The damping is provided by the next term in the series. This is the reason for the p+2 instead of the p+1. Moreover we observe the coefficient before the power of θ has a stronger dependence with respect to ν (ν^2 instead of ν). It highlights that for small CFL number, the damping in (5.7) is not sufficient to counterbalance the oscillations caused by (5.8)–(5.9). For example, it is an explanation why the Lax–Wendroff scheme highly oscillates in the regime $\nu \approx 0$.

Since the p = 2k + 2 case is treated in section (A.3) by comparison with the case p = 2k with trick $\nu' = 1 - \nu$, therefore the same conclusion holds true. But we get in the case p = 2k + 2 the estimates

$$(5.10) |\lambda(\theta)| \le 1 - \widetilde{m_1(k)}\nu(1-\nu)^2\theta^{p+2}$$

(5.11)
$$|\sigma'(\theta)| \le \widetilde{m_2(k)}\nu(1-\nu)\theta^p,$$

and

$$(5.12) |\sigma''(\theta)| \le \widetilde{m_3(k)}\nu(1-\nu)\theta^{p-1}.$$

For this family the damping (5.10) is very small in the regime $\nu \approx 1$. Numerical observations (not reproduced in this paper) confirm this behavior; see [8] for the Beam–Warming scheme. See also the discussion in the section on numerical results.

6. Incremental formulation. The incremental or finite volume formulation is

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + a \frac{u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n}{\Delta x} = 0.$$

To obtain this incremental formulation, one has to reorganize all contributions in the general finite difference formulation (1.2). A necessary condition is of course $\sum \alpha_r(\nu) = 1$. This is easy to check.

But there is also the "dissipation" issue. Indeed if the flux $u_{j+\frac{1}{2}}^n$ contains a fractional dependence with respect to the CFL number ν , which is possible as well because we divide by Δt to pass from the finite difference formulation to the incremental formulation, then the scheme degenerates in some sense in the limit $\nu \to 0^+$. It corresponds to a loss of consistency in this regime. The typical example is the Lax–Friedrichs scheme, which is known to be very dissipative for small time steps. This is why it is important to verify that the flux $u_{j+\frac{1}{2}}^n$ can be defined as a polynomial with respect to ν , without any $\frac{1}{\nu}$ dependence.

Proposition 10. For p = 2k+1, p = 2k, and p = 2k+2, the flux is a polynomial with respect to ν .

The correspondence between the amplification factor and the Fourier symbol $\mu(\theta)$ of the flux is $\lambda(\theta) = 1 - \nu \mu(\theta) (1 - e^{i\theta}) \iff \mu(\theta) = \frac{1}{\nu} \times \frac{1 - \lambda(\theta)}{1 - e^{i\theta}}$. The flux is a polynomial function of the CFL number ν if and only if its Fourier symbol $\mu(\theta)$ is polynomial with respect to ν . From (1.6) and (3.1) it is clear that if $\nu = 0$, then $\alpha_{k,p} = 0$ for $p \geq k$ (which is our case): then $\lambda(\theta) - 1 = 0$ for $\nu = 0$. Therefore $\lambda(\theta) - 1$ is a polynomial with respect to ν , with the root 0. So division by ν is not singular. Finally the flux μ is a polynomial with respect to ν . This ends the proof. \square

For (p,k) = (1,0) one gets back the upwind flux $u_{j+\frac{1}{2}} = u_j$. For (p,k) = (3,1) one gets the flux of the O3 scheme studied in [8]:

$$u_{j+\frac{1}{2}} = u_j + \frac{(2-\nu)}{6}(1-\nu)(u_{j+1} - u_j) + \frac{1+\nu}{6}(1-\nu)(u_j - u_{j-1}).$$

This scheme can also be defined as a convex combination of the Lax–Wendroff scheme (coefficient $\frac{2-\nu}{3}$) and of the Beam–Warming scheme (coefficient $\frac{1+\nu}{3}$).

The case (p, k) = (5, 2) is the next one in the list p = 2k + 1. In finite difference form, one has

$$\begin{split} u_{j}^{n+1} &= u_{j+2}^{n} + (2+\nu)(u_{j+1}^{n} - u_{j+2}^{n}) + \frac{(2+\nu)(1+\nu)}{2}(u_{j}^{n} - 2u_{j+1}^{n} + u_{j+2}^{n}) \\ &\quad + \frac{(2+\nu)(1+\nu)\nu}{6}(u_{j-1}^{n} - 3u_{j}^{n} + 3u_{j+1}^{n} - u_{j+2}^{n}) \\ &\quad + \frac{(2+\nu)(1+\nu)\nu(\nu-1)}{24}(u_{j-2}^{n} - 4u_{j-1}^{n} + 6u_{j}^{n} - 4u_{j+1}^{n} + u_{j+2}^{n}) \\ &\quad + \frac{(2+\nu)(1+\nu)\nu(\nu-1)(\nu-2)}{120} \\ &\quad \times (u_{j-3}^{n} - 5u_{j-2}^{n} + 10u_{j-1}^{n} - 10u_{j}^{n} + 5u_{j+1}^{n} - u_{j+2}^{n}). \end{split}$$

The finite volume formulation of this scheme is

$$\begin{split} u_{j+\frac{1}{2}} &= u_{j+2}^n + \frac{\nu+3}{2}(u_{j+1}^n - u_{j+2}^n) + \frac{(2+\nu)(1+\nu)}{6}(u_j^n - 2u_{j+1}^n + u_{j+2}^n) \\ &\quad + \frac{(2+\nu)(1+\nu)(\nu-1)}{24}(u_{j-1}^n - 3u_j^n + 3u_{j+1}^n - u_{j+2}^n) \\ &\quad + \frac{(2+\nu)(1+\nu)(\nu-1)(\nu-2)}{120}(u_{j-2}^n - 4u_{j-1}^n + 6u_j^n - 4u_{j+1}^n + u_{j+2}^n). \end{split}$$

7. Numerical examples. In order to illustrate the interest for practical computations of the strong stability properties of the advection schemes p=2k+1, we give the results of some numerical examples. The initial condition that we consider is singular. It is a Dirac mass at the origin $u(0,\cdot)=\delta(\cdot)$. The exact solution is of course $u(t,\cdot)=\delta(\cdot-at)$. We use 100 cells $(\Delta x=\frac{1}{100})$, so the numerical at t=0 is a discrete Dirac function of amplitude 100 in the central cell. We use the schemes p=2k+1 and p=2k for k=0 to k=4.

A first series of results ($\nu = 0.5$) is displayed on Figure 7.1. The results for p = 2k + 1 are clearly less oscillating than for p = 2k. It is not surprising since the odd order schemes are TVB (1.8).

We plot in Figure 7.2 the results computed with a small CFL number $\nu=0.05$. The strong dependence of the even order schemes p=2k with respect to ν is visible. On the contrary the L^1 stability and TVB stability of the odd order schemes uniformly with respect to the CFL number ν is also evident.

It is also worthwhile to derive more quantitative estimates of the diffusion and dispersion coefficients of the schemes in order to have a better understanding of the numerical results. These coefficients are the first two nontrivial coefficients in the Taylor expansion of

$$\lambda(\theta)e^{-\nu \mathbf{i}\theta} = \left[1 - \mathbf{i}^{p+1}\alpha_{k,p}2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} e^{(k - \frac{p}{2} + \nu)\mathbf{i}(\varphi - \theta)} d\varphi\right].$$

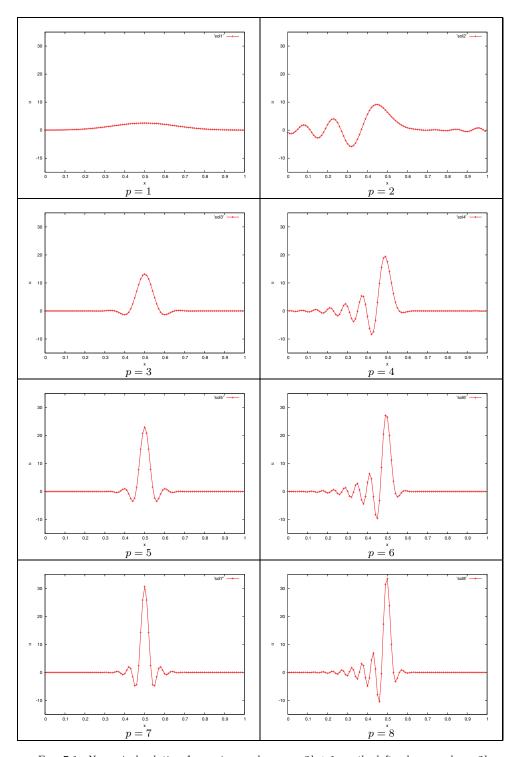


Fig. 7.1. Numerical solution for various orders: p=2k+1 on the left column and p=2k on the right column. The CFL number is $\nu=0.5$.

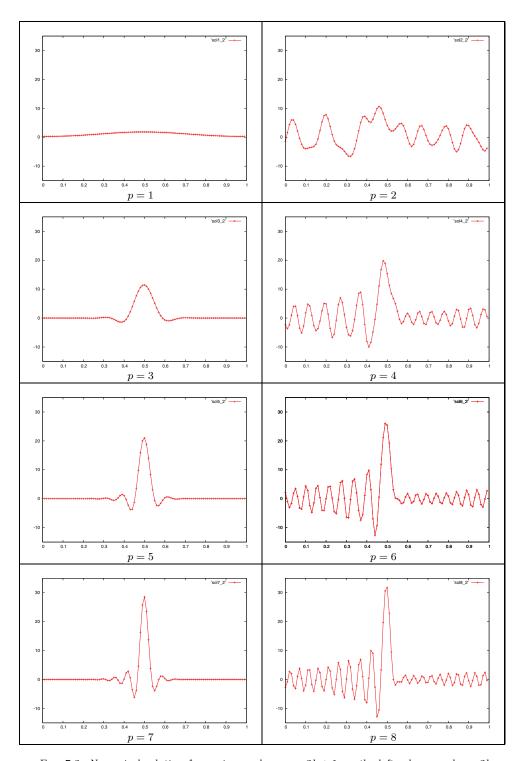


Fig. 7.2. Numerical solution for various orders: p=2k+1 on the left column and p=2k on the right column. The CFL number is small, $\nu=0.05$.

Since $0 \le \varphi \le \theta$, one has the second order Taylor expansions $\sin \frac{\theta - \varphi}{2} = \frac{\theta - \varphi}{2} + o(\theta^2)$ and $e^{(k - \frac{p}{2} + \nu)\mathbf{i}(\varphi - \theta)} = 1 + (k - \frac{p}{2} + \nu)\mathbf{i}(\varphi - \theta) + o(\theta^2)$. Therefore we immediately obtain the following:

• If p = 2k + 1, then

$$\lambda(\theta)e^{-\nu\mathbf{i}\theta}=1-(-1)^{k+1}\alpha_{k,2k+1}\left(\frac{\theta^{p+1}}{p+1}-\mathbf{i}\left(\nu-\frac{1}{2}\right)\frac{\theta^{p+2}}{p+2}\right)+o\left(\theta^{p+2}\right).$$

The diffusion coefficient is the real coefficient in this expansion, and the dispersion coefficient is the imaginary one. In this case the diffusion coefficient is the leading order term $-(-1)^{k+1}\frac{\alpha_{k,2k+1}}{p+1}<0$. The dispersion coefficient is the second one, $(-1)^{k+1}\left(\nu-\frac{1}{2}\right)\frac{\alpha_{k,2k+1}}{p+2}\mathbf{i}$. In order to minimize the phase error of the scheme, a reasonable prescription is to diminish the dispersion coefficient as much as possible, that is, to choose $\nu=\frac{1}{2}$.

• If p = 2k, then

$$\lambda(\theta)e^{-\nu \mathbf{i}\theta} = 1 - \mathbf{i}(-1)^k \alpha_{k,2k} \left(\frac{\theta^{p+1}}{p+1} - \mathbf{i}\nu \frac{\theta^{p+2}}{p+2} \right) + o\left(\theta^{p+2}\right).$$

In this case the diffusion coefficient is the second one and the dispersion coefficient is the leading one. Moreover one sees that dispersion outperforms diffusion for small CFL number $\nu \approx 0^+$. This explains why the results become worse in the right column of Figure 7.2.

• If p = 2k + 2, then

$$\lambda(\theta)e^{-\nu \mathbf{i}\theta} = 1 - 1 - \mathbf{i}(-1)^{k+1}\alpha_{k,2k+2} \left(\frac{\theta^{p+1}}{p+1} - \mathbf{i}(\nu - 1)\frac{\theta^{p+2}}{p+2}\right) + o\left(\theta^{p+2}\right).$$

In this case dispersion is also the leading term. Dispersion outperforms diffusion for a CFL number close to one, $\nu \approx 1$.

Appendix. The case p even. Assume that p is even. Then $|\lambda(\theta)| \leq 1$ is equivalent to

$$(A.1) |\lambda(\theta)|^2 = \left(1 + (-1)^{\frac{p}{2}} \alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} \sin\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2 + \left(\alpha_{k,p} 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2 \le 1.$$

A.1. Necessary condition: p even. By inspection of (A.1) one has to study the sign of the integral

$$\int_0^\theta \sin^p \frac{\theta - \varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) (\varphi - \theta) \ d\varphi = -\int_0^\theta \sin^p \frac{\varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi$$

for all $\theta \in [0, \pi]$. Assume for a while that $|k - \frac{p}{2} + \nu| \ge 2$. Define $\overline{\theta} = \frac{2\pi}{|k - \frac{p}{2} + \nu|} \in [0, \pi]$. Then

$$\begin{split} & \int_0^{\overline{\theta}} \sin^p \frac{\varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi \\ & = \int_0^{\frac{\overline{\theta}}{2}} \sin^p \frac{\varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi + \int_{\frac{\overline{\theta}}{2}}^{\overline{\theta}} \sin^p \frac{\varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi. \end{split}$$

By construction $-\sin(k-\frac{p}{2}+\nu)(\varphi+\frac{\overline{\theta}}{2})=\sin(k-\frac{p}{2}+\nu)\varphi\geq 0$. So

$$\begin{split} & \int_0^{\overline{\theta}} \sin^p \frac{\varphi}{2} \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi \\ & = \int_0^{\frac{\overline{\theta}}{2}} \left(\sin^p \frac{\varphi}{2} - \sin^p \frac{\varphi + \frac{\overline{\theta}}{2}}{2} \right) \sin \left(k - \frac{p}{2} + \nu \right) \varphi \ d\varphi \end{split}$$

has the sign of $k - \frac{p}{2} + \nu$. For small θ the sign is the opposite. This means the scheme is not stable.

Therefore the necessary condition is

(A.2)
$$\left|k - \frac{p}{2} + \nu\right| < 2 \quad \forall \nu \in [0, 1].$$

That is, $-2 < k - \frac{p}{2} \le k - \frac{p}{2} + \nu \le k - \frac{p}{2} + 1 < 2$. Since p is even by the hypothesis, the solutions are $k - \frac{p}{2} = -1 \iff p = 2k + 2$ and $k - \frac{p}{2} = 0 \iff p = 2k$.

A.2. Sufficient condition: p = 2k. Consider $B(\theta) = |\lambda(\theta)|^2$. Straightforward calculations show that $B'(\theta) = |\alpha_{k,2k}| 2^{2k} \sin^{2k} \frac{\theta}{2} h(\theta)$, with

(A.3)
$$h(\theta) = |\alpha_{k,2k}| \, 2^{2k} \int_0^\theta \sin^{2k} \frac{\theta - \varphi}{2} \cos \nu \varphi \, d\varphi - \sin \nu \theta.$$

The function h is the difference of two increasing functions. It is not possible to conclude directly about the sign of the derivative of h. However, we have $h(\pi) = 0$ since $B'(\pi) = 0$ is true. Therefore

(A.4)
$$|\alpha_{k,2k}(\nu)| \, 2^{2k} \int_0^\pi \sin^{2k} \frac{\varphi}{2} \cos \nu (\pi - \varphi) \, d\varphi - \sin \nu \pi = 0.$$

This formula is true for all Courant numbers $\nu \in [0,1]$. Note that $\alpha_{k,2k}(\nu)$ is by definition a function of ν ; this will be important at the end of the proof. Consider $\theta \in [0,\pi]$ and define

$$\widetilde{\nu} = \nu \frac{\theta}{\pi} \in [0, 1].$$

Writing (A.4) for $\widetilde{\nu}$ one gets

$$|\alpha_{k,2k}(\widetilde{\nu})| \, 2^{2k} \int_0^{\pi} \sin^{2k} \frac{\varphi}{2} \cos(\widetilde{\nu}\theta - \widetilde{\nu}\varphi) \, d\varphi - \sin(\widetilde{\nu}\pi) = 0,$$

that is, after a change of variable, $\varphi \leftarrow \frac{\theta}{\pi} \varphi$ in the integral,

(A.5)
$$\frac{\pi}{\theta} |\alpha_{k,2k}(\widetilde{\nu})| \, 2^{2k} \int_0^{\theta} \sin^{2k} \frac{\frac{\pi}{\theta} \times \varphi}{2} \cos(\nu\theta - \nu\varphi) \, d\varphi - \sin\nu\theta = 0.$$

Let us compare the left-hand side of (A.5) with the definition of h (A.3). If one shows that

$$A_1: \sin \frac{\varphi}{2} \le \sin \frac{\frac{\pi}{\theta} \times \varphi}{2}$$

and

$$A_2: |\alpha_{k,2k}(\nu)| \le \frac{\pi}{\theta} |\alpha_{k,2k}(\widetilde{\nu})|,$$

then one concludes that $h(\theta) \leq h(\pi) = 0$, which turns into the stability of the scheme. It remains to prove A_1 and A_2 .

Proof of A_1 . One has by construction $\frac{\varphi}{2} \leq \frac{\pi}{\theta} \times \frac{\varphi}{2} \leq \frac{\pi}{2}$. This proves A_1 . Proof of A_2 . By definition of $\alpha_{k,2k}$ (see (1.6)), one has

$$\frac{\theta \alpha_{k,2k}(\nu)}{\pi \alpha_{k,2k}(\widetilde{\nu})} = \prod_{q=1}^k \frac{q^2 - \nu^2}{q^2 - \widetilde{\nu}^2}.$$

Since $0 < \widetilde{\nu} \le \nu \le 1$, then $\frac{\theta \alpha_{k,2k}(\nu)}{\pi \alpha_{k,2k}(\widetilde{\nu})} \le 1$. The proof of A_2 is finished. Therefore the scheme is L^2 stable for p = 2k.

A.3. Sufficient condition: p = 2k + 2. The proof is by comparison with the previous case. Set k' = k + 1. We make use of inequality (A.1) with the parameters p = 2k' and $\nu' = 1 - \nu$, that is,

$$(A.6) \qquad \left(1 + (-1)^{\frac{p}{2}} \alpha_{k',p}(\nu') 2^p \int_0^{\theta} \sin^p \frac{\theta - \varphi}{2} \sin\left(k' - \frac{p}{2} + \nu'\right) (\varphi - \theta) d\varphi\right)^2$$

$$+ \left(\alpha_{k',p}(\nu') 2^p \int_0^{\theta} \cos^p \frac{\theta - \varphi}{2} \cos\left(k' - \frac{p}{2} + \nu'\right) (\varphi - \theta) d\varphi\right)^2 \le 1.$$

One has

$$k' - \frac{p}{2} + \nu' = k' - \frac{p}{2} + 1 - \nu = 1 - \nu = -\left(k - \frac{p}{2} - \nu\right)$$

and

$$\alpha_{k',p}(\nu') = \frac{\prod_{q=0}^{2k+2} (k' + \nu' - q)}{p!} = \frac{\prod_{q=0}^{2k+2} (k+1+1-\nu - q)}{p!}$$
$$= -\frac{\prod_{q=0}^{2k+2} (k+\nu - q)}{p!} = -\alpha_{k,p}(\nu).$$

Inserting the previous inequality, one gets directly

$$(A.7) \qquad \left(1 + (-1)^{\frac{p}{2}} \alpha_{k,p}(\nu) 2^p \int_0^\theta \sin^p \frac{\theta - \varphi}{2} \sin\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2$$

$$+ \left(\alpha_{k,p}(\nu) 2^p \int_0^\theta \cos^p \frac{\theta - \varphi}{2} \cos\left(k - \frac{p}{2} + \nu\right) (\varphi - \theta) d\varphi\right)^2 \le 1.$$

This proves inequality (A.1) for p = 2k + 2 and $\nu \in [0, 1]$.

It is striking to observe that (A.1) is true for a larger interval of stability $\nu \in [0, 2]$. The reason is that the left-hand side of (A.6) is a symmetric function of ν' because (recall $k' - \frac{p}{2} = 0$) $\alpha_{k',2k'}(\nu')\sin(\nu'(\varphi - \theta)) = \alpha_{k',2k'}(-\nu')\sin(-\nu'(\varphi - \theta))$. Then inequality (A.6) is true for $-1 \le \nu' \le 1$, that is, $\nu = 1 - \nu' \in [0, 2]$.

REFERENCES

- L. Anné, P. Joly, and Q.H. Tran, Construction and analysis of higher order finite difference schemes for the 1D wave equation, Comput. Geosci., 4 (2000), pp. 207–249.
- [2] C. Bailly and C. Bogey, Numerical solution of acoustic propagation problems using linearized Euler equations, AIAA J., 38 (2000), pp. 22–29.
- [3] H. Boïng, R. Jeltsch, and P. Kiani, Convergence and stability of multi time level schemes for the advection equations, in Finite Approximations in Fluid Mech. II, Notes Numer. Fluid Mech. 25, Vieweg, Braunschweig, 1989, pp. 30–42.
- [4] P. Brenner, V. Thomée, and L. Wahlbin, Besov Spaces and Applications to Difference Methods for Initial Value Problems, Lecture Notes in Math. 434, Springer-Verlag, Berlin, New York, 1975.
- [5] V. Daru and C. Tenaud, High order one-step monotonicity-preserving schemes for unsteady compressible flow calculations, J. Comput. Phys., 193 (2004), pp. 563-594.
- [6] S. DEL PINO AND H. JOURDREN, Arbitrary high-order schemes for linear advection and waves equations: Application to hydrodynamics and aeroacoustic, C. R. Math. Acad. Sci. Paris, 342 (2006), pp. 441–446.
- [7] S. Del Pino, B. Després, P. Havé, H. Jourdren, and P.F. Piserchia, 3D finite volume simulation of acoustic waves in the earth atmosphere, Comput. & Fluids, 38 (2009), pp. 765–777.
- [8] B. DESPRÉS, Finite volume transport schemes, Numer. Math., 108 (2008), pp. 529-556.
- [9] S.K. Godunov and V.S. Ryaben'kıı, Difference Schemes: An Introduction to the Underlying Theory, North-Holland, Amsterdam, 1987.
- [10] A. HARTEN, B. ENGQUIST, S. OSHER, AND S. CHAKRAVARTHY, Uniformly high-order accurate essentially nonoscillatory schemes, III, J. Comput. Phys., 71 (1987), pp. 231–303.
- [11] A. ISERLES AND S.P. NORSETT, Order Stars, Chapman & Hall, London, 1991.
- [12] A. ISERLES AND G. STRANG, The optimal accuracy of difference schemes, Trans. Amer. Math. Soc., 277 (1983), pp. 779–803.
- [13] R. Jeltsch, R.A. Renaut, and J.H. Smit, On the Courant-Friedrichs-Lewy condition equipped with order for hyperbolic differential equations, in Proceedings of the Fifth International Conference on Hyperbolic Problems: Theory, Numerics, Applications, World Scientific, River Edge, NJ, 1996, pp. 30–42.
- [14] P.D. LAX AND B. WENDROFF, On the stability of difference schemes, Comm. Pure Appl. Math., 15 (1962), pp. 363–371.
- [15] B.P. LEONARD, The ultimate conservative difference scheme applied to unsteady one-dimensional advection, Comput. Methods Appl. Mech. Engrg., 88 (1991), pp. 17–74.
- [16] R.D. RICHTMYER AND K.W. MORTON, Difference Methods for Initial Value Problems, 2nd ed., John Wiley and Sons, New York, 1967.
- [17] W. Rudin, Real and Complex Analysis, 3rd ed., MacGraw-Hill, New York, 1987.
- [18] T. SCHWARTZKOPFF, M. DUMBSER, AND C.D. MUNZ, Fast high order ADER schemes for linear hyperbolic equations and their numerical dispersion and dissipation, J. Comput. Phys., 197 (2004), pp. 532–538.
- [19] C.W. Shu, Essentially Nonoscillatory and Weighted Essentially Nonoscillatory Schemes for Hyperbolic Conservation Laws, ICASE Report 97-65, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA, 1997.
- [20] G. Strang, Trigonometric polynomials and difference methods of maximum accuracy, J. Math. Phys., 41 (1962), pp. 147–154.
- [21] V. Thomée, Stability of difference schemes in the maximum-norm, J. Differential Equations, 1 (1965), pp. 273–292.
- [22] V.A. TITAREV AND E.T. TORO, High order ADER schemes for scalar advection-diffusionreaction equations, J. Comput. Fluid Dynamics, 12 (2003), pp. 1-6.
- [23] R. VICHNEVETSKI AND J. BOWLES, Fourier Analysis of Numerical Approximations of Hyperbolic Equations, SIAM, Philadelphia, 1982.
- [24] R.F. WARMING AND R.M. BEAM, Recent advances in the development of implicit schemes for the equations of gas dynamics, in Proceedings of the Seventh International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics, Springer, 1981, pp. 429–433.