Numerische Mathematik 7, 73—77 (1965)

The Near-Stability of the Lax-Wendroff Method*

By

G. W. HEDSTROM

In discussing finite difference methods for the solution of hyperbolic partial differential equations, STETTER [1] used estimates on some absolutely convergent Fourier series to prove stability and instability with respect to uniform convergence. If f, a complex valued function on the circle, has an absolutely convergent Fourier series, then the n-th power of f also has an absolutely convergent Fourier series:

$$f^n(x) = \sum_k c_{kn} e^{ikx}, \qquad \|f^n\| = \sum_k |c_{kn}| < \infty, \qquad n = 1, 2, \ldots.$$

A difference scheme determines a corresponding f. If this f is such that the $\|f^n\|$ are bounded, then the difference scheme is stable in the uniform norm; and if f is a polynomial with $\|f^n\|$ unbounded, then it is unstable [1, p. 407].

In this paper we supply a proof of the instability, but near-stability of the Lax-Wendroff method. STETTER [1, p. 421] has shown that this is a consequence of the following theorem.

Theorem. Let $f \in C^2$ on the circle; let $|f(t)| < 1$, $t \neq 0$; and let

$$f(t) = \exp\{i\alpha t + \varphi(t)\}, \qquad -\varrho < t < \varrho, \tag{1}$$

for some $\varrho > 0$, α real, and $\varphi(t)$ analytic in $|t| < \varrho$ such that

$$\varphi(t) = i\beta t^{2N-1} - \gamma t^{2N} + O(t^{2N+1}), \qquad t \to 0, \tag{2}$$

where β is real, $\beta \neq 0$, $\gamma > 0$, and N is one of $2, 3, 4, \ldots$. Then there exist constants C', C, depending on N, such that

$$C' n^{1/4N} < \|f^n\| < C n^{1/4N}. \tag{3}$$

Remarks. When a $(2N-1)$-point Lax-Wendroff average is used, the f is a polynomial of this type [2, pp. 147—148], and thus there is a mild instability. For the case $N = 2$,

$$F(t) = 1 - \alpha^2(1 - \cos t) + i\alpha \sin t,$$

with $0 < |\alpha| < 1$, STETTER [1, p. 423] showed that $\|F^n\| < C n^{\frac{1}{4}}$ and gave experimental evidence to indicate that $\|F^n\| > C' n^{\frac{1}{12}}$.

Proof of the theorem. We give a proof for the case $\beta > 0$.

We need estimates for

$$c_{kn} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^n(t)\, e^{-ikt}\, dt. \tag{4}$$

* This research was supported in part by NSF Grant 06016.

To get them we first rewrite c_{kn} in the form

$$c_{kn}=\frac{1}{2\pi}\int_{-\pi}^{\pi} g^n(t)\, e^{-in\omega t}\,dt,$$

where $g(t)=f(t)\,e^{-i\alpha t}$ and

$$\omega=(k-n\,\alpha)/n. \tag{5}$$

Integration by parts twice gives

$$|c_{kn}|\leqq C/n\,\omega^2 \tag{6}$$

for some $C>0$.

When $|\omega|$ is small, we use a different estimate based on

$$c_{kn}=\frac{1}{2\pi}\left\{\int_{-\pi}^{-\varrho}+\int_{-\varrho}^{\varrho}+\int_{\varrho}^{\pi}\right\} g^n\, e^{-in\omega t}\,dt=I_1+I_2+I_3. \tag{7}$$

For I_1+I_3 we have

$$|I_1+I_3|\leqq 2\pi \max_{\varrho\leqq|t|\leqq\pi}|f^n(t)|=2\pi\,\vartheta^n, \tag{8}$$

for some ϑ, $0<\vartheta<1$. To estimate I_2 we use a contour through 2 saddle points of

$$g^n(t)\,e^{-in\omega t}=\exp n\left(\varphi(t)-i\,\omega\,t\right), \qquad |t|<\varrho.$$

For $\omega\to 0+$, the saddles we use satisfy

$$t_j=(-1)^j\lambda\,\omega^{1/2(N-1)}-i\,\gamma\,N(N-1)^{-1}\lambda^{2N}\omega^{1/(N-1)}+O\left(\omega^{3/2(N-1)}\right), \tag{9}$$

$j=1, 2$, where

$$\lambda=[\beta(2N-1)]^{-1/2(N-1)}, \qquad \lambda>0. \tag{10}$$

Since the saddle points depend continuously on ω, it is possible to find an A such that $|t_j|\leqq\varrho/2$ if $0\leqq\omega\leqq A$. It will be important that

$$\begin{aligned}\varphi(t_j)-i\,\omega\,t&=-i(-1)^j\lambda(N-1)\left(N-\tfrac{1}{2}\right)^{-1}\omega^{(N-\frac{1}{2})/(N-1)}-\gamma\,\lambda^{2N}\omega^{N/(N-1)}+\\&\quad+O\left(\omega^{(N+\frac{1}{2})/(N-1)}\right), \qquad \omega\to 0+, \qquad j=1, 2.\end{aligned} \tag{11}$$

The Taylor series about a saddle is of the form

$$\varphi(t)-i\,\omega\,t=\varphi(t_j)-i\,\omega\,t_j+\sum_{k=2}^{\infty}a_k(\omega)\,(t-t_j)^k, \qquad j=1, 2. \tag{12}$$

For some positive constants M_k, independent of ω,

$$|a_2(\omega)|=M_2\,\omega^{(N-\frac{3}{2})/(N-1)}+O(\omega), \qquad \omega\to 0+;$$

$$|a_k(\omega)|\leqq M_k\,\omega^{(2N-k-1)/2(N-1)}, \qquad 0\leqq\omega\leqq A, \qquad k=3, 4, \ldots, 2N-2;$$

$$a_{2N-1}=i\beta+O\left(\omega^{1/2(N-1)}\right), \qquad \omega\to 0+;$$

$$|a_k(\omega)|\leqq M_k, \qquad 0\leqq\omega\leqq A, \qquad k=2N, 2N+1, \ldots.$$

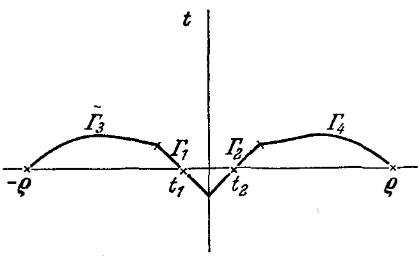

Fig. 1. This picture explains the last paragraph of p. 74.

We make a contour of 4 arcs in the disk, $|t|<\varrho$. We choose Γ_j through t_j, $j=1, 2$, along the path of steepest descent of

$$\exp\{n\,a_2(\omega)\,(t-t_j)^2\}.$$

We connect the beginning of Γ_2 to the end of Γ_1 and require that Γ_j has length less than $3|t_j|$, $j=1, 2$. Γ_4 connects the end of Γ_2 to the point ϱ in such a manner that near Γ_2 it is in the direction of the steepest descent of

$$\exp\{n\, i\, \beta\, (t-t_2)^{2N-1}\}.$$

Γ_3 connects $-\varrho$ to the beginning of Γ_1 in a similar manner.

Let

$$J_k = \int_{\Gamma_k} \exp n\left(\varphi(t) - i\,\omega\, t\right) dt, \qquad k=1, \ldots, 4. \tag{13}$$

Then by the saddle point method [3, pp. 66—69] it follows that

$$J_j = C\, n^{-\frac{1}{2}}\, \omega^{-(N-\frac{3}{2})/2(N-1)} \exp n\left(\varphi(t_j) - i\,\omega\, t_j\right) \{1 + O(n^{-\frac{1}{2}}\, \omega^{-(N-\frac{1}{2})/2(N-1)})\}, \tag{14}$$

$j=1, 2$, if $n\to\infty$ and $\omega\to 0+$ in such a manner that

$$B(n)\, n^{-(N-1)/(N-\frac{1}{2})} \leqq \omega \leqq \delta(n), \tag{15}$$

where $B(n)\to\infty$ and $\delta(n)\to 0$. Later, when we estimate $\|f^n\|$, we shall want also that

$$\begin{aligned} &\delta(n)\, n^{(N-1)/N} \to \infty, \\ &B(n)\, n^{-(N-\frac{3}{2})/4N(N-\frac{1}{2})} \to 0, \qquad n\to\infty. \end{aligned} \tag{16}$$

Using (11) we find that for some positive constants C, C_1, C_2,

$$\begin{aligned} J_1 + J_2 = {} & C\, n^{-\frac{1}{2}}\, \omega^{-(N-\frac{3}{2})/2(N-1)} \exp\{-C_1\, n\, \omega^{N/(N-1)}\} \times \\ & \times \cos\{C_2\, n\, \omega^{(N-\frac{1}{2})/(N-1)} (1+o(1))\}\{1+o(1)\}, \end{aligned} \tag{17}$$

uniformly if $n\to\infty$ and ω satisfies (15).

In any case

$$|J_j| \leqq 3|t_j| \exp\{n \operatorname{Re}\left(\varphi(t_j) - i\,\omega\, t_j\right)\}, \qquad j=1, 2,$$

and for some positive constants C, C_1,

$$|J_1 + J_2| \leqq C\, \omega^{1/2(N-1)} \exp\{-C_1\, n\, \omega^{N/(N-1)}\}, \tag{18}$$

$0 \leqq \omega \leqq A$.

To estimate $|J_3+J_4|$ we use the fact that on Γ_j, $j=3, 4$, there are positive constants C, C_1, such that

$$|\exp n\left(\varphi(t) - i\,\omega\, t\right)| \leqq C \exp\{-C_1 n\, |t-t_{j-2}|^{2N-1}\}, \tag{19}$$

$j=3, 4$, $0\leqq\omega\leqq A$. Consequently, we find that for some $a>0$

$$|J_3 + J_4| \leqq 2 \int_{a\,\omega^{1/2(N-1)}}^{\infty} C \exp\{-C_1\, n\, v^{2N-1}\}\, dv.$$

Treating the integrand as a square, and using the 2-nd mean value theorem, we find that for some other positive C, C_1,

$$|J_3 + J_4| \leqq C\, n^{-1/(2N-1)} \exp\{-C_1\, n\, \omega^{(N-\frac{1}{2})/(N-1)}\}, \tag{20}$$

$0 \leqq \omega \leqq A$.

There are some differences when $\omega < 0$. In the first place the saddle points we use are different from those in (9). We have as $\omega \to 0-$

(21) $$t_j = \lambda_j |\omega|^{1/2(N-1)} + O(|\omega|^{1/(N-1)}), \qquad j = 1, 2.$$

Here

(22) $$\lambda_j = [-\beta(2N-1)]^{-1/2(N-1)},$$

chosen such that

(23) $$\arg \lambda_1 = \pi[1 - 1/2(N-1)], \qquad \arg \lambda_2 = \pi/2(N-1).$$

This time in place of (11) we have

(24) $$\varphi(t_j) - i\omega t_j = i\lambda_j (N-1)(N-\tfrac{1}{2})^{-1} |\omega|^{(N-\frac{1}{2})/(N-1)} + O(|\omega|^{N/(N-1)}), \quad j = 1, 2,$$

$\omega \to 0-$. Note that now for some $C_1 > 0$

$$\mathrm{Re}(\varphi(t_j) - i\omega t_j) < C_1 |\omega|^{(N-\frac{1}{2})/(N-1)}, \qquad j = 1, 2,$$

$-A \leqq \omega \leqq 0$, where A is again chosen so that $|t_j| < \varrho/2$ if $|\omega| < A$.

For $N = 3, 4, \ldots$ we use a contour similar to that used when $\omega > 0$:

$$c_{kn} = \int_{-\pi}^{-\varrho} + \int_{\Gamma_3} + \int_{\Gamma_1} + \int_{\Gamma_2} + \int_{\Gamma_4} + \int_{\varrho}^{\pi} = I_1 + J_3 + J_1 + J_2 + J_4 + I_3.$$

For $N = 2$ the saddle points coincide, and $J_1 + J_2$ collapses to some J_1.

The estimate (8) still holds for $|I_1 + I_3|$. In place of (17) and (18) we use an estimate similar to (18): for some positive constants C, C_1,

(25) $$|J_1 + J_2| \leqq C |\omega|^{1/2(N-1)} \exp\{-C_1 n |\omega|^{(N-\frac{1}{2})/(N-1)}\},$$

$-A \leqq \omega \leqq 0$. Since the estimate (19) also holds on Γ_3, Γ_4 for $-A \leqq \omega \leqq 0$, we find that for some positive C, C_1,

(26) $$|J_3 + J_4| \leqq C n^{-1/(2N-1)} \exp\{-C_1 n |\omega|^{(N-\frac{1}{2})/(N-1)}\},$$

$-A \leqq \omega \leqq 0$.

We now can estimate $\|f^n\|$. We split it up

(27) $$\|f^n\| \leqq \sum_{|\omega| > A} |c_{kn}| + \sum_{|\omega| \leqq A} |I_1 + I_3| + \sum_{|\omega| \leqq A} |J_1 + J_2| + \sum_{|\omega| \leqq A} |J_3 + J_4| = \sum_{k=1}^{4} S_k.$$

The main contribution comes from S_3. We split S_3 into

$$\sum_{|\omega| \leqq A} |J_1 + J_2| = \sum_{-A \leqq \omega \leqq 0} + \sum_{0 < \omega < B(n) n^{-p}} + \sum_{B(n) n^{-p} \leqq \omega \leqq \delta(n)} + \sum_{\delta(n) < \omega \leqq A},$$

where

(28) $$p = (N-1)/(N-\tfrac{1}{2}).$$

Making use of (16), we estimate these sums with (25), (18), (17), and (18), respectively, and find that for some constant C

(29) $$S_3 \leqq C n^{1/4N}, \qquad n = 1, 2, \ldots.$$

Note that we can also get an estimate from below. In fact, it follows from (16) and (17) that for some $C'>0$

(30) $$S_3 \geqq \sum_{B(n)\,n^{-p} \leqq \omega \leqq \delta(n)} |J_1+J_2| \geqq C' n^{1/4N}, \qquad n=1,2,\ldots.$$

From (6) we see that for some constant C

(31) $$S_1 \leqq C, \qquad n=1,2,\ldots.$$

To estimate S_2 we use (8) and find that for some constant C

(32) $$S_2 \leqq 4\pi A n \vartheta^n \leqq C, \qquad n=1,2,\ldots.$$

We split S_4 into

$$S_4 = \Big(\sum_{-A\leqq\omega<0} + \sum_{0<\omega\leqq A}\Big) |J_3+J_4|,$$

and then we use (26) and (20) to obtain for some constant C

(33) $$S_4 \leqq C, \qquad n=1,2,\ldots.$$

If we add together (29), (31), (32), and (33), we find that for some C depending on N

(34) $$\|f^n\| \leqq \sum_1^4 S_k \leqq C\, n^{1/4N}, \qquad n=1,2,\ldots.$$

On the other hand we find from (30), (31), (32), and (33) that for some $C'>0$

(35) $$\|f^n\| \geqq S_3 - (S_1+S_2+S_4) \geqq C' n^{1/4N}, \qquad n=1,2,\ldots.$$

These are the required estimates.

If $\beta<0$, the analysis is similar; it essentially amounts to replacing ω with $-\omega$ in several of the estimates.

References

[1] STETTER, H. J.: Maximum bounds for the solutions of initial value problems for partial difference equations. Numerische Math. **5**, 399—424 (1963).

[2] STRANG, W. G.: Trigonometric polynomials and difference methods of maximum accuracy. J. Mat. Phys. **41**, 147—154 (1962).

[3] BRUIJN, N. G. DE: Asymptotic methods in Analysis. Amsterdam: North-Holland Publishing Co. and Groningen: P. Noordhoff Ltd. 1961.

The University of Michigan,
Dept. of Mathematics,
Ann Abor, Michigan 48104 (USA)

(Received July 6, 1964)