SIAM J. NUMER. ANAL.
Vol. 27, No. 5, pp. 1142-1159, October 1990

THE RATE OF CONVERGENCE OF SPECTRAL-VISCOSITY METHODS FOR PERIODIC SCALAR CONSERVATION LAWS*

S. SCHOCHET†

Abstract. A rate of convergence is proven for spectral-viscosity methods for periodic scalar conservation laws. This rate is obtained by showing the discretization error to be small enough that the difference between the solutions of the spectral-viscosity method and the ordinary viscosity method tends to zero in L^1 as the number of discrete modes tends to infinity and the viscosity simultaneously tends to zero at the appropriate rate. The method is also used to obtain an L^∞ bound for the spectral-viscosity approximations to the elasticity equations; convergence, although without a rate, then follows from compensated compactness theory.

Key words. conservation law, spectral method, artificial viscosity, rate of convergence

AMS(MOS) subject classification. 65M60

1. Introduction. The solution of the scalar conservation law

$$\partial_t u + \partial_x f(u) = 0, \tag{1.1}$$

$$u(0, x) = u_0(x) \tag{1.2}$$

can develop discontinuities even if the initial data u_0 and the flux f are smooth functions [9]. A solution exists in the weak sense even after the time when discontinuities form, although one of various equivalent entropy conditions must be added to (1.1)-(1.2) in order for the weak solution to be unique [9]. This weak solution has the property that its spatial L^2 norm decays in time after the first shock forms, even though (1.1) preserves the L^2 norm of smooth solutions. The L^2 norm is also preserved in time by the classical spectral (Fourier) approximation method for (1.1)-(1.2) with periodic u_0,

$$\partial_t u^N + \partial_x S_N f(u^N) = 0, \tag{1.3}$$

$$u^N(0, x) = S_N u_0, \tag{1.4}$$

where u^N is a trigonometric polynomial of degree N and S_N is the projection into the space of such polynomials. Hence, even though (1.3), (1.4) formally approximates (1.1), (1.2) to arbitrary order, u^N cannot converge strongly in L^2 to u when the latter contains shocks; in fact, if f is genuinely nonlinear, then u^N cannot even converge weakly to u [12], [13]. In order to overcome this difficulty while retaining formal infinite-order accuracy, Tadmor [12] added artificial viscosity to the high Fourier modes of (1.3), and proved [12], [10], [13] that such spectral-viscosity (and pseudospectral-viscosity) methods for periodic scalar conservation laws converge to the correct solution of (1.1), (1.2). (In [13] this result is extended to a wide variety of other semidiscrete-viscosity methods and to certain 2×2 systems.) Because the convergence proof appeals to the theory of compensated compactness [14], no rate of convergence is obtained.

A rate of convergence has been demonstrated for monotone approximations of (1.1), (1.2), however (see [11] and its references). Although the spectral viscosity method is not monotone, the results of [12] suggest that, for the special case when f is quadratic, the spectral-viscosity method is close to the pure (nondiscretized) viscosity

* Received by the editors October 6, 1988; accepted for publication (in revised form) October 29, 1989. This research was partially supported by the Center for Absorption in Science of the Department of Immigrant Absorption of the State of Israel.

† School of Mathematical Sciences, Beverly and Raymond Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv 69978, Israel.

method, which is monotone. It will be shown here that these methods are in fact close for general fluxes f, and that this fact can be used to obtain a rate of convergence for the spectral- (and pseudospectral-) viscosity method for (1.1), (1.2).

For sufficiently smooth u_0 and f and certain choices of modes to be damped by viscosity, the difference between the spectral-viscosity and pure viscosity approximations is smaller than the error of the latter, so that the rate of convergence of the spectral-viscosity approximation u^N is determined solely by the rate at which the viscosity parameter ε_N can be allowed to vanish. The requirement

$$\varepsilon_N \geqq O\left(\frac{\log N}{\sqrt{N}}\right) \tag{1.5}$$

of [12], [10], [13] will be improved here for smooth f to

$$\varepsilon_N \geqq O\left(\frac{1}{N^{1-\delta}}\right) \quad \text{for arbitrary } \delta > 0, \tag{1.6}$$

which yields the convergence rate [11]

$$\sup_{0 \leqq t \leqq T} \|u - u^N\|_{L^1} \leqq C N^{-(1-\delta)/2}. \tag{1.7}$$

The rate (1.6) is almost optimal in the following sense. If u_0 is the space BV of functions of bounded variation, then the BV norm of u is bounded for all time by that of u_0, and integration by parts then shows that $|\hat{u}(t, k)| \leqq c/(1+|k|)$ since the derivative of a BV function is a bounded measure. It will be shown here that if u_0 and f are smooth enough, then $|k|\hat{u}^N(k)$ is likewise bounded on fixed time intervals $[0, T]$. If the method with

$$\varepsilon_N = o\left(\frac{1}{N}\right) \tag{1.8}$$

had this same property, then the energy equation

$$\frac{1}{2}\|u^N(t)\|_{L^2}^2 + \int_0^t \varepsilon_N \|\nabla u^N\|_{L^2}^2 = \frac{1}{2}\|u_0^N\|_{L^2}^2$$

and the estimate

$$\varepsilon_N \|\nabla u^N\|_{L^2}^2 \leqq c\varepsilon_N \sum_{|k| \leqq N} \left[\frac{k}{1+|k|}\right]^2 \leqq cN\varepsilon_N \to 0$$

would imply that the L^2 norm was asymptotically preserved, which as noted above precludes the convergence of u^N to u. Hence ε_N is restricted to be $O(1/N)$, which is approached arbitrarily closely by (1.6).

After reviewing the definition of the spectral-viscosity and pseudospectral-viscosity approximations in § 2, we will estimate the decay rate of the Fourier coefficients of u^N in § 3, obtain an L^∞ bound in § 4, and prove the rate of convergence in § 5. The decay rate is also valid for systems provided an a priori L^∞ bound is assumed, but the proofs of the L^∞ bound and the rate of convergence use the maximum principle and the L^1 contraction property, respectively, and so are valid only in the scalar case.

There are some special systems of conservation laws, such as the elasticity equations, for which L^∞ bounds hold for the exact system and its pure viscosity-method approximation because of the existence of compact invariant regions [2]. It will be pointed out in § 6 that the method used here easily adapts to show L^∞ bounds for the spectral-viscosity approximation of the elasticity equations. Now it was shown in [4] that L^∞-bounded approximate solutions of this system converge to an entropy solution, provided they satisfy certain other conditions that, as noted in [13], are readily verifiable for the spectral-viscosity method. Hence it follows that the spectral-viscosity method for the elasticity equations converges to an admissible solution. The method of [13] could also be used to obtain the L^∞ bound for this case if the stronger condition (1.5) is imposed instead of (1.6).

2. The approximating equations. The spectral projection operator s_N is defined by

$$[S_N u](t,x)=\sum_{|k|\leqq N}\hat{u}(t,k)\,e^{ikx}, \tag{2.1}$$

where

$$\hat{u}(t,k)=\frac{1}{2\pi}\int_0^{2\pi}u(t,x)\,e^{-ikx}. \tag{2.2}$$

In order to take into account the error produced in approximating the integral in (2.2), we often use instead the pseudospectral projection operator $(PS)_N$ based on interpolation at the $2N+1$ points $x_j=2\pi(j/(2N+1))$, $j=0,1,2,\cdots,2N$:

$$[(PS)_N u](t,x)=\sum_{|k|\leqq N}\tilde{u}(t,k)\,e^{ikx}, \tag{2.3}$$

where

$$\tilde{u}(t,k)=\frac{1}{2N+1}\sum_{j=0}^{2N}u(t,x_j)\,e^{-ikx_j}. \tag{2.4}$$

Because the pseudospectral projection differs from the spectral one only by the "aliasing" error [10],

$$A_N u=\sum_{|k|\leqq N}\left[\sum_{j\neq 0}\hat{u}(t,k+j[2N+1])\,e^{ikx}\right], \tag{2.5}$$

we can conveniently refer to both projections by using the notation

$$P_N=S_N+aA_N, \tag{2.6}$$

with $a=0$ or 1 in the spectral or pseudospectral cases, respectively.

We can now define a (pseudo)spectral-viscosity approximation to the periodic conservation law (1.1) by

$$\partial_t u^N+\partial_x P_N f(u^N)=\varepsilon_N\partial_x^2 u^N. \tag{2.7}$$

We will also consider the case when viscosity is applied to high Fourier modes only, as for example in the equation [12]

$$\partial_t u^N+\partial_x P_N f(u^N)=\varepsilon_N\partial_x^2[u^N-R_{m(N)}*u^N], \tag{2.8}$$

where $*$ denotes convolution, $m(N)\leqq N^\beta$ for some $\beta<1$ to be chosen later, and R_m satisfies

$$\begin{aligned}\hat{R}_m(k)&=1, && |k|\leqq m,\\ 0\leqq\hat{R}_m(k)&\leqq 1, && m<|k|\leqq 2m,\\ \hat{R}_m(k)&=0, && |k|>2m.\end{aligned} \tag{2.9}$$

These two possibilities can be treated together using (2.8) with either (2.9) or else

$$R_m\equiv 0. \tag{2.10}$$

Besides our approximation (2.8) to the differential equation (1.1), we also need an approximation to the initial condition (1.2). The natural approximation $u^N(0, x) = S_N u_0 = D_N(x) * u_0$ need not be bounded in L^∞ uniformly in N, so it is convenient to replace the Dirichlet kernel $D_N(x) = \sum_{|k| \leqq N} e^{ikx}$ by a polynomial summability kernel [7], i.e., we set

$$u^N(0, x) = K_N u_0 \equiv K_N(x) * u_0, \tag{2.11}$$

where $K_N(x)$ satisfies

$$\begin{aligned} &\frac{1}{2\pi}\int_0^{2\pi} K_N(x) = 1, \qquad \frac{1}{2\pi}\int_0^{2\pi} |K_N(x)| \leqq c_1 \quad \text{independently of } N, \\ &\lim_{N\to\infty}\int_\delta^{2\pi-\delta} K_N(x) = 0, \qquad K_N(x) = \sum_{|k|\leqq N} \hat{K}(k)\, e^{ikx}. \end{aligned} \tag{2.12}$$

Possible choices for K_N include Fejér's kernel $F_N(x) = \sum_{|k| \leqq N} (1 - k/N)\, e^{ikx}$ and de la Vallée Poussin's kernel

$$V_N(x) = 2F_{2[N/2]}(x) - F_{[N/2]}(x) = \sum_{|k| \leqq [N/2]} e^{ikx} + \sum_{[N/2] < k \leqq 2[N/2]} \left(2 - \frac{k}{[N/2]}\right) e^{ikx}.$$

The former has the advantage of neither increasing the maximum nor decreasing the minimum of a function, which could be important if the flux f is defined only in a small neighborhood of $[\min u_0, \max u_0]$, while the latter has the advantage of being an approximation of arbitrary high order of accuracy when u_0 is smooth, and increases the L^∞ norm by no more than a factor of $\| V_N(x) \|_{L^1([0,2\pi); dx/2\pi)}$, which equals approximately 1.436. For typical BV initial data with discontinuities, so that

$$|\hat{u}_0(k)| \leqq \frac{c_0}{1+|k|} \quad \text{but } |\hat{u}_0(k_j)| \geqq \frac{\tilde{c}_0}{1+|k_j|} \quad \text{for some sequence } k_j \to \infty,$$

the kernels F_N, V_N, and D_N all yield approximations to u_0 with L^2-error of order $N^{-1/2}$, which is less than the dynamic error (1.7), so that the choice of f_N versus V_N does not affect the proven convergence rate of the method. In a practical computation with N finite, however, we are apt to obtain better results by taking a closer approximation of the initial data, i.e., by using V_N or even D_N. Also, if the initial data u_0 is sufficiently smooth (i.e., in H^s with $s > \frac{1}{2}$), then the theorems to be proven will apply to both $u^N(0, x) = S_N u^0$ and $u^N(0, x) = (PS)_N u_0$ since $u^N(0, x)$ is then uniformly in L^∞.

Approximation (2.8), (2.11) can be defined for systems by allowing P_N and R_M to act on each component.

3. Decay of the high Fourier modes. Everything in this section is valid for systems except where explicitly noted to the contrary. In order to determine the behaviour of the high Fourier modes of u_N, we follow the method of [12], [10], and [8] and apply the operator $I - S_k$ to (2.8), with $k > 2m(N)$, to obtain

$$\partial_t (I - S_k) u^N = \varepsilon_N \partial_x (I - S_k) u_x^N + (I - S_k) \partial_x P_N f. \tag{3.1}$$

Multiplying (3.1) by $(I - S_k) u^N$ yields

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt} \|(I - S_k) u^N\|^2 + \varepsilon_N \|(I - S_k) u_x^N\|_{L^2}^2 \\ &\quad \leqq \|(I - S_k) u^N\| \, \|S_N (I - S_k) f_x\| + a \|(I - S_k) u_x^N\| \, \|A_N f\|, \end{aligned} \tag{3.2}$$

where $\| \ \|$ denotes the L^2 norm. To proceed further we need an estimate for the operator A_N.

LEMMA 3.1. *For any $p > \frac{1}{2}$ and any w in H^p, $\|A_N w\| \leqq (c(p)/N^p)\|(I-S_N)\partial_x^p w\|$.*

Proof.

$$\|A_N w\|^2 = \sum_{|k| \leqq N} \left| \sum_{j \neq 0} \hat{w}(k+j[2N+1]) \right|^2$$

$$\leqq \sum_{|k| \leqq N} \left[\sum_{j \neq 0} \left(\frac{1}{|j|N} \right)^p |(k+j[2N+1])^p \hat{w}(k+j[2N+1])| \right]^2$$

$$\leqq \sum_{|k| \leqq N} \left(\sum_{j \neq 0} \frac{1}{j^{2p} N^{2p}} \right) \left(\sum_{j \neq 0} |(k+j[2N+1])^{2p} \hat{w}(k+j[2N+1])|^2 \right)$$

$$= \frac{c(p)}{N^{2p}} \sum_{|L| > N} L^{2p} |\hat{w}(L)|^2 = \frac{c(p)}{N^{2p}} \|(I-S_N)\partial_x^p w\|^2. \qquad \square$$

Now $\|(I-S_k)u_x^N\| \geqq k\|(I-S_k)u^N\|$ since $I-S_k$ kills off the modes with |wavenumber| less than k, and $\|u_x^N\| \leqq N\|u^N\|$, so by using Lemma 3.1 with $p=1$ in (3.2) we obtain

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|(I-S_k)u^N\|^2 + \varepsilon_N k^2 \|(I-S_k)u^N\|^2 \\ &\qquad \leqq c\|(I-S_k)u^N\|[\|S_N(I-S_k)f_x\| + \|(I-S_N)f_x\|] \\ &\qquad \leqq c\|(I-S_k)u^N\| \ \|f(u^N)_x\| \\ &\qquad \leqq c\|(I-S_k)u^N\|[\sup_{|u| \leqq c_2} |f'(u)|]\|u_x^N\| \\ &\qquad \leqq c[\sup_{|u| \leqq c_2} |f'(u)|]\|(I-S_k)u^N\|\{k\|S_k u^N\| + N\|(I-S_k)u^N\|\}, \end{aligned} \tag{3.3}$$

where c_2 is an assumed bound on $\|u^N\|_{L^\infty}$. Removing the common factor of $\|(I-S_k)u^N\|$ from (3.3) we get

$$\frac{d}{dt}\|(I-S_k)u^N\| + \{\varepsilon_N k^2 - cN[\sup_{|u| \leqq c_2} f'(u)]\}\|(I-S_k)u^N\| \leqq ck[\sup_{|u| \leqq c_2} |f'(u)|]\|u^N\|, \tag{3.4}$$

so that for

$$k > \tilde{c}\sqrt{N/\varepsilon_N} \tag{3.5}$$

Gronwall's lemma yields

$$\|(I-S_k)u_N\| \leqq \frac{c}{\varepsilon_N k} + e^{-c\varepsilon_N k^2 t}\|(I-S_k)u^N(0,x)\|. \tag{3.6}$$

Since $\|(I-S_k)u^N\|$ measures the energy in the Fourier modes higher than k, (3.6) is an estimate of the decay of the Fourier modes for high k. Now from (3.3) we see that

$$\frac{d}{dt}\|(I-S_k)u^N\| + \varepsilon_N k^2 \|(I-S_k)u^N\| \leqq c[\|S_N(I-S_k)f_x\| + \|(I-S_N)f_x\|] \tag{3.7}$$

so that if we can estimate the decay of $\hat{f}(k)$ from that of $\hat{u}(k)$ we can use (3.6) in (3.7) to improve the decay rate of $\hat{u}(k)$, and repeat this process to obtain ever higher rates of decay. If u^N obeyed an estimate of the form

$$\|(I-S_k)u^N\| \leqq ck^{-q}, \qquad k \geqq k_0 \tag{3.8}$$

with c and k_0 independent of N, then we could use the facts that (i) (3.8) implies that u is in H^r for $r<q$; (ii) H^r is an algebra for $r>\frac{1}{2}$ so that $f(u)$ is in H^r provided $f\in C^r$; and (iii) $f\in H^r$ implies $\|(I-S_k)f\|\leqq ck^{-r}$ to obtain such estimates for $f(u^N)$. However, even in the case when u^N is smooth, estimate (3.6) holds only for $k\geqq k_0(N)$ given by (3.5), so that estimates obtained by the method just mentioned have an unfavorable dependence on N. Instead we will seek a direct link between $\|(I-S_k)f(u)\|$ and $\|(I-S_{\tilde{k}})u\|$ so as to take into account restriction (3.5) on estimate (3.6). Such estimates have been obtained straightforwardly in [12] and [10] for quadratic f; the following lemmas provide the needed estimates for general (sufficiently smooth) f. The first of these, a variant of a lemma in [12], deals with the quadratic case.

LEMMA 3.2. *For all u, w in* L^∞,

$$\|(I-V_k)uw\|\leqq w\|_{L^\infty}\|(I-V_{[k/4]})u\|+3\|u\|_{L^\infty}\|(I-V_{[k/4]})w\|.$$

Proof.

$$\|(I-V_k)uw\|=\|(I-V_k)\{w(I-V_{[k/4]})u+(V_{[k/4]}u)(I-V_{[k/4]})w+(V_{[k/4]}u)(V_{[k/4]}w)\}\|.$$

Since $(I-V_k)$ annihilates the polynomials of order less than or equal to $[k/2]$, the last term inside the norm vanishes. Because $I-V_k$ is bounded by 1 on L^2, the terms that remain are dominated by $\|w\|_{L^\infty}\|(I-V_{[k/4]})u\|+\|V_{[k/4]}u\|_{L^\infty}\|(I-V_{[k/4]})w\|$. But Young's inequality says that $V_{[k/4]}$ is bounded on L^∞ by $c_2=\|V_{[k/4]}(x)\|_{L^1}$, which is less than or equal to 3 since $\|F_k\|_{L^1}=1$. □

Repeated use of Lemma 3.2 allows us to estimat($\|(I-V_k)P(u)\|$ in terms of $\|(I-V_{\tilde{k}})u\|$ when $P(u)$ is any positive integer power of u, and hence also, by addition, when P is any polynomial. By approximating any sufficiently smooth f in some systematic way by polynomials, we want to obtain a similar estimate for $\|(I-V_k)f(u)\|$ in terms of $\|(I-V_{\tilde{k}})u\|$ and an arbitrarily small error term. Because we do not want to assume that f is analytic, we will approximate f by a series of orthogonal polynomials rather than by a power series. However, if we had to break down each monomial u^m into an orthogonal polynomial separately using Lemma 3.2 and take the sum over all such monomials, then most of the advantages of orthogonality would be lost. We would therefore like our orthogonal polynomials to be such that any polynomial in the series is equal to the product of two polynomials of lower order from the same series, so that we could use Lemma 3.2 repeatedly in an efficient manner. The next lemma is based on the fact that the Chebyshev polynomials almost satisfy this condition.

LEMMA 3.3. *Let* T_j *be the jth Chebyshev polynomial. Then for any u with* $\|u\|_{L^\infty}\leqq 1$,

$$\begin{aligned}&\|(I-V_k)T_j(u)\|\\ &\leqq 2\|(I-V_{[k/4]})T_{[j/2]}(u)\|+6\|(I-V_{[k/4]})T_{j-[j/2]})(u)\|+\|(I-V_k)T_{j-2[j/2]}(u)\|.\end{aligned}\tag{3.9}$$

Proof. From the definition [1] $T_j(u)=\cos(j[\cos^{-1}(u)])$ and the trigonometric identity $\cos(\alpha+\beta)=2\cos\alpha\cos\beta-\cos(\alpha-\beta)$ we have, on setting $\alpha=[j/2]$ and $\beta=j-[j/2]$,

$$T_j(u)=2T_{[j/2]}(u)T_{j-[j/2]}(u)-T_{j-2[j/2]}(u).$$

Since $\|T_j(u)\|_{L^\infty}\leqq 1$, Lemma 3.2 implies that (3.9) holds. For later use, note that for $k\geqq 2$ the term $(I-V_k)T_{j-2[j/2]}(u)=0$ unless j is odd, in which case it equals $(I-V_k)T_1(u)=(I-V_k)u$. □

Using Lemma 3.3, we can express the energy in the high modes of $f(u)\hat{}$ in terms of the energy in the high modes of $\hat{u}$.

LEMMA 3.4. *If $\tilde{f}$ is in C^s with $s \geqq 2$ and $\tilde{u} \in L^\infty$ is a scalar, then for any positive integer p*

$$\|(I-S_k)\tilde{f}(\tilde{u})\| \leqq c(s, \|\tilde{f}\|_{C^s}, \|\tilde{u}\|_{L^\infty}) \left[\sum_{j \leqq p} (j^{3-s} \|(I - S_{[k/32j^2]})\tilde{u}\|) + p^{1-s} \right]. \tag{3.10}$$

When $\tilde{u}$ is an n-vector then an analogous estimate holds, although the minimum smoothness required and the various constants appearing in (3.10) *must be changed to values that depend on n.*

Proof. First assume that u is a scalar. Let $u = \tilde{u}/\|\tilde{u}\|_{L^\infty}$ and $f(\cdot) = \tilde{f}(\|\tilde{u}\|_{L}{}^{\infty} \times \cdot)$ so that $\tilde{f}(\tilde{u}) = f(u)$. Since f is in C^2, its Chebyshev expansion converges to itself; i.e., $f(u) = \frac{1}{2} f_0 + \sum_j f_j T_j(u)$, where $f_j = \int_0^{2\pi} f(\cos\theta) \cos j\theta \, d\theta$ [1]. Furthermore, integration by parts shows that $|f_j| \leqq c_s j^{-s}$ because f is in C^s. Now $\|T_j\| \leqq c$ since $\|T_j\|_{L^\infty} \leqq 1$, so for any positive integer p

$$\begin{aligned} \|(I-V_k)f(u)\| &\leqq \sum_{j \leqq p} |f_j| \, \|(I-V_k)T_j(u)\| + c \sum_{j > p} |f_j| \\ &\leqq c_s \sum_{j \leqq p} j^{-s} \|(I-V_k)T_j(u)\| + \tilde{c}_s p^{1-s}. \end{aligned} \tag{3.11}$$

Next, if $4^{k_0} \leqq k < 4^{k_0+1}$ and $2^{j_0} < j \leqq 2^{j_0+1}$, then Lemma 3.3 says that $\|(I-V_k)T_j(u)\|$ is less than or equal to the sum of eight terms of the form $\|(I-V_{\tilde{k}})T_{\tilde{j}}(u)\|$ with $\tilde{k} \geqq 4^{k_0-1}$ and $\tilde{j} \leqq 2^{j_0}$, plus at most one term with $\tilde{k} = k$ and $\tilde{j} = 1$. By iterating we can reduce all the $\tilde{j}$'s to 1 in at most j_0+1 steps, and hence all the final $\tilde{k}$'s are at least $4^{k_0-(j_0+1)}$. Thus, since the number of terms that reduce directly to $\tilde{j} = 1$ at each stage is not greater than the number of terms present at the previous stage,

$$\|(I-V_k)T_j(u)\| \leqq 2(8^{j_0+1}) \|(I - V_{4^{k_0-(j_0+1)}})u\| \tag{3.12}$$

because $\|(I-V_{k_1})u\| \leqq \|(I-V_{k_2})u\|$ if $k_2 < k_1$.

Finally, $j_0 + 1 \leqq (\log_2 j) + 1$, so $8^{j_0+1} \leqq 8 j^{\log_2 8} = 8j^3$. Similarly, $4^{k_0-(j_0+1)} \geqq 4^{k_0-1} j^{-2} \geqq k/16j^2$, and, as noted above, we can freely replace $I - V_{k_1}$ by $I - V_{k_2}$ with $k_2 \leqq k_1$, so plugging our modified (3.12) into (3.11) shows that (3.10) holds if we replace S by V on both sides and k by $2k$ on the right. Using $\|(I-S_{2k})u\| \leqq \|(I-V_{2k})u\| \leqq \|(I-S_k)u\|$ then yields (3.10) as written.

When $\tilde{u}$ is a vector then f can be expanded in a series whose terms are products of Chebyshev polynomials of the components of $\tilde{u}$, and the analysis proceeds similarly. □

Remark. Lemma 3.4 could be used to show that if $u \in L^\infty$ is a scalar and $\|(I - S_k)u\| \leqq ck^{-q}$ then $\|(I-S_k)f(u)\| \leqq \tilde{c}k^{-q}$ provided $f \in C^r$ with $r > 2q+4$. Just use (3.10) with $p = \sqrt{k}/64$.

We are now ready to use (3.7) and (3.10) to improve the decay rate (3.6).

THEOREM 1. *Let u^N be the solution of the approximation* (2.8), (2.11) *with*

$$\varepsilon_N \geqq c_3 N^{\alpha-1}, \qquad \alpha > 0 \tag{3.13}$$

$$m(N) \leqq cN^\beta, \qquad \beta \leqq 1 - \alpha. \tag{3.14}$$

Suppose that

$$\|(I-S_k)u^N(0,x)\| \leqq c_4 k^{-r}, \qquad r \geqq 0, \tag{3.15}$$

$$\|u^N(t,\cdot)\|_{L^\infty} \leqq c_2 \quad \text{on } [0, T] \tag{3.16}$$

for some $T > 0$.

Assume finally that f is in C^s with

$$s \geqq s_0 \equiv \max(2r+6, 16L(L+1)+2), \tag{3.17}$$

where L is a positive integer (except that s_0 must be replaced by $s_0(n)$ if u is an n-vector). Then for $0 \leqq i \leqq L$ and $t \in [0, T]$,

$$\|(I-S_k)u^N\| \leqq c\left[\frac{1}{\varepsilon_N k}\left(\frac{N}{\varepsilon_N k^2}\right)^i + N^{-(L+1)\alpha} + k^{-r}\exp\left(-\varepsilon_N k^2 t/cN^{(i\alpha/4L)}\right)\right] \tag{3.18}$$

for $k \geqq \tilde{c}(L)\sqrt{N/\varepsilon_N}\, N^{i\alpha/4L}$;

$$\begin{aligned}&\|(I-S_k)f(u^N)\|\\ &\quad\leqq c\left[\frac{1}{\varepsilon_N k}\left(\frac{N}{\varepsilon_N k^2}\right)^i + N^{-(L+1)\alpha} + k^{-r}\exp\left(-\varepsilon_N k^2 t/cN^{((i+1)\alpha/4L)}\right)\right]\end{aligned} \tag{3.19}$$

for $k \geqq \hat{c}(L)\sqrt{N/\varepsilon N}\, N^{(i+1)\alpha/4L}$;

$$\begin{aligned}\|(I-S_k)f(u^N)\| \leqq c\Bigg[& N^{-(L+1)\alpha}\left(\frac{k}{N}\right)^{-8L^2}\\ &+\left(\frac{k}{N}\right)^{-(s-2r-5)/2} k^{-r}\exp\left(-ctN^{(1+\alpha/4)}\right)\Bigg]\end{aligned} \tag{3.20}$$

for $k \geqq N$. Furthermore, (3.19) *and* (3.20) *also hold with f replaced by f'.*

Proof. First, assume that u is a scalar. For $i=0$, (3.18) follows from (3.6). To obtain (3.19) for $i=0$, plug (3.6) into (3.10) and set p there to $[cN^{\alpha/16L}]$. There results

$$\begin{aligned}\|(I-S_k)f(u^N)\| \leqq c\Bigg\{&\sum_{j\leqq[cN^{\alpha/16L}]} j^{3-s}\left(\frac{32j^2}{\varepsilon_N k}\right)\\ &+\sum_{j\leqq[cN^{\alpha/16L}]} j^{3-s}\exp\left(-\varepsilon_N k^2 t/j^4\right)\left[\frac{k}{32j^2}\right]^{-r} + cN^{(\alpha/16L)(1-s)}\Bigg\}.\end{aligned} \tag{3.21}$$

Note that for $k \geqq \tilde{c}(L)\sqrt{N/\varepsilon_N}\, N^{\alpha/4L}$, $[k/32j^2] \geqq \tilde{c}\sqrt{N/\varepsilon_N}$ for $j \leqq [cN^{\alpha/16L}]$ so that the use of (3.6) in (3.10) is justified. Next, since $s>7$, the first sum on the right side of (3.21) is

$$\leqq \frac{c}{\varepsilon_N k}\sum_j j^{5-s} \leqq \frac{\tilde{c}}{\varepsilon_N k};$$

similarly, since $s > 16L(L+1)+1$,

$$N^{(\alpha/16L)(1-s)} \leqq cN^{-(L+1)\alpha}.$$

Finally, the second term on the right side of (3.21) is

$$\begin{aligned}&\leqq ck^{-r}\exp\left(-\varepsilon_N k^2 t/N^{\alpha/4L}\right)\sum_j j^{2r+3-s}\\ &\leqq \tilde{c}k^{-r}\exp\left(-\varepsilon_N k^2 t/N^{\alpha/4L}\right)\end{aligned}$$

since $s \geqq 2r+5$, so that (3.19) holds for $i=0$. Because the right side of (3.17) is at least one greater than was needed, (3.19) also holds for $i=0$ with f replaced by f'.

For $i>0$ we will proceed by induction, so assume that (3.18), (3.19) hold for $i \leqq i_0$. Recall (3.7):

$$\frac{d}{dt}\|(I-S_k)u^N\| + \varepsilon_N k^2\|(I-S_k)u^N\| \leqq c[\|S_N(I-S_k)f_x\| + \|(I-S_N)f_x\|]. \tag{3.22}$$

Now

$$\|S_N(I-S_k)f_x\| \leqq N\|(I-S_k)f\|, \tag{3.23}$$

while

$$\begin{aligned}\|(I-S_N)f_x\| &= \|(I-S_N)f'(u^N)u_x^N\| \leqq \|(I-V_N)f'(u^N)u_x^N\|\\ &\leqq \|f'(u^N)\|_{L_\infty}\|(I-V_{[N/4]})u_x^N\|\\ &\quad+3\|u_x^N\|_{L^\infty}\|(I-V_{[N/4]})f'(u^N)\|\end{aligned} \tag{3.24}$$

by Lemma 3.2. Although the bound $\|u_x^N\|_{L^2} \leqq N\|u^N\|_{L^2}$ follows immediately from $\|u^N\|_{L^2} = \|\hat{u}^N\|_{L^2}$, it is a little less obvious that with an extra constant factor thrown in the same bound holds in any L^p space.

LEMMA 3.5. *If u^N is an N-trigonometric polynomial, then $\|u_x^N\|_{L^p} \leqq cN\|u^N\|_{L^p}$ with c independent of u^N, N, and p.*

Proof. Since $S_N u^N = u^N$, $V_{2N} u^N = u^N$. Hence

$$\begin{aligned}\|\partial_x u^N\|_{L^p} &= \|\partial_x(V_{2N} * u^N)\|_{L^p} = \|(\partial_x V_{2N}) * u^N\|_{L_p}\\ &\leqq \|\partial_x V_{2N}\|_{L^1}\|u^N\|_{L^p} = \|\partial_x(2F_{2N} - F_N)\|_{L^1}\|u^N\|_{L^p},\end{aligned}$$

which implies that it suffices to show that

$$\|\partial_x F_N\|_{L^1} \leqq cN. \tag{3.25}$$

But the footnote to Chap. XI, Equations 1.9–1.10 of [15] says that $|\partial_x F_N(x)| \leqq cN^2/(1+N|x|)^2$, for $|x| \leqq \pi$, and integrating this from $-\pi$ to π yields (3.25). □

Using Lemma 3.5 and our assumed L^∞ bound in (3.24), we get

$$\begin{aligned}\|(I-S_N)f_x\| &\leqq cN[\|(I-V_{[N/4]})u^N\| + \|(I-V_{[N/4]})f'(u^N)\|]\\ &\leqq cN[\|(I-S_{[N/8]})u^N\| + \|(I-S_{[N/8]})f'(u^N)\|],\end{aligned} \tag{3.26}$$

and inserting (3.23), (3.26) into (3.22) yields

$$\frac{d}{dt}\|(I-S_k)u^N\| + \varepsilon_N k^2\|(I-S_k)u^N\| \leqq cN \sum_{g=I,f,f'} \|(I-S_k)g(u^N)\|. \tag{3.27}$$

Note that we are free to replace $I - S_{[N/8]}$ in (3.26) by $(I - S_k)$ in (3.27); the resulting restriction $k \leqq [N/8]$ can be avoided by noting that (3.18) holds trivially for $k \geqq N$ and that if it holds for $k \leqq [N/8]$ it also holds, with slightly larger values of c, for $k \leqq N$.

By plugging (3.18), (3.19), and (3.20) with f replaced by f', all for the case $i = i_0$, into (3.27) and applying Gronwall's lemma, we obtain (3.18) for $i = i_0 + 1$ provided that we note $N/\varepsilon_N k^2 \leqq 1$ for the values of k under consideration and that, for the case $i_0 = 0$, this inequality implies $Nt \exp(-\varepsilon_N k^2 t) \leqq c \exp(-\varepsilon_N k^2 t/N^{\alpha/4L})$. Estimate (3.19) for $i = i_0 + 1$ then follows from (3.18) and (3.10) as in the case $i = 0$.

Finally, since $\|(I - S_{\tilde{k}})u^N\| = 0$ when $\tilde{k} \geqq N$, for $k \geqq N$ we obtain from (3.10) that for any positive integer p.

$$\|(I-S_k)f(u^N)\| \leqq \left[\sum_{\sqrt{k/32(N+1)}} \leqq j \leqq p j^{3-s}\|(I - S_{[k/32j^2]})u^N\| + p^{1-s}\right]. \tag{3.28}$$

In order to use estimate (3.18) with $i = L$ in (3.28), p must satisfy

$$\frac{k}{32p^2} \geqq \tilde{c}(L)\sqrt{N/\varepsilon_N}\, N^{\alpha/4}. \tag{3.29}$$

By our assumption on ε_N, the right side of (3.29) is $\leqq cN^{1-\alpha/4}$, so if we pick

$$p = c\sqrt{k/N^{1-\alpha/4}} \tag{3.30}$$

with appropriate constant c, then (3.29) will be satisfied. Inserting (3.30) and (3.18) with $i = L$ into (3.28), we obtain

$$\begin{aligned}&\|(I-S_k)f(u^N)\|\\ &\leqq c\left[\sum_{\tilde{c}\sqrt{k/N} \leqq j \leqq \hat{c}\sqrt{k/N}\, N^{\alpha/8}} j^{3-s}\left((32j^2)^{2L+1}\left(\frac{N}{k}\right)^L \frac{1}{(\varepsilon_N k)^{L+1}} + \frac{1}{N^{\alpha(L+1)}}\right.\right.\\ &\qquad\left. + \left(\frac{k}{32j^2}\right)^{-r} \exp(-\varepsilon_N k^2 t/\tilde{c}j^4 N^{\alpha/4})\right)\\ &\qquad\left. + c\left(\frac{k}{N}\right)^{(1-s)/2} N^{(\alpha/8)(1-s)}\right].\end{aligned} \tag{3.31}$$

Since $k \geqq N$, $N\varepsilon_N \geqq cN^{\alpha}$, and $j \leqq c\sqrt{k/N}\,N^{\alpha/8}$, the right side of (3.31) is less than or equal to

$$(3.32)\qquad c \sum_{\tilde{c}\sqrt{k/N} \leqq j \leqq \hat{c}\sqrt{k/N}\,N^{\alpha/8}} \Big(j^{4L+5-s} N^{-\alpha(L+1)} + j^{2r+3-s} k^{-r} \cdot \exp\left(-\varepsilon_N k^2 t / c(k/N)^2 N^{\alpha/2+\alpha/4}\right)\Big) + c\left(\frac{k}{N}\right)^{(1-s)/2} N^{(\alpha/8)(1-s)}.$$

Using the facts that $j \geqq \tilde{c}\sqrt{k/N}$, s satisfies (3.17), and ε_N obeys (3.13), we find that (3.32) is less than or equal to

$$(3.33)\qquad c\left[\left(\frac{k}{N}\right)^{-8L^2} N^{-\alpha(L+1)} + \left(\frac{k}{N}\right)^{-(s-2r-5)/2} k^{-r} \exp\left(-ctN^{(1+\alpha/4)}\right)\right].$$

In fact, in obtaining estimate (3.33) for $\|(I-S_k)f(u^N)\|$ we have only used the fact that $s \geqq s_0 - 1$, where s_0 is defined in (3.17), and so estimate (3.20) that we have deduced is also valid for f'.

If u is an n-vector, then the proof is the same except that the modified version of (3.10) valid for vectors must be used. □

Remark. The smoothness requirement (3.17) is given only for the sake of concreteness and is by no means optimal. Improved estimates of the amount of smoothness required can be obtained by using a kernel more complicated than V_N, a more precise estimate of the kernel's L^1 norm, or more delicate estimates in the proof of the theorem.

4. The L^∞ bound. Although Theorem 1 assumes the existence of an L^∞ bound, in the scalar case that theorem can in fact be used to justify such a bound a posteriori.

THEOREM 2. *Assume that* (1.1) *is a scalar equation, and that the hypotheses of Theorem* 1 *hold, except possibly the L^∞ bound* (3.16). *Assume further that $m(N)$ satisfies the stricter bound*

$$(4.1)\qquad m^2\|R_m(x)\|_{L^1} \leqq \frac{c}{\varepsilon_N N^{\beta_1}} \quad \textit{with } \beta_1 \geqq 0$$

(*where R_m is defined in* (2.9) *or* (2.10)). *If either*

(i) $r \geqq \frac{1}{2}$ *in* (3.15), $\alpha(L+1) > \frac{3}{2}$, *and* $s \geqq s_0 + 1$

or

(ii) $r \leqq \frac{1}{2}$ *and* $L+1 > 3/2\alpha + 2/\alpha^2$,

then for some $\gamma > 0$,

$$(4.2)\qquad \begin{aligned} &\max_x u^N(t,x) - \max_x u^N(0,x) \leqq (1+t)\left(\frac{c}{N^{\beta_1}} + \frac{\hat{c}}{N^{\gamma}}\right), \\ &\min_x u^N(t,x) - \min_x u^N(0,x) \geqq -(1+t)\left(\frac{c}{N^{\beta_1}} + \frac{\hat{c}}{N^{\gamma}}\right) \end{aligned}$$

on $0 \leqq t \leqq \tilde{c}N^{\min(\beta_1,\gamma)}$. *If* $\beta_1 = 0$, $\tilde{c}$ *can be taken arbitrarily large for* N *greater than or equal to some* $N_0(\tilde{c})$. *Furthermore,* (4.2) *holds even when* (1.1) *is a system, provided that an estimate of the form* (4.4) *below can be obtained, e.g., by the method of invariant regions.*

Remark. If the kernel K_N in (2.11) is chosen to be Fejér's kernel F_N, then

$$\min_x u^0(x) \leqq \min_x u^N(0,x) \leqq \max_x u^N(0,x) \leqq \max_x u_0{}'x),$$

and hence if $\beta_1 > 0$, then (4.2) states that for big enough N and any fixed time interval $[0, T]$ the approximation $u^N(t, x)$ remains within an arbitrarily small distance of $[\min u_0, \max u_0]$. Hence the convergence theorem to be proven in the next section will apply even when f is defined on a small neighborhood of that set. For example, if $f(u)$ is defined for $u > 0$, then u_0 need only be larger than an arbitrarily small positive constant.

Proof of Theorem 2. Given an assumed bound B for $\|u^N\|_{L^\infty}$, we will prove a bound of the form (4.2) with $\hat{c} = \hat{c}(B)$. If B is large enough, then (4.2) implies that $\|u^N\|_{L^\infty}$ is actually smaller than B for $t \leqq O(N^{\min(\beta_1, \gamma)})$. By the continuity in time of $\|u^N(t, \cdot)\|_{L^\infty}$ and the fact that (4.2) holds at time zero, it follows that (4.2) indeed holds at least up to times of this order. If $\beta_1 > 0$, then the maximum allowable time clearly tends to infinity with N, but even if $\beta_1 = 0$ we can make $\hat{c}$ as large as desired by choosing B and N big enough. Hence we may assume that $\|u^N\|_{L^\infty} \leqq B$ so that the conclusion of Theorem 1 holds. Also, since the solution u^N_{new} to (2.8), (2.11) with u_0 replaced by $u_0 + c_n$ and f replaced by $f(\cdot - c_n)$ is $u^N_{\text{old}} + c_n$, by picking $c_n = -\frac{1}{2}[\max u^N(0, x) + \min u^N(0, x)]$ we can reduce (4.2) to

$$\|u^N(t, x)\|_{L^\infty} - \|u^N(0, x)\|_{L^\infty} \leqq (1+t)\left(\frac{c}{N^{\beta_1}} + \frac{\hat{c}}{N^\gamma}\right). \tag{4.3}$$

But

$$\begin{aligned} &\|u^N(t, x)\|_{L^\infty} - \|u^N(0, x)\|_{L^\infty} \\ &\quad \leqq \int_0^t [\|\partial_x(I - P_N)f(u^N)\|_{L^\infty} + \varepsilon_N \|\partial_x^2 R_m(N) * u^N\|_{L^\infty}], \end{aligned} \tag{4.4}$$

as can be seen by applying the maximum principle to $\partial_t u^N + f(u^N)_x = \varepsilon_N u_{xx} + g$ or by using energy estimates to bound the change of the L^p norm and letting $p \to \infty$.

The second term of the integrand in (4.4) is less than or equal to

$$c\varepsilon_N m^2 \|R_m\|_{L^1} \|u^N\|_{L^\infty} \leqq \tilde{c} B N^{-\beta_1}$$

by two applications of Lemma 3.5 plus assumption (4.1). Integrating this from zero to t gives $ctN^{-\beta_1}$, which complies with (4.3). The treatment of the first term of the integrand depends on which of the alternative hypotheses is assumed, although in either case we start with the inequality

$$\|\partial_x(I - P_N)f(u^N)\|_{L^\infty} \leqq \|(I - S_N)f(u^N)_x\|_{L^\infty} + \|\partial_x A_N f(u^N)\|_{L^\infty}. \tag{4.5}$$

If alternative (i) holds, then by the Gagliardo-Nirenberg inequality $\|u_x\|_{L^\infty} \leqq c\|u_{xx}\|^{1/2}\|u_x\|^{1/2}$ (see [6]), the right side of (4.5) is dominated by

$$\begin{aligned} &c\|(I - S_N)f(u^N)_{xx}\|^{1/2}\|(I - S_N)f(u^N)_x\|^{1/2} \\ &\quad + c\|\partial_x^2 A_N f(u^N)\|^{1/2}\|\partial_x A_N f(u^N)\|^{1/2} \\ &\leqq c\|(I - V_N)[f'(u^n)u^N_{xx} + f''(u^N)u^N_x u^N_x]\|^{1/2} \\ &\quad \cdot \|(I - V_N)f'(u^N)u^N_x\|^{1/2} + cN^{3/2}\|A_N f(u^N)\| \\ &\leqq cN^{3/2} \sum_{g = I, f', f''} \|(I - V_{[N/k]})g(u^N)\| + cN^{1/2}\|(I - S_N)f(u^N)_x\| \\ &\leqq \tilde{c}N^{3/2} \sum_{g = I, f', f''} \|(I - V_{[N/k]})g(u^N)\| \\ &\leqq \tilde{c}N^{3/2} \sum_{g = I, f', f''} \|(I - S_{[N/2k]})g(u^N)\| \end{aligned} \tag{4.6}$$

for some k, where we have made several uses of Lemmas 3.1, 3.2, and 3.5. Since $s \geqq s_0+1$, estimate (3.19) of Theorem 1 also holds when f is replaced by f'', and hence the last line in (4.6) is less than or equal to

$$\begin{aligned} &cN^{3/2}\left[\left(\frac{1}{N\varepsilon_N}\right)^{L+1} + N^{-\alpha(L+1)} + N^{-r}\exp\left(-c\varepsilon_N N^{2-\alpha/4}t\right)\right] \\ &\leqq cN^{3/2}[N^{-\alpha(L+1)} + N^{-r}\exp\left(-ctN^{(1+3\alpha/4)}\right)], \end{aligned} \tag{4.7}$$

and hence the integral from 0 to t of the first term of the integrand in (4.4) is less than or equal to $\hat{c}(1+t)[N^{3/2-\alpha(L+1)} + N^{3/2-r-1-3\alpha/4}] \leqq \hat{c}N^{-\gamma}$ with $\gamma = \min(\alpha(L+1)-\frac{3}{2}, r+(3\alpha/4)-\frac{1}{2}) > 0$ by assumption. Thus (4.3) holds as claimed in this case.

If the alternative (ii) is assumed instead, then we will use the Gagliardo-Nirenberg inequality [6] $\|u_x\|_{L^\infty} \leqq c_{(m)}\|\partial_x^m u\|_{L^2}^{a}\|u\|_{L^\infty}^{1-a}$ for $m \geqq 2$ with $a = (m-\frac{1}{2})^{-1}$. Apply this to (4.5) to get

$$\begin{aligned} &\|\partial_x(I-P_N)f(u^N)\|_{L^\infty} \\ &\leqq c(m)[\|\partial_x^m(I-S_N)f(u^N)\|^{1/(m-1/2)}\|(I-S_N)f(u^N)\|_{L^\infty}^{1-1/(m-1/2)} \\ &\qquad + \|\partial_x^m A_N f(u^N)\|^{1/(m-1/2)}\|A_N f(u^N)\|_{L^\infty}^{1-1/(m-1/2)}] \\ &\leqq c(m)c(B)\log N[\|\partial_x^m(I-S_N)f(u^N)\|^{1/(m-1/2)} \\ &\qquad + N^{(m/m(m-1/2))}\|A_N f(u^N)\|^{1/(m-1/2)}] \\ &\leqq c(m)c(B)\log N\|\partial_x^m(I-S_N)f(u^N)\|^{1/(m-1/2)}, \end{aligned} \tag{4.8}$$

where we have used Lemma 3.1 and the fact that the operators S_N and PS_N, and hence also $A_N = PS_N - S_N$, are bounded in L^∞ with norm less than or equal to $c \log N$. (For the case of PS_N the L^∞ norm must be interpreted with care, since $PS_N w$ depends on the values of w at only a finite number of points, but since A_N is only applied to continuous functions in (4.8) we have no problem.) The bound for S_N follows from the formula $S_N w(y) = \int D_N(y-x)w(x)\,dx$ upon estimating the L^1 norm of D_N [15], while the bound for PS_N follows from a similar estimate for the formula

$$PS_N w(y) = \sum_{j=0}^{2N} D_N\left(y - \frac{2\pi j}{2N+1}\right) w\left(\frac{2\pi j}{2N+1}\right),$$

which is just a Riemann sum for the above integral.

Now, since $r \leqq \frac{1}{2}$ by assumption, Theorem 1 says that for $k \geqq N$, $\|(I-S_k)f(u^N)\| \leqq c(k/N)^{-8L^2}[N^{-(L+1)\alpha} + \exp(-ctN^{(1+\alpha/4)})]$. Therefore,

$$\begin{aligned} \|\partial_x^m(I-S_N)f(u^N)\| &\leqq \sum_{j=0}^{\infty} \|\partial_x^m S_{2^{j+1}N}(I-S_{2^j N})f(u^N)\| \\ &\leqq \sum_{j=0}^{\infty} (2^{(j+1)}N)^m\|(I-S_{2^j N})f(u^N)\| \\ &\leqq c\sum_{j=0}^{\infty} 2^{(j+1)m}N^m(2^j)^{-8L^2}[N^{-(L+1)\alpha} + \exp(-ctN^{1+\alpha/4})] \\ &\leqq c\sum_{j=0}^{\infty} (2^j)^{m-8L^2}[N^{m-(L+1)\alpha} + N^m\exp(-ctN^{1+\alpha/4})]. \end{aligned} \tag{4.9}$$

Because $(A_1+A_2)^{1/m} \leqq A_1^{1/m}+A_2^{1/m}$ for $A_i \geqq 0$ and $m \geqq 1$, by inserting (4.9) into (4.8) and integrating from zero to t we obtain

$$\begin{aligned} &\int_0^t \|\partial_x(I-P_N)f(u^N)\|_{L^\infty} \\ &\leqq \int_0^t c \log N\left[\sum_{j=0}^{\infty}(2^j)^{m-8L^2}\right]^{1/(m-1/2)} \\ &\quad\cdot\{N^{[m-(L+1)\alpha]/(m-1/2)}+N^{m/(m-1/2)}\exp(-ctN^{1+\alpha/4}/(m-\tfrac{1}{2}))\} \\ &\leqq c(1+t)\log N\{N^{[m-(L+1)\alpha]/(m-1/2)}+N^{[m/(m-1/2)]-(1+\alpha/4)}\} \end{aligned} \tag{4.10}$$

provided that

$$m<8L^2 \tag{4.11}$$

so that the sum converges.

Finally, hypothesis (ii) implies that $\alpha(L+1)>\frac{3}{2}+2/\alpha$, and hence there exists a positive integer m such that

$$\alpha(L+1)>\frac{3}{2}+\frac{2}{\alpha} \geqq m>\frac{1}{2}+\frac{2}{\alpha}, \tag{4.12}$$

which implies that both powers of N in (4.10) are negative. Except for the uninteresting case where $\alpha>4$, (4.12) implies that $m \geqq 2$, as required by the Gagliardo-Nirenberg inequality cited above, and also that (4.11) holds since L is a positive integer. Hence (4.10) complies with (4.3) since $\log N<N^{\delta}$ for any δ. Even if $\alpha>4$, all inequalities needed to obtain (4.3) will hold if we redefine $m=2$.

Estimate (4.4) cannot be derived by the maximum principle when (1.1) is a system, since that principle does not hold for systems. However, if we assume (4.4) then the rest of the proof is still valid for systems, since it uses essentially only the Gagliardo-Nirenberg inequalities and the results of § 3. □

5. The convergence theorem.

THEOREM 3. *Assume that* (1.1) *is a scalar equation. Let* u^N *be the solution of the approximating equations* (2.8), (2.11). *Assume that* ε_N *and* $m(N)$ *in* (2.8) *satisfy*

$$\varepsilon_N \geqq cN^{\alpha-1} \quad \textit{with } \alpha>0, \tag{5.1}$$

$$m(N)^2\|R_{m(N)}(x)\|_{L^1} \leqq \frac{1}{\varepsilon_N N^{\beta_1}} \quad \textit{with } \beta_1>0. \tag{5.2}$$

Let $r \geqq 0$ *be a number such that*

$$\sup_k k^r\|(I-S_k)u^N(0,x)\|<\infty. \tag{5.3}$$

For nonnegative integers L let

$$s_0(L,r)=\max(2r+6,16L(L+1)+2), \tag{5.4}$$

and assume either that

$$r \geqq \frac{1}{2} \quad \textit{and} \quad f \in C^s \quad \textit{with} \quad s \geqq s_0(L,r)+1 \quad \textit{and} \quad \alpha(L+1)>\frac{3}{2} \tag{5.5i}$$

or

$$r \leqq \frac{1}{2} \quad \textit{and} \quad f \in C^s \quad \textit{with} \quad s \geqq s_0(L,r) \quad \textit{and} \quad (L+1)>\frac{3}{2\alpha}+\frac{2}{\alpha^2}. \tag{5.5ii}$$

Finally, let v^N *be the solution of*

$$v_t^N + f(v^N)_x = \varepsilon_N v_{xx}^N, \qquad v^N(0,x) = u_0(x). \tag{5.6}$$

Then for $t \leqq O(N^\gamma)$ *with* $\gamma > 0$,

$$\|u^N - v^N\|_{L^1} \leqq \|u^N(0,x) - u_0\|_{L^1} + c(1+t)\left[\frac{1}{N^{\beta_1}} + \frac{1}{N^{\alpha(L+1)-1}} + \frac{1}{N^{r+\alpha/2}}\right]. \tag{5.7}$$

Proof.

$$\begin{aligned} &\partial_t(u^N - v^N) + \partial_x[f(u^N) - f(v^N)] \\ &\quad = \varepsilon_N \partial_x^2(u^N - v^N) + \partial_x(I - P_N)f(u^N) + \varepsilon_N \partial_x^2 R_m * u^N. \end{aligned} \tag{5.8}$$

Multiplying (3.8) by $\operatorname{sgn}(u^N - v^N)$ and integrating yields [9]

$$\begin{aligned} &\|u^N - v^N\|_{L^1} - \|u^N(0,x) - u_0\|_{L^1} \\ &\leqq \int_0^t \|\partial_x(I - P_N)f(u^N)\|_{L^1} + \varepsilon_N \|\partial_x^2 R_m * u^N\|_{L^1} \\ &\leqq c\int_0^t \|\partial_x(I - P_N)f(u^N)\| + \varepsilon_N \|\partial_x^2 R_m * u^N\| \\ &\leqq c\int_0^t N[\|(I - S_N)f(u^N)\| + \|A_N f(u^N)\|] + c\int_0^t \varepsilon_N m^2 \|R_m(x)\|_{L^1} \|u^N\| \\ &\leqq c\int_0^t N\|(I - S_N)f(u^N)\| + \|(I - S_N)f(u^N)_x\| + cN^{-\beta_1}\int_0^t \|u^N\| \\ &\leqq c\int_0^t N \sum_{g = I, f, f'} \|(I - S_{[N/8]}g(u^N))\| + cN^{-\beta_1}\int_0^t \|u^N\|, \end{aligned} \tag{5.9}$$

where we have made use of Lemmas 3.1 and 3.2 and the relationship between S and V. Since the hypotheses ensure that Theorems 1 and 2 hold for $t \leqq cN^\gamma$, the last line in (5.9) is less than or equal to

$$\begin{aligned} &cN\int_0^t \left[\left(\frac{1}{\varepsilon_N N}\right)^{L+1} + N^{-(L+1)\alpha} + N^{-r}\exp\left(-c\tilde{t}\varepsilon_N N^{2-((L+1)\alpha/4L)}\right)\right] d\tilde{t} + ctN^{-\beta_1} \\ &\quad \leqq c(1+t)\left[\frac{1}{N^{\alpha(L+1)-1}} + \frac{1}{N^{r+\alpha/2}} + \frac{1}{N^{\beta_1}}\right]. \qquad \square \end{aligned} \tag{5.10}$$

Since $\|u - v^N\|_{L^1} \leqq ct\varepsilon_N^{1/2}$ [8], where u is the solution of (1.1), (1.2), the following corollaries hold.

COROLLARY 1. *Under the hypotheses of Theorem* 3,

$$\|u - u^N\|_{L^1} \leqq \|u_0 - u^N(0,x)\|_{L^1} + c(1+t)\left[\frac{1}{N^{\alpha(L+1)-1}} + \frac{1}{N^{r+\alpha/2}} + \frac{1}{N^{\beta_1}} + \frac{1}{N^{(1-\alpha)/2}}\right]$$

for $t \leqq O(N^\gamma)$.

COROLLARY 2. *If the hypotheses of Theorem* 3 *hold with* $r = \frac{1}{2}$ *and* $\beta_1 \geqq \frac{1}{2}$, $f \in C^{36/\alpha^2}$, *and* $u_0 \in \mathrm{BV}$, *and if* K_N *in* (2.11) *is Fejer's kernel, de la Vallée Poussin's kernel, or any other kernel that approximates* BV *data to* $O(N^{-1/2})$, *then on any finite time interval* $[0, T]$

$$\|u - u^N\|_{L^1} \leqq \frac{c(T)}{N^{(1-\alpha)/2}}. \tag{5.11}$$

Theorem 3 and its corollaries give the convergence results mentioned in the Introduction. As is clear from Corollary 1, there are five sources of error in the approximation of u by u^N: approximation of the initial data; limited smoothness of f; limited smoothness of u_0; inclusion of the nonlocal second-order term $\varepsilon_N \partial^2 R_m * u^N$ for the purpose of achieving formal infinite-order accuracy; and the error of the viscous approximation v^N. Because of the last of these, the accuracy of the (pseudo) spectral-viscosity method is limited to (5.11) even when the other four are negligible. Of course, we might think that the fourth source of error only appears to cause error due to the method of proof, but in fact reduces the error below that obtained in (5.11). While the first half of this conjecture may be true, we can see that the second is false, at least for certain values of $m(N)$. If we take f and u_0 smooth, let $R_m = F_m$ and $K_n = V_N$, for example, and set $m(N) = N^{1/8}$. Then our method is formally infinite-order accurate, but the error $\|u^N - v^N\|_{L^1} \sim N^{-3/4}$ is of a smaller order than the error $\|u - v^N\|_{L^1} \sim N^{-(1/2-\delta)}$, so that the error $\|u - u^N\|_{L^1}$ is essentially the same as that of $\|u - v^N\|_{L^1}$.

Finally, let us show that for f and u_0 sufficiently smooth and $R_m = 0$,

$$|\hat{u}^N(t,k)| \leqq \frac{c(T)}{1+|k|} \quad \text{on } [0,T], \tag{5.12}$$

as claimed in the Introduction. Let u_h^N be the solution to (2.8), (2.11) with u_0 replaced by $u_0(x+h)$.

By the method used to prove Theorem 3, we obtain on $[0, T]$ that

$$\begin{aligned}\|u_h^N - u^N\|_{L^1} &\leqq \|K_N * \{u_0(x+h) - u_0(x)\}\|_{L^1} + c(T)/N \\ &\leqq c\left[h + \frac{1}{N}\right].\end{aligned}$$

But $u_h^N(x) = u^N(x+h)$ since there is no explicit x dependence in (2.8), so

$$\begin{aligned}|e^{ikh} - 1||\hat{u}^N(k)| &= |\hat{u}_h^N - \hat{u}^N(k)| \\ &\leqq \|u_h^N - u^N\|_{L^1} \leqq c\left(h + \frac{1}{N}\right).\end{aligned}$$

Since $\hat{u}^N(k) = 0$ for $|k| > N$, and $|e^{ikh} - 1| \geqq c|kh|$ for $|kh| \leqq \frac{1}{2}$, by letting $h = N/2$ we obtain

$$\begin{aligned}|\hat{u}^N(k)| &\leqq c\left(h + \frac{1}{N}\right) / |e^{ikh} - 1| \\ &\leqq \frac{c(1/N)}{|k|/N} \leqq \frac{c}{|k|}.\end{aligned}$$

Since $|\hat{u}^N(0) \leqq \|u^N\|_{L^1} \leqq c\|u^N\|_{L^\infty} \leqq \tilde{c}$, therefore (5.12) holds.

6. The elasticity equations. The elasticity equations are [4]

$$u_t - \sigma(v)_x = 0, \tag{6.1}$$

$$v_t - u_x = 0, \tag{6.2}$$

where σ satisfies

$$\sigma'(v) > c > 0, \tag{6.3}$$

$$v\sigma''(v) > 0 \quad \text{for } v \neq 0. \tag{6.4}$$

Although Theorem 3 is inapplicable to (6.1), (6.2), the spectral-viscosity approximants for this system can be shown to satisfy an a priori inequality of the form (4.4), so that those approximants are bounded in L^∞ by Theorem 2. Now DiPerna [4] has proven via compensated compactness that L^∞-bounded approximants of (6.1), (6.2) actually converge to a solution of (6.1), (6.2) provided that the approximants satisfy an additional condition, which he shows to hold for the pure viscosity method. Because the estimate for the right side of (4.4) obtained in the proof of Theorem 2 suffices to show that this additional condition also holds for the spectral-viscosity method, we can use his result to obtain a convergence theorem for the elasticity equations. Before stating the theorem, let us review some definitions that will be used in the statement and proof of the theorem.

First of all, an entropy for system (6.1), (6.2) is a function $\eta(u, v)$ for which there exists a function $q(u, v)$ such that all smooth solutions of (6.1), (6.2) satisfy the additional conservation law $\eta_t + q_x = 0$. For example, the function $\eta^* = \frac{1}{2}u^2 + \int_0^v \sigma(s)\, ds$ is one entropy for (6.1), (6.2), for which $q = -u\sigma(v)$; in [9] it is shown that there exist many entropies for 2×2 systems. Next, an entropy is called convex if $\eta(u, v)$ is a convex function. Finally, a weak solution of (6.1), (6.2) is called admissible if it satisfies the condition $\eta_t + q_x \leqq 0$ in the weak sense for every convex entropy.

THEOREM 4. *Let $U^N = (u^N, v^N)$ be the solution of the spectral-viscosity approximation* (2.8), (2.11) *to system* (6.1), (6.2) *with initial data (u_0, v_0) in L^∞. Assume that the viscosity parameter ε_N satisfies* (3.13) *and tends to zero as $N \to \infty$, that* (4.1) *holds, and that σ is in C^s with s sufficiently large. Then for some subsequence N_j, U^{N_j} converges in L^p, $1 \leqq p < \infty$, to an admissible solution $U = (u, v)$ of* (6.1), (6.2) *with initial data (u_0, v_0). Furthermore, if there exists an admissible piecewise-Lipschitz solution to* (6.1), (6.2) *with initial data (u_0, v_0), then U^N converges to that solution without the need to restrict to a subsequence.*

Proof. As noted in [4] and its references, the theory of [2] implies that the viscosity approximation for (6.1), (6.2) obtained by adding εu_{xx} and εv_{xx}, respectively, to the right side of those equations has invariant regions

$$\{(u, v) | \max_{\pm} |w^\pm(u, v)| \leqq k\}, \tag{6.5}$$

where

$$w^\pm(u, v) \equiv u \pm \int_0^v \sqrt{\sigma'(s)}\, ds \tag{6.6}$$

are the Riemann invariants of (6.1), (6.2). This invariance follows from

$$\partial_t w^\pm \pm \sqrt{\sigma'}\,\partial_x w^\pm = \varepsilon \partial_x^2 w^\pm \mp \varepsilon \frac{\sigma''}{2\sqrt{\sigma'}} v_x^2 \tag*{(6.7)$_\pm$}$$

satisfied by the Riemann invariants, since (6.4) together with the maximum principle for $(6.7)_\pm$ show that the maximum of w^+ cannot increase if $v \geqq 0$ nor can its minimum decrease if $v \leqq 0$, while for w^- the same holds with the sign of v reversed.

Now, the spectral-viscosity method is obtained by the further addition of $-(I - P_N)\partial_x \sigma(v) - \partial_x^2 R_m(N) * u$ and $-\partial_x^2 R_m(N) * v$, respectively, to the right sides of (6.1), (6.2), adding

$$-(I - P_N)\partial_x \sigma(v) - \partial_x^2 R_M(N) * u \mp \sqrt{\sigma'(v)}\,\partial_x^2 R_m(N) * v$$

to the right side of $(6.7)_{\pm}$, so the maximum principle for the modified $(6.7)_{\pm}$ yields

$$\max_{\pm} \|w^{\pm}(u^N(T,\cdot), v^N(T,\cdot))\|_{L^\infty} - \max_{\pm} \|w^{\pm}(u^N(0,\cdot), v^N(0,\cdot))\|_{L^\infty}$$

$$\leqq \int_0^T [\|(I-P_N)\partial_x\sigma(v^n)\|_{L^\infty} + \varepsilon_N \|\partial_x^2 R_m(N) * u^N\|_{L^\infty} + \varepsilon_N \|\sqrt{\sigma'(v^N)}\|_{L^\infty} \|\partial_x^2 R_m(N) * v^N\|_{L^\infty}]. \tag{6.8}$$

Inequality (6.8) is essentially of the form (4.4) since an L^∞ bound on w^- and w^+ implies an L^∞ bound for u and v. Note also that the factor $\|\sqrt{\sigma'(v^N)}\|_{L^\infty}$ in one term on the right of (6.8) causes no trouble, since the rest of the term is less than or equal to $C(T)N^{-\delta}$ for some $\delta > 0$, so that any assumed bound on v^N used to bound this factor can be justified a posteriori. Hence Theorem 2 says that U^N is bounded uniformly in L^∞ on any bounded time interval, and the proof of this theorem shows that

(6.9) The right side of (6.8) is less than or equal to $c(1+T)N^{-\delta}$

for some $\delta > 0$.

In order to apply DiPerna's theorem we must show, besides the L^∞ bound for U^N, that for any entropy η of (6.1)-(6.4), $\partial_t\eta(u^N, v^N) + \partial_x q(u^N, v^N) = A + B$, where A converges in $H^{-1}(t, x)$ and B is uniformly bounded in $L^1_{\text{loc}}(t, x)$. Just as in the pure-viscosity case, the first step towards proving this is to use the entropy η^* to show that

$$\varepsilon_N \int_0^T \|(u_x^N(t,\cdot), v_x^N(t,\cdot))\|^2 \leqq c(T). \tag{6.10}$$

If we multiply the equations for u^N and v^N by u^N and $\sigma(v^N)$, respectively, we obtain

$$\partial_t\left\{\frac{1}{2}(u^N)^2 + \int_0^{v^N} \sigma(S)\, ds\right\} - \partial_x\{u^N\sigma(v^N)\} + \varepsilon_N\{(u_x^N)^2 + \sigma'(v^N)(v_x^N)^2\} = -u^N\partial_x(I-P_N)\sigma(v^N). \tag{6.11}$$

Integrating (6.11) over x and from $t=0$ to $t=T$ yields (6.10) since $\eta^*(u^N, v^N)$ is bounded and

$$\int_0^t \int_0^{2\pi} |u^N\partial_x(I-P_N)\sigma(v^N)|\, dx\, dt \leqq 2\pi\|u^N\|_{L^\infty} \int_0^T \|\partial_x(I-P_N)\sigma(v^N)\|_{L^\infty} \leqq c(1+T)N^{-\delta}$$

by (6.9) and the L^∞ bound for u^N. Now let η be any entropy of (6.1), (6.2). Then

$$\partial_t\eta(u^N, v^N) + \partial_x q(u^N, v^N) = -\varepsilon_N[\eta_{uu}(u_x^N)^2 + \eta_{vv}(v_x^N)^2] + \sqrt{\varepsilon_N}\,\partial_x\sqrt{\varepsilon_N}\,(\eta_u u_x + \eta_v v_x) - \eta_u\partial_x(I-P_N)\sigma(v^N). \tag{6.12}$$

As in [4], the first term on the right of (6.12) is bounded uniformly in $L^1_{\text{loc}}(t, x)$ by (6.10) and the second is $\sqrt{\varepsilon_N}$ times the derivative of a function bounded uniformly in L^2_{loc}, again by (6.10). The last term on the right of (6.12) is also bounded uniformly in L^1_{loc} because of (6.9) plus the L^∞ bound on (u^N, v^N).

Hence the conditions of DiPerna's theorem hold, so by [4] there exists subsequence N_j for which U^{N_j} converges to a solution $U = (u, v)$ of (6.1), (6.2). As noted in [5], this convergence takes place in L^p for any finite p. Furthermore, $|f(U^N)-f(U)| \leqq |\sup_{|w| \leqq |U^N|_{L^\infty}} |f'(w)|| U^N - U|$ shows that for any C^1 function f, $f(U^N)$ converges in L^1 to $f(U)$. We can therefore take a weak limit in (6.12) and obtain that $\eta_t(u, v) + q_x(u, v) \leqq 0$ for any convex entropy, i.e., (u, v) is an admissible solution.

Finally, to obtain convergence without the need to consider a subsequence when the additional condition mentioned in the theorem holds, we note that DiPerna has proved in [3] that under this condition (6.1), (6.2) has a unique admissible solution U. Since any sequence ε_{N_j} will have a subsequence for which U^{N_j} converges to this same U, U^N must itself converge to U. □

REFERENCES

[1] E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1986.

[2] K. CHUEH, C. CONLEY, AND J. SMOLLER, *Positively invariant regions for systems of nonlinear diffusion equations*, Indiana Univ. Math. J., 26 (1977), pp. 373-392.

[3] R. J. DIPERNA, *Uniqueness of the solutions of nonlinear hyperbolic systems of conservation laws*, Indiana Univ. Math. J., 28 (1979), pp. 137-188.

[4] ———, *Convergence of approximate solutions to conservation laws*, Arch. Rational Mech. Anal., 82 (1983), pp. 27-70.

[5] ———, *Measure-valued solutions to conservation laws*, Arch. Rational Mech. Anal., 88 (1985), pp. 223-270.

[6] A. FRIEDMAN, *Partial Differential Equations*, Krieger, New York, 1976.

[7] Y. KATZNELSON, *An Introduction to Harmonic Analysis*, Dover, New York, 1976.

[8] H. O. KREISS, *Fourier Expansions of the solutions of the Navier-Stokes equations and their exponential decay rate*, in Analyse mathématique et appliquée, Gauthier-Villars, Paris, 1988, pp. 245-262.

[9] P. D. LAX, *Hyperbolic Systems of Conservation Law and the Mathematical Theory of Shock Waves*, Society for Industrial and Applied Mathematics, Philadelphia, 1973.

[10] Y. MADAY AND E. TADMOR, *Analysis of the spectral vanishing viscosity method for periodic conservation laws*, SIAM J. Numer. Anal., 26 (1989), pp. 854-870.

[11] R. SANDERS, *On convergence of monotone finite difference schemes with variable spatial differences*, Math. Comp., 40 (1983), pp. 91-106.

[12] E. TADMOR, *Convergence of spectral methods for nonlinear conservation laws*, SIAM J. Numer. Anal., 26 (1989), pp. 30-44.

[13] ———, *Semi-discrete approximation to nonlinear systems of conservation laws: consistency and L^∞ stability imply convergence*, Math. Comp., to appear.

[14] L. TARTAR, *Compensated compactness and applications to partial differential equations*, in Nonlinear Analysis and Mechanics, Heriot-Watt Symposium 4, R. J. Knopps, ed., Res. Notes in Math., Pitman, Boston, 1975, pp. 136-211.

[15] A. ZYGMUND, *Trigonometric Series*, Vols. I and II, Cambridge University Press, Cambridge, U.K., 1959.