The Fourier Method for Nonsmooth Initial Data

Author(s): Andrew Majda, James McDonough and Stanley Osher

Source: *Mathematics of Computation,* Vol. 32, No. 144 (Oct., 1978), pp. 1041-1081

Published by: American Mathematical Society

Stable URL: https://www.jstor.org/stable/2006332

Accessed: 17-09-2019 16:27 UTC

The Fourier Method for Nonsmooth Initial Data

By Andrew Majda,* James McDonough and Stanley Osher*

Abstract. Application of the Fourier method to very general linear hyperbolic Cauchy problems having nonsmooth initial data is considered, both theoretically and computationally. In the absence of smoothing, the Fourier method will, in general, be globally inaccurate, and perhaps unstable. Two main results are proven: the first shows that appropriate smoothing techniques applied to the equation gives stability; and the second states that this smoothing combined with a certain smoothing of the initial data leads to infinite order accuracy away from the set of discontinuities of the exact solution modulo a very small easily characterized exceptional set. A particular implementation of the smoothing method is discussed; and the results of its application to several test problems are presented, and compared with solutions obtained without smoothing.

Introduction. In recent years the Fourier method for the numerical approximation of solutions to hyperbolic initial value problems has been used quite successfully. In fact, if the initial function is C^∞ and the coefficients of the equation are constant the method converges arbitrarily fast, i.e. is limited in practice only by the method of time discretization which is chosen. This is the reason that the Fourier method is called "infinite order" accurate.

However, the situation is drastically different when the initial function is not smooth. We take as a model the one space dimension scalar problem $u_t = u_x$ to be solved for 2π periodic u on the interval $-\pi \leqslant x \leqslant \pi$ with initial values $\varphi(x)$, having a simple jump discontinuity at $x = 0$, but otherwise smooth and 2π periodic.

In this simple example the rate of convergence is *globally only second order.* (In fact, if any value for $\varphi(0)$ except the average of the right and left limits is chosen the method degenerates further to be globally only first order.) This means that even in regions where the exact solution is smooth, i.e. away from the line $x = t$, the error is $O(h^2)$, where h is the mesh width. The analysis of this and related examples is carried out in Section 1 of this paper. There we show that such a large global error occurs in general situations. However, we note here that in certain constant coefficient problems special time discretizations may give better accuracy on mesh points than predicted for the semidiscrete problem. We analyze this phenomenon, which we believe is limited to constant coefficient problems, in the next section. Thus, even in the simplest cases, discontinuous initial data causes a large error unless we modify the Fourier method.

Even in the variable coefficient scalar one dimensional case $u_t = a(x)u_x$ numerical experiments indicate that the Fourier method is unstable for appropriate $a(x)$ when $a(x)$ passes through zero. Analytic examples of this phenomenon have also been discussed by Kreiss and Oliger [5]. (However, the proof of instability given there is incorrect. They do give a valid procedure for stabilizing a certain class of problems.)

In this paper we first discuss smoothing modifications of the Fourier method using only the fast Fourier transform (FFT) which guarantee stability. Similar smoothing procedures to stabilize the Fourier method have been developed independently and concurrently by Kreiss and Oliger in [15]. Unfortunately, these easily implemented smoothing procedures, while guaranteed to be stable, give large regions where the rate of convergence is still only second order. For example, there is a large region of low accuracy for the one dimensional problem

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_x + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_{t=0} = \begin{pmatrix} \varphi \\ 0 \end{pmatrix},$$

where $\varphi(x)$ has a jump discontinuity at $x = 0$ but is otherwise C^∞. We define $\varphi(0) = \frac{1}{2}(\varphi(0^+) + \varphi(0^-))$. This is the problem which we approximated using a dissipative finite difference method in [7]. We discuss this example in Section 2, showing there that the Fourier method using the above mentioned smoothing gives, as in [7], an error of $O(h^2)$, throughout most of the range of influence of the initial discontinuity at $x = 0$. This simple smoothing does work well for the special case of a scalar, constant coefficient equation $u_t = au_x$.

We show in Section 2 that in this case the rate of convergence is infinite off the singular support of the exact solution under very general conditions on $\varphi(x)$. But whenever there is coupling—either through lower order terms, variable coefficients, or multi-dimensions—we expect very restricted rates of convergence if we use this type of smoothing.

In order to make the Fourier method a genuine infinite order method in regions where the exact solution is smooth, a more severe smoothing technique must necessarily be introduced. We define the method in the preliminary section and prove in Section 3 that away from the singular set this modified smoothing method has infinite rate of convergence.

The techniques which we shall use to prove stability and convergence of the Fourier method with smoothing are discrete analogues of standard methods in the theory of linear hyperbolic equations. These include pseudodifferential operators, wave front sets, and Egorov's theorem. Much of the technical work involves translating these methods to a grid. This is implemented with the help of the Poisson summation formula and the integration by parts procedure of the method of stationary phase.

Finally, our rate of convergence and stability results were verified when we per-

formed several numerical experiments on the IBM 360/91 computer at UCLA. One interesting aspect of this work is the construction of certain functions used to implement the smoothing method. Theoretically, we required that they be C^∞. This is too much to require realistically for computing purposes. However, the results with C^2, C^4 and C^6 functions were very good numerically. Moreover, even the Gibbs phenomenon of large error near the discontinuity was reduced using our methods.

The authors would like to thank Michael Taylor for some interesting discussions of the spurious singularity phenomenon discussed in the preliminary section and analyzed in Section 4.

0. Preliminary Section. We shall consider the hyperbolic system

$$(0.1) \qquad \frac{\partial u}{\partial t} = \sum_{\nu=1}^{d} A_\nu(x) \frac{\partial u}{\partial x_\nu} + B(x)u = Lu,$$

where the A_ν and B are C^∞ square $n \times n$ matrices, the unknown u is an n vector

$$u = \begin{pmatrix} u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_d \end{pmatrix}.$$

For the stability result, Theorem 1, we require that (0.1) be symmetrizable hyperbolic. However, for the propagation of error result, Theorem 2, we shall require strict hyperbolicity in the usual sense.

For simplicity of exposition we have taken time independent coefficients.

The initial function is

$$(0.2) \qquad u(x, 0) = \varphi(x),$$

where $\varphi(x)$ is a distribution.

The problem is to be solved for $t > 0$ on the cube Ω_d, i.e. the set of x such that $-\pi \leqslant x_i \leqslant \pi$, $i = 1, 2, \ldots, d$; with $\varphi(x)$, $u(x, t)$, and the coefficient matrices all periodic with period 2π in each x_i.

The Fourier method usually considered consists of replacing the x_ν derivatives on a grid of uniform spacing h by the derivative of the interpolating trigonometric polynomial at the grid points. This can be implemented using two fast Fourier transforms for each u_j, $j = 1, 2, \ldots, n$, and each x_ν, $\nu = 1, \ldots, d$. This leads to a system of differential-difference equations which can be solved numerically by introducing a time step k and using various possible time discretizations (specific examples are shown in Section 4).

This method was originally suggested by Kreiss and Oliger [5], and Orszag [9], [10], and was later considered by Fornberg [2], [3].

The usual initial function one takes for this method must have a discrete Fourier transform

$$(0.3) \qquad F(\varphi)(\zeta) = \widetilde{\varphi}(j) = (h/2\pi)^d \sum_{x \in Y_d} e^{-ijx} \varphi(x),$$

where Y_d is the d-dimensional lattice with spacing h: $x \in Y_d$ if and only if $x = (x_1, \ldots, x_d)^T$ and $x_\nu \in \{\mu h\}_{\mu=-N}^{N}$, with $(2N + 1)h = 2\pi$. For $j \in \mathbf{Z}^d$ we define $j \cdot x = \Sigma_{\nu=1}^{d} j_\nu x_\nu$.

For this discretization $\varphi(x)$ should make sense on lattice points. We shall show below after introducing more general smoothing techniques that this restriction is unnecessary.

In the standard Fourier method one uses the interpolating polynomial

$$(0.4) \qquad I^h \varphi = \sum_{jh \in Y_d} \widetilde{\varphi}(j) e^{ij \cdot x}$$

as the initial function for the difference-differential equation defined for $x \in Y_d$, $t > 0$:

$$(0.5) \qquad u_t^h = \left[\sum_{\nu=1}^{d} i A_\nu(x) F^{-1} j_\nu F + B(x) \right] u^h = L^h u^h.$$

The first question is one of stability: Is the semidiscrete problem of (0.5) stable? Precisely, we define the H_s norm of a lattice function $w(x)$ to be

$$(0.6) \qquad \|w\|_s = \left(\sum_{jh \in Y_d} |\widetilde{w}(j)|^2 (1 + |j|^2)^{s/2} \right)^{1/2}$$

This is just the standard H_s norm for the trigonometric interpolant of $w(x)$.

We wish an estimate for all h sufficiently small:

$$(0.7) \qquad \|u^h(\cdot, t)\|_s \leqslant C_s e^{kt} \|\varphi\|_s.$$

Here we introduce the convention that C, k, δ, ϵ, with or without sub or superscripts, are universal positive constants which may differ from relation to relation.

We shall prove stability of a smoothing version of (0.5). Let $\rho(x)$ be a C^∞ function which vanishes identically when any x_ν is near $\pm \pi$. Let $\rho(x) \equiv 1$ in some neighborhood of the origin, and $0 \leqslant \rho \leqslant 1$. Such cutoff functions will be used throughout this work.

In the Fourier method with smoothing we replace (0.5) with the following semidiscrete problem to be solved on Y_d for $t > 0$,

$$(0.8) \qquad u_t^h = \left(\sum_{\nu=1}^{d} A_\nu(x) F^{-1} i j_\nu + B(x) F^{-1} \right) \rho(jh) \widetilde{u}(j) = L^{h,\rho} u^h.$$

We assume that the operator $\Sigma_{\nu=1}^{d} A_\nu(x) j_\nu$ is smoothly symmetrizable, i.e. that there exists a smooth invertible matrix function $T(x, \zeta)$, homogeneous of degree zero in ζ for $|\zeta|$ large, such that $T^{-1}(x, \zeta)(\Sigma_{\nu=1}^{d} A_\nu(x) \zeta_\nu) T(x, \zeta)$ is symmetric. This assumption is satisfied by strictly hyperbolic and of course symmetric hyperbolic problems.

We shall prove the following stability theorem in Section 2.

THEOREM 1. *The solution to* (0.8) *with initial data* φ^h *satisfies, for any s and h sufficiently small, the estimate*:

(0.9) $$\|u^h(\cdot,\ t)\|_s \leqslant C_s e^{kt}\|I^h\varphi\|_s.$$

We also have:

COROLLARY 1. *The error between the solutions u and u^h to the Cauchy problems (0.1) and (0.8), respectively, when $u(x,\ 0) = \varphi(x)$ on Ω_d and $u^h(x,\ 0) = \varphi(x)$ on Y_d with $\varphi \in C^\infty$ satisfies*

$$\|u(\cdot,\ t) - u^h(\cdot,\ t)\|_s \leqslant C_{s,\lambda}h^\lambda \quad \forall s,\ \lambda.$$

From now on we use the notation $v \approx 0$ for a vector v obeying the estimates satisfied by $u - u^h$ above. We use similar notation for operators between any two Sobolev spaces.

Next we address ourselves to the question of the convergence of this method.

Let $\sigma(x)$ be a cutoff function with the additional requirement that $\rho(x)\sigma(x) \equiv \sigma(x)$.

A first attempt at getting optimal rates of convergence away from the singular support of $u(x,\ t)$ might be to replace the initial data (0.4) by:

(0.10) $$\sigma(hD)\varphi^h = \sum_{jh\in Y_d} e^{ij\cdot x}\sigma(jh)\widetilde{\varphi}(j).$$

However, we have pointed out in the introduction the very restricted worth of such an initial smoothing; and we give the details in Section 2.

Thus, we need a different initial function for the semidiscrete problem (0.8). We proceed as follows. Let the Fourier coefficients of a distribution $u(x)$ be defined as:

(0.11) $$\hat{u}(j) = (Fu)(j) = \left(\frac{1}{2\pi}\right)^d \int_{\Omega_d} e^{-ix\cdot j}u(x)\,dx.$$

We define

(0.12) $$\sigma(hD)\varphi = \sum_{j=-\infty}^{\infty} e^{ij\cdot x}\sigma(jh)\hat{\varphi}(j),$$

and we solve (0.8) with this as our initial function.

Our main result concerns the error between the solution to (0.1), (0.2) and the solution to (0.8) with initial function

(0.13) $$u(x,\ 0) = \sigma(hD)\varphi.$$

Our propagation of error result is very close to the following statement—The rate of convergence of u^h to u is arbitrary in any compact region in which $u(x,\ t)$ is smooth. It is not quite that—viz. the spurious singular set discussed below. However, for the reader who is unfamiliar with wave front sets and the propagation of singularity results of Hörmander [4], the statement above serves as a good basis for understanding the result in most cases.

We first recall the notion of singular support of a distribution $u(x)$. To say that a point x_0 is not in the singular support means that u is C^∞ in a neighborhood of x_0. Or, there exists a function $v(x)$ with compact support which agrees with u in a neighborhood of x_0 and is C^∞, i.e. its Fourier transform satisfies:

for any N, $(1 + |\xi|)^N (Fv)(\xi)$ is bounded as $|\xi| \to \infty$.

We say that a point (x_0, ξ^0), $\xi^0 \in R^d\backslash 0$, is not in the wave front set of $u(x)$ (denoted WF u) if there exists a function $u(x)$ with compact support which agrees with $u(x)$ in a neighborhood of x_0 and whose Fourier transform satisfies the following: there is a cone R in ξ-space,

$$R: \left| \frac{\xi}{|\xi|} - \frac{\xi_0}{|\xi_0|} \right| < \epsilon$$

such that for every N, $(1 + |\xi|)^N (Fv)(\xi)$ is bounded in R as $|\xi| \to \infty$.

Thus, WF u is a closed conical set in (x, ξ) space.

To make our results precise we consider the set WF φ which is the union of the wave front sets of each component of

$$\varphi = \begin{pmatrix} \varphi_1 \\ \cdot \\ \cdot \\ \cdot \\ \varphi_n \end{pmatrix}.$$

The results of [4] indicate that WF $u(x, t)$ is contained in the set obtained in the following manner:

Take the scalar dth order differential operator which is the principal symbol of

(0.14) $$\det [D_t - \sum A_\nu(x)D_{x_\nu}] = p(x, D_x, D_t).$$

Let $(x_0, \xi^0) \in$ WF φ.

By strict hyperbolicity the equation

(0.15) $$p(x_0, \xi^0, \tau) = 0$$

has the distinct roots, $\tau = \tau_j(x_0, \xi^0), j = 1, \ldots, d$; and hence, there are d distinct null bicharacteristics $\Gamma_j, j = 1, \ldots, d$, associated with (x_0, ξ^0).

Bicharacteristics are the curves in (x, t, ξ, τ) space satisfying the Hamilton-Jacobi equations:

$$dx_j/ds = \partial p/\partial \xi_j, \quad j = 1, 2, \ldots, d,$$

$$dt/ds = \partial p/\partial \tau,$$

$$d\xi_j/ds = -\partial p/\partial x_j, \quad j = 1, 2, \ldots, d,$$

$$d\tau/ds = -\partial p/\partial t.$$

It turns out that p is constant on each of these curves; one on which p vanishes is called a null bicharacteristic of p.

Let the union of the set of all projections onto x space of all the null bicharacteristics of p through $(x_0, 0, \xi^0, \tau_j(x_0, \xi^0))$, $j = 1, \ldots, d$, for fixed t be denoted by $S(t)$. Given $\delta > 0$, $T > 0$, define:

$$(0.16) \qquad R_\delta = \{(x, t) \in \Omega_d \times (0, T) | \text{ distance } (x, S(t)) > \delta\}.$$

We can now state our convergence theorem which is proven in Section 3.

THEOREM 2. *For any* $\lambda > 0$, $|s| \geqslant 0$, $\exists$ *a constant* $C_{\delta,s,\lambda}$ *such that*:

$$\sup_{(x,t)\in R_\delta} |\partial_x^s(u - u^h)| \leqslant C_{\delta,s,\lambda} h^\lambda.$$

(Here and below $\partial_x^s u^h$ means the derivative of the trigonometric interpolant of $u^h = I^h u^h$.)

We notice that the set $S(t)$ is, in general, only slightly larger than the singular support of $u(x, t) = SSu(x, t)$. In fact, the total decoupling procedure of Taylor [13] gives us the exact singular support of u. However, we believe that the rate of convergence is not good on the spurious singular set, $SpSu(x, t) = S(t) - SSu(x, t)$. This was noticed by Smoller and Taylor [11] in a related problem. We, in fact, give an alternate proof in Section 3 that the rate of convergence is not infinite for a term similar to theirs on $SpSu(x, t)$.

1. **Examples of Large Regions of Low Accuracy for the Fourier Method.** In this section we present simple examples to show that the Fourier method without smoothing is of limited accuracy globally. We then show that for the time discretized problem it is possible in special cases to do better than predicted for the semidiscrete problem and improve the still limited accuracy by a power of h. This phenomenon, which we believe occurs only for constant coefficients, resembles the collocation method used in finite element problems.

We approximate the Cauchy problem for the scalar differential equation

$$(1.1) \qquad \partial u/\partial t = \partial u/\partial x \quad \text{on } \Omega_1 \times [0, T]$$

with

$$(1.2) \qquad u(x, 0) = \Phi(x).$$

The semidiscrete approximation is a very simple version of (0.5).

$$(1.3) \qquad \partial u^h/\partial t = F^{-1}ijFu^h$$

with initial function defined on the lattice Y_1 to be

$$(1.4) \qquad u(x, 0) = \varphi(x).$$

We are interested in functions $\varphi(x)$ of the type

$$(1.5) \qquad \varphi_\gamma(x) = \rho(x)x_+^\gamma, \qquad \gamma > 0, \text{ where } x_+ = \begin{cases} x, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

and $\rho(x)$ is a cutoff function of the type discussed in the introduction.
We also define:

$$(1.6) \qquad\qquad \varphi_0(x) = \rho(x)H(x),$$

where $H(x)$ is the Heaviside function:

$$
\begin{aligned}
H(x) &= 1, & x &> 0,\\
H(0) &= 1/2,\\
H(x) &= 0, & x &< 0.
\end{aligned}
$$

Each of these functions is in the class $F_e(S^{-1-\gamma})$. This means that its Fourier transform

$$(1.7) \qquad\qquad \hat{\varphi}(\xi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\xi x}\varphi(x)\,dx_{\cdot}$$

has the property

$$\hat{\varphi}(\xi) \in S_{1,0}^{-1-\gamma}(R^1)$$

or

$$(1.8) \qquad\qquad |\partial_\xi^\alpha \hat{\varphi}(\xi)| \leqslant C_\alpha(1 + |\xi|)^{-1-\gamma-\alpha}$$

and

$$\varphi_\gamma \notin F_e(S^{-1-\mu}) \quad \text{if } \mu > \gamma.$$

Recall the definition of R_δ in the previous section. In this case:

$$(1.9) \qquad\qquad R_\delta = \{(x, t) \in \Omega_1 \times [0, t] \,|\, |x + t| > \delta\}.$$

We now have:

PROPOSITION 1. *Suppose that the initial function* $\varphi \in F_e(S^{-1-\gamma})$ *for some* γ
> 0. *Then*
    (1)

$$(1.10) \qquad\qquad \max_{(x,t)\in R_{\delta,h}} |u(x,\, t) - u^h(x,\, t)| \leqslant C_\delta h^{1+\gamma}.$$

    (2) *The conclusion is sharp in the sense that we can find some* $\varphi \in F_e(S^{-1-\gamma})$
so that, with this initial data:

$$(1.11) \qquad\qquad \overline{\lim_{h\searrow 0}}\ \max_{(x,t)\in R_\delta \cap \theta} \left| \frac{u(x,\, t) - u^h(x,\, t)}{h^{1+\gamma}} \right| \geqslant C > 0,$$

where θ *is any region with compact closure in* R_δ.

An anomalous situation occurs for $\varphi = \varphi_0$. The rate of convergence is the same
as for $\varphi \in F_e(S^{-2})$ which is a class of functions less singular than φ_0. Hence, the
convergence is better than for the smoother functions $\varphi \in F_e(S^{-1-\gamma})$ with $1 > \gamma > 0$.
This anomaly was noticed in our previous paper [7] in a similar connection. It is

explained by the fact that, although $\hat{\varphi}_0(\zeta)$ behaves like $C_1/|\zeta|$ for $\zeta \rightarrow \infty$, the discrete Fourier coefficient $\widetilde{\varphi}(j)$ behaves like C_2/j^2 not $C_2/|j|$, for $|j| \rightarrow \pi/h$.

We have the following.

PROPOSITION 2. *Suppose that the initial data $\varphi = \varphi_0$. Then*
(1)

$$(1.12) \qquad \max_{(x,t) \in R_\delta} |u(x, t) - u^h(x, t)| \leqslant C_\delta h^2.$$

(2) *This conclusion is sharp for φ_0 in the same sense as in Proposition* 1.

Moreover, if any value of $\varphi_0(0)$ except $\varphi_0(0) = 1/2$ is chosen, the statements (1) *and* (2) *above are valid with h^2 replaced by h, i.e. the method is globally only first order accurate.*

Next we consider a fully discrete problem

$$(1.13) \qquad u^{h,\lambda}(\cdot, t + k) = F^{-1}q(jk)FU^{h,\lambda}(\cdot, t)$$

defined for $t = \nu k$, $\nu = 0, 1, 2, \ldots$.

Here $\lambda = k/h > 0$ is fixed independent of t.

The function $q(y)$ is C^∞ for $|y| < \pi\lambda + \delta^1$ for some $\delta^1 > 0$, and not, in general, periodic.

The difference scheme is assumed to be rth order accurate, which means

$$(1.14) \qquad q(jk) - e^{ijk} = O((jk)^{r+1})$$

as $jk \rightarrow 0$ and stable, which means

$$(1.15) \qquad \sup_{|y| \leqslant \pi\lambda} |q(y)| \leqslant 1.$$

We also make a certain generic assumption about $q(y)$ which we shall state precisely below.

Let

$$Y_{d,\lambda} = \{(x, t) \in \Omega_d \times [0, t] \,|\, x \in Y_d, \, t = \nu\lambda h, \, \nu = 0, 1, 2, \ldots\}$$

and

$$R_{\delta,\lambda} = R_\delta \cap Y_{d,\lambda}.$$

We now have two propositions which are discrete analogs of Propositions 1 and 2, with the improved convergence rate on $R_{\delta,\lambda}$ under certain circumstances.

PROPOSITION 1'. *Suppose the initial data $\varphi \in F_e(S^{-1-\gamma})$ for some $\gamma > 0$. Then*
(1)

$$(1.16) \qquad \max_{(x,t) \in R_{\delta,\lambda}} |u(x, t) - u^{h,\lambda}(x, t)| \leqslant C_{\delta,\lambda} h^{\min(1+\gamma,r)}.$$

(2) *The conclusion* (1.16) *is sharp in the same sense as in Proposition* 1 *unless $q(\pi\lambda) = q(-\pi\lambda)$ and $q'(\pi\lambda) = q'(-\pi\lambda)$ in which case the upper bound in* (1) *may be replaced by $C_{\delta,\lambda} h^{\min(2+\gamma,r)}$.*

PROPOSITION 2'. *Suppose the initial data $\varphi = \varphi_0$. Then*

(1)

(1.17) $$\max_{(x,t)\in R_{\delta,h,\lambda}} |u(x,\ t) - u^{h,\lambda}(x,\ t)| \leqslant C_{\delta} h^{\min(2,r)}.$$

(2) *This conclusion is sharp for φ_0 in the same sense as Proposition 2.*

Finally, any choice of $\varphi_0(0)$ except $\varphi_0(0) = 1/2$ will degrade the error by one power of h unless $q(\pi\lambda) = q(-\pi\lambda)$ and $q'(\pi\lambda) = q'(-\pi\lambda)$ in which case the estimate in (1) is valid.

Proof of Propositions 1 *and* 2. The key fact in the proof of Proposition 1 and the second half of Proposition 2 is:

LEMMA 1. *For* $\varphi \in F_e S^{-1-\gamma}$

$$\max_{(x,t)\in R_\delta} \left| u - u^h + \widetilde{\varphi}(N) \frac{\sin(\pi(x + t)/h)}{\sin((x + t)/2)} \right| \leqslant Ch^{2+\gamma}.$$

We have, if $\gamma > 0$,

(1.18) $$\widetilde{\varphi}(j) = \sum_{\mu=-\infty}^{\infty} \hat{\varphi}(j + \mu(2N + 1))$$

(see Kreiss and Oliger [6]); and, thus, it is easy to see that for such γ:

(1.19) $$|\widetilde{\varphi}(N)| \leqslant C_1 h^{1+\gamma}.$$

Moreover, for the function in $F_e(S^{-1-\gamma})$

(1.20) $$\varphi(x) = \sum_{j=-\infty}^{\infty} e^{ijx}(1 + |j|)^{-(1+\gamma)},$$

it is easy to show that

$$\lim_{h\downarrow 0} \frac{\widetilde{\varphi}(N)}{h^{1+\gamma}} = C_2 \neq 0.$$

Thus, the proof of Proposition 1 is immediate.

We also have the above-mentioned anomaly:

LEMMA 2.

(1.21) $$\lim_{h\to 0} \frac{\widetilde{\varphi}_0(N)}{h^2} = C > 0,$$

in fact

$$\widetilde{\varphi}_0(j) = \frac{h}{2\pi}\left[\frac{1}{2i}\frac{\cos(jh/2)}{\sin(jh/2)} + O\left(\frac{h^\lambda}{(\sin(jh/2))^\lambda}\right)\right] \quad \forall \lambda > 0.$$

The second part of Proposition 2 follows from the fact that if we add to $\varphi_0(x)$ the trigonometric interpolant of a function which is 1 at $x = 0$ and zero on the other mesh points, the associated finite Fourier transform changes by h. This result together

with the previous two lemmas gives us the desired result.

The first half of Proposition 2 follows from the previous two lemmas and the following aesthetically annoying result.

LEMMA 3. *For* $\varphi = \varphi_0$ *we have*

$$
\max_{(x,t)\in R_{\delta,h}} \left| u(x,\ t) - u^h(x,\ t) + \widetilde{\varphi}(N)\frac{\sin(\pi(x\ +\ t)/h)}{\sin((x\ +\ t)/2)} \right.
$$

$$
+ \frac{e^{i(N+1)(x+t)}}{(e^{i(x+t)}-1)^2}(\widetilde{\varphi}(N) - \widetilde{\varphi}(N-1))
$$

(1.22)

$$
\left. + \frac{e^{i(N+1)(x+t)}}{(e^{i(x+t)}-1)^2}(\widehat{\varphi}(-N) - \widehat{\varphi}(-(N+2)) - \widetilde{\varphi}(-(N+1)) + \widetilde{\varphi}(-(N+2))) \right|
$$

$$
\leqslant C_{\delta,0}h^3.
$$

For any fixed x, the three last functions above are linearly independent functions of t. Thus, using Lemmas 1 and 2 we have proven a sharpened version of Proposition 2.

Proof of Lemmas 1 and 3. We have

$$
(1.23) \qquad u(x,\ t) = \sum_{j=-\infty}^{\infty}\widehat{\varphi}(j)e^{ij(x+t)}, \qquad u^h(x,\ t) = \sum_{j=-N}^{N}\widetilde{\varphi}(j)e^{ij(x+t)}.
$$

Then, we write

$$
(1.24) \qquad u - u^h = \sum_{|j|\geqslant N+1}\widehat{\varphi}(j)e^{ij(x+t)} + \sum_{j=-N}^{N}(\widehat{\varphi}(j) - \widetilde{\varphi}(j))e^{ij(x+t)}.
$$

Next, we sum by parts in a standard fashion. Let $D_{\pm j}U(j) = \pm[U(j\pm1) - U(j)]$. We have in R_{δ}:

$$
\sum_{|j|\geqslant N+1}\widehat{\varphi}(j)e^{ij(x+t)}
$$

$$
= \frac{1}{(e^{i(x+t)}-1)}\sum_{j\geqslant N+1}[\widehat{\varphi}(j)D_{+j}e^{ij(x+t)} - \widehat{\varphi}(-j)D_{-j}e^{-ij(x+t)}]
$$

(1.25)

$$
= -e^{i(N+1)(x+t)}\widehat{\varphi}(N) + e^{-iN(x+t)}\widehat{\varphi}(-(N+1))
$$

$$
- \frac{1}{(e^{i(x+t)}-1)}\sum_{j\geqslant N+1}[e^{ij(x+t)}D_{-j}\widehat{\varphi}(j) - e^{-ij(x+t)}D_{+j}\widehat{\varphi}(-j)]
$$

and

$$\sum_{j=-N}^{N} (\hat{\varphi}(j) - \widetilde{\varphi}(j))e^{ij(x+t)}$$

$$= (e^{i(x+t)} - 1)^{-1} \sum_{j=-N}^{N} (\hat{\varphi}(j) - \widetilde{\varphi}(j))D_{+j}e^{ij(x+t)}$$

(1.26)

$$= (e^{i(x+t)} - 1)^{-1}[e^{i(N+1)(x+t)}(\hat{\varphi}(N) - \widetilde{\varphi}(N))$$

$$- e^{-iN(x+t)}(\hat{\varphi}(-(N+1)) - \widetilde{\varphi}(-(N+1)))]$$

$$- (e^{i(x+t)} - 1)^{-1} \sum_{j=-N}^{N} e^{ij(x+t)}D_{-j}(\hat{\varphi}(j) - \widetilde{\varphi}(j)).$$

Using the obvious fact that $\widetilde{\varphi}(j) = \widetilde{\varphi}(j + \mu(2N+1))$ for any integer μ gives us, after another summation by parts,

$$u - u^h = -\widetilde{\varphi}(N)\frac{\sin(\pi(x+t)/h)}{\sin((x+t)/2)} + e^{i(N+1)(x+t)}(e^{i(x+t)} - 1)^{-2}D_{-j}\widetilde{\varphi}(N)$$

$$+ e^{-iN(x+t)}(e^{i(x+t)} - 1)^{-2}[D_{+j}\hat{\Phi}(-(N+1))$$

$$+ D_{-j}\hat{\varphi}(-(N+1)) + D_{-j}\widetilde{\varphi}(-(N+1))]$$

(1.27)

$$+ (e^{i(x+t)} - 1)^{-2}\left(\sum_{j \geqslant N+1} (e^{ij(x+t)}D_{-j}^2\hat{\varphi}(j) + e^{-ij(x+t)}D_{+j}^2\hat{\varphi}(-j))\right.$$

$$\left. - \sum_{j=-N}^{N} e^{ij(x+t)}D_{-j}^2\widetilde{\varphi}(j)\right).$$

The proofs of the lemmas now follow from the triangle inequality, Lemma 2, the definition of the classes $F_e(S^{-1-\gamma})$, and (1.18).

Proof of Lemma 2.

$$\widetilde{\varphi}_0(j) = \frac{h}{2\pi}\left[\sum_{\mu=1}^{N} e^{-ij\mu h}\rho(\mu h) + \frac{1}{2}\right]$$

(1.28)

$$= \frac{h}{2\pi}\left[(e^{-ijh} - 1)^{-1}\left(\sum_{\mu=1}^{N}(h\Delta_+ e^{-ijh})\rho(\mu h)\right) + \frac{1}{2}\right].$$

Here

$$h\Delta_{\pm}f(\mu h) = \pm(f((\mu \pm 1)h) - f(\mu h))$$

$$= \frac{h}{2\pi}\left[\frac{1}{2} - \frac{e^{-ijh}}{(e^{-ijh} - 1)} - (e^{-ijh} - 1)^{-1}\sum_{\mu=1}^{N} e^{-ij\mu h}h\Delta_-\rho(\mu h)\right]$$

$$= \frac{h}{2\pi}\left[\frac{1}{2i}\frac{\cos(jh/2)}{\sin(jh/2)} + O\left(\frac{h^\lambda}{(\sin(jh/2))^\lambda}\right)\right] \quad \forall\lambda \geqslant 1.$$

Next, we treat the fully discrete problem.

The technical assumption we make on the function $q(y)$ is:

$$(1.29) \qquad \inf_{\substack{|y| \leq \pi\lambda \\ (x,t) \in R_\delta}} \left| ix + \frac{tq'(y)}{q(y)} \right| > 0.$$

The proofs of Propositions $1'$ and $2'$ follow from discrete analogues of Lemmas 1 and 3.

First, we have:

LEMMA $1'$. *For $\varphi \in Fe(S^{-1-\gamma})$*

$$\max_{(x,t) \in R_{\delta,\lambda}} \left| u - u^h + \frac{\widetilde{\varphi}(N)}{2i} \left[e^{i(N+1)x}(q((N+1)k))^{t/k}((e^{ix}(q(N+1)k)/q(N))^{t/k} - 1)^{-1} \right. \right.$$

$$\left. \left. - e^{iNx}(q(-Nk))^{t/k}\left(e^{ix}\left(\frac{q(-Nk)}{q(-(N+1)k)}\right)^{t/k} - 1\right)^{-1} \right] \right|$$

$$\leq C_\delta h^{2+\gamma}.$$

This result follows from summation by parts in the same way as in the proof of Lemma 2. We note that the coefficient of $\widetilde{\varphi}(N)$ will be $O(h)$ if and only if

$$(1.30) \qquad [q(\pi\lambda + k/2)]^{t/k} = [q(-\pi\lambda + k/2)]^{t/k} + O(k)$$

or if and only if

$$(1.31) \qquad q(\pi\lambda) = q(-\pi\lambda) \quad \text{and} \quad q'(\pi\lambda) = q'(-\pi\lambda).$$

We also have a messy looking discrete analogue of Lemma 3 which we obtain by summing by parts once more, as in the proof of that lemma. We omit the details, except to say that the cancellation occurring in Lemma $1'$ is impossible here.

2. Stability of the Smoothed Fourier Method and the Necessity of the Proper Initial Smoothing. In this section we shall prove Theorem 1, stability of the smoothed Fourier method (0.8). It might then be thought that convergence in R_δ follows for initial data smoothed as follows:

$$(2.1) \qquad u(x, 0) = \sigma(hD)\varphi^h = \sum_{j \in Y_d} e^{ij \cdot x}\sigma(jh)\widetilde{\varphi}(j).$$

We present here a proof that the example mentioned in the introduction which we discussed in detail in a related connection in [7] leads to a *large* region of low accuracy. In fact we have:

PROPOSITION 3. *For (0.8) approximating the above-mentioned example with initial data (2.1) for*

$$\varphi = \begin{pmatrix} \varphi_0 \\ 0 \end{pmatrix},$$

we have

    (1)

(2.2)
$$\max_{(x,\,t)\in R_{\delta}^{(1)}} |u(x,\,t) - u^h(x,\,t)| \leqslant C_{\delta} h^2,$$

where $R_{\delta}^{(1)} = R_{\delta} \cap \{(x,\,t)| - t < x < t\}$.

    (2) *The conclusion is sharp in the same sense as in Proposition* 1.

We note that for more complicated examples having jump discontinuities in higher dimensions it is likely that the error will be worse than second order in a large region if we use this smoothing of the initial data.

Proof. The proof is an easier version of that of the main theorem of [7] and is typical of the approach used in this analysis. We write

$$u - u^h = \sum_{j=-\infty}^{\infty} e^{ijx} e^{(A\,ij+B)t}(1 - \sigma(jh))\hat{\varphi}(j)$$

(2.3)
$$+ \sum_{|j|\leqslant N} e^{ijx}(e^{(A\,ij+B)t} - e^{(A\,ij+B)\rho(jh)t})\sigma(jh)\hat{\varphi}(j)$$

$$- \sum_{|j|\leqslant N} e^{ijx}(e^{(A\,ij+B)\rho(jh)t})\sigma(jh) \sum_{|\nu|\geqslant 1} \hat{\varphi}(j + \nu(2N + 1))$$

(using (1.19))

$$= [E_{\mathrm{I}}] + [E_{\mathrm{II}}] + [E_{\mathrm{III}}],$$

but $[E_{\mathrm{II}}] \equiv 0$ since $\rho(jh) \equiv 1$ on the support of $\sigma(jh)$.

    Now Lemma 2.1 of [7] tells us

(2.4)
$$\max_{(x,\,t)\in R_{\delta}} |[E_{\mathrm{I}}(x,\,t,\,h)]| \leqslant C_{\delta,\lambda} h^{\lambda} \quad \forall \lambda > 0,$$

while $[E_{\mathrm{III}}] + O(h^3) = $ the solution of the Cauchy problem for this hyperbolic system with initial data:

$$u(x,\,0) = h^2 c_0 \begin{pmatrix} \delta'(0) \\ 0 \end{pmatrix},$$

since

$$\sum_{|\nu|\geqslant 1} \hat{\varphi}(j + \nu(2N + 1)) = \begin{pmatrix} c_0 h^2 ij + O(h^3) \\ 0 \end{pmatrix},$$

with c_0 a nonzero constant and δ', the derivative of the Dirac measure. The result is immediate using (3.24) of [7].

    For completeness we mention:

PROPOSITION 4. *For the single scalar equation*

(2.5)
$$u_t = cu_x$$

with $u(x,\,0) = \varphi(x) \in F_e(S^{-1-\gamma})$, $\gamma > 0$ *or* $\varphi = \varphi_0(x)$ *and the approximation* (0.8)

with initial data (2.1) *we have*

(2.6)
$$\max_{(x,t)\in R_\delta} |u(x,\ t) - u^h(x,\ t)| \leqslant C_{\delta,\gamma,\lambda}h^\lambda \quad \forall \lambda > 0.$$

The proof follows from the decomposition of (2.3), using a by now standard summation of parts.

In order to prove Theorems 1 and 2 we need to introduce a class of pseudo-differential operators defined on the lattice Y_d. For any lattice function u we define:

(2.7)
$$P^h(x,\ D,\ h)u = \sum_{jh\in Y_d} e^{ij\cdot x}p(x,\ j,\ h)\widetilde{u}(j),$$

where $p(x,\ \xi,\ h)$ is a C^∞ function with the usual 2π period in each x_i, vanishing identically if $|\xi_i| > \pi/h \,\forall i$, and having certain other properties discussed below.

Recall, on R^d, Psdops are defined by

(2.8)
$$P(x,\ D)u = \int e^{i\xi\cdot x}p(x,\ \xi)(Fu)(\xi)\,d\xi;$$

and an often used class of symbols (see [12]) is

$$S_{1,0}^m = \{p(x,\ \xi)|\ |\partial_x^\alpha\partial_\xi^\beta p(x,\ \xi)| \leqslant C_{\alpha,\beta}(1 + |\xi|)^{m-|\beta|}\}.$$

We shall define a related class of symbols to be used on lattice functions. These have the two properties:

(1) $\exists\ m_1,\ m_2$ real numbers with $m_2 \geqslant 0$

$$|\partial_x^\alpha\partial_\xi^\beta p(x,\ \xi,\ h)| \leqslant C_{\alpha,\beta}(1 + |\xi|)^{m_1-|\beta|}h^{m_2}.$$

(2) For each $p(x,\ \xi,\ h)\ \exists\ \epsilon > 0$ such that

$$|\partial_x^\alpha\partial_\xi^\beta p(x,\ \xi,\ h)| \leqslant C_{\alpha,\beta,\lambda}h^\lambda$$

$\forall\lambda > 0$ if $|\xi_i| > (\pi - \epsilon)/h\ \forall i$.

Call this class of symbols $S_{1,0}^{m_1,m_2}$.

We also shall need a subclass of these symbols, which have the property $p(x,\ \xi,\ h) \in S_{1,0}^{m_1,m_2}$ and

(3) $|\partial_x^\alpha\partial_\xi^\beta p(x,\ \xi,\ h)| \leqslant C_{\alpha,\beta}h^{m_2+|\beta|}(1 + |\xi|)^{m_1}.$

Call this subclass $T_{1,0}^{m_1,m_2}$. Notice $\rho(\xi h) \in T_{1,0}^{0,0}$.

We now list some of the basic properties of these operators. We shall indicate the proof of the main results in the appendix. We merely remark here that the main tool in the proof is the Poisson summation formula, which reduces the problem to a standard one involving the usual pseudo-differential operator calculus on R^n; modulo a "smoothing" error discussed below.

Properties of the Psdops on Y_d:

(1) $P^h(x,\ D,\ h)$ maps C^∞ functions into $\bigcap_{s=0}^\infty H_s$ with a bound independent of h.

(2) The class of operators whose symbols are in $\bigcup_{m_1,m_2} S_{1,0}^{m_1,m_2}$ form an algebra as does the symbol subclass $\bigcup_{m_1,m_2} T_{1,0}^{m_1,m_2}$.

If $p \in S_{1,0}^{m_1,m_2}$, $q \in S_{1,0}^{m_1',m_2'}$, then the symbol of $Q^hP^h \in S_{1,0}^{m_1+m_1',m_2+m_2'}$ is determined by Leibniz' formula

(3) $r(x, \xi, h) \sim \Sigma_\alpha (1/\alpha!) [\partial_\xi^\alpha q(x, \xi, h)] [D_x^\alpha p(x, \xi, h)]$ (where $D_x = \partial_x/i$), the series is asymptotic in the sense that:

$$\left| r(x, \xi, h) - \sum_{|\alpha|=0}^{|m|} \frac{1}{\alpha!} [\partial_\xi^\alpha q(x, \xi, h)] [D_x^\alpha p(x, \xi, h)] \right|$$

(2.10a)

$$\leq C_m (1 + |\xi|)^{m_1 + m_1' - |m| - 1} h^{m_2 + m_2'}.$$

If in addition $q \in T_{1,0}^{m_1', m_2'}$, then

$$\left| r(x, \xi, h) - \sum_{|\alpha|=0}^{|m|} \frac{1}{\alpha!} [\partial_\xi^\alpha q(x, \xi, h)] [D_x^\alpha p(x, \xi, h)] \right|$$

(2.10b)

$$\leq C_m (1 + |\xi|)^{m_1 + m_1'} h^{m_2 + m_2' + |m| + 1}$$

Note that the second inequality implies that the H_s norm of the error goes to zero as a power of h for $|m|$ sufficiently large, and this power increases with $|m|$.

   Remark 1. As a consequence of (2.10a) we see that if the supports of the symbols of P^h and Q^h are disjoint, then $R^h = Q^h P^h$ is infinitely smoothing. It maps any function u for which $|\widetilde{u}(\eta)| \leq C(1 + |\eta|)^k$ for some fixed k, $C > 0$ independent of h, into an element of $\bigcap_{s=1}^\infty H_s$. In fact, we then have

(2.11) $\| (I^h - \rho^h(hD)) Q^h(x, D, h) P^h(x, D, h) u \|_s \leq C_{s,\lambda} h^\lambda \quad \forall s, \lambda.$

   (4) For any Psdop P^h $\exists$ an "adjoint" Psdop $(P^h)^*$ satisfying

$$(P^h u, v) = (u, (P^h)^* u) \quad \text{for } u, v \in C^\infty$$

(we take $(,)$ as the H_0 inner product), and $(P^h)^*$ has symbol

$$p^*(x, \xi, h) \sim \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha D_x^\alpha \overline{p(x, \xi, h)}^t.$$

Here t denotes the transpose if $p(x, \xi, h)$ is a matrix.

   Finally,

   (5) Suppose $p(x, \xi, h) \in S_{1,0}^{m_1, m_2}$ has an asymptotic expansion of the type

$$p(x, \xi, h) \sim \sum_{m=-m_1}^\infty p_{-m}(x, \xi, h).$$

Here each element of p_m is a sum of terms of the form

$$\sum_{j=0}^{j_m} h^{m_2 + j} p_{m,j}^{(1)}(x, \xi) a_{m,j}(h\xi),$$

where the $p_{k,j}^{(1)}(x, \xi)$ are homogeneous of degree $j + k$ in ξ for $|\xi|$ large, and each $a_{k,j}(x) \equiv 0$ if x_i is near $\pm \pi$ for any i.

   We then say that the principal symbol of P^h is p_{m_1} and the principal symbol of $P^h - P_{m_1}^h$ is p_{m_2}, etc.

Moreover, given a sequence $\{p_m\}$ of the type above, $\exists p \in S_{1,0}^{m_1,m_2}$ such that $p \sim \Sigma_{m=-m_1}^{\infty} p_{-m}(x, \xi, h)$.

The subclass of $S_{1,0}^{m_1,m_2}$ having this expansion will be useful in the proof of Theorem 2.

In order to prove Theorem 1 and Corollary 1 we shall need a lattice version of Gårding's Inequality.

LEMMA 4 (*Lattice Gårding's Inequality*). *Let* $p(x, \xi, h) \in S_{1,0}^{\circ m}$ *have the property that as a matrix,* Re $p(x, \xi, h) \geqslant c |\xi|^m$ *for* ξ *large. Then for any real* s' *and all* h *sufficiently small*

$$(2.12) \quad \mathrm{Re}(P^h(x, D, h)\rho^h(hD)u, \rho^h(hD)u) \geqslant C_0\|\rho^h(hD)u\|_{m/2}^2 - C_1\|\rho^h(hD)u\|_{s'}^2,$$

with C_0, C_1 *independent of* u.

Proof. The proof is a simple modification of the standard one on R^d (see, e.g. Taylor [12]).

We shall prove it for $m = 0$, the general case presents no new difficulties.

Following [12], we may construct a Psdop on R^d, $b(x, \xi, h) \in S_{1,0}^0$ such that Re $P(x, D, h) - B^*(x, D, h)B(x, D, h)$ is infinitely smoothing. Now let $\rho_1(x)$ be a cutoff function with $\rho_1(x) \equiv 1$ on the support of ρ. Then

$$(2.13) \quad \mathrm{Re}(P^h\rho^h u, u) \approx \mathrm{Re}(\rho_1 P^h\rho_1\rho^h u, \rho^h u).$$

It is easy to see that the symbol of $\rho_1^h(hD)P^h(x, D, h)\rho_1^h(hD) \in S_{1,0}^{0;0}$ as does the symbol of $\rho_1^h(hD)B^h(x, D, h)\rho_1^h(hD)$ and its adjoint.

Moreover, if $\rho_2(x)$ is a cutoff function with $\rho_1(x) \equiv 1$ on the support of $\rho_2(x)$, we may use the functional calculus to show

$$[\mathrm{Re}\,\rho_1^h P^h\rho_1^h - (\rho_1^h B^h\rho_1^h)^*(\rho_1^h B^h\rho_1^h)] \rho_2^h = S_{-\infty}^h,$$

where $S_{-\infty}^h$ is an infinitely smoothing operator.

Thus:

$$(2.14) \quad \mathrm{Re}(P^h\rho^h u, \rho^h u) = (\rho_1^h B^h\rho_1^h\rho^h u^h, \rho_1^h B^h\rho_1^h\rho^h u^h) + \mathrm{Re}((S_{-\infty}^h)\rho^h u, \rho^h u).$$

The rest of the proof is standard.

Proof of Theorem 1. We shall again only prove it for $s = 0$, the general case is similar.

We may rewrite the semidiscrete Cauchy problem (0.8) as

$$\text{(a)} \quad u_t^h = (P^h(x, D)\rho^h(hD) + B^h(x, D)\rho^h(hD))u^h \quad \text{with initial data on } Y_d$$

(2.15)

$$\text{(b)} \quad u^h(x, 0) = \varphi(x).$$

Here $p(x, \xi) = \Sigma_{\nu=1}^d A_\nu(x)i\xi_\nu$.

We have assumed $\exists$ a symmetric positive definite symbol $M_0(x, \xi)$ homogeneous of degree zero in ξ if $|\xi|$ is large, so that

$$(2.16) \quad \mathrm{Re}\, M_0(x, \xi)p(x, \xi) = 0.$$

Using the functional calculus and the lattice Gårding's Inequality, we may con-

struct a pseudo-differential operator with symbol in $S_{1,0}^{0,0}$ which differs from $\rho_1(\xi h)M_0(x, \xi)\rho_1(\xi h)$ by an operator with symbol in $S_{1,0}^{-1,0}$. The associated operator $\rho_1^h(hD)M^h(x, D, h)\rho_1^h(hD)$ is symmetric, and has the property

$$(2.17) \qquad (\rho_1^h M^h \rho_1^h \rho^h u^h, \rho^h u^h) \geqslant C\|\rho^h u^h\|_0^2.$$

Then, as in the continuous case

$$(2.18) \quad \frac{\partial}{\partial t}(\rho_1^h M^h \rho_1^h \rho^h u^h, \rho^h u^h) = 2\,\mathrm{Re}(\rho_1^h M^h \rho_1^h \rho_1^h(P^h + B^h)\rho^h u^h, \rho^h u^h)$$
$$\leqslant C_1\|\rho^h u^h\|_0^2 \leqslant C_2(\rho_1^h M^h \rho_1^h \rho^h u^h, \rho^h u^h).$$

Thus, it follows $\exists C_3,\, k > 0$, such that

$$(2.19) \qquad \|\rho^h u^h\|_0 \leqslant C_3 e^{kt}\|\rho^h \varphi^h\|_0.$$

Next we multiply (2.15) by $I^h - \rho_1^h$

$$(2.20) \qquad \partial(I^h - \rho_1^h)u^h/\partial t = (I^h - \rho_1^h)(P^h + B^h)\rho^h u^h = T_{-\infty}^h \rho^h u^h,$$

where $\|T_{-\infty}^h\|_s \leqslant C_{s,\lambda}h^\lambda\ \forall s, \lambda$. It thus follows $\forall \lambda > 0$

$$(2.21) \qquad \|(I^h - \rho_1^h)u^h(t)\|_0 \leqslant C_{4,\lambda}e^{kt}(h^\lambda\|\rho^h\varphi\|_0 + \|(I^h - \rho_1^h)\varphi\|_0).$$

Next, we multiply (2.15) a) by $\rho_1^h - \rho^h$ obtaining:

$$(2.22) \qquad (\rho_1^h - \rho^h)u_t^h = (P^h + B^h)\rho^h(\rho_1^h - \rho^h)u^h + D\rho^h u^h,$$

where D is an operator bounded on H_0 with norm bounded independently of h.

Let ρ_0 be a cutoff function with $\rho_0 \equiv 1$ on support of ρ_1.
We now multiply (2.22) by $\rho_0^h M^h \rho_0^h$, easily, obtaining

$$(2.23) \qquad \|(\rho_1^h - \rho^h)u^h(t)\|_0 \leqslant C_4 e^{kt}\|(\rho_1^h - \rho^h)\varphi\|_0.$$

The theorem now follows from adding (2.19), (2.21), and (2.23), and using the triangle inequality.

Proof of Corollary 1. Since $u \in C^\infty$, we know that

$$(2.24) \qquad \|u(x, t) - I^h u(x, t)\|_s \leqslant C_{s,\lambda}h^\lambda \quad \forall s, \lambda.$$

Thus, we need only show that

$$(2.25) \qquad \|I^h u(x, t) - u^h(x, t)\|_s \leqslant C_{s,\lambda}h^\lambda.$$

We consider

$$(2.26) \quad \begin{aligned}(\partial_t - L^{h,\rho})(u - u^h) &= (\partial_t - L^{h,\rho})u = (L - L^{h,\rho})u \\ &= (P^h(x, D) + B^h(x))(I^h - \rho^h(hD))u \\ &\quad + ((P(x, D) + B(x)) - (P^h(x, D) + B^h(x)))u.\end{aligned}$$

Both terms above obey an estimate like (2.25) because $u \in C^\infty$. Moreover, $I^h u(x, t)$ and $u^h(x, t)$ agree at $t = 0$. Thus, (2.25) follows from the previous theorem and Duhamel's principle.

3. **Proof of the Convergence Theorem.** In this section we shall prove the main result of this paper—Theorem 2.

Recall u solves the strictly hyperbolic Cauchy problem (0.1), (0.2) on $\Omega_d \times [0, T]$; u^h solves the Fourier approximation with smoothing (0.8) for smoothed initial data (0.13) on the region $Y_d \times [0, T]$.

Let T_t be the solution operator to the problem (0.1), (0.2), i.e. $T_t u(x, 0) = u(x, t)$.

Let $T_t^{h,\rho}$ be the solution operator to (0.8) with initial data $u^h(x, 0)$, i.e. $T_t^{h,\rho} u^h(x, 0) = u^h(x, t)$.

We now have the expected decomposition:

$$u - u^h = T_t \varphi - T_t^{h,\rho} \sigma(hD)\varphi$$

(3.1)
$$= T_t(I - \sigma(hD))\varphi + (T_t - T_t^{h,\rho})\sigma(hD)\varphi$$

$$= [E_I] + [E_{II}].$$

Notice that the term

$$[E_{III}] = T_t^{h,\rho} \sigma(hD)(\varphi^h - \varphi),$$

which caused the large region of low accuracy in Section 2 and in [7] is not present because of our initial smoothing.

We shall first obtain the estimate:

(3.2)
$$\sup_{(x,t)\in R_\delta} |\partial_x^s[E_I]| \leqslant C_{\delta,s,\lambda} h^\lambda.$$

In fact, the same estimate will hold on the wave front set level. We shall discuss this below.

For technical reasons we do the following: use a partition of unity to write the initial function $\varphi = \varphi_1 + \varphi_2$, where φ_1 has support in $|x_i| \leqslant \pi - \epsilon_i$, $\epsilon_i > 0$, $i = 1, \ldots, d$, φ_2 has support near each $x_i \pm \pi$, and both functions are periodic.

For the Cauchy problem corresponding to initial data φ_2, we may translate each x_i by π. Thus, we need consider initial data with support of the same type as φ_1, and we restrict $t \leqslant t_1$ to be small enough so that the support of $u(x, t)$ stays away from $x_i = \pm \pi$. This obviously is no restriction, since for $t = t_1$ we just repeat this procedure. This has the effect of simplifying certain technical aspects of the proof.

In particular, it turns out that if the symbol of $P(x, D, h)$ is in $S_{1,0}^{m_1, m_2}$ and $p(x, \xi, h) \equiv 0$ near $x_i = \pm \pi \, \forall i$, then the operator defined by:

(3.3)
$$P^L(x, D, h)\varphi = \frac{1}{(2\pi)^d} \sum_{j=-\infty}^{\infty} \int_{\Omega_d} e^{ij\cdot(x-x_0)} p(x, j, h)\varphi(x_0)\, dx_0$$

differs from the standard Psdop

(3.4)
$$p(x, D, h)\varphi = \frac{1}{(2\pi)^d} \int_{-\infty}^{\infty} d\xi \int_{\Omega_d} e^{i\xi\cdot(x-x_0)} p(x, \xi, h)\varphi(x_0)\, dx_0$$

by an infinitely smoothing operator. Moreover, if $p \in T_{1,0}^{m_1, m_2}$, the difference has norm $O(h^\lambda) \, \forall \lambda$ between any two H_s spaces.

The proof is a simple exercise in the Poisson summation formula and will be carried out in the Appendix.

The operator $\sigma(hD)$ defined in (0.10) should be called $\sigma^L(hD)$. However, for the class of functions that we are now dealing with

$$(3.5) \qquad \sigma^L(hD)\varphi = \sigma^L(hD)\rho_3(x)\varphi(x) \approx \sigma(hD)\rho_3(x)\varphi(x) = \sigma(hD)\varphi,$$

where $\rho_3(x)$ is a cutoff function and $\rho_3\varphi \equiv \varphi$.

Proof of estimate 3.2. Let $\mu(x, \xi) \in S^0_{1,0}$ and have the property that $\mu \equiv 0$ on WFφ, $\mu \equiv 1$ on a set whose distance from WFφ is bounded by $\delta_1 > 0$, and $0 \leqslant \mu \leqslant 1$.

In the rest of the paper we shall let $S_{-\infty}$, $S^h_{-\infty}$ denote smoothing operators on Ω_d and Y_d, respectively. These may vary from relation to relation.

Taylor [13] has shown that $\exists$ two zero order elliptic Psdops $M_1(x, D), M_2(x, D)$ having the properties

$$(3.6) \qquad M_1M_2 = I + S_{-\infty}, \quad M_2M_1 = I + S_{-\infty}, \quad M_1LM_2 = i\Lambda(x, D) + S_{-\infty},$$

where

$$\Lambda(x, D) = \begin{pmatrix} \lambda_1(x, D) & 0 \cdots 0 \\ & \ddots & \\ 0 \cdots 0 & \lambda_n(x, D) \end{pmatrix}$$

and each $\lambda_j(x, D)$ is first order with real principal symbol. Let

$$(3.7) \qquad Q^{(1)}(x, D) = M_2QM_1 + S_{-\infty},$$

where

$$Q(x, D) = \begin{pmatrix} Q_1(x, D) & 0 \cdots 0 \\ 0 & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & Q_n(x, D) \end{pmatrix}$$

and each $q_j(x, \xi) \in S^m_{1,0}$.

Taylor in [14], which is as yet unpublished, has given a new proof of Egorov's Theorem. We shall need both his method of proof and the theorem itself in our proof of Theorem 2. A short discussion of this follows.

Let $q(x, \xi)$ be a scalar symbol of a Psdop on R^d in the class $S^m_{1,0}$ with principal symbol $q_0(x, \xi)$. Let $\lambda(x, \xi)$ be in $S^1_{1,0}$ with real principal symbol $\lambda_1(x, \xi)$. Define R_t to be the solution operator to the Cauchy problem for $u_t = i\lambda(x, D)u$, and let $Q(x, D, t) = R_tQ(x, D)R_{-t}$. Egorov's theorem states: $Q(x, D, t)$ is a smooth family of Psdops defined for all t whose symbols belong to $S^m_{1,0}$. The principal symbol of $Q(x, D, t)$ at a point (x_0, ξ^0) is equal to $q(y_0, \eta^0)$ where (y_0, η^0) is obtained from (x_0, ξ^0) by following the flow generated by the time dependent vector field

$$H_{\lambda_1} = \sum_{j=1}^d \left(\frac{\partial \lambda_1}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial \lambda_1}{\partial x_j} \frac{\partial}{\partial \xi_j} \right).$$

Taylor's proof, which we mimic on a lattice in the proof of Lemma 5 in the Appendix, gives an algorithm for computing all the succeeding symbols of $Q(x, D, t)$ via integration along the above vector field.

It is easy to modify this scalar result to show that:

$$(3.8) \qquad T_t Q^{(1)}(x, D) T_{-t} = M_2 Q(x, D, t) M_1 + S_{-\infty}(t),$$

where the Psdops above depend smoothly on t. We have

$$Q(x, D, t) = \begin{pmatrix} Q_1(x, D, t) & 0 \cdots 0 \\ \cdot & \cdot & \\ 0 & \cdot & Q_n(x, D, t) \end{pmatrix};$$

and each $Q_j(x, D, t)$ is computed, as in [14], by solving the scalar operator O.D.E.

$$dQ_j/dt = i[\lambda_j(x, D), Q_j] + S_{-\infty}(t),$$

(3.10)

$$Q_j(x, D, 0) = Q_j(x, D),$$

where $[A, B]$ denotes the commutator of A and B. We may now define:

$$(3.11) \qquad \mu(x, D, t) = T_t M_1 \mu(x, D) I M_2 T_{-t}.$$

This gives us

$$\mu(x, D, t)[E_I] = T_t M_1 \mu(x, D) I M_2 (I - \sigma(hD))\varphi$$

$$\approx T_t((I - \sigma(hD))\mu(x, D) - M_1 [M_2(I - \sigma(hD)), \mu(x, D)])\varphi.$$

Thus,
$$(3.13) \qquad \|\mu(x, D, t)[E_I]\|_s \leqslant C_{s,\lambda} h^\lambda$$

because $\nu(x, D)\varphi \in C^\infty$ if $\nu(x, \xi) \equiv 0$ on WF φ.

Now any smoothing operator applied to $(I - \sigma(hD))\varphi \approx 0$. This together with (3.13) gives us estimate (3.2).

The algorithm for computing the symbol mentioned above (and outlined on a lattice in the proof of Lemma 5) as well as the definition of $S(t)$ tells us that the symbol of $\mu(x, D, t)$ is, for each t, essentially the identity on a set in (x, ξ) space, which is in general bigger than the set whose projection onto $R^d = R_\delta$. Hence, the justification of the remark after (3.2).

We now justify the remark we made after the statement of Theorem 2. I.e., in general, on the set of spurious singularities the rate of convergence of $[E_I]$ to zero is not infinite. This part can be skipped for the reader who is only interested in the main theorem.

Using the operators M_1, M_2, Λ, defined in (3.6), we can show that the Cauchy problem (0.1), (0.2), is equivalent, modulo smoothing, to

$$(3.14) \qquad w_t = i\Lambda(x, D)w, \qquad w(x, 0) = M_1 u(x, 0).$$

The wave front set of each component w_ν of w is found by following the Ham-

iltonian flow of the νth scalar hyperbolic operator, passing through $(x_0, 0, \xi^0, \lambda_\nu(x_0, \xi^0))$, where $(x_0, \xi^0) \in WF w_\nu(x, 0)$.

Let $\mu_\nu(x, \xi)$, $\nu = 1, 2, \ldots, n$, be a function of the same type as $\mu(x, \xi)$ above, with $\mu_\nu(x, \xi) = 0$ on $WF w_\nu(x, 0)$, etc.

Now, consider the family of operators

$$\mu^{(1)}(x, D, t) = T_t M_2(x, D) \begin{pmatrix} \mu_1(x, D) & & \\ & \ddots & \\ & & \mu_n(x, D) \end{pmatrix} M_1(x, D) T_{-t}.$$

It is again clear from [14] as above that the symbol of $\mu^{(1)}(x, D, T)$ differs from the identity by the symbol of an infinitely smoothing operator on a set close to $WF u$.

Now

$$\mu^{(1)}(x, D, T)[E_I] \approx T_t M_2 \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} (I - \sigma(hD)) w(x, 0)$$

(3.15)
$$+ T_t M_2 \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} M_1 [(I - \sigma(hD)), M_2] w(x, 0)$$

$$\approx T_t M_2 \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} M_1 [(I - \sigma(hD)), M_2] w(x, 0).$$

This term is not ≈ 0 because, in general, the symbol of the Psdop above need not annihilate $WF w(x, 0)$.

Thus, on the spurious singular set the rate of convergence of $[E_I]$ to zero is limited.

Now we consider $[E_{II}]$.

The proof of the estimate involving $[E_{II}]$ looks more complicated than it actually is because of the need to decouple $L^{h,\rho}$.

It follows from the discrete calculus of the previous section that Taylor's [13] total decoupling construction can be simply modified to apply to $L^{h,\rho}$.

We construct the two symbols $N_1(x, \xi, h)$, $N_2(x, \xi, h) \in S^{0;0}_{1;0}$ with $N_i(x, \xi, h) = M_i(x, \xi)$, $i = 1, 2, \ldots$, for $h\xi$ on the set in which $\rho(h\xi) \equiv 1$. Each N_i has the asymptotic expansion mentioned in (5), Section 2. Thus, if $\bar{\rho}(x)$ is any of the usual cutoff functions with $\bar{\rho}\rho \equiv \rho$, we have:

$$N_1^h(x, D, h) N_2^h(x, D, h) \bar{\rho}^h(hD) = \bar{\rho}^h(hD) + S^h_{-\infty}(x, D, h),$$

(3.16) $$N_2^h(x, D, h) N_1^h(x, D, h) \bar{\rho}^h(hD) = \bar{\rho}^h(hD) + S^h_{-\infty}(x, D, h),$$

$$N_1^h(x, D, h) L^{h,\rho} N_2^h(x, D, h) = i\Lambda^{h,\rho}(x, D, h) + S^h_{-\infty}(x, D, h)$$

where

$$\Lambda^{h,\rho}(x, D, h) = \begin{pmatrix} \lambda_1^{h,\rho}(x, D, h) & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & \lambda_n^{h,\rho}(x, D, h) \end{pmatrix}.$$

Here, each $\lambda_i^{h,\rho}(x, \xi, h) \approx \lambda_i(x, \xi)$ for ξh in the set for which $\rho(h\xi) \equiv 1$, $\lambda_i^{h,\rho} \approx 0$ outside the support of $\rho(h\xi)$, and $\lambda_i^{h,\rho}$ has the asymptotic expansion mentioned above, with real principal symbol.

We have

(3.17) $\qquad [E_{II}] = (T_t - T_t^{h,\rho})\sigma(hD)\varphi \approx (T_t - T_t^{h,\rho})N_2^h\sigma_1^h N_1^h \sigma\varphi,$

where σ_1 is a cutoff function satisfying $\sigma(x)\sigma_1(x) \equiv \sigma(x)$ and $\rho\sigma_1 \equiv \sigma_1$. This follows because

(3.18) $\qquad N_2^h\sigma_1^h N_1^h \sigma \approx \sigma + N_2^h[\sigma_1, N_1]\sigma \approx \sigma.$

Thus, we may write

$$[E_{II}] \approx (T_t - T_t^{h,\rho})N_2^h\mu^h\sigma_1^h N_1^h \sigma\varphi + (T_t - T_t^{h,\rho})N_2^h[I^h - \mu^h]\sigma_1^h N_1^h \sigma\varphi$$
$$= [E_{II}^{(1)}] + [E_{II}^{(2)}].$$

Now the symbol modulo smoothing of the operator $N_2^h\mu^h\sigma_1^h N_1^h$ is the same as that of the operator without the h superscript. This operator $N_2\mu\sigma_1 N_1\varphi$ has symbol which vanishes on WF φ. We can show easily since $I - \sigma_1$ and σ have disjoint supports:

(3.20) $\qquad N_2\mu\sigma_1 N_1\sigma\varphi \approx N_2\mu N_1\varphi \in C^\infty.$

But under such circumstances it follows easily that:

(3.21) $\qquad N_2\mu\sigma_1 N_1\sigma\varphi \approx (N_2\mu\sigma_1 N_1)^h\sigma\varphi \approx N_2^h\mu^h\sigma_1^h N_1^h \sigma\varphi \approx N_2\mu N_1\varphi.$

Thus, $[E_{II}^{(1)}] \approx 0$ by Corollary 1. Thus, we need only show that

(3.22) $\qquad \sup_{(x,t)\in R_\delta} |\partial_x^s[E_{II}^{(2)}]| \leq C_{\delta,s,\lambda}h^\lambda.$

For $t = \bar{t} \leq T$ fixed, we let $\chi(x, \bar{t}) \equiv 1$ in a neighborhood of a fixed point $\bar{x}$ for which $u(x, t) \in C^\infty$, and let $\chi(x, \bar{t}) \equiv 0$ for all x such that the symbol of $I - \mu(x, D, \bar{t})$ is not asymptotic to zero. We can arrange things so $R_\delta \subset$ union of all such $\bar{x}, \bar{t}$.

Regarding $\chi(x, \bar{t})$ as a multiplication operator, and hence a Psdop, we consider

(3.23) $\qquad T_{t-\bar{t}}M_2\chi(x, \bar{t})M_1 T_{\bar{t}-t}[E_{II}^{(2)}] = v(t).$

We shall show:

(3.24) $\qquad \|v(t)\|_s \leq C_{s,\lambda}h^\lambda \quad \forall s, \lambda.$

We may then use an elementary compactness argument to obtain (3.22) from (3.24).

Now

$$v(t) = T_{t-\bar{t}} M_2 \chi(x, \bar{t}) M_1 [T_{\bar{t}} - T_{\bar{t}-t} T_t^{h,\rho}]$$

(3.25)

$$\cdot N_2^h [I^h - \mu^h(x, D) I^h] \sigma_1^h(hD) N_1^h \sigma(hD) \varphi,$$

so $v(0) = 0$. We have

(3.26) $\partial v / \partial t - Lv = T_{t-\bar{t}} M_2 \chi M_1 T_{\bar{t}-t} [L - L^{h,\rho}] T_t^{h,\rho} N_2^h [I^h - \mu^h I^h] \sigma_1^h N_1^h \sigma \varphi.$

We shall show

(3.27) $\|\partial v / \partial t - Lv\|_s \leqslant C_{s,\lambda} h^\lambda \quad \forall s, \lambda.$

This will then complete the proof of the main theorem.

We obtain (3.27) from the following three results.

LEMMA 5 (*Discrete Egorov's Theorem*). *Let R_t^h be the solution operator of the scalar lattice hyperbolic equation on $Y_d \times (-\infty, \infty)$*

(3.28) $u_t = i \lambda^h(x, D, h) u,$

where $\lambda(x, \xi, h) \in S_{1,0}^{1;0}$ and has the asymptotic expansion mentioned in (5), Section 2. Define

(3.29) $Q^h(t) = R_t^h Q^h(x, D, h) R_{-t}^h,$

where $q^h(x, \xi, h) \in S_{1,0}^{m_1,m_2}$, and has the above-mentioned asymptotic expansion. Then $Q^h(t)$ differs by a C^∞ family of smoothing operators from a lattice Psdop whose symbol is in $S_{1,0}^{m_1,m_2}$ for all t. Moreover, it can be constructed exactly as in [14] for the usual scalar Psdops on R^d.

COROLLARY 2. *The family of operators*

$$T_t^{h,\rho} N_2^h [I^h - \mu^h(x, D) I^h] \sigma_1^h(hD) N_1^h T_{-t}^{h,\rho}$$

differs by a C^∞ family of smoothing operators from $v^h(x, D, h, t)$, where $v^h(t)$ is a C^∞ family of lattice Psdops with symbol in $S_{1,0}^{0;0}$. Moreover, the projection of the support of these symbols onto x space is disjoint from that of the symbols of the family $T_{t-\bar{t}} M_2 \chi M_1 T_{\bar{t}-t}$.

LEMMA 6. *Let $p(x, \xi) \in S_{1,0}^n$, $q(x, \xi, h) \in S_{1,0}^{n_1,n_2}$. Let $m = (m_1, m_2, \ldots, m_d)$ with each m_j an integer and $m \neq (0, 0, \ldots)$. Suppose $p(x, \xi + 2\pi m/h)$ and $q(x, \xi, h)$ have disjoint supports for all such m. Then, for any u such that $(F(\partial_x^\alpha u))(\eta) = O(h^\lambda) \forall \lambda, \alpha$ if $|\eta| > (\pi + \epsilon)/h \forall \epsilon > 0$, we have:*

(3.30) $\|P(x, D)[Q(x, D, h) - Q^h(x, D, h)] u\|_s \leqslant C_{s,\lambda} h^\lambda \forall s, \lambda.$

(This lemma is more general than the one we need for the proof of the main theorem.)

The proof of these results will be given in the Appendix.

Proof of estimate (3.27). Using Corollary 2, we may write

(3.31) $\partial u / \partial t - Lv \approx T_{t-\bar{t}} M_2 \chi M_1 T_{\bar{t}-t} (L - L^{h,\rho}) v^h(t) T_t^{h,\rho} \sigma \varphi.$

We shall show below that $\forall \alpha$

(3.32) $(\mathsf{F}(\partial_x^\alpha (T_t^{h,\rho} \sigma(hD)\varphi)))(\eta) = O(h^\lambda) \; \forall \lambda \quad \text{if } |\eta| > (\pi + \epsilon)/h \; \forall \epsilon > 0.$

Thus, we can invoke Lemma 6 and Corollary 2 to show

(3.33) $\partial u / \partial t - Lv = T_{t-\bar{t}} M_2 \chi M_1 T_{\bar{t}-t} L(1 - \rho) v(x_1, D, h, t) T_t^{h,\rho} \sigma^h(hD)\varphi.$

Finally, we again notice that the supports of $T_{t-\bar{t}} M_2 \chi M_1 T_{\bar{t}-t}$ and $v(x, D, h, t)$ are disjoint. The result is now immediate.

Proof of (3.32).

(3.34) $$\partial_x^\alpha T_t^{h,\rho} \sigma(hD) u = \sum_{j=-n}^{n} u_j^{(\alpha)} e^{ij \cdot x}$$

with $|u_j^{(\alpha)}| \leqslant C(1 + |j|)^{\beta + |\alpha|}$ for some fixed β and C as $h \downarrow 0$.

Using a partition of unity lets us consider

$$\left(\frac{1}{2\pi} \right)^d \int_{\Omega_d} e^{-i\eta \cdot x + ij \cdot x} \psi(x) \, dx,$$

where ψ is a C^∞ periodic function. The result is immediate from the integration by parts procedure of the method of stationary phase.

4. Numerical Experiments. In this section we discuss the results of a variety of numerical experiments designed to demonstrate the above theory. The principal aim of the theoretical treatment was to show the utility, and indeed in most cases the necessity, of using smoothing techniques when the Fourier method is applied to problems having discontinuous initial data. We will present numerical results obtained both with, and without, smoothing for the following Cauchy problems: (i) scalar, constant coefficient, (ii) constant coefficient system of two equations in a single space variable, (iii) scalar, variable coefficient in one space dimension, and (iv) scalar variable coefficient in two space dimensions. In addition, we will provide a short discussion on the computer implementation of smoothing techniques.

In all problems except the two space dimension problem, the initial function is

(4.1) $$\Phi(x) = \begin{cases} 1, & x \leqslant 0, \\ 0, & x > 0; \end{cases}$$

calculations are for the interval $[-3.2, 3.2]$ (approximately $[-\pi, \pi]$) in these cases.

We note that in all the following tables the numbers are written as exponentials, e.g. $3.8894 - 3 = 3.8894 \times 10^{-3}$.

Tables 1 and 2 show the results of solving $u_t = u_x$ using the scheme

(4.2) $$u^h(t + k) = F^{-1}(e^{i\omega k} F u^h(t)),$$

where F is the discrete Fourier transform (implemented as a fast Fourier transform), $\omega = \pi j / L$, $|j| \leqslant N$, with L being the half interval length, and $2N + 1$ is the number of lattice points. In the smoothed case $u^h(0)$ is the truncated Fourier series of $\Phi(x)$.

Table 2 requires some explanation. The first three columns of this table show the anomalously accurate results which occur due to collocation effects, as discussed

TABLE 1

Smoothed initial data; $k/h = 0.75$, $t = 0.3$

x	ERROR			
	k = 0.075	k = 0.0375	k = 0.01875	k = 0.009375
-0.5	-4.9890 - 2	3.8894 - 3	-7.0314 - 6	3.6030 - 6
-0.4	5.3010 - 2	-4.8936 - 2	3.6401 - 3	-5.5788 - 6
-0.3	5.0000 - 1	5.0000 - 1	5.0000 - 1	5.0000 - 1
-0.2	-5.3010 - 2	4.8936 - 2	-3.6407 - 3	4.8070 - 6
-0.1	4.9890 - 2	-3.8890 - 3	7.2624 - 6	-3.4742 - 6
0.0	2.8426 - 3	-1.0497 - 5	-3.1571 - 6	-3.1001 - 6
0.1	-4.3736 - 3	1.2167 - 5	-2.0758 - 6	-1.8599 - 6
0.2	7.6252 - 4	1.4673 - 6	-1.3210 - 6	-1.2600 - 6
0.3	6.1409 - 5	1.9813 - 6	-4.3396 - 7	-6.1489 - 7
0.4	-6.5268 - 5	2.6122 - 6	1.0054 - 7	-3.3954 - 7
0.5	2.6429 - 5	3.6611 - 6	9.3540 - 7	2.5505 - 7

in Section 1. Notice that the accuracy is independent of k/h for the chosen sequence of k's until $k/h \neq$ integer. Then the unsmoothed case shows an extreme decrease in accuracy. Only the last column of Table 2 should be compared with Table 1, since these are the only unsmoothed results free of the collocation effects. A direct comparison cannot be made; but in terms of time step sizes, the last column of Table 2 lies between the third and fourth columns of Table 1. The increased accuracy due to smoothing is obvious. One other feature to be noticed in Table 1 is the localization of the region of high error in the smoothed case. In the last column, only the point of discontinuity, itself, exhibits a significant error.

We note that (4.2) is an infinite order, unconditionally stable scheme. It is generally not possible to develop such methods for variable coefficient problems. However, high-order, unconditionally stable schemes can be constructed in such cases by replacing $e^{i\omega k}$ with an appropriate Padé approximant.

We next consider the constant coefficient system given earlier in the introduction:

$$(4.3) \qquad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_x + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

with initial data

$$(4.4) \qquad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_{t=0} = \begin{pmatrix} \Phi(x) \\ 0 \end{pmatrix}.$$

TABLE 2

Unsmoothed initial data; $h = 0.025$, $t = 0.3$

	ERROR			
	$k = 0.1$	$k = 0.05$	$k = 0.025$	$k = 0.0125$
-0.5	1.1009 - 8	1.1009 - 8	1.1009 - 8	7.8125 - 3
-0.4	-4.3282 - 7	-4.3282 - 7	-4.3281 - 7	7.8121 - 3
-0.3	-1.0124 - 7	-1.0124 - 7	-1.0124 - 7	7.8124 - 3
-0.2	8.8447 - 8	8.8447 - 8	8.8447 - 8	7.8126 - 3
-0.1	3.3105 - 7	3.3105 - 7	3.3105 - 7	7.8128 - 3
0.0	-2.4569 - 7	-2.4569 - 7	-2.4569 - 7	7.8123 - 3
0.1	4.5625 - 8	4.5625 - 8	4.5625 - 8	7.8125 - 3
0.2	-9.1595 - 8	-9.1595 - 8	-9.1595 - 8	7.8124 - 3
0.3	-2.3477 - 8	-2.3478 - 8	-2.3478 - 8	7.8125 - 3
0.4	-5.5973 - 8	-5.5973 - 8	-5.5974 - 8	7.8124 - 3
0.5	-3.5402 - 13	-5.9459 - 13	-9.8428 - 13	7.8125 - 3

TABLE 3

Exact solution to Eq. (4.3) at $t = 0.3$

x	u_1	u_2
-0.5	9.553365 - 1	-2.955202 - 1
-0.4	9.553365 - 1	-2.955202 - .1
-0.3	9.553365 - 1	-2.955202 - 1
-0.2	-3.095798 - 2	-2.458528 - 1
-0.1	-1.980069 - 2	-1.966826 - 1
0.0	-1.114488 - 2	-1.477601 - 1
0.1	-4.962601 - 3	-9.883758 - 2
0.2	-1.244540 - 3	-4.966739 - 2
0.3	0.0	0.0
0.4	0.0	0.0
0.5	0.0	0.0

TABLE 4

Error in solution of (4.3) *with smoothed initial data* $k/h = 0.75$, $t = 0.3$

x	ERROR			
	u_1		u_2	
	k = 0.0375	k = 0.01875	k = 0.0375	k = 0.01875
-0.5	3.8858 - 3	-7.0348 - 6	-1.0983 - 5	1.6484 - 8
-0.4	-4.8791 - 2	3.6388 - 3	4.8614 - 4	-3.8571 - 6
-0.3	4.9847 - 1	4.9924 - 1	-5.1103 - 3	-2.5356 - 3
-0.2	4.9082 - 2	-3.6418 - 3	4.8625 - 4	-3.8439 - 6
-0.1	-3.8922 - 3	7.2428 - 6	-1.0908 - 5	-2.9210 - 8
0.0	-1.1217 - 5	-3.1485 - 6	1.3332 - 7	1.7460 - 7
0.1	1.2318 - 5	-2.0788 - 6	1.0895 - 6	9.7857 - 8
0.2	1.6040 - 6	-1.2606 - 6	-4.8618 - 4	3.9963 - 6
0.3	2.1666 - 6	-4.1114 - 7	5.1103 - 3	2.5357 - 3
0.4	2.3108 - 6	3.7681 - 8	-4.8603 - 4	4.0194 - 6
0.5	3.3841 - 6	9.3247 - 7	1.0986 - 5	2.7831 - 8

We express (4.3) in matrix form:

$$U_t = AU_x + BU;$$

then the Fourier method takes the form

(4.5) $$U^h(t + k) = F^{-1}(e^{(i\omega A + B)k}FU^h(t)).$$

Again, for the smoothed case $U^h(0)$ is the truncated Fourier series of (4.4).

The exact solution to (4.3) with initial data (4.4) is given by Apelkrans [1]. Table 3 contains the exact values needed for the comparisons we will make here.

Table 4 gives the results obtained with smoothed initial data. These contrast sharply with the unsmoothed case shown in Table 5, both in terms of absolute accuracy, and the convergence rate. Table 6 provides data computed using the incorrect smoothing technique discussed in Section 2. Notice that outside the characteristics, this method of smoothing does, in fact, give good results; but inside the region R_δ the results are only slightly better than for the unsmoothed case (for the second component they are actually worse than the unsmoothed case). Thus, we see that smoothing is a very effective technique for improving the accuracy of the solution, but only when applied properly.

We now discuss the variable coefficient problem results. We note at the outset that the implementation of smoothing in this case differs somewhat from the constant coefficient cases. In particular, a frequency cutoff function $\rho(\omega h)$ must be applied at

every time step, rather than just to the initial data, since the variable coefficients may cause a regeneration of the undesirable high frequencies; and these must be eliminated before the succeeding time step.

We first consider the problem

(4.6) $$u_t + (a(x)u)_x = 0,$$

with initial data (4.1). For the calculations reported here

$$a(x) = \sin \frac{2\pi}{L} x.$$

We notice that $a(0) = 0$, and $a(\pm L/2) = 0$; and in light of discussions in [2] and [5], this might be expected to lead to numerical instability of the Fourier method, unless some precaution is taken. In [5] it is shown, for the semidiscrete problem, that stability can be maintained by rearranging (4.6) to a specific form before spatial discretization. But the required form leads to a significant increase in the numerical operation count, and so is not a very effective remedy. However, we have shown in Theorem 1 that the stability problem can be easily overcome by applying smoothing operators. We remark that the technique requires only a small amount of extra computation, the amount being dependent on the particular time discretization employed. (For the method we give below, no additional computation is required after the second time step, since the smoothed Fourier frequencies can be stored in place of the original ones at this time, and used thereafter.)

TABLE 5

Error in solution of (4.3) with unsmoothed initial data; $k/h = 0.75$, $t = 0.3$

x	ERROR			
	u_1		u_2	
	k = 0.0375	k = 0.01875	k = 0.0375	k = 0.01875
-0.5	7.4078 - 3	3.9030 - 3	1.4928 - 4	5.5809 - 5
-0.4	7.4677 - 3	3.9317 - 3	3.3999 - 4	1.1686 - 4
-0.3	7.7710 - 3	4.1147 - 3	1.3470 - 3	7.2482 - 4
-0.2	1.0263 - 2	5.4049 - 3	1.1446 - 2	5.9699 - 3
-0.1	4.7501 - 3	5.1174 - 3	1.1708 - 2	6.0059 - 3
0.0	9.1654 - 3	4.8146 - 3	1.1750 - 2	6.0060 - 3
0.1	8.5679 - 3	4.5112 - 3	1.1753 - 2	6.0107 - 3
0.2	7.9631 - 3	4.2055 - 3	1.1674 - 2	6.0005 - 3
0.3	7.3550 - 3	3.8984 - 3	1.0823 - 2	5.4700 - 3
0.4	7.3045 - 3	3.8852 - 3	5.9360 - 4	1.5328 - 4
0.5	7.3078 - 3	3.8865 - 3	2.1888 - 4	6.4966 - 5

TABLE 6
*Error in solution of (4.3) with "incorrectly" smoothed
initial data; $k/h = 0.75$, $t = 0.3$*

x	ERROR			
	u_1		u_2	
	k = 0.0375	k 0.01875	k = 0.0375	k = 0.01875
-0.5	-1.0502 - 4	1.4224 - 5	1.9388 - 5	-4.8778 - 9
-0.4	-2.2265 - 2	-2.8466 - 4	-4.3262 - 5	9.6793 - 6
-0.3	2.2479 - 1	2.2264 - 1	-4.3989 - 5	-4.2684 - 6
-0.2	5.7923 - 2	-6.1195 - 3	1.3536 - 2	6.1389 - 3
-0.1	-5.2143 - 3	1.2876 - 3	1.2179 - 2	6.1269 - 3
0.0	2.3544 - 3	9.4844 - 4	1.2220 - 2	6.1103 - 3
0.1	1.3255 - 3	6.3860 - 4	1.2227 - 2	6.1244 - 3
0.2	6.8014 - 4	3.2921 - 4	1.2379 - 2	6.1600 - 3
0.3	9.9985 - 5	2.7527 - 5	1.2506 - 2	6.2448 - 3
0.4	-3.5457 - 8	4.5456 - 6	-1.1756 - 3	3.5951 - 5
0.5	8.6983 - 6	4.4958 - 6	7.7666 - 5	-5.2321 - 8

Table 7a, b presents results obtained for the smoothed and unsmoothed cases, respectively. The time discretization used is a two-level, second order scheme, analogous to Heun's method for ordinary differential equations, due to Gottlieb and Turkel [16]. It is given by the following. Let

$$f_x(u^h(t)) = -F^{-1}\left[\rho(\omega h)\left(\frac{e^{i\mu\omega k}-1}{\mu}\right)F(au^h(t))\right],$$

$$f_x(u^h(t+k^*)) = -F^{-1}\left[\rho(\omega h)\left(\frac{1-e^{-i\mu\omega k}}{\mu}\right)F(au^h(t+k^*))\right],$$

where

$$\mu > \max_{x\in[-L,L]}|a(x)|;$$

and

$$u^h(t+k^*) = u^h(t) + kf_x(u^h(t)).$$

Then

(4.7) $$u^h(t+k) = u^h(t) + \frac{k}{2}[f_x(u^h(t)) + f_x(u^h(t+k^*))].$$

<div align="center">

TABLE 7

Solutions to (4.7) at $t = 0.1$ ($h = 0.025$)

</div>

x	SOLUTIONS					
	(a) Smoothed			(b) Unsmoothed		
	k = 0.05	k = 0.025	k = 0.0125	k = 0.05	k = 0.025	k = 0.0125
-1.0	1.061460-0	1.060923-0	1.060796-0	1.060407-0	1.059749-0	1.059687-0
-0.8	9.808670-1	9.809524-1	9.809878-1	9.809856-1	9.809232-1	9.809962-1
-0.6	9.108518-1	9.113701-1	9.114987-1	9.127343-1	9.130233-1	9.132257-1
-0.4	8.576136-1	8.580321-1	8.581253-1	8.624822-1	8.624770-1	8.627549-1
-0.2	8.202939-1	8.203211-1	8.203238-1	8.326650-1	8.315115-1	8.322071-1
0.0	4.137983-1	4.116699-1	4.110694-1	-8.435581-2	8.695430-2	-8.737234-2
0.2	1.175060-2	1.170288-2	1.169012-2	-5.747359-4	2.510335-4	3.065230-5
0.4	4.604301-3	4.604535-3	4.604539-3	-2.397436-4	3.500402-5	3.786343-5
0.6	1.732821-3	1.732797-3	1.732781-3	-1.311246-4	6.779353-6	2.343124-5
0.8	2.651869-5	2.649095-5	2.647747-5	-7.835093-5	2.024904-6	-4.192288-6
1.0	-1.112882-3	-1.112900-3	-1.112912-3	-5.005815-5	6.237049-6	-4.613823-6

As usual, for the smoothed case, we use $u^h(0)$ equal to the truncated Fourier series of $\Phi(x)$. Results obtained using (4.7) both with, and without, smoothing are given in Table 7. The main point to observe regarding Table 7 is the degree of consistency of the convergence rate, from point to point, exhibited by the results obtained with smoothing, and the corresponding lack of such consistency for the results computed without smoothing. In particular, nearly all of the points computed by the smoothed method show roughly a second order convergence rate, while the unsmoothed method produces results which do not appear to converge at all, at certain grid points. Although this erratic behavior (which continued, even for further mesh refinements) of the unsmoothed calculations would seem to be indicative of possible instabilities, this was not borne out by long time integrations. It should also be noted that with further mesh refinement, second order convergence was not maintained, even for the smoothed case near the point of discontinuity.

As a final example, we show that the theoretical results of the earlier sections also apply in higher space dimensions by considering a second order splitting method for the two space variable problem

$$(4.8) \qquad u_t = (a(x, y)u)_x + (b(x, y)u)_y \qquad \text{on } [-1.6, 1.6] \times [-1.6, 1.6],$$

where

$$a(x, y) = -\sin \frac{2\pi}{L}x \sin \frac{2\pi}{L}y \qquad \text{and} \qquad b(x, y) = -e^{\gamma(x^2+y^2)},$$

with $\gamma = -7.5$ for the results displayed in Table 8. For initial data we use the checkerboard pattern, shown in Figure 1, and given by

$$(4.9) \qquad \Psi(x, y) = \Phi(x)\Phi(y) + (1 - \Phi(x))(1 - \Phi(y)).$$

TABLE 8

Solutions to (4.8) at t = 0.1 (h = 0.05)

x	SOLUTIONS					
	k = 0.05		k = 0.025		k = 0.0125	
	y = -0.3	y = 0.3	y = -0.3	y = 0.3	y = -0.3	y = 0.3
-0.5	9.699919-1	2.362806-3	9.704543-1	2.774479-3	9.706911-1	2.999335-3
-0.4	9.734857-1	5.895654-3	9.745015-1	6.720165-3	9.750242-1	7.159855-3
-0.3	9.721103-1	1.188208-2	9.734723-1	1.287873-2	9.741895-1	1.340468-2
-0.2	9.644197-1	2.393817-2	9.663444-1	2.536328-2	9.673653-1	2.601167-2
-0.1	9.352408-1	4.765305-2	9.387540-1	5.037910-2	9.405873-1	5.176501-2
0.0	4.947608-1	5.063481-1	4.946418-1	5.063733-1	4.945939-1	5.063746-1
0.1	5.471994-2	9.641620-1	5.103118-2	9.615326-1	4.913242-2	9.601736-1
0.2	2.580417-2	9.865284-1	2.384192-2	9.853471-1	2.282085-2	9.847137-1
0.3	1.601895-2	9.995212-1	1.472681-2	9.987544-1	1.405752-2	9.983327-1
0.4	8.857457-3	1.010681-0	7.930383-3	1.010063-0	7.460291-3	1.009720-0
0.5	3.611647-3	1.023265-0	3.197425-3	1.022997-0	2.990576-3	1.022839-0

Splitting of (4.8) was done in the following way, to maintain second order accuracy (cf. Marchuk [8]). Denote the right side of (4.7) by $L_a(k, x)u^h(t)$. We then compute $u^h(t + k)$ from

$$(4.10) \qquad u^h(t + k) = L_a(k/2, x) \circ L_b(k, y) \circ L_a(k/2, x)u^h(t).$$

For the particular problem considered, (4.10) was stable for both smoothed and un-smoothed cases; the former is shown in Table 8.

To this point, we have treated the smoothing procedure theoretically; and we have also shown numerical results of applying it to specific examples. We now give a brief discussion concerning implementational details. The basic idea is, of course, elimination of the high Fourier frequencies which arise from discontinuous data. In the past, this has been accomplished by merely setting to zero all of the frequency spectrum beyond some prescribed magnitude. With such a procedure, the frequency cutoff function is, itself discontinuous, making precise mathematical treatment difficult, as one might infer from the nature of the various proofs given above. Moreover, as is proven in Section 1, and demonstrated by the numerical solution of (4.3) (see Table 6), simply cutting off frequencies, even in a smooth manner, is not sufficient for maintaining the order of accuracy of the numerical method on the whole grid, except in special cases (see Proposition 4). In general, two smoothing operations must be utilized: one for the initial data, and one at subsequent time steps. Before treating each of these, we make the following general remarks.

It seems likely that the precise behavior of an optimal smoothing procedure must be very problem-dependent. Thus, in practice, we should be content to imple-
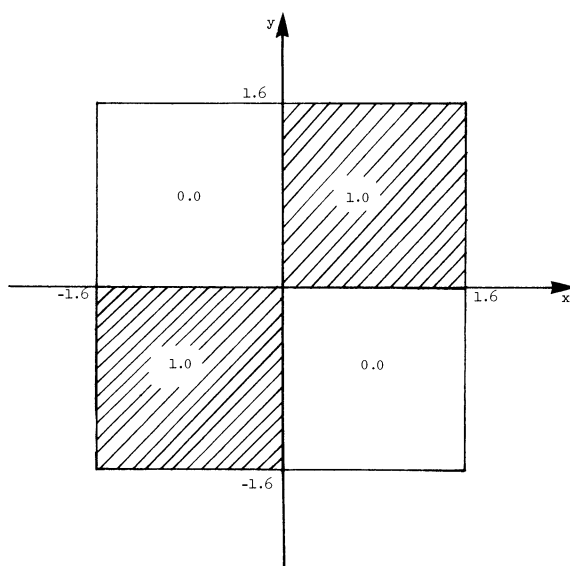
FIGURE 1

Initial data for the two space dimension problem

ment a procedure which works well, but which may be slightly suboptimal with respect to accuracy, and possibly convergence rate. The following would appear to be requisite qualities of such a smoothing method: the method should

(1) produce results in substantial agreement (with respect to theoretical accuracy and stability) with theoretical predictions;

(2) work well over a wide range of problems;

(3) be easily programmed, and inexpensive to compute.

Smoothing of initial data is accomplished by first replacing the original (discontinuous) data by its Fourier series, truncated after a certain number of terms. We emphasize that the true Fourier series, and *not* the discrete Fourier transform, must be used at this step, as indicated by Proposition 3. The number of terms to be used in the truncated Fourier series is related to the mesh size, h, of the spatial discretization in the following way. We associate this initial smoothing with the smoothing operator σ, discussed above, and identify ρ with smoothing at subsequent time steps. We will show below how ρ is related to mesh size through the discrete Fourier frequencies. But in the proof of Theorem 2 we must have supp $\sigma \subset$ supp ρ. Thus, the number of terms carried in the truncated Fourier series of the initial data must be less than the number of frequencies arising from the use of the discrete Fourier transform in the spatial discretization. All results presented herein were obtained using $2N$ terms, where $2N + 1$ is the number of points in the spatial grid (= number of discrete Fourier frequencies).

In Figure 2 we show a sketch of the initial smoothing of (4.1). The main purpose of this figure is to provide a clear indication of the cause of the large error exactly at the point(s) of discontinuity of the smoothed solutions.
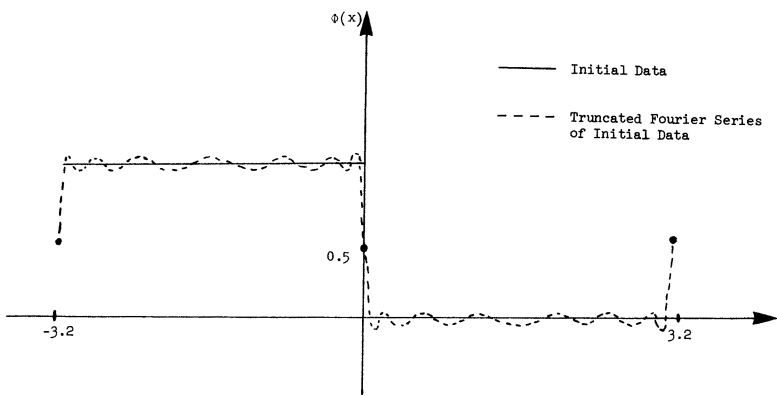
FIGURE 2

Truncated Fourier series representation of initial data

The difference between the Fourier series and the actual initial data at the points of discontinuity is, in this case, exactly 0.5; and this difference propagates in time. But as the tabulated results show, when smoothing is used the error is localized to a small neighborhood of the discontinuity; whereas, with no smoothing, the errors may be quite large over the entire grid (e.g., compare Tables 4 and 5).

In the second step in smoothing, we employed partitions of unity to obtain the theoretical results. Thus, frequencies are multiplied by functions of the type depicted in Figure 3, rather than brutally chopped off. It is easy to construct such functions, analytically, by using tails of the general form

$$\exp\left[\frac{-1}{\xi - \xi_{\min}} \exp\left(\frac{1}{\xi - \xi_0}\right)\right], \quad \xi_{\min} \leqslant \xi \leqslant \xi_0;$$

but computationally, this can lead to difficulties. If the tails of the cutoff function drop off too sharply, the computer is unable to distinguish such a C^∞-function from a discontinuous one. Thus, a more controllable characteristic decay is necessary. For this reason we chose to use "generalized" Gaussians, centered at $\pm \omega_0$, to form the tails of the cutoff functions used in the numerical studies. That is, the smoothing functions were of the form

$$\rho(\omega) = \begin{cases} e^{-\alpha(\omega + \omega_0)^{2m}}, & \omega < -\omega_0, \\ 1, & -\omega_0 \leqslant \omega \leqslant \omega_0, \\ e^{-\alpha(\omega - \omega_0)^{2m}}, & \omega_0 < \omega, \end{cases}$$

where m is a positive integer. In the work described here $m = 1, 2,$ or 3; and we shall refer to the corresponding functions ρ as the 2nd, 4th, or 6th-order cutoff func-

tions, respectively. It is clear that these functions are not C^∞; moreover, as the order of continuous differentiability is increased, the tail decays to zero more slowly for a fixed value of α. It was found that the 4th-order cutoff function gave the best results overall, although they were not significantly different from those obtained for 2nd and 6th-order cutoff functions. To obtain the most accurate results near discontinuities it was found that $\omega_0 = 0$ (i.e., an actual Gaussian) should be used. However, higher accuracy could be obtained away from discontinuities with $\omega_0 \cong 0.3\omega_{max}$ for the problems tested. Here $\omega_{max} = L/h$. The value of α was chosen so that $\rho(\eta\omega_{max}) = 10^{-4}$, with η in the range 0.5 to 0.9.

These are not claimed to be optimal parameter values for the particular cutoff functions used, nor is the form of the cutoff function necessarily optimal. But the numerical results of the preceding tables indicate excellent agreement with theoretical predictions; the same form of smoothing was used in all problems; and the method is easily implemented, and inexpensively computed. Thus, all criteria proposed earlier have been satisfied.
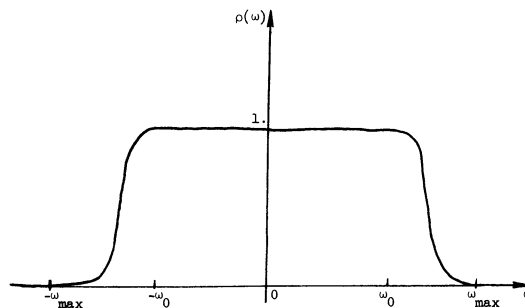

FIGURE 3

Typical C_0^∞-frequency cutoff function

Appendix. We begin by verifying the properties 1 to 5 of the discrete pseudo-differential calculus of Section 2.

If $u \in C^\infty$, it is easy to show that $\forall \lambda > 0$ $|\widetilde{u}(j)| \leqslant C_\lambda (1 + |j|)^{-\lambda}$. Thus

$$(A.1) \qquad |D_x^\alpha P^h(x, D, h)u| \leqslant C_{\alpha,\lambda} h^{m_2} \sum_{jh \in Y_d} (1 + |j|)^{-\lambda + |m_1|};$$

and property (1) is immediate if we take λ sufficiently large.

We also have $Q^h(x, D)P^h(x, D) = R^h(x, D)$ for

$$(A.2) \qquad R^h(x, D)u = \sum_{j=-N}^{N} e^{ij \cdot x} r(x, j)\widetilde{u}(j)$$

(supressing the h dependence in thisargument), where

$$(A.3) \qquad r^h(x, \eta) = \left(\frac{h}{2\pi}\right)^d \sum_{z \in Y_d} \sum_{\xi h \in Y_d} e^{i(x-z)\cdot(\xi-\eta)} q(x, \xi) p(z, \eta).$$

By means of a partition of unity we may write $p(z, \eta) = p_1(z, \eta) + p_2(z, \eta)$, where $p_1 \equiv 0$ near $z_i = \pm \pi \ \forall i$, $p_2(z, \eta) \equiv 0$ near $z_i = 0 \ \forall i$, and p_1 and p_2 are still each in $S_{1,0}^{m_1, m_2}$. Moreover, by periodicity in z and the fact that we are working on a lattice $z, x, \xi h, \eta h \in Y_d$, we can change the region of summation for the sum involving p_2 to

$$
(A.4) \qquad r^h(x, \eta) = \left(\frac{h}{2\pi}\right)^d \sum_{\substack{z_i - Nh \in Y_d \\ i=1,\ldots,d}} \sum_{\xi h \in Y_d} e^{i(x-z)\cdot(\xi-\eta)} q(x, \xi) p_2(z, \eta).
$$

Thus, without loss of generality we may assume that in what follows, $p(z, \eta) \equiv 0$ near each $z_i = \pm \pi$.

By the Poisson sum formula we have

$$
(A.5) \quad r^h(x, \eta) = \left(\frac{1}{2\pi}\right)^d \sum_{\xi h \in Y_d} \sum_{m=-\infty}^{\infty} \int_{-\pi}^{\pi} dz \cdot e^{i(x-z)\cdot(\xi-\eta)+2\pi i m \cdot z} q(x, \xi) p(z, \eta).
$$

For $m \neq 0$ the z gradient of the phase functions in the above integrands vanishes if and only if $\eta_i - \xi_i = -2\pi m_i/h$. This is impossible if any $|m_i| > 1$. However, if $|m_i| \leqslant 1$ for all i, Eq. (2.9) shows us that the contribution of these integrals for $\xi_i h$ near $\pm \pi \approx 0$.

Thus, we may integrate by parts on all the $m \neq 0$ integrals, sum over m, and get a contribution ≈ 0.

The $m = 0$ integral is treated by applying the Poisson sum formula in ξ, obtaining

$$
(A.6) \quad r^h(x, \eta) \approx \left(\frac{1}{2\pi}\right)^d \sum_{m=-\infty}^{\infty} \int d\xi \int_{-\pi}^{\pi} dz \, e^{i(x-z)\cdot(\xi-\eta)+2\pi i m \cdot \xi} q(x, \xi) p(z, \eta).
$$

We note that the ξ gradient of the phase function vanishes when $z - x = 2\pi m$; hence, only the $m = 0$ integral is critical.

Now we write

$$
(A.7) \qquad q(x, \xi) = q(x, \eta) + \sum_{|\alpha|=0}^{|\alpha|=k-1} \frac{1}{\alpha!} (\partial_\eta^\alpha q(x, \eta))(\xi - \eta)^\alpha
$$

$$
+ \frac{1}{k!} \int_0^1 (1 - t)^k \left(\frac{\partial}{\partial t}\right)^k q(x, \xi t + \eta(1 - t)) \, dt.
$$

For the $m \neq 0$ integrals we need only consider, modulo harmless constants,

$$
(A.8) \qquad \int_{-\infty}^{\infty} d\xi \int_{-\pi}^{\pi} \sum_{|\alpha|=k} (\xi - \eta)^\alpha e^{i(x-z)\cdot(\xi-\eta)+2\pi i m \cdot \xi}
$$

$$
\cdot \int_0^1 (1 - t)^k \partial_\xi^\alpha q(x, \xi) \Big|_{\xi=\xi t + \eta(1-t)} p(z, \eta) \, dt.
$$

First integrate by parts in each z_i a total of α_i times. Next integrate by parts a

large number of times in ξ, each time picking up a factor $(x - z + 2\pi m_i)^{-1}$. Sum over $m \neq 0$. It is clear that we can take k large enough so that we obtain a converging series of operators in $S^{\overline{m}, m'_2 + m'_2}$ for any $\overline{m}$.

For $m = 0$, we are concerned with

(A.9)
$$\left(\frac{1}{2\pi}\right)^d \int_{-\infty}^{\infty} d\xi \int_{-\pi}^{\pi} dz e^{i(x-z)\cdot(\xi-\eta)} q(x, \xi) p(z, \eta).$$

We now use (A.7). The composition formula of Leibniz follows in the standard fashion.

The result if q is in the subclass $T^{m'_1, m'_2}_{1,0}$ follows as above, using a slightly refined analysis of the remainder in the Taylor series expansion (A.7).

The only nontrivial thing left to prove is the adjoint formula. It is easy to show that

(A.10) $\quad (P^h(x, D, h))^* u = \left(\frac{h}{2\pi}\right)^d \sum_{jh \in Y_d} \sum_{x_0 \in Y_d} \overline{p(x_0, j, h)}^t e^{ij \cdot (x - x_0)} u(x_0).$

This result now follows in an analogous fashion to the previous one, i.e. only one critical integral arises when we Poisson sum in j.

We now prove the remark corresponding to (3.3). We have

$$P^L(x, D, h)\varphi = \left(\frac{1}{2\pi}\right)^d \sum_{\xi=-\infty}^{\infty} \int_{-\pi}^{\pi} e^{i\xi \cdot (z-z_0)} p(x, \xi, h)\varphi(x_0)\, dx_0$$

$$= \left(\frac{1}{2\pi}\right)^d \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \int_{-\pi}^{\pi} dx_0 \int_{-\infty}^{\infty} d\xi e^{i\xi \cdot (x - x_0) + 2\pi m i \cdot \xi} p(x, \xi, h)\varphi(x_0)$$

$$+ P(x, D, h)\varphi$$

by the Poisson summation formula.

For each integral corresponding to $m \neq 0$, we may integrate by parts arbitrarily often in ξ, and because $p(x, \xi, h) \equiv 0$ near $x_i = \pm \pi$, none of these integrals are critical. We then add over $m \neq 0$. The result follows.

Proof of Lemma 5 (Discrete Egorov's Theorem). Our proof is just a copy of Taylor's [14] in the continuous case with a few simple modifications. However, it is so fundamental to this work that we include it for completeness.

We note that $Q^h(t)$ solves the operator differential equation

(A.12)
$$\partial Q^h(t)/\partial t = i[\lambda^h(x, D, h), Q^h],$$

$$Q^h(0) = Q^h(x, D, h).$$

We are looking for a solution to

(A.13)
$$dN^h/dt = i[\lambda^h(x, D, h), N^h] + S^h_{-\infty}(t),$$

$$N^h(0) = Q^h(x, D, h),$$

where, here and below, $S^h_{-\infty}(t)$ is a C^∞ family of smoothing operators on Y_d.

We will obtain a lattice Psdop solution $N^h(t)$ with symbol in $S^{m_1, m_2}_{1,0}$, and then show $N^h(t) - Q^h(t)$ is a smoothing operator ≈ 0.

We shall assume $m_1 = 0 = m_2$. The general case presents no difficulties.

Following [14] we write:

$$n(x, \xi, h, t) \sim n_0(x, \xi, h, t) + n_1(x, \xi, h, t) + \cdots,$$

(A.14)

$$\lambda(x, \xi, h, t) \sim \lambda_1(x, \xi, h, t) + \lambda_0(x, \xi, h, t) + \cdots.$$

The symbol of $i[\lambda^h, N^h]$ is asymptotic to

(A.15) $$H_{\lambda_1} n + \{\widetilde{\lambda}_0, n\} + \sum_{|\alpha| \geqslant 2} \frac{i^{|\alpha|-1}}{\alpha!} (\lambda^{(\alpha)} n_{(\alpha)} - n^{(\alpha)} \lambda_{(\alpha)}),$$

where $\widetilde{\lambda}_0 = \lambda - \lambda_1 = \lambda_0 + \lambda_{-1} + \cdots$, $\{\ ,\ \}$ denotes Poisson brackets, and H_{λ_1} is the Hamiltonian vector field which depends on h

$$H_{\lambda_1} = \sum_{j=1}^{n} \left(\frac{\partial \lambda_1}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial \lambda_1}{\partial x_j} \frac{\partial}{\partial \xi_j} \right).$$

We thus define $n_0(x, \xi, h, t)$ by the equation

(A.16) $$(\partial/\partial t - H_{\lambda_1}) n_0(x, \xi, h, t) = 0, \qquad n_0(x, \xi, h, 0) = q_0(x, \xi, h).$$

We can solve this for all t. n_0 is constant on the integral curves of $\partial/\partial t - H_{\lambda_1}$.

We determine the other $n_j(x, \xi, h, t)$, recursively, as follows. If $N^h_0(t) = N^h_0(x, D, h, t)$, then from (A.15) we have

$$\text{symbol of } i[\lambda, n^h_0(t)] - \frac{\partial n^h_0(t)}{\partial t} \sim i\{\widetilde{\lambda}_0, n_0\}$$

(A.17)

$$+ \sum_{|\alpha| \geqslant 2} \frac{i^{|\alpha|-1}}{\alpha!} (\lambda^{(\alpha)} n_{0(\alpha)} - n^{(\alpha)}_0 \lambda_{(\alpha)})$$

$$= a_0(x, \xi, h, t) \in S^{-1,0}_{1,0}.$$

In view of this, we let $n_{-1}(x, \xi, h, t)$ solve

(A.18) $$(\partial/\partial t - H_{\lambda_1}) n_{-1}(x, \xi, h, t) = -a_0(x, \xi, h, t), \quad n_{-1}(x, \xi, h, 0) = 0.$$

It follows that $n_{-1}(x, \xi, h, t) \in S^{-1,0}_{1,0}$. We continue in this fashion obtaining $n_{-j}(x, \xi, h, t) \in S^{-j,0}_{0,0}$, and we have found $n(x, \xi, h, t) \in S^{0,0}_{1,0}$ such that

$$n(x, \xi, h, t) \sim \sum_{j=0}^{+\infty} n_{-j}(x, \xi, h, t).$$

Our assertion that $N^h(t) - Q^h(t)$ is smoothing is equivalent to showing that $\|v(t)\|_s$ is bounded independently of h, where

(A.19) $$v(t) = (R^h_t Q^h(x, D) - N^h(t) R^h_t) u$$

for any u defined on Y_d. Now $n(0) = 0$ and

$$(A.20) \qquad \frac{dv}{dt} = i\lambda^h(x, D, h)v + S^h_{-\infty}(t)v.$$

The result now follows from a simple modification of Theorem 1.

Proof of Corollary 2. As above, we let $v^h(t)$ solve the differential equation

$$(A.21) \qquad \partial v^n(t)/\partial t = i[L^{h,\rho}, v^h] + S^h_{-\infty}(t),$$

$$v^h(0) = N^h_2[I^h - \mu^h I^h]\sigma^h_1 N^h_1.$$

Let $N^h_1 v^h N^h_2 = \gamma^h$. Then γ^h satisfies:

$$(A.22) \qquad \partial\gamma^h(t)/\partial t = i[\Lambda^{h,\rho}, \gamma^h] + S^h_{-\infty}(t),$$

$$\gamma^h(0) \approx \sigma^h_1 - \mu^h \sigma^h_1.$$

Since this problem is essentially decoupled, we can solve it modulo smoothing as in the previous lemma and obtain, again modulo smoothing, $v^h = N^h_2 \gamma^h N^h_1$.

Taylor's construction shows us that the symbol of each component of $\gamma^h(t)$ is constant along the Hamiltonian flow associated with the corresponding component of $\Lambda^{h,\rho}$. But this is the same as that for Λ if at $t = 0$ (x, ξ) is in the set for which $\rho(h\xi) \equiv 1$. The symbol of $\sigma^h_1 - \mu^h \sigma^h_1$ vanishes outside this set. Thus, the symbol of v^h is the same as that of

$$\mu(x, D, t) = T_t M_2 [I - \mu(x, D)I]\sigma_1 M_1 T_{-t}.$$

At $t = \bar{t}$ the symbol of $\mu(x, D, t)$ has projection on x space disjoint from that of $M_2 \chi M_1$. This means, again by Taylor's construction, that these sets stay disjoint for all time.

Proof of Lemma 6. We have

$$(A.23) \qquad P(x, D)Q^h(x, D, h)u = \sum_{\eta=-N}^{N} e^{ix\cdot\eta}\widetilde{u}(\eta)r(x, \eta, h),$$

$$P(x, D)Q(x, D, h)u = \int_{-\pi/h}^{\pi/h} d\eta\, e^{ix\cdot\eta}\hat{u}(\eta)s(x, \eta, h).$$

Here

$$(A.24) \qquad r(x, \eta, h) = \left(\frac{h}{2\pi}\right)^d \int_{-\infty}^{\infty} d\xi \sum_{z\in Y_d} e^{i(x-z)\cdot(\xi-\eta)}p(x, \xi)q(z, \eta, h),$$

$$s(x, \eta, h) = \left(\frac{1}{2\pi}\right)^d \int_{-\infty}^{\infty} d\xi \int_{\Omega_d} dz\, e^{i(x-z)\cdot(\xi-\eta)}p(x, \xi)q(z, \eta, h).$$

We shall show

$$(A.25) \qquad r(x, \eta, h) - s(x, \eta, h) \approx 0.$$

Having done this, it follows from the composition formula that both symbols ≈ 0 for η_i near $\pm\pi \,\forall i$.

By (A.23), (A.25), and the Poisson sum formula, we need only show

$$(A.26) \qquad \int_{-\pi/h}^{\pi/h} \sum_{\substack{m\neq 0 \\ m=-\infty}}^{\infty} \int e^{i(x-x_0)\cdot\eta + 2\pi i m \cdot x_0/h} r(x, \eta, h)u(x_0)\,dx_0\,d\eta = 0.$$

For each m we apply the method of stationary phase in x_0 and we may, when $|m_i| \leqslant 1$, for all i, consider the integrand as vanishing near $\eta_i = \pm\pi$. Then using the fact that $(F(\partial_x^\alpha u))(\eta) = O(h^\lambda)\ \forall\lambda$ if $|\eta| > (\pi + \epsilon)/h\ \forall\epsilon > 0$, we sum over m; we can thus show that the term in (A.26) ≈ 0.

Proof of (A.25). By the Poisson sum formula we have

$$
\begin{aligned}
r - s &= \left(\frac{1}{2\pi}\right)^d \sum_{m\neq 0} \int_{-\infty}^{\infty}d\xi \int_{\Omega_d}dz\, e^{i((x-z)\cdot(\xi-\eta) + 2\pi m\cdot z/h)}p(x, \xi)q(z, \eta, h) \\
(A.27) & \\
&= \sum_{m\neq 0} I_m.
\end{aligned}
$$

We notice that each of these integrals is critical in z because $\xi - \eta = 2\pi m/h$ in this range of integration. (This is the only time in this entire paper that we ever have to deal with an infinite number of critical integrals.)

Take a particular integral and make a simple change of variables, arriving at

$$(A.28)\quad I_m = \left(\frac{1}{2\pi}\right)^d e^{2\pi m i\cdot x/h}\int_{-\infty}^{\infty}d\xi\int_{\Omega_d}dz\, e^{i(x-z)\cdot(\xi-\eta)}\cdot p\left(x, \xi + \frac{2\pi m}{h}\right)q(x, \eta, h).$$

Now we apply Taylor's theorem on $p(x, \xi + 2\pi m/h)$, expanding it at $\xi = \eta$. By the disjoint support hypothesis, we arrive at (modulo harmless constants)

$$
\begin{aligned}
I_m &= \left(\frac{1}{2\pi}\right)^d e^{2\pi m i\cdot x/h}\int_{-\infty}^{\infty}d\xi\int_{\Omega_d}\sigma(h(\xi - \eta))(\xi - \eta)^{\alpha+1}e^{i(x-z)\cdot(\xi-\eta)}dz \\
&\quad \cdot\int_0^1 dt(1-t)^\alpha \partial_\xi^\alpha p\left(x, \xi + \frac{2\pi m}{h}\right)\Big|_{\xi=\eta+t(\xi-\eta)}q(z, \eta) \\
(A.29)\qquad &+ \left(\frac{1}{2\pi}\right)^d e^{2\pi m i\cdot x/h}\int_{-\infty}^{\infty}d\xi\int_{\Omega_d}(I - \sigma(h(\xi -\eta)))(\xi - \eta)^{\alpha+1}e^{(x-z)(\xi-\eta)}dz \\
&\quad \cdot\int_0^1(1 - t)^\alpha\partial_\xi^\alpha p\left(x, \xi + \frac{2\pi}{h}\right)\Big|_{\xi=\eta+t(\xi-\eta)}q(z, \eta) \\
&= I_1^m + I_2^m.
\end{aligned}
$$

For I_1^m we choose $|\alpha|$ sufficiently large so that when we integrate by parts in z we get

$$|\partial_x^\beta D_\eta^\gamma I_1^m| \leqslant C_{\lambda,\beta,\gamma}h^\lambda|m|^{-\lambda}\quad \text{for }\lambda\text{ arbitrary and fixed.}$$

For I_2^m we may integrate by parts in z, dividing by $(\xi_i - \eta_i)$ which is now non-vanishing. Do this sufficiently often, and we wind up with an estimate

$$|\partial_x^\beta D_\eta^\gamma I_2^m| \leqslant C_{\lambda,\beta,\gamma} \int_{|\xi-\eta|\geqslant\delta} {}^{(2)}/h \frac{d\xi}{|\xi-\eta|^\lambda} \left[\left(1+\left|\xi-\frac{2\pi m}{h}\right|\right)^{-\lambda} + \left|\eta-\frac{2\pi m}{h}\right|^{-\lambda}\right]$$

$$\leqslant C'_{\lambda,\beta,\gamma} \frac{h^\lambda}{|m|^\lambda} \quad \text{for } \lambda \text{ arbitrary.}$$

We now sum over m and the result follows.

Department of Mathematics
University of California
Los Angeles, California 90024

1. M. Y. T. APELKRANS, "Some properties of difference schemes with discontinuities and a new method with almost quadratic convergence," Report #15A, Department of Computer Science, Uppsala University, Uppsala, 1969.

2. B. FORNBERG, "On high order approximations of hyperbolic partial differential equations by a Fourier method," Rep. No. 39, Department of Computer Science, Uppsala University, Uppsala, 1972.

3. B. FORNBERG, "On a Fourier method for the integration of hyperbolic equations," SIAM J. Numer. Anal., v. 12, 1975, pp. 509—528.

4. L. HÖRMANDER, "On the existence and the regularity of solutions of linear pseudo-differential equations," Enseignement Math., v. 17, 1971, pp. 99—163.

5. H.-O. KREISS & J. OLIGER, "Comparison of accurate methods for the integration of hyperbolic equations," Tellus, v. 24, 1972, pp. 199—215.

6. H.-O. KREISS & J. OLIGER, "Methods for the approximate solution of time dependent problems," GARP Publications Series no. 10, 1973.

7. A. MAJDA & S. OSHER, "Propagation of error into regions of smoothness for accurate difference approximations to hyperbolic equations," Comm. Pure Appl. Math., v. 30, 1977, pp. 671—705.

8. G. I. MARCHUK, Numerical Methods in Weather Prediction, Academic Press, New York, 1974.

9. S. A. ORSZAG, "Numerical simulation of incompressible flows within simple boundaries, I: Galerkin (spectral) representations," Studies in Appl. Math., v. 50, 1971, pp. 293--327.

10. S. A. ORSZAG, "Comparison of pseudospectral and spectral approximations," Studies in Appl. Math., v. 51, 1972, pp. 253—269.

11. J. SMOLLER & M. TAYLOR, "Wave front sets and the viscosity method," Bull. Amer. Math. Soc., v. 79, 1973, pp. 431—436.

12. M. TAYLOR, Pseudodifferential Operators, Lecture Notes in Math., vol. 410, Springer-Verlag, Berlin, 1974.

13. M. TAYLOR, "Reflection of singularities of solutions to systems of differential equations," Comm. Pure. Appl. Math., v. 28, 1975, pp. 457—478.

14. M. TAYLOR, Pseudodifferential Operators, Vol. 2. (To appear.)

15. H.-O. KREISS & J. OLIGER, "Stability of the Fourier method," Stanford Computer Science Report 77-616, 1977.

16. D. GOTTLIEB & E. TURKEL, "On time discretizations for spectral methods," ICASE Report #78-1, 1978.