Chapter 5.
Dissipation, Dispersion, and Group Velocity

Things fall apart; the center cannot hold;
Mere anarchy is loosed upon the world.
— *W. B. YEATS, The Second Coming (190–)*

It would be a fine thing if discrete models calculated solutions to partial differential equations exactly, but of course they do not. In fact in general they could not, even in principle, since the solution depends on an infinite amount of initial data. Instead, the best we can hope for is that the errors introduced by discretization will be small when those initial data are reasonably well-behaved.

This chapter is devoted to understanding the behavior of numerical errors. From truncation analysis we may have a bound on the magnitude of discretization errors, depending on the step sizes h and k, but much more can be said, for the behavior of discretization errors exhibits great regularity, which can be quantified by the notions of numerical dissipation and dispersion. Rounding errors too, though introduced essentially at random, propagate in the same predictable ways.

So long as we can estimate the magnitude of the discretization and rounding errors, what is the point in trying to investigate their behavior in more detail? There are several answers to this question. One is that it is a good idea to train the eye: a practitioner familiar with artificial numerical effects is less likely to mistake spurious features of a numerical solution for mathematical or physical reality. Another is that in certain situations it may be advantageous to design schemes with special properties—low dissipation, for example, or low dispersion. A third is that in more complicated circumstances, the magnitude of global errors may depend on the behavior of local errors in ways that ordinary analysis of discretization and rounding errors cannot predict. In particular, we shall see in the next chapter that the stability of boundary conditions for hyperbolic partial differential equations depends upon phenomena of numerical dispersion.

One might say that this chapter is built around an irony: finite difference approximations have a more complicated "physics" than the equations they are designed to simulate. The irony is no paradox, however, for finite differences are used not because the numbers they generate have simple properties, but because those numbers are simple to compute.

5.1. Dispersion relations

Any time-dependent scalar, linear partial differential equation with constant coefficients on an unbounded space domain admits plane wave solutions

$$u(x,t) = e^{i(\xi x+\omega t)}, \qquad \xi \in \mathbb{R}, \tag{5.1.1}$$

where ξ is the **wave number** and ω is the **frequency**. (Vector differential equations admit similar modes multiplied by a constant vector; the extension to multiple space dimensions is described at the end of this section.) For each ξ, not all values of ω can be taken in (5.1.1). Instead, the PDE imposes a relationship between ξ and ω,

$$\omega = \omega(\xi), \tag{5.1.2}$$

which is known as the **dispersion relation**, mentioned already in §3.1. In general each wave number ξ corresponds to m frequencies ω, where m is the order of the differential equation with respect to t, and that is why (5.1.2) is called a relation rather than a function. For most purposes it is appropriate to restrict attention to values of ξ that are real, in which case ω may be real or complex, depending on the PDE. The wave (5.1.1) decays as $t \to \infty$ if $\operatorname{Im}\omega > 0$, and grows if $\operatorname{Im}\omega < 0$.

For example, here again are the dispersion relations for the model equations of §3.1, and also for the second-order wave equation:

$$u_t = u_x : \qquad \omega = \xi, \tag{5.1.3}$$

$$u_{tt} = u_{xx} : \qquad \omega^2 = \xi^2, \text{ i.e., } \omega = \pm\xi, \tag{5.1.4}$$

$$u_t = u_{xx} : \qquad i\omega = -\xi^2, \tag{5.1.5}$$

$$u_t = iu_{xx} : \qquad \omega = -\xi^2. \tag{5.1.6}$$

These relations are plotted in Figure 5.1.1. Notice the double-valuedness of the dispersion relation for $u_{tt} = u_{xx}$, and the dashed curve indicating complex values for $u_t = u_{xx}$.

More general solutions to these partial differential equations can be obtained by superimposing plane waves (5.1.1), so long as each component satisfies the dispersion relation; the mathematics behind such Fourier synthesis

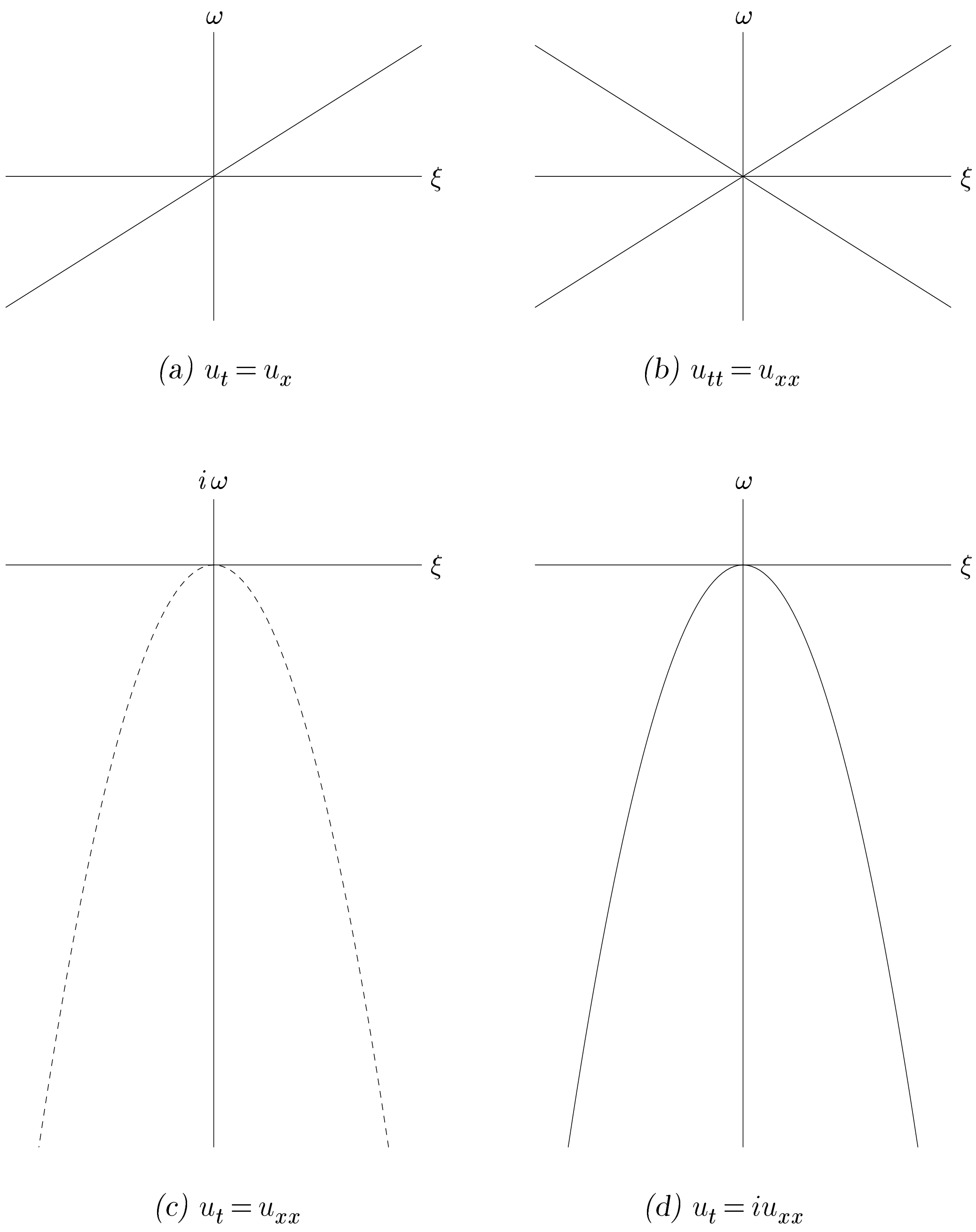


Figure 5.1.1. Dispersion relations for four model partial differential equations. The dashed curve in *(c)* is a reminder that ω is complex.

was described in Chapter 2, and examples were given in §3.1. For a PDE of first order in t, the result is

$$u(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(\xi x+\omega(\xi)t)}\hat{u}_0(\xi)\,d\xi. \tag{5.1.7}$$

Since most partial differential equations of practical importance have variable coefficients, nonlinearity, or boundary conditions, it is rare that this integral representation is exactly applicable, but it may still provide insight into local behavior.

Discrete approximations to differential equations also admit plane wave solutions (5.1.1), at least if the grid is uniform, and so they too have dispersion relations. To begin with, let us discretize in x only so as to obtain a semidiscrete formula. Here are the dispersion relations for the standard centered semidiscretizations of (5.1.3)–(5.1.6):

$$u_t = \delta_0 u: \qquad \omega = \frac{1}{h}\sin\xi h, \tag{5.1.8}$$

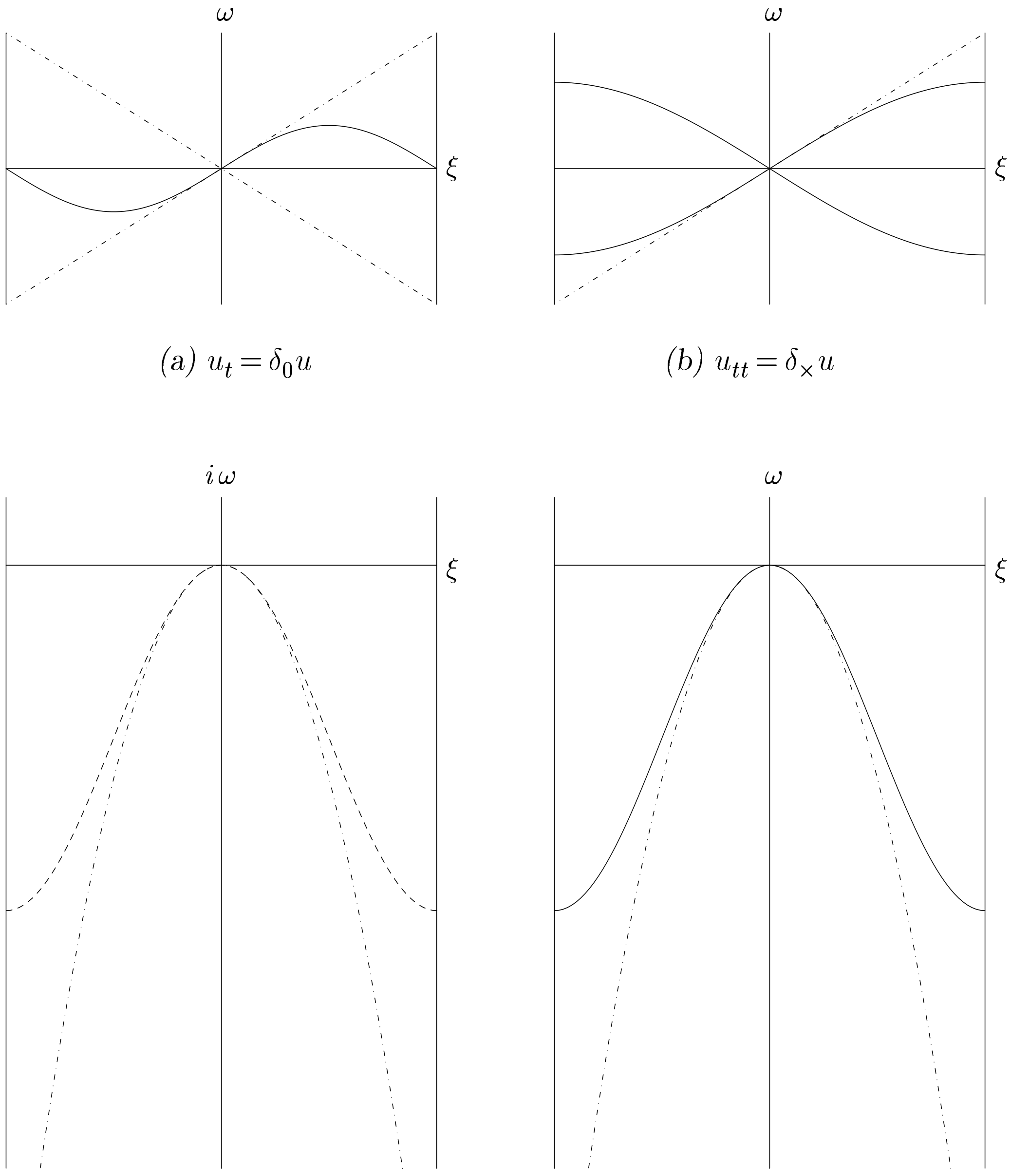
$$u_{tt} = \delta_\times u: \qquad \omega^2 = \frac{4}{h^2}\sin^2\frac{\xi h}{2}, \tag{5.1.9}$$

$$u_t = \delta_\times u: \qquad i\omega = -\frac{4}{h^2}\sin^2\frac{\xi h}{2}, \tag{5.1.10}$$

$$u_t = i\delta_\times u: \qquad \omega = -\frac{4}{h^2}\sin^2\frac{\xi h}{2}. \tag{5.1.11}$$

These formulas are obtained by substituting (5.1.1) into the finite difference formulas with $x = x_j$. In keeping with the results of §2.2, each dispersion relation is $2\pi/h$-periodic in ξ, and it is natural to take $\xi \in [-\pi/h, \pi/h]$ as a fundamental domain. The dispersion relations are plotted in Figure 5.1.2, superimposed upon dotted curves from Figure 5.1.1 for comparison.

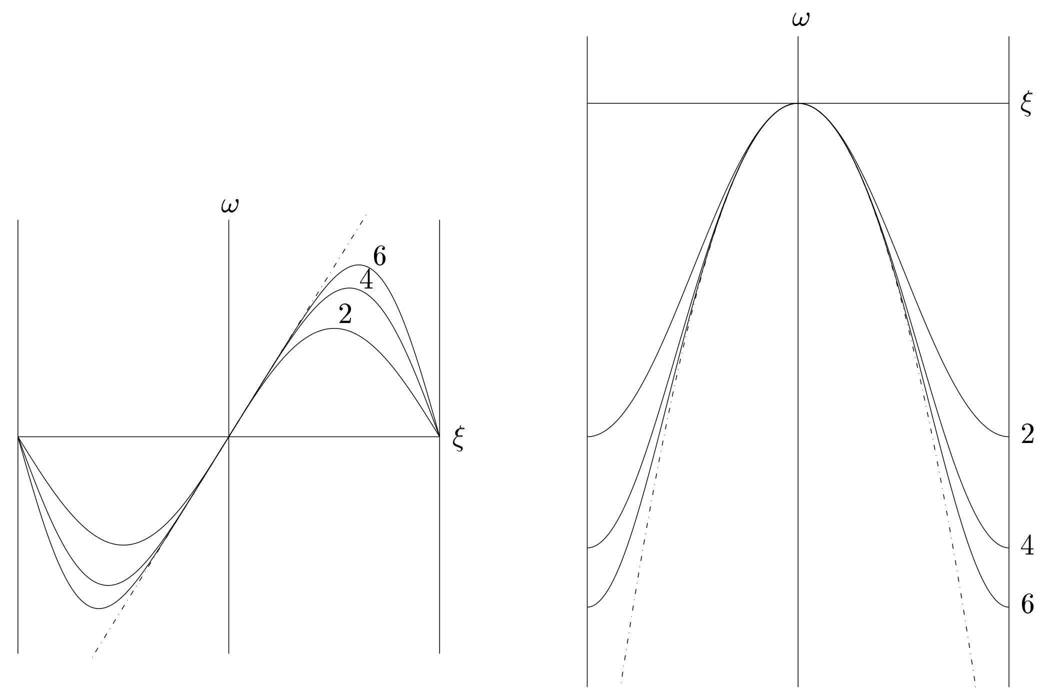
Stop for a moment to compare the continuous and semidiscrete curves in Figure 5.1.2. In each case the semidiscrete dispersion relation is an accurate approximation *when ξ is small*, which corresponds to *many grid points per wavelength.* (The number of points per spatial wavelength for the wave (5.1.1) is $2\pi/\xi h$.) In general, the dispersion relation for a partial differential equation is a polynomial relation between ξ and ω, while a discrete model amounts to a trigonometric approximation. Although other design principles are possible, the standard discrete approximations are chosen so that the trigonometric function matches the polynomial to as high a degree as possible at the origin $\xi = \omega = 0$. To illustrate this idea, Figure 5.1.3 plots dispersion relations for the standard semidiscrete finite difference approximations to $u_t = u_x$ and $u_t = iu_{xx}$ of orders 2, 4, and 6. The formulas were given in §3.3.



(a) $u_t = \delta_0 u$ (b) $u_{tt} = \delta_\times u$

(c) $u_t = \delta_\times u$ (d) $u_t = i\delta_\times u$

Figure 5.1.2. Dispersion relations for centered semidiscrete approximations to the four model partial differential equations. Each function is $2\pi/h$-periodic in ξ; the plots show the fundamental domain $\xi \in [-\pi/h, \pi/h]$.



(a) $u_t = u_x$ (b) $u_t = iu_{xx}$

Figure 5.1.3. Dispersion relations for semidiscrete centered difference approximations to $u_t = u_x$ and $u_t = iu_{xx}$ of orders 2, 4, 6.

Now let us turn to fully discrete finite difference formulas: discrete in time as well as space. The possibilities become numerous. For example, substituting the plane wave

$$v_j^n = e^{i(\xi x_j + \omega t_n)} = e^{i(\xi jh + \omega nk)}$$

into the leap frog approximation of $u_t = u_x$ yields the dispersion relation

$$e^{i\omega k} - e^{-i\omega k} = \lambda(e^{i\xi x} - e^{-i\xi x}),$$

where $\lambda = k/h$, that is,

$$\sin \omega k = \lambda \sin \xi h. \qquad (5.1.12)$$

Similarly, the Crank-Nicolson approximation of $u_t = u_{xx}$ has the dispersion relation

$$e^{i\omega k} - 1 = \frac{\sigma(e^{i\omega k} + 1)}{2}\left[e^{i\xi h} - 2 + e^{i\xi h}\right],$$

which reduces to

$$i \tan \frac{\omega k}{2} = -2\sigma \sin^2 \frac{\xi h}{2}. \qquad (5.1.13)$$

These dispersion relations are $2\pi/h$-periodic in ξ and $2\pi/k$-periodic in ω. With the use of inverse trigonometric functions it is possible to solve such equations for ω, so as to exhibit the functional dependence explicitly, but the resulting formulas are less appealing and often harder to work with. Equations like (5.1.12) and (5.1.13) have a certain elegance—one sees at a glance that the time and space derivatives have been replaced by trigonometric analogs.

Tables 5.1.1 and 5.1.2* consider once again the various finite difference approximations to $u_t = u_x$ and $u_t = u_{xx}$ that appeared in Tables 3.2.1/3.2.2 and 4.4.1/4.4.2. In each case the dispersion relation is both listed and plotted. Since h and k are independent parameters, there is now a range of possible plots; we have arbitrarily taken $\lambda = k/h = 0.5$ in the first table and $\sigma = k/h^2 = 0.5$ in the second. That is why each plot occupies a rectangle of aspect ratio 2 in wave number space.† Notice that the multistep (leap frog) formulas contain two branches of ω values.

For partial differential equations in several space dimensions, the notion of dispersion relation generalizes just as §2.6 would suggest: a plane wave takes the form

$$u(x,t) = e^{i(\xi\cdot x + \omega t)}, \qquad \xi \in \mathbb{R}, \tag{5.1.14}$$

where ξ and x are vectors, and (5.1.2) becomes a scalar function (or relation) of a vector argument. For example, the wave equation

$$u_{tt} = u_{xx} + u_{yy}$$

has the dispersion relation

$$\omega^2 = \xi^2 + \eta^2, \tag{5.1.15}$$

if the vector ξ is written (ξ,η), so that the lines of constant ω in the (ξ,η) plane are concentric circles. On the other hand the leap frog approximation

$$v_{ij}^{n+1} - 2v_{ij}^n + v_{ij}^{n-1} = \lambda^2 (v_{i+1,j}^n + v_{i-1,j}^n + v_{i,j+1}^n + v_{i,j-1}^n - 4v_{ij}^n),$$

appropriate to a uniform grid with $h = \Delta x = \Delta y$, has the dispersion relation

$$\sin^2 \frac{\omega k}{2} = \lambda^2 \Big[\sin^2 \frac{\xi h}{2} + \sin^2 \frac{\eta h}{2}\Big], \tag{5.1.16}$$

which is plotted in Figure 5.1.4 for $\lambda \approx 0$. Once again the dispersion relation is accurate for small wave numbers but diverges dramatically elsewhere.

EXERCISES

▷ *5.1.1.* What are the coefficients as trigonometric functions of the dispersion relations plotted in Figure 5.1.3?

▷ *5.1.2.* Sketch the dispersion relation for the leap frog model of $u_t = u_x$ with $\lambda > 1$—say, $\lambda = 2$. How is the instability of this finite difference formula reflected in your sketch?

*Not yet written.

†analogous to a **Brillouin zone** in crystallography (C. Kittel, *Introduction to Solid State Physics*, Wiley, 1976).

BE_x = Backward Euler	$-i(1-e^{-i\omega k}) = \lambda \sin \xi h$
CN_x = Crank-Nicolson	$2 \tan \frac{\omega k}{2} = \lambda \sin \xi h$
LF = Leap Frog	$\sin \omega k = \lambda \sin \xi h$
BOX_x = Box	$\tan \frac{\omega k}{2} = \lambda \tan \frac{\xi h}{2}$
LF4 = 4th-order Leap Frog	$\sin \omega k = \frac{4}{3}\lambda \sin \xi h - \frac{1}{6}\lambda \sin 2\xi h$
LXF = Lax-Friedrichs	$e^{i\omega k} = \cos \xi h + i\lambda \sin \xi h$
UW = Upwind	$e^{i\omega k} - 1 = \lambda(e^{i\xi h} - 1)$
LW = Lax-Wendroff	$-i(e^{i\omega k} - 1) = \lambda \sin \xi h + 2i\lambda^2 \sin^2 \frac{\xi h}{2}$

Table 5.1.1. Dispersion relations for various finite difference approximations to $u_t = u_x$ with $\lambda = k/h = 0.5$. See Tables 3.2.1 and 4.4.1. The dashed lines indicate the slope $d\omega/d\xi$ at isolated points where ω is real.

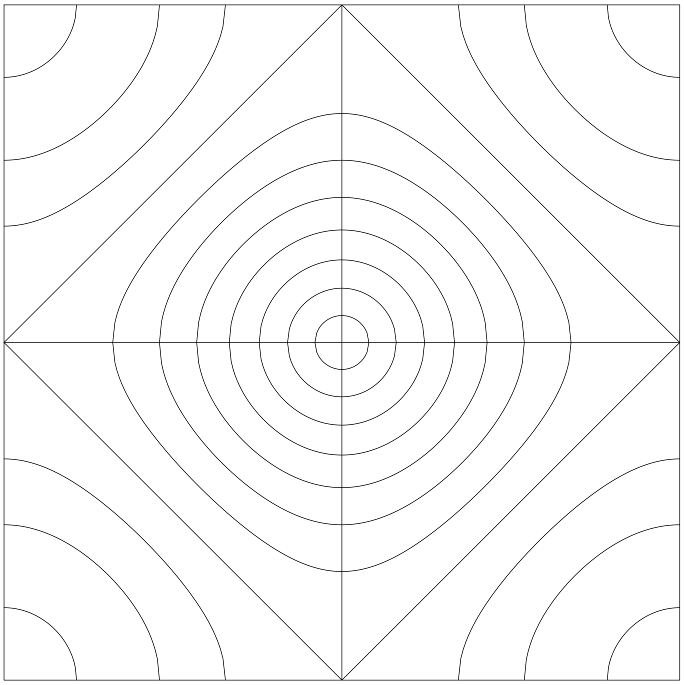


Figure 5.1.4. Dispersion relation for the leap frog model of $u_{tt} = u_{xx} + u_{yy}$ in the limit $\lambda \to 0$. The region shown is the domain $[-\pi/h, \pi/h]^2$ of the (ξ, η) plane. The concentric curves are lines of constant ω for $\omega h = \frac{1}{4}, \frac{1}{2}, \ldots, \frac{11}{4}$.

5.2. Dissipation

Even though a partial differential equation may conserve energy in the L^2 norm, its finite difference models will often lose energy as t increases, especially in the wave numbers comparable to the grid size. This property is **numerical dissipation**, and it is often advantageous, since it tends to combat instability and unwanted oscillations. In fact, **artificial dissipation*** is often added to otherwise nondissipative formulas to achieve those ends. An example of this kind appeared in Exercise 3.2.1.

To make the matter quantitative, suppose we have a linear partial differential equation or finite difference approximation that admits waves (5.1.1) with ω given by a dispersion relation (5.1.2). Since ξ is assumed to be real, it follows that the wave has absolute value

$$|e^{i(\xi x + \omega t)}| = e^{-t\,\mathrm{Im}\,\omega} \tag{5.2.1}$$

*In computational fluid dynamics one encounters the more dramatic term **artificial viscosity**.

as a function of t, and thus decays exponentially if $\operatorname{Im}\omega > 0$. By Parseval's equality, the L^2 norm of a superposition of waves (5.1.1) is determined by a superposition of such factors:

$$\|u(\cdot,t)\| = \frac{1}{2\pi}\left\|e^{-t\operatorname{Im}\omega(\xi)}\hat{u_0}(\xi)\right\|. \tag{5.2.2}$$

As an extreme case the heat equation $u_t = u_{xx}$, with dispersion relation $\omega = i\xi^2$, dissipates nonzero wave numbers strongly—and indeed it does nothing else; at time t, only wave numbers $\xi = O(\sqrt{t})$ remain with close to their initial amplitudes. But it is principally the dissipation introduced by finite difference formulas that we are concerned with here.

The following definitions are standard:

> *A finite difference formula is* **nondissipative** *if* $\operatorname{Im}\omega = 0$ *for all* ξ. *It is* **dissipative** *if* $\operatorname{Im}\omega > 0$ *for all* $\xi \neq 0$. *It is* **dissipative of order** $2r$ *if* ω *satisfies*
>
> $$\operatorname{Im}\omega k \ge \gamma_1(\xi h)^{2r}, \qquad \text{i.e.,} \qquad |e^{i\omega k}| \le 1 - \gamma_2(\xi h)^{2r} \tag{5.2.3}$$
>
> *for some constants* $\gamma_j > 0$. *In each of these statements,* ξ *varies over the interval* $[-\pi/h, \pi/h]$, *and* ω *represents all possible values* ω *corresponding to a given* ξ. *For problems in multiple space dimensions,* (ξh) *is replaced by* $\|\xi h\|$ *in any norm.*

For example, the leap frog and Crank-Nicolson models of $u_t = u_x$ are nondissipative, while the upwind and Lax-Wendroff formulas are dissipative.

Dissipative and nondissipative are mutually exclusive, but not exhaustive: a finite difference formula can be neither dissipative nor nondissipative. See Exercise 5.2.1.

According to the definition, no consistent finite difference approximation of $u_t = u_x + u$ could be dissipative, but customary usage would probably use the term dissipative sometimes for such a problem anyway. One could modify the definition to account for this by including a term $O(k)$ in (5.2.3).

A more serious problem with these standard definitions arises in the case of multistep formulas, for which each ξ corresponds to several values of ω. In such cases the definition of dissipative, for example, should be interpreted as requiring $\operatorname{Im}\omega > 0$ for every value of ω that corresponds to some $\xi \neq 0$. The difficulty arises because for multistep formulas, that condition ensures only that the formula dissipates oscillations in space, not in time. For example, the leap frog model becomes dissipative if a small term such as $k\delta_\times v^n$ is added to it, according to our definitions, yet the resulting formula still admits the wave (5.1.1) with $\xi = 0$, $\omega = \pi/h$, which is sawtoothed in time. To exclude possibilities of that kind it is sometimes desirable to use a stronger definition:

> *A finite difference formula is* **totally dissipative** *if it is dissipative and in addition,* $\operatorname{Im}\omega = 0$ *implies* $\omega = 0$.

EXERCISES

▷ *5.2.1.* Determine whether each of the following models of $u_t = u_x$ is nondissipative, dissipative, or neither. If it is dissipative, determine the order of dissipativity.
(a) Lax-Wendroff, *(b)* Backward Euler, *(c)* Fourth-order leap frog, *(d)* Box, *(e)* Upwind.

5.3. Dispersion and group velocity

[This section is not yet properly written. The next few pages contain a few remarks, followed by an extract from my paper "Dispersion, dissipation, and stability."]

The general idea. Whereas dissipation leads to decay of a wave form, dispersion leads to its gradual separation into a train of oscillations. This phenomenon is a good deal less obvious intuitively, for it depends upon constructive and destructive interference of Fourier components. It is of central importance in finite difference modeling, because although many partial differential equations are nondispersive, their discrete approximations are almost invariably dispersive. (Spectral methods are an exception.)

Caution. Phase and group velocity analysis depend upon the problem being linear and nondissipative—i.e., ω must be real when ξ is real. (However, similar predictions hold if there is a sufficiently slight amount of dissipation.)

Phase velocity. Suppose that a PDE or finite difference formula admits a solution $e^{i(\omega t+\xi x)}$. It's then a triviality to see that any individual "wave crest" of this wave, i.e., a point moving in such a way that the quantity inside the parentheses has a constant value (the phase), moves at the velocity

$$\textbf{Phase velocity:} \qquad c(\xi,\omega) = -\frac{\omega}{\xi}. \tag{5.3.1}$$

Group velocity. However, early in the twentieth century it was realized that wave energy propagates at a different velocity,

$$\textbf{Group velocity:} \qquad c_g(\xi,\omega) = -\frac{d\omega}{d\xi}. \tag{5.3.2}$$

The algebraic meaning of this expression is that we differentiate the dispersion relation with respect to ξ. (In a plot in ξ–ω space, c is minus the slope of the line through (ξ,ω) and the origin, and c_g is minus the slope of the line tangent to the dispersion relation at (ξ,ω).) The physical meaning is that, for example, a **wave packet**—a smooth envelope times an oscillatory carrier wave with parameters (ξ,ω)—will move approximately at the velocity c_g. The same goes for a wave front and for any other signal that can carry information.

Dispersive vs. nondispersive systems. If the dispersion relation is linear, i.e., $\omega = \text{const}\cdot\xi$, then (5.3.1) and (5.3.2) are equal and the system is **nondispersive**. If the dispersion relation is nonlinear, the system is **dispersive**. Finite difference formulas are almost always dispersive, since their dispersion relations are periodic and therefore nonlinear. (However, see Exercise 5.3.2.)

Precise meaning of group velocity. The meaning of group velocity can be made precise in various asymptotic fashions, for example by considering the limit $t \to \infty$. The mathematics behind this is usually a **stationary phase** or **steepest descent** argument.

Simple explanations of group velocity. There are a number of intuitive ways to understand where the derivative (5.3.2) comes from. One is to superimpose two waves with nearby parameters (ξ_1, ω_1) and (ξ_2, ω_2). It is then readily seen that the superposition consists of a smooth envelope times a carrier wave at frequency $\frac{1}{2}(\omega_1+\omega_2)$ and wave number $\frac{1}{2}(\xi_1+\xi_2)$, and the envelope moves at velocity $-(\omega_2-\omega_1)/(\xi_2-\xi_1)$, which approaches (5.3.2) in the limit $\xi_2 \to \xi_1$, $\omega_2 \to \omega_1$. Another approach is to take a pure exponential $e^{i(\omega t+\xi x)}$, with ξ and ω real, and change ξ slightly to a complex value $\xi + i\Delta\xi$ "in order to visualize which way the envelope is moving." If the effect on ω is to make it change to $\omega + i\Delta\omega$, it is readily calculated that the resulting evanescent wave has an envelope that moves laterally at the velocity $-\Delta\omega/\Delta\xi$, and this again approaches (5.3.2) in the limit $\Delta\xi \to 0$, $\Delta\omega \to 0$. A third, more "PDE-style" explanation is based upon advection of local wave number according to a simple hyperbolic equation with coefficient c_g; see Lighthill.

Group velocity in multiple dimensions. If there are several space dimensions, the group velocity becomes the gradient of ω with respect to the vector ξ, i.e., $c_g = -\nabla_\xi \omega$.

Phase and group velocities on a grid. There is another sense in which group velocity has more physical meaning than phase velocity on a finite difference grid: the former is well-defined, but the latter is not. On a periodic grid, any Fourier mode can be represented in terms of infinitely many possible choices of ξ and ω that are indistinguishable physically, and according to (5.3.1), each choice gives a different phase velocity. What's going on here is that naturally one can't tell how fast a pure complex exponential wave is moving if one sees it at only intermittent points in space or time, for one wave crest is indistinguishable from another. By contrast, the group velocity is well-defined, since it depends only on the slope, which is a local property; formula (5.3.2) has the same periodicity as the dispersion relation itself.

Computation of a group velocity on a grid. To compute the group velocity for a finite difference formula, differentiate the dispersion relation implicitly and then solve for $c_g = -d\omega/d\xi$. For example, the wave equation $u_t = u_x$ has $c = c_g = -1$ for all ξ. For the leap frog approximation the dispersion relation is $\sin\omega k = \lambda \sin\omega\xi$, which implies $k\cos\omega k\, d\omega = h\lambda \cos\omega\xi\, d\xi$, and since $k = h\lambda$, $c_g(\xi,\omega) = -\cos\xi h/\cos\omega k$.

Parasitic waves. Many finite difference formulas admit parasitic waves as solutions, i.e., waves that are sawtoothed with respect to space or time. These correspond to $\xi = \pm\pi/h$, $\omega = \pm\pi/k$, or both. It is common for such waves to have group velocities opposite in sign to what is correct physically. In the example of the leap frog formula, all four parasitic modes 1, $(-1)^j$, $(-1)^n$, and $(-1)^{j+n}$ are possible, with group velocities -1, 1, 1, and -1, respectively.

Spurious wiggles near interfaces and boundaries. It is common to observe spurious wiggles in a finite difference calculation, and they appear most often near boundaries, interfaces, or discontinuities in the solution itself. The explanation of where they appear is usually a matter of group velocity. Typically a smooth wave has passed through the discontinuity and generated a small reflected wave of parasitic form, which propagates backwards into the domain because its group velocity has the wrong sign. More on this in the next chapter.

Waves in crystals. Dispersion relations for vibrations in crystals are also periodic with respect to ξ. As a result, sound waves in crystals exhibit dispersive effects much like those associated with finite difference formulas, including the existence of positive and negative group velocities.

Figure 5.3.1. Dispersion under the leap frog model of $u_t = u_x$ with $\lambda = 0.5$. The lower mesh is twice as fine as the upper.

EXERCISES

▷ *5.3.1. The Box formula.*

(a) Write out the BOX_x formula of Table 3.2.1 in terms of v_j^n, v_{j+1}^n, etc.

(b) Determine the dispersion relation (expressed in as simple a form as possible).

(c) Sketch the dispersion relation.

(d) Determine the group velocity as a function of ξ and ω.

▷ *5.3.2. Schrödinger equation.*

(a) Calculate and plot the dispersion relation for the Crank-Nicolson model of $u_t = iu_{xx}$ of Exercise 3.2.1*(f)*.

(b) Calculate the group velocity. How does it compare to the group velocity for the equation $u_t = iu_{xx}$ itself?

▷ *5.3.3. A paradox.* Find the resolution of the following apparent paradox, and be precise in stating where the mistake is. Draw a sketch of an appropriate dispersion relation to explain your answer.

One the one hand, if we solve $u_t = u_x$ by the leap frog formula with $\lambda = 1$, the results will be exact, and in particular, no dispersion will take place.

On the other hand, as discussed above, the dispersion relation on any discrete grid must be periodic, hence nonlinear—and so dispersion must take place after all.