DIFFICULTY OF HIGHER DIMENSIONS

The chapter after this introduces the so-called "splitting" method of solving partial-differential equations. Despite the fact that it is a perfectly legitimate method, many students find it aesthetically objectionable. But it turns out to be a very useful method for dealing with lateral velocity variation. Furthermore, it turns out to be an economic necessity with so-called "3-D" seismic migrations. In this section we see why multi-dimensional problems are tougher in kind as well as in scale. If you find that you are not interested in 3-D and don't object to dirty tricks, then you could skip this section and leap forward to the offending method in the next chapter.

So far we have had no trouble obtaining cheap, safe, and accurate difference methods for solving ***partial-differential equations*** (PDE's). The implicit method has met all needs. But in space dimensions higher than one the implicit method becomes prohibitively costly. For the common example of problems in which ∂_{xx} becomes generalized to $\partial_{xx} + \partial_{yy}$ we will learn the reason why. The simplest case is the heat-flow equation for which the Crank-Nicolson method gave us

$$\left(1 - \frac{\sigma \Delta t}{2C\Delta x^2}\,\delta_{xx}\right) T_{t+1} = \left(1 + \frac{\sigma \Delta t}{2C\Delta x^2}\,\delta_{xx}\right) T_t \qquad (1)$$

The nested expression on the left represents a tridiagonal matrix. The critical stage is when we solve the tridiagonal simultaneous equations for the vector of unknowns T_{t+1}. Luckily there is a special algorithm for this solution, and the cost increases only linearly with the size of the matrix. Now turn from the one-dimensional physical space of x to two-dimensional (x,y)-space. Letting b denote the numerical constant in (1) we find we are supposed to step forward in time with

$$\left[1 - b\left(\delta_{xx} + \delta_{yy}\right)\right] T_{t+1} = \left[1 + b\left(\delta_{xx} + \delta_{yy}\right)\right] T_t \qquad (2)$$

The unknowns T_{t+1} are a two-dimensional function of x and y which can be denoted by a matrix. Next we must interpret the bracketed expression on the

left side. It turns out to be a four-dimensional matrix!

To clarify the meaning of this we illustrate a mapping from two dimensions to one. Take the temperature T to be defined on a 4-by-4 mesh. A natural way of numbering the points on the mesh is

$$\begin{matrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{matrix} \tag{3}$$

For algebraic purposes these 16 numbers may be mapped into a vector. There are many ways to do this. A simple way would be to associate the locations in (3) with vector components by the column arrangement

$$\begin{matrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{matrix} \tag{4}$$

The second difference operator has the following star in the x-y plane

$$\begin{array}{|c|c|c|} \hline & 1 & \\ \hline 1 & -4 & 1 \\ \hline & 1 & \\ \hline \end{array} \tag{5}$$

We may now lay this star down in the (x,y)-plane (4) and move it around. Unfortunately with just 16 points, much of what you see is dominated by edges and corners. We will try every position of the star which allows the center -4 to overlay one of the 16 points. Never mind the 1's going off the sides. Start with the -4 over the 1 in the upper left corner. Observe 1's on the 2 and the 5. Then put the -4 over the 2. Observe 1's on the 1, 3, and 6. Then put the -4 over the 3. Observe 1's on the 2, 4, and 7. Continuing likewise the results may be tabulated in the following 16-by-16 square table which is conveniently numbered 1 thru 16 along the top and along the right.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
-4	1	.	.	1	.	.	.	.	.	.	.	.	.	.	.	1
1	-4	1	.	.	1	.	.	.	.	.	.	.	.	.	.	2
.	1	-4	1	.	.	1	.	.	.	.	.	.	.	.	.	3
.	.	1	-4	.	.	.	1	.	.	.	.	.	.	.	.	4
1	.	.	.	-4	1	.	.	1	.	.	.	.	.	.	.	5
.	1	.	.	1	-4	1	.	.	1	.	.	.	.	.	.	6
.	.	1	.	.	1	-4	1	.	.	1	.	.	.	.	.	7
.	.	.	1	.	.	1	-4	.	.	.	1	.	.	.	.	8
.	.	.	.	1	.	.	.	-4	1	.	.	1	.	.	.	9
.	.	.	.	.	1	.	.	1	-4	1	.	.	1	.	.	10
.	.	.	.	.	.	1	.	.	1	-4	1	.	.	1	.	11
.	.	.	.	.	.	.	1	.	.	1	-4	.	.	.	1	12
.	.	.	.	.	.	.	.	1	.	.	.	-4	1	.	.	13
.	.	.	.	.	.	.	.	.	1	.	.	1	-4	1	.	14
.	.	.	.	.	.	.	.	.	.	1	.	.	1	-4	1	15
.	.	.	.	.	.	.	.	.	.	.	1	.	.	1	-4	16

(6)

Now that (6) has been constructed we can return to the interpretation of equation (2). The matrix of unknowns T_{t+1} has been mapped into a 16-point column vector, and the bracketed expression multiplying T_{t+1} can be mapped into a 16-by-16 matrix. Clearly, the matrix contains zeros everywhere that (6) contains zeros. It seems fortunate that (6) contains many zeros and we can hope for a rapid solution method for the simultaneous equations. The bad news is that no good method has ever been found, despite the efforts of many professionals. The best methods seem to be proportional to N^3 where in this case N=4. Based on our experience in one dimension we had hoped for a method proportional to N^2, which is the cost of an explicit method, essentially, the cost of computing the right side of (2). Even all the nice features of implicit methods do not justify a factor of N of additional cost. The next best thing is the splitting method, described in the next chapter.

chapters where we saw that in two space dimensions the Crank-Nicolson method becomes prohibitively costly, while the explicit methods raise stability concerns and can also become very costly.

On first sight, some people regard splitting as obviously valid while other people are troubled by it. Perhaps it is related to the old paradox of the length of a stairway compared to the length of a ramp. In the limit of smaller and smaller step size, should these not have the same length? After discussion of applications we will return to a proof of the validity of splitting, thus showing the irrelevance of the ramp-and-stairway paradox.

Full Separation

Splitting can turn out to be much more accurate than you might imagine. In many important cases there is ***no*** loss of accuracy. Then we can take the method to an extreme limit. Think about a radical approach to equations (2a) and (2b) in which, instead of alternating back and forth between them at alternate time steps, what is done is to march (2a) through all time steps. Then this intermediate result is used as an initial condition for (2b), which is marched through all time steps to produce a final result. It might seem surprising that this radical method can produce the correct solution to equation (1). But in fact, if σ is a constant function of x and y, this radical method does produce the right answer. The process is depicted in figure 1 for an impulsive initial disturbance. A differential equation like (1) is said to be ***fully separable*** when the correct solution is obtainable by the radical method. It should not be too surprising that full separation works when σ is a constant, because then Fourier transformation may be used, and it is obvious that $\exp\left[\sigma(k_x^2 + k_y^2)t\right]$ equals the product $\exp(\sigma k_x^2 t)\ \exp(\sigma k_y^2 t)$. It turns out, and we will later show, that the condition required for applicability of full separation is that $\sigma\partial_{xx}$ should commute with $\sigma\partial_{yy}$. Technically there is also a boundary condition requirement, but there is no problem when the disturbance dies out before reaching a boundary.

Surprisingly, no notice is made of full separability in many textbooks on numerical solutions. Perhaps this is because the total number of additions and multiplications is the same whether a solution is found by splitting or by full separation. But as a practical matter, costs for large problems do not