Introduction to non-standard finite-difference modelling

Peter M. Manning and Gary F. Margrave

ABSTRACT

This report introduces some of the principles and uses of non-standard finite-difference modelling. It begins with a review of the elements of finite-differences, and how continuous derivatives are replaced with differences, and incorporated into a 'scheme'. It shows examples of scheme selection by testing. Finally, it demonstrates some more advanced strategies of standard finite-differencing for stability or higher accuracy. Then the non-standard ideas are introduced. A set of standard finite-differences is compared with their equivalent 'exact' finite-differences. It finishes with some examples where exact difference forms are incorporated into non-standard schemes.

INTRODUCTION

The laws of nature are usually best described by differential equations which, for the macroscopic world, apply to continuous materials, and usually have continuous solutions. Analytic results (from calculus) give elegant solutions for these cases, but for a quite limited set of problems. That leaves many problems which must have their solutions estimated by using arithmetic operations on arrays of numbers. The point of the finite-difference literature is to minimize the difficulties that occur when continuous functions are represented by arrays of samples spaced at finite intervals. Most of the effort within finite-difference studies is focused on 'schemes' which solve problems posed in the form of the differential equations. These schemes are then evaluated for efficiency and accuracy by theoretical means, and then with test runs. This study concentrates on these 'schemes', and how they are derived.

CLASSIC FINITE-DIFFERENCE MODELLING

First derivatives

The essential difficulty of finite-difference procedures becomes obvious when a finite-difference option must be chosen to represent a first derivative. A first derivative, or slope, can only be defined as a difference in values between two points. Within calculus the slope at a point is obtained by using the definition of limits to shrink the difference between the sample points, and the differences of the values at those points, down to infinitesimal amounts (see equation 1).

$$\frac{df}{dx} = \lim_{\Delta x \downarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$ (1)

This procedure is not allowed with finite differences, because only the samples as given are available. An equivalent finite-difference equation is given in equation 2, where the difference is made as small as possible, or one fixed sample length.

$$D_x f = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$ (2)

This same equation is given in 3, where the input and output values are given in terms of sample numbers instead of their locations along the x axis.

$$D_x f_n = \frac{f_{n+1} - f_n}{\Delta x}. \tag{3}$$

One of the finite-difference first derivative options available is the forward difference (to the point ahead) in equation 3. Also valid is a backward difference (to the point behind, equation 4), or a central difference (from the point ahead to the point behind, equation 5). The central difference clearly has less bias, but depends on values that are further (in total) from the point where it will be used.

$$D_x f_n = \frac{f_n - f_{n-1}}{\Delta x}. \tag{4}$$

$$D_x f_n = \frac{f_{n+1} - f_{n-1}}{2\Delta x}. \tag{5}$$

The result of using the forward and backward finite-differences are shown graphically in Figure 1. The analytic curve which has been sampled is shown in black, and the derivative of this curve at the same point is shown in blue. It is obvious that the finite-differences
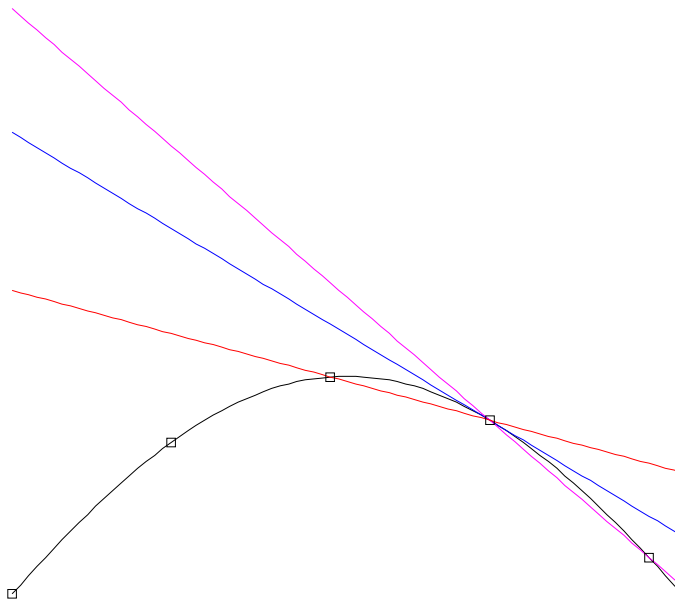

FIG. 1. The backward finite-difference slope is shown in red, and the forward finite-difference in magenta. These may be compared with the true analytic derivative plotted in blue.

shown are biased toward the slopes in the direction from which they were calculated.

The central-difference calculation of equation 5 is much closer to the analytic result, as shown in Figure 2.
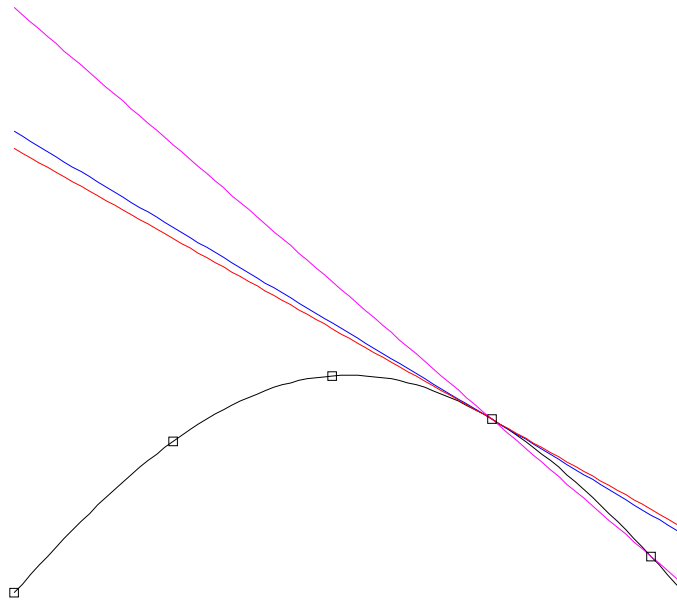
FIG. 2. The central finite-difference slope is shown in red, but is shifted to go through the point at which it applies. For comparison, the forward finite-difference in magenta is also shown. The central finite-difference is obviously much closer to the true analytic derivative plotted in blue.

Second derivative

A second derivative does not have a similar range of options, because the difference of differences falls naturally at the positions of the original samples, see equation 6.

$$D_x^2 f_n = \frac{f_{n+1} - 2f_n + f_{n-1}}{(\Delta x)^2}. \tag{6}$$

Elementary schemes

The most simple finite-difference 'scheme' then has a selected type of estimate for each term of the differential equation, which makes it into a difference equation. The desired output is usually the function throughout space projected to an advanced time (a time-stepped result). In many cases (the explicit cases), the difference equation may be manipulated to solve for an advanced time term. In some cases the advanced time terms may be implicit, which means the advanced time appears in ways that are impossible to solve for directly. Special techniques may then be required for the solution.

In either case, the functions must be successively stepped through further increments in time to get a full picture through space and time.

The final requirement is to test the scheme. The most important test is for stability. Here amplitudes at a range of frequencies are examined to make sure that these amplitudes never become infinite. In linear stability analysis this amounts to the exponential growth of at least some frequencies, which soon swamp all other features of the model.

Another test is for consistency, to make sure that a decrease in the step size of all variables leads to reduced errors.

Further tests usually entail comparisons with certain known solutions, usually analytical solutions, to see if the results are close enough to what is expected.

An example of this methodology is given in Aki and Richards (1980) where, for an initial attempt, the wave equation represented by the simultaneous equation pair

$$\frac{\partial \dot{u}}{\partial t} = \frac{1}{\rho(x)} \frac{\partial \tau}{\partial x}, \tag{7}$$

$$\frac{\partial \tau}{\partial t} = E(x) \frac{\partial \dot{u}}{\partial x} \tag{8}$$

is translated to a scheme with forward finite differences for t-derivatives, and central differences for x-derivatives. Although this appears to be a very natural choice for an explicit scheme, it turns out to be unstable. A scheme with central differences for both t and x is then tried and found to be stable. The details of this scheme are laid out as follows,

$$\frac{\dot{u}_n^{j+1} - \dot{u}_n^{j-1}}{2\Delta t} = \frac{1}{\rho_n} \frac{\tau_{n+1}^j - \tau_{n-1}^j}{2\Delta x}, \tag{9}$$

$$\frac{\tau_n^{j+1} - \tau_n^{j-1}}{2\Delta t} = E_n \frac{\dot{u}_{n+1}^j - \dot{u}_{n-1}^j}{2\Delta x}. \tag{10}$$

Here the time sample variations are annotated as superscripts instead of subscripts. This explicit scheme may be solved for the advanced time terms $\dot{u}^{j+1}$ or τ^{j+1} as a function of past states and the material parameters.

Smoothing

Several additional operations may be utilized to supplement the basic finite-difference choices. One which may be incorporated is smoothing in the space domain, a technique introduced by Lax. An example of its use is given in Press et al. (1992), where the one way wave-equation

$$\frac{\partial u}{\partial t} = -v \frac{\partial u}{\partial x} \tag{11}$$

is represented by a forward finite-difference in time, and a central finite-difference in space. This equation is

$$\frac{u_n^{j+1} - u_n^j}{\Delta t} = -v \left(\frac{u_{n+1}^j - u_{n-1}^j}{2\Delta x} \right). \tag{12}$$

This may be solved for the advanced time term u_n^{j+1}, but as with the Aki and Richards example, it is shown to be unstable. A smoother applied to the u_n^j term on the time side (a gapped running average) gives a result called the Lax-Friedrichs scheme

$$\frac{u_n^{j+1} - 1/2(u_{n+1}^j + u_{n-1}^j)}{\Delta t} = -v \left(\frac{u_{n+1}^j - u_{n-1}^j}{2\Delta x} \right), \tag{13}$$

and with appropriate sample rates this is shown to be stable.

High accuracy derivatives

A second technique used is to improve the accuracy of the differentiation, usually of the spatial co-ordinates. This is done by using a range of samples that extend further than normal from the estimation point. For example, an estimated second derivative may make use of 5 samples rather than the minimal number of 3. Strikwerda (2004) shows how the 'difference calculus' may be used to derive an appropriate weighting for these samples.

Dablain (1986) also shows how spatial derivatives may be calculated to a higher order by incorporating samples further from the estimation position. He uses a more direct technique, and so the principle may be a little easier to see.

Dablain starts with the formulation of the Taylor series given as

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + f''(x)\frac{(\Delta x)^2}{2!} + f'''(x)\frac{(\Delta x)^3}{3!} + \dots \quad . \tag{14}$$

If this is added to a similar expansion for $-\Delta x$

$$f(x - \Delta x) = f(x) - f'(x)\Delta x + f''(x)\frac{(\Delta x)^2}{2!} - f'''(x)\frac{(\Delta x)^3}{3!} + \dots \tag{15}$$

the result is

$$f(x + \Delta x) + f(x - \Delta x) = 2f(x) + 2f''(x)\frac{(\Delta x)^2}{2!} + 2f''''(x)\frac{(\Delta x)^4}{4!} + \dots \quad . \tag{16}$$

This may be rearranged so that

$$f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2} - 2f''''(x)\frac{(\Delta x)^2}{4!} + \dots \quad . \tag{17}$$

Comparing this with equation 6 shows that the usual expression for the second derivative is really the first term of an infinite series, and is accurate to second order (fourth order derivatives are required for higher accuracy).

The next stage of accuracy requires that larger Δx's of the Taylor series be utilized, and so doubling the difference used in equation 16 gives

$$f(x + 2\Delta x) + f(x - 2\Delta x) = 2f(x) + 2f''(x)\frac{(2\Delta x)^2}{2!} + 2f''''(x)\frac{(2\Delta x)^4}{4!} + \dots . \tag{18}$$

Equation 18 may then be divided by 16 and subtracted from equation 16 to eliminate the f'''' term. This equation is then solved for the f'' term, as for equation 17, but the result will be accurate to fourth order. In terms of sample numbers, the equation becomes

$$D_x^2 f_n = \frac{-f_{n+2} + 16f_{n+1} - 30f_n + 16f_{n-1} - f_{n-2}}{12(\Delta x)^2}. \tag{19}$$

High accuracy time-stepping

The final finite-difference extension shown here demonstrates how the time stepping side of the equation may be enhanced. This can be done if the problem equation itself can be differentiated to give time derivatives of the unknown function in terms of spatial derivatives. An example of this is given in Strikwerda (2004), where the one way equation with a source function term,

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} + f, \tag{20}$$

is differentiated in time to get coefficients for a Taylor series expanded in time. The steps to derive the second derivative with respect to the t term are as follows:

$$\frac{\partial^2 u}{\partial t^2} = -a\frac{\partial^2 u}{\partial t \partial x} + \frac{\partial f}{\partial t}, \tag{21}$$

$$\frac{\partial^2 u}{\partial t^2} = -a\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t}\right) + \frac{\partial f}{\partial t}, \tag{22}$$

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2} - a\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t}. \tag{23}$$

The Taylor series in time may be written as

$$u(t + \Delta t, x) = u(t, x) + \Delta t \frac{\partial u(t, x)}{\partial t} + \frac{(\Delta t)^2}{2!}\frac{\partial^2 u(t, x)}{\partial t^2} + \dots \quad . \tag{24}$$

Then equation 23, and the original equation 20 are used to get

$$u(t + \Delta t, x) = u + \Delta t\left[-a\frac{\partial u}{\partial x} + f\right] + \frac{(\Delta t)^2}{2!}\left[a^2\frac{\partial^2 u}{\partial x^2} - a\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t}\right] + \dots \quad , \tag{25}$$

$$u(t + \Delta t, x) = u - a\Delta t\frac{\partial u}{\partial x} + \Delta t f + a^2\frac{(\Delta t)^2}{2!}\frac{\partial^2 u}{\partial x^2} - a\frac{(\Delta t)^2}{2!}\frac{\partial f}{\partial x} + \frac{(\Delta t)^2}{2!}\frac{\partial f}{\partial t} + \dots \quad . \tag{26}$$

The third and sixth terms above may be combined, and with $\lambda = \Delta t/\Delta x$, the finite-difference equation is

$$u_n^{j+1} = u_n^j - \frac{a\lambda}{2}(u_{n+1}^j - u_{n-1}^j) + \frac{(a\lambda)^2}{2}(u_{n+1}^j - 2u_n^j + u_{n-1}^j) -$$
$$\frac{a\Delta t\lambda}{4}(f_{n+1}^j - f_{n-1}^j) + \frac{\Delta t}{2}(f_n^{j+1} + f_n^j). \tag{27}$$

This results in the Lax-Wendorf scheme, which is often much better than the equivalent Lax-Friedrichs scheme. Dablain (1986) uses the same technique to obtain a two-dimensional fourth order accurate time step from fourth order spatial derivatives in x and z.

NON-STANDARD FINITE-DIFFERENCE MODELLING

The basic theory of non-standard finite-difference (NSFD) modelling was set up by Mickens, with some of his papers from the eighties. An important reference is the book (Mickens, 2000) which he edited, and where he provided the first chapter.

Exact finite-difference schemes

Mickens shows that an ordinary differential equation with an analytical solution may be matched exactly by a difference equation with a sampled solution. This equation is what he calls an exact finite-difference solution. This polynomial, with difference and other terms, comes from the determinant of a matrix which includes a minimal set of samples from all the independent analytic solutions of the differential equation.

Use of the analytic solution as a guide for the finite-difference solution is nothing new. Standard finite-difference (SFD) texts always plot their solutions against analytical solutions as a final quality check. Mickens' 'exact schemes', however, use the analytic solutions directly. As it turns out, these schemes have close parallels, but significant deviations from the SFD equations.

An incomplete set of Mickens' SFD and exact solutions is shown in Table

Eq.	Differential equation	Finite-difference equation	Exact FD equation
1	$\frac{du}{dt} = -\lambda u$	$\frac{u_{k+1}-u_k}{h} = -\lambda u_k$	$\frac{u_{k+1}-u_k}{\left(\frac{1-e^{-\lambda h}}{\lambda}\right)} = -\lambda u_k$
2	$\frac{du}{dt} = -u^2$	$\frac{u_{k+1}-u_k}{h} = -u_k^2$	$\frac{u_{k+1}-u_k}{h} = -u_{k+1}u_k$
3	$\frac{du}{dt} = -u^3$	$\frac{u_{k+1}-u_k}{h} = -u_k^3$	$\frac{u_{k+1}-u_k}{h} = -\left(\frac{2u_{k+1}}{u_{k+1}+u_k}\right)u_{k+1}u_k^2$
4	$\frac{du}{dt} = \lambda_1 u - \lambda_2 u^2$	$\frac{u_{k+1}-u_k}{h} = \lambda_1 u_k - \lambda_2 u_k^2$	$\frac{u_{k+1}-u_k}{\left(\frac{e^{\lambda_1 h}-1}{\lambda}\right)} = \lambda_1 u_k - \lambda_2 u_{k+1}u_k$
5	$\frac{d^2u}{dt^2} = \lambda\frac{du}{dt}$	$\frac{u_{k+1}-2u_k+u_{k-1}}{h^2} = \lambda\frac{u_k-u_{k-1}}{h}$	$\frac{u_{k+1}-2u_k+u_{k-1}}{\left(\frac{e^{\lambda h}-1}{\lambda}\right)h} = \lambda\frac{u_k-u_{k-1}}{h}$
6	$\frac{d^2u}{dt^2} + \omega^2 u = 0$	$\frac{u_{k+1}-2u_k+u_{k-1}}{h^2} + \omega^2 u_k = 0$	$\frac{u_{k+1}-2u_k+u_{k-1}}{\left(\frac{4}{\omega^2}\right)\sin^2\left(\frac{\omega h}{2}\right)} + \omega^2 u_k = 0$
7	$\frac{du}{dt} = w$	$\frac{u_{k+1}-u_k}{h} = w_k$	$\frac{u_{k+1}-\cos(\omega h)u_k}{\left(\frac{\sin(\omega h)}{\omega}\right)} = w_k$
	$\frac{dw}{dt} = -\omega^2 u$	$\frac{w_{k+1}-w_k}{h} = -\omega^2 u_k$	$\frac{w_{k+1}-\cos(\omega h)w_k}{\left(\frac{\sin(\omega h)}{\omega}\right)} = -\omega^2 u_k$

Table 1. Exact finite-difference schemes for ODEs, compared with standard finite-difference schemes.

For all these examples, the two solutions have the same number of terms in roughly the same forms. The deviations tend to be of three types:

- The denominators of the derivative terms are usually not h ($h = \Delta t$), but instead some function of h, depending on the original equation. This is the case for all of these examples except for the second and third. The function is close to h (to the order of h^2). In other words, the smaller h is set, the less the modified function is required.

- Sometimes one of the terms in the numerator of a derivative is multiplied by a function. This is shown in the seventh example. The modification is close to 1 (to the order of h^2).

- Nonlinear terms are represented by non-local expressions, or non-local difference expressions. This may be seen in examples 2, 3, and 4, where x^2 and x^3 do not appear as x_k^2 and x_k^3, but as more complex expressions.

Non-standard finite-difference schemes

The rationale of the NSFD method, then, is to improve general finite-difference schemes by including some of the 'exact' forms. Exact results would not be expected from the new schemes, but higher efficiency and better accuracy would be. As with standard schemes, testing would determine whether the new forms were better. This testing has been encouraging for many cases.

Mickens encapsulates his experiences with NSFD schemes into 6 rules which he has found to be useful. He then goes on to show examples of ordinary and partial differential-equations where NSFD based schemes have given superior results, and gives the rationale for his choice of particular schemes.

He usually modifies the denominators of the derivative terms according to the expected form of the solution. For harmonic type equations, denominators of the type in example 6 are used; and for exponential type equations, denominators of the type in example 1 are used.

Mickens discovered that important benefits can be found with the use of appropriate 'non-local' terms. The obvious symmetry of example 2 shows how the correct choice of the non-derivative expression makes the usual adjustment of the derivative expression unnecessary. However, example 3 shows that the expression may become quite complex.

A successful example of a NSFD equation is one set up for a combustion model

$$\frac{du}{dt} = u^2(1 - u). \tag{28}$$

The SFD equation would be

$$\frac{u_{k+1} - u_k}{h} = u_k^2 - u_k^3, \tag{29}$$

whereas Mickens' NSFD equation is

$$\frac{u_{k+1} - u_k}{\left(1 - e^{-h}\right)} = 2(u_k)^2 - u_{k+1}u_k - u_{k+1}(u_k)^2. \tag{30}$$

To the right of the equal sign, the net of the first two terms makes up a square with amplitude 1, and the last term is a cube with amplitude 1, just as the SFD scheme has. The exact formulation for these terms is designed to give a finite-difference solution with the same zeros as the analytic solution. In particular, when u reaches 1, the slope is exactly zero. Note that this is an explicit scheme, it is easily solved for u_{k+1}.

The second example of a NSFD scheme is a solution for a linear advection-diffusion equation (normalized). This partial differential equation is

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = b\frac{\partial^2 u}{\partial x^2}, \tag{31}$$

and the scheme is

$$\frac{u_m^{j+1} - u_m^j}{\Delta t} + \frac{u_m^j - u_{m-1}^j}{\Delta x} = b\left[\frac{u_{m+1}^j - 2u_m^j + u_{m-1}^j}{b(e^{\Delta x/b} - 1)\Delta x}\right]. \tag{32}$$

This was constructed quite simply, because

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \tag{33}$$

has an already exact solution

$$\frac{u_m^{j+1} - u_m^j}{\Delta t} + \frac{u_m^j - u_{m-1}^j}{\Delta x} = 0, \tag{34}$$

and

$$\frac{\partial u}{\partial x} = b\frac{\partial^2 u}{\partial x^2}, \tag{35}$$

has an exact solution

$$\frac{u_m - u_{m-1}}{\Delta x} = b\left[\frac{u_{m+1} - 2u_m + u_{m-1}}{b(e^{\Delta x/b} - 1)\Delta x}\right], \tag{36}$$

and the two solutions were just combined.

Non-standard finite-difference wave equations

The lead in using NSFD schemes for the wave equation has been taken by Cole. Chapter 3 in Mickens' book was written by Cole, and shows how he has used NSFD techniques to model the wave equation in one, two, and three dimensions.

The starting point for wave equation NSFD modelling is the one-dimensional one, where it is shown that for a given wavenumber and frequency, the sample rate Δt may be replaced by $sin(\omega\Delta t/2)$, and the sample rate Δx by $sin(k\Delta x/2)$. The full equation is

$$\frac{\bar{d}_t^2\psi(x,t)}{4\sin^2(\omega\Delta t/2)/\omega^2} - v(x)^2\frac{\bar{d}_x^2\psi(x,t)}{4\sin^2(k\Delta x/2)/k^2} = 0, \tag{37}$$

where $\bar{d}_t$ is the central finite difference. This conforms with Mickens' first observation of his exact scheme changes. Cole shows how a truncated single frequency wave train propagated through a narrow high velocity barrier appears much closer to expectations with the NSFD algorithm compared to the SFD algorithm.

Cole continues to two dimensions by combining the normal Laplacian with another, oriented at 45 degrees. The combination begins with a variable weighting, but is then optimized. The accuracy of this combination he calculates to be of order 8, as compared to order 2 for the SFD scheme. In a similar fashion he extends the method to the full three dimensions.

At the end of the chapter, Cole shows how the SFD Yee algorithm for Maxwell's equations may be improved by using a NSFD version. The discussion here is for an harmonic (single frequency) case, but apparently works very well for the limited range of frequencies that are required to make up a signal with a finite length in time. He also claims that NSFD techniques may be used to improve the propagation of broad band signals as well.

CONCLUSIONS

- Mickens and Cole have proved that, at least in some cases, NSFD modelling provides stability, accuracy, and efficiency benefits over SFD modelling.

- These techniques could have wide applicability.

- At this stage, NSFD techniques are as much an art as science, requiring even more of the 'experiments' characteristic of FD scheme design.

- The ideas are comparatively recent, and improvements can be expected.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative Seismology: W.H. Freeman and Co., first edn.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, No. 01, 54–66.

Mickens, R. E., 2000, Application of nonstandard finite difference schemes: World Scientific Publishing Co., Singapore, first edn.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, Numerical recipies in C: Cambridge University Press, first edn.

Strikwerda, J. C., 2004, Finite difference schemes and partial differential equations: Society for Industrial and Applied Mathematics, Philadelphia, second edn.