Notes: Associated Equation to a Numerical Scheme

R. R. Rosales (MIT, Math. Dept., room 2-337, Cambridge, MA 02139)

April 26, 2022

Contents

1 The associated equation to a FD scheme. 1

1.1 Example: the "bad" scheme from the lectures. 2

1.2 Example: the "good" scheme from the lectures. 3

1.3 Example: implicit forward differences for $u_t + u_x = 0$. 4

1.4 Example: a nonlinear equation. 5

1.5 **Not yet written.** Example: backward differences for $u_t + u_x = u$. 6

List of Figures

1 The associated equation to a FD scheme.

Every finite differences numerical scheme for a pde has an associated equation (some times called *the model equation*). This equation describes the behavior of the scheme for *long wave* "inputs" — that is: their space dependence is on scales that much larger than the scheme grid size Δx. As long as this property is maintained by the scheme evolution,[1] the grid values of the numerical solution can be expanded in Taylor series of Δt and Δx. Substituting these expansions into the scheme equation(s), and keeping the leading order *beyond* consistency, provides an equation which is a "small" perturbation of the equation that the scheme is intended to solve — *the AENS.* In this equation, the small terms kept beyond consistency give an *indication of what is the behavior of the leading order error produced by the numerical scheme.* If these terms amplify short scale perturbations, then the scheme is unstable — but we can use the knowledge thus gained to correct the scheme, and make it stable. Even if the scheme is stable, knowledge of the type of leading error is useful as a guide on how to improve the method.

We illustrate the points made above with a few examples in what follows. Note that

1. The associated equation to a numerical scheme **does not provide an absolute test of numerical stability.** It is possible for a scheme to be such that high frequencies are amplified (instability), while the low frequencies that the AENS explores behave properly. On the other hand, *if the long wave AENS analysis shows instability . . . then the scheme is unstable.*

2. In spite of the above, the associated equation often tells us a lot about what is wrong with a numerical scheme, and provides good hints as to how an unstable scheme can be fixed, or a stable one be improved.

3. Advantages of the AENS approach over the von Neumann stability analysis are:
3a. *It is simpler to do.*
3b. *It can be used with both nonlinear and with variable coefficient equations.*
Of these, the second is by far the most important advantage.

Of course, the **disadvantage** is that the AENS furnishes less information about a scheme than a von Neumann stability analysis.

[1] It can be shown that this is, generally, true for stable schemes. For unstable schemes, which (generally) amplify the short (grid) scales modes, this property is, of course, quickly lost.

1.1 Example: the "bad" scheme from the lectures.

Consider the "bad" scheme used in the GBNS_lecture script of the 18.311 MatLab Toolkit. This scheme is supposed to solve the wave equation, written in the non-dimensional form (with wave speed $c = 1$)

$$u_t = v \qquad \text{and} \qquad v_t = u_{x\,x}, \tag{1.1}$$

Scheme description: Pick a *uniform grid in space and time,* with "small" grid spacings Δx and Δt — that is: $\boldsymbol{x_{n+1} = x_n + \Delta x}$ and $\boldsymbol{t_{m+1} = x_m + \Delta t}$. On this grid the *scheme approximates the solution by the grid functions* u_n^m *and* v_n^m

$$u(x_n,\, t_m) \approx u_n^m \quad \text{and} \quad v(x_n,\, t_m) \approx v_n^m, \tag{1.2}$$

where the grid functions satisfy the following [2] discretized version of the equations in (1.1)

$$\frac{u_n^{m+1} - u_n^m}{\Delta t} = v_n^m \qquad \text{and} \qquad \frac{v_n^{m+1} - v_n^m}{\Delta t} = \frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{(\Delta x)^2}. \tag{1.3}$$

Remark 1.1 *The idea behind this is as follows: if* $u_n^m = U(x_n,\, t_m)$ *and* $v_n^m = V(x_n,\, t_m)$, *for some "smooth" functions* $U = U(x,\, t)$ *and* $V = V(x,\, t)$, *then the equations in (1.3) indicate* [3] *that* U *and* V *(approximately) satisfy the wave equation system (1.1) — with errors that vanish as* Δt *and* Δx *vanish.* **At least, this is the idea, but it does not always work.** *The fact that:*

> *substituting* U, V *as above and expanding in Taylor series in the small quantities* Δx, Δt, *yields the equations we want to solve (plus small errors vanishing as* Δx *and* Δt *vanish)*

is a **necessary condition** *for the scheme to work* **(consistency)**. *But it is* **not sufficient.** ♣

To see what can go wrong, next we calculate the leading order error terms in the Taylor expansions mentioned in remark **1.1**. These will show that U and V above, *if they existed,* would satisfy a wave equation system (i.e.: (1.1)) with "small" correction terms, whose effect happens to be rather dramatic. This corrected system of equations, (1.10 – 1.11), are the **Associated Equations to the Numerical Scheme** in (1.3).

Remark 1.2 *Note that we say in the paragraph above "if they existed," when referring to* U *and* V. *This is a key fact: the argument in remark* ***1.1*** *to show that the equations in (1.3) provide an approximation to the wave equation system (1.1), makes sense only if we have such smooth* U *and* V. *However, as we show below using the associated equation approach, this cannot be true!* ♣

We proceed now with **constructing the associated equations for (1.3).** Thus we assume that we can write $u_n^m = U(x_n\, t_m)$ and $v_n^m = V(x_n,\, t_m)$, for some smooth functions $U = U(x,\, t)$ and $V = V(x,\, t)$ — for the implications of this assumption, see remark **1.3**. It then follows that

$$u_n^m = U, \tag{1.4}$$

$$u_n^{m+1} = U + (\Delta t)\, U_t + \frac{1}{2!}\, (\Delta t)^2\, U_{t\,t} + \dots, \tag{1.5}$$

$$u_{n+1}^m = U + (\Delta x)\, U_x + \frac{1}{2!}\, (\Delta x)^2\, U_{x\,x} + \frac{1}{3!}\, (\Delta x)^3\, U_{x\,x\,x} + \frac{1}{4!}\, (\Delta x)^4\, U_{x\,x\,x\,x} + \dots, \tag{1.6}$$

$$u_{n-1}^m = U - (\Delta x)\, U_x + \frac{1}{2!}\, (\Delta x)^2\, U_{x\,x} - \frac{1}{3!}\, (\Delta x)^3\, U_{x\,x\,x} + \frac{1}{4!}\, (\Delta x)^4\, U_{x\,x\,x\,x} + \dots, \tag{1.7}$$

where U and its derivatives are all evaluated at $(x_n,\, t_m)$ (similar expansions can be done for the v's). Substituting these expansions into the equations in (1.3), we obtain:

$$U_t + \frac{1}{2!}\, (\Delta t)\, U_{t\,t} + O((\Delta t)^2) = V, \tag{1.8}$$

$$V_t + \frac{1}{2!}\, (\Delta t)\, V_{t\,t} + O((\Delta t)^2) = U_{x\,x} + \frac{2}{4!}\, (\Delta x)^2\, U_{x\,x\,x\,x} + O((\Delta x)^4). \tag{1.9}$$

[2] This scheme is called the "bad" numerical scheme in the MatLab scripts.

[3] Expand $u_n^{m+1} = U(x_n,\, t_m + \Delta t)$, $u_{n+1}^m = U(x_n + \Delta x,\, t_m)$, etc., in a Taylor series centered at $(x_n,\, t_m)$.

Now we use that $U_t = V + O(\Delta t)$ and $V_t = U_{x\,x} + O(\Delta t,\, (\Delta x)^2)$, as given by these last equations, to write

$$U_{t\,t} = U_{x\,x} + O(\Delta t,\, (\Delta x)^2) \quad \text{and} \quad V_{t\,t} = V_{x\,x} + O(\Delta t,\, (\Delta x)^2).$$

Substituting these equalities into the equations, we obtain the **Associated Equations to the Numerical Scheme in (1.3)**

$$U_t = V - \frac{1}{2!}\,\Delta t\, U_{x\,x} + O((\Delta t)^2,\, (\Delta t)(\Delta x)^2), \tag{1.10}$$

$$V_t = U_{x\,x} - \frac{1}{2!}\,\Delta t\, V_{x\,x} + O((\Delta t)^2,\, (\Delta x)^2). \tag{1.11}$$

Now, the key point:

> the perturbation to the wave equation system (1.1) in these equations is, at leading order, a **negative diffusion!**

That is: the terms $U_{x\,x}$ and $V_{x\,x}$ are multiplied by a negative coefficient. This means that *these terms induce a behavior that is equivalent to the situation that would arise in a universe where heat flowed from cold to hot,* instead of from hot to cold. In this universe any small nonuniformity in the temperature distribution would grow immediately without bounds, in a run-away process leading to the formation of singularities in very short time.[4] Thus we see that, generally, **there are no smooth functions U and V satisfying the assumptions made in remark 1.1** — see remark **1.2**. If there were, they would have to satisfy (1.10 – 1.11) — which (generally) does not have smooth solutions (or even solutions!).

> Systems such as this are called **ill-posed.** We will talk more about them later in the course. Main point is: whenever one occurs, something, somewhere, is very wrong!

In **conclusion:** the analysis above shows that *the scheme given by (1.3) cannot possibly work.* It also suggests a way to fix the problem: add terms that "neutralize" the negative diffusion that shows up in the associated system of equations (1.10 – 1.11) — an example is given in § 1.2.

1.2 Example: the "good" scheme from the lectures.

Modify the scheme in (1.3) as follows

$$\frac{u_n^{m+1} - u_n^m}{\Delta t} = v_n^m + \frac{1}{2}\,\delta\,\Delta t\,\frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{(\Delta x)^2}, \tag{1.12}$$

$$\frac{v_n^{m+1} - v_n^m}{\Delta t} = \frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{(\Delta x)^2} + \frac{1}{2}\,\delta\,\Delta t\,\frac{v_{n+1}^m - 2v_n^m + v_{n-1}^m}{(\Delta x)^2}, \tag{1.13}$$

where $\delta > 1$ is a constant.[5] Here we have added **numerical dissipation to stabilize the scheme.** This works: provided $\lambda = \Delta t/\Delta x$ is kept below some critical value λ_c (which depends on δ), this last scheme is stable.[6] Since it is also consistent, its solutions converge to the solutions of the wave equation as Δt and Δx vanish.

The associated equations for this new scheme are easy to construct, since we only need to add to (1.10–1.11) the contributions that result from the extra terms added to (1.3) to obtain (1.12–1.13). At leading order these extra terms yield $\frac{1}{2}\,\delta\,\Delta t\, U_{x\,x}$ and $\frac{1}{2}\,\delta\,\Delta t\, V_{x\,x}$, which yields (1.14–1.15).

Remark 1.3 *Earlier we mentioned that the associated equation approach provides only a partial answer to the issue of numerical stability, because it only looks at the behavior of "long" waves in the numerical solution. By long waves here we mean: disturbances whose wavelength and period are much larger than the grid spacings Δx and Δt. That this is so should be obvious, for once we assume that $u_n^m = U(x_n,\, t_m)$ and $v_n^m = V(x_n,\, t_m)$ [for some "smooth"*

[4] In fact, arbitrarily short!

[5] For $\delta = 2$ this scheme is the "good" scheme used in the GBNS_lecture script of the 18.311 MatLab Toolkit.

[6] This is shown in the introduction to the vNSA problem series (von Neumann Stability Analysis).

*functions $U = U(x,\, t)$ and $V = V(x,\, t)$ — as done in remark **1.1**] we exclude from consideration any possibility of "high" frequency oscillations in the numerical solution, with wavelengths and periods close to Δx and Δt.*

It is possible to have an unstable numerical scheme, with unbounded growth of high frequency oscillations, such that the associated equation is completely "oblivious" to this fact. An example is provided by the scheme above in (1.12 – 1.13). The associated equations for this scheme, given by

$$U_t = V + \frac{1}{2}(\delta - 1)\,\Delta t\, U_{x\,x} + O((\Delta t)^2,\, (\Delta t)(\Delta x)^2), \tag{1.14}$$

$$V_t = U_{x\,x} + \frac{1}{2}(\delta - 1)\,\Delta t\, V_{x\,x} + O((\Delta t)^2,\, (\Delta x)^2), \tag{1.15}$$

show no difficulties as long as $\delta > 1$. However, unless one takes $\lambda = \Delta t/\Delta x < \lambda_c$ — for some $\lambda_c = \lambda_c(\delta) > 0$ — the scheme is unstable!

How can we see that the instability mentioned in the prior paragraph exists, in a fast and easy way? *The answer is simple: just as the associated equation approach looks at the behavior of a scheme in the "long wave" limit, we can look directly at the other end (short waves). It is clear that the shortest waves a numerical grid can support are those where the solution alternates (in space) between two values. Thus we look for special solutions for the scheme in (1.12 – 1.13), of the form:*

$$u_n^m = U\, G^m\, (-1)^n \quad \text{and} \quad v_n^m = V\, G^m\, (-1)^n\,, \tag{1.16}$$

where U, V, and G, are constants to be determined. Substituting into (1.12 – 1.13), we get the eigenvalue problem:

$$\left.\begin{array}{rcrcr} G\,U &=& (1 - 2\,\delta\,\lambda^2)\,U &+& \Delta t\, V \\ G\,V &=& -4\,\nu\, U &+& (1 - 2\,\delta\,\lambda^2)\,V \end{array}\right\}, \tag{1.17}$$

where $\lambda = \Delta t/\Delta x$ and $\nu = \Delta t/(\Delta x)^2$. Thus $G = 1 - 2\,\delta\,\lambda^2 \pm 2\,i\,\lambda$, which yields

$$|G^2| = 1 - 4\,(\delta - 1)\,\lambda^2 + 4\,\delta^2\,\lambda^4 < 1 \;\; \textit{as long as} \;\; \delta^2\lambda^2 < \delta - 1.$$

This turns out to be, precisely, the stability condition for the scheme (1.12 – 1.13). That is:

$$\lambda < \lambda_c = \frac{\sqrt{\delta - 1}}{\delta}, \quad \textit{where} \quad \delta > 1. \tag{1.18}$$

The approach in (1.16) is a special case of the von Neumann stability analysis. However, by restricting the solutions to the shortest possible wave length, it turns out that this technique can be generalized to some situations that are not linear and constant coefficients. We will not pursue these extensions here. ♣

1.3 Example: implicit forward differences for $u_t + u_x = 0$.

Here we consider the scheme

$$\frac{u_n^{m+1} - u_n^m}{\Delta t} + \frac{u_{n+1}^{m+1} - u_n^{m+1}}{\Delta x} = 0. \tag{1.19}$$

This scheme results from replacing the time and space partial derivatives in the equation

$$u_t + u_x = 0, \tag{1.20}$$

by the approximations $u_t(x,\, t) \approx \frac{u(x,\,t) - u(x,\,t-\Delta t)}{\Delta t}$ and $u_x(x,\, t) \approx \frac{u(x+\Delta x,\,t) - u(x,\,t)}{\Delta x}$.

We now compute **the associated equation to (1.19).** Thus we assume that, for some smooth function $U = U(x,\, t)$, we can write $u_n^m = U(x_n\, t_m)$. It follows that

$$u_n^m = U - (\Delta t)\, U_t + \tfrac{1}{2}\,(\Delta t)^2\, U_{t\,t} + O((\Delta t)^3), \tag{1.21}$$

$$u_n^{m+1} = U, \tag{1.22}$$

$$u_{n+1}^{m+1} = U + (\Delta x)\, U_x + \tfrac{1}{2}\,(\Delta x)^2\, U_{x\,x} + O((\Delta x)^3), \tag{1.23}$$

where U and its derivatives are all evaluated at $(x_n,\, t_{m+1})$ — see note **1.1**.

Substituting these expansions into the scheme equation in (1.19), we obtain:

$$U_t + U_x = \frac{1}{2}\,\Delta t\, U_{t\,t} - \frac{1}{2}\,\Delta x\, U_{x\,x} + O((\Delta t)^2,\, (\Delta x)^2). \tag{1.24}$$

Now we use that $U_t = -U_x + O(\Delta t,\, \Delta x) \implies U_{t\,t} = U_{x\,x} + O(\Delta t,\, \Delta x)$, as follows by this last equation,[7] to eliminate $U_{t\,t}$ from (1.24). This yields the **the associated equation to (1.19)**

$$U_t + U_x = \frac{1}{2}\,(\Delta t - \Delta x)\, U_{x\,x} + O((\Delta t)^2,\, (\Delta t)^2,\, \Delta t\, \Delta x). \tag{1.25}$$

In order *for this equation to be well posed,* and avoid a backward heat equation situation, we require

$$\Delta t > \Delta x. \tag{1.26}$$

This, as it happens, is exactly *the stability condition for the scheme in (1.19)*, as shown in the introduction to the *vNSA problem series* (von Neumann Stability Analysis). Note also that *(1.26) is the CFL condition.*

Note 1.1 *Exactly which point in a scheme's stencil is used to center the Taylor expansions at does not matter in terms of the final answer. However, some choices are better than others in terms of the amount of calculation needed to get the AENS.*

1.4 Example: a nonlinear equation.

Here we consider the scheme

$$\frac{u_n^{m+1} - u_n^m}{\Delta t} + \frac{u_n^m + (u_n^m)^3 - u_{n-1}^m - (u_{n-1}^m)^3}{\Delta x} = 0. \tag{1.27}$$

This scheme results from replacing the time and space partial derivatives in the equation

$$u_t + (\underbrace{u + u^3}_{f})_x = 0, \tag{1.28}$$

by the approximations $u_t(x,\, t) \approx \frac{u(x,\, t+\Delta t) - u(x,\, t)}{\Delta t}$ and $f_x(x,\, t) \approx \frac{f(x,\, t) - f(x-\Delta x,\, t)}{\Delta x}$.

We now compute **the associated equation to (1.27).** As usual, assume that (for some smooth function $U = U(x,\, t)$) we can write $u_n^m = U(x_n\, t_m)$. It follows that

$$\begin{aligned} u_n^{m+1} &= U + (\Delta t)\, U_t + \tfrac{1}{2}\,(\Delta t)^2\, U_{t\,t} + O((\Delta t)^3), && (1.29)\\ u_n^m &= U, && (1.30)\\ u_{n-1}^m &= U - (\Delta x)\, U_x + \tfrac{1}{2}\,(\Delta x)^2\, U_{x\,x} + O((\Delta x)^3), && (1.31) \end{aligned}$$

where U and its derivatives are all evaluated at $(x_n,\, t_m)$.

Substituting these expansions into the equation in (1.27), after a bit of work, we obtain:

$$U_t + (U + U^3)_x = -\frac{1}{2}\,\Delta t\, U_{t\,t} + \frac{1}{2}\,\Delta x\,(U + U^3)_{x\,x} + O((\Delta t)^2,\, (\Delta x)^2). \tag{1.32}$$

Clearly, we can write $U_t = -(1 + 3\,U^2)\, U_x + O(\Delta t,\, \Delta x)$. Writing the derivative of this with respect to both t and x, and eliminating $U_{t\,x}$ from the resulting equations, yields $U_{t\,t} = ((1 + 3\,U^2)^2\, U_x)_x + O(\Delta t,\, \Delta x)$. Substituting this into (1.32) yields the **the associated equation to (1.27)**

$$U_t + U_x = \left(\underbrace{\frac{1}{2}\,(1 + 3\,U^2)\,\Big(\Delta x - \Delta t\,(1 + 3\,U^2)\Big)}_{\mu}\, U_x\right)_x + O((\Delta t)^2,\, (\Delta t)^2,\, \Delta t\, \Delta x). \tag{1.33}$$

[7] Differentiate $U_t = -U_x + \ldots$, to obtain $U_{t\,t} = -U_{x\,t} + \ldots$ and $U_{t\,x} = -U_{x\,x} + \ldots$. Then eliminate $U_{t\,x}$.

In order *for this equation to be well posed,* and avoid a backward heat equation type of situation, the diffusion coefficient μ must be positive (everywhere). Thus

$$\Delta t < \frac{1}{\max(1 + 3\,U^2)}\,\Delta x. \tag{1.34}$$

Since $1 + 3\,u^2$ is the characteristic speed for (1.28), *this is the CFL condition for the scheme in (1.27).* It can also be shown that this condition guarantees stability.

Remark 1.4 *As an example generalization of the calculation in (1.16–1.18) to a nonlinear problem: It is easy to see that the scheme equations (1.27) have "short wave" solutions of the form $u_n^m = (-1)^n\,g_m$, where the g_m are determined by*

$$g_{m+1} = \underbrace{\big(1 - 2\,\lambda\,(1 + g_m^2)\big)}_{\xi_m}\,g_m, \quad \textit{with } \lambda = \frac{\Delta t}{\Delta x}. \tag{1.35}$$

For these solutions to be damped (needed for stability), it must be that $-1 < \xi_m < 1$, that is:

$$\Delta t \le \frac{\Delta x}{1 + g_m^2}. \tag{1.36}$$

Since $1 + g_m^2 \le 1 + 3\,g_m^2$, (1.34) guarantees (1.34). ♣

1.5 Example: backward differences for $u_t + u_x = u$.

This subsection is yet to be written.

THE END.