M2AN - Modélisation mathématique et analyse numérique

Romuald Carpentier
Armel de La Bourdonnaye
Bernard Larrouturou

On the derivation of the modified equation for the analysis of linear numerical methods

M2AN - Modélisation mathématique et analyse numérique, tome 31, n^{o} 4 (1997), p. 459-470
<http://www.numdam.org/item?id=M2AN_1997__31_4_459_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 31, n° 4, 1997, p 459 à 470)

ON THE DERIVATION OF THE MODIFIED EQUATION FOR THE ANALYSIS OF LINEAR NUMERICAL METHODS (*)

by Romuald CARPENTIER (1), Armel de LA BOURDONNAYE (1) and Bernard LARROUTUROU (1), (2)

Abstract — *The modified equation is a powerful tool for the error analysis of the numerical solution of partial differential equations We present here a method which considerably simplifies the derivation of this equation in the linear case Our method uses formal expansions, with no elimination step , it keeps the same simplicity when multistep Runge-Kutta schemes are used and in any space dimensions*

Key words Evolution Partial Differential Equations, Modified equations, Error analysis
AMS subject classifications : 65M06, 65M12, 76M20

Abstract — *L'équation équivalente est un outil puissant d'analyse d'erreur pour la résolution numérique d'équations aux dérivées partielles Nous présentons une méthode qui simplifie considérablement l'obtention de cette équation dans le cas linéaire La méthode présentée utilise des séries formelles et ne nécessite aucune étape d'élimination , elle garde la même simplicité lorsque l'on utilise des schémas de Runge-Kutta et quelque soit la dimension spatiale*

1. INTRODUCTION

The modified equation technique, which was introduced by Warming and Hyett [12], is a powerful tool for the analysis of the accuracy and stability of a numerical method aimed at solving a time-dependent problem governed by an evolution partial differential equation. For constant-coefficients linear partial differential equations, it allows a detailed analysis of the truncation error of the numerical methods. In particular, the effect, either dissipative or dispersive, of each error term can be interpreted using the modified equation, so that it allows detailed comparisons between different numerical methods ; it may also sometimes be used as a tool for designing new numerical schemes (see e.g. [1, 12]). Lastly, the modified equation may also be used for the numerical analysis of some constant-coefficients nonlinear equations (see e.g. [7, 8, 9]), although the interpretation of the truncation error terms is less easy in the nonlinear case.

(*) Manuscript received September 27, 1995 , revised May 6, 1996
(1) CERMICS, B P 93, 06902 Sophia-Antipolis Cedex France
(2) École polytechnique, 91128 Palaiseau Cedex France

Let us briefly recall how the modified equation is derived, on a very simple example. Consider the explicit first-order upwind scheme :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = - c \frac{u_j^n - u_{j-1}^n}{\Delta x}, \tag{1}$$

for the solution of the wave equation $w_t = - c w_x$, with $c > 0$; in (1), j and n are the spatial and temporal indices respectively, Δx and Δt are the mesh size and the time step, so that u_j^n is an approximation of $w(j\,\Delta x, n\,\Delta t)$.

The modified equation for the scheme (1) is a *formal* partial differential equation, which is derived from the difference equation :

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = - c \frac{u(x, t) - u(x - \Delta x, t)}{\Delta x}, \tag{2}$$

which mimics (1). Assuming that u is $\mathscr{C}^\infty$ in (2), one can deduce from (2) the following Taylor expansions at point (x, t) :

$$u_t + \frac{\Delta t}{2} u_{tt} + \frac{\Delta t^2}{6} u_{ttt} + \cdots = - c\left(u_x - \frac{\Delta x}{2} u_{xx} + \frac{\Delta x^2}{6} u_{xxx} + \cdots \right). \tag{3}$$

We will call this equation the *unresolved* modified equation. The goal is now to transform this equation (3), for Δt and Δx small, by replacing the time derivatives, except the first one, by spatial derivatives, using successive differentiations and substitutions. For instance, taking the partial derivative of (3) with respect to t and x, we obtain :

$$u_{tt} + \frac{\Delta t}{2} u_{ttt} + \mathcal{O}(\Delta t^2) = - c\left(u_{xt} - \frac{\Delta x}{2} u_{xxt} + \mathcal{O}(\Delta x^2) \right), \tag{4}$$

$$u_{tx} + \frac{\Delta t}{2} u_{ttx} + \mathcal{O}(\Delta t^2) = - c\left(u_{xx} - \frac{\Delta x}{2} u_{xxx} + \mathcal{O}(\Delta x^2) \right), \tag{5}$$

and we can eliminate the mixed derivative u_{tx} from these relations to get :

$$u_{tt} = c^2 u_{xx} + \frac{\Delta t}{2} (c u_{ttx} - u_{ttt}) + \frac{\Delta x}{2} (c u_{xxt} - c^2 u_{xxx}) + \mathcal{O}(\Delta t, \Delta x)^2, \tag{6}$$

from which, setting $\nu = \frac{c\,\Delta t}{\Delta x}$, we deduce a first form of the modified equation :

$$u_t = - c u_x + \frac{c\,\Delta x}{2} (1 - \nu) u_{xx} + \mathcal{O}(\Delta t, \Delta x)^2. \tag{7}$$

Differentiating again (4) and (5) with respect to time and space makes it possible to further eliminate the mixed time-space derivatives. After several steps, we finally obtain :

$$u_t = -cu_x + \frac{c\,\Delta x}{2}(1-\nu)\,u_{xx} - \frac{c\,\Delta x^2}{6}(2\nu^2 - 3\nu + 1)\,u_{xxx} + \frac{c\,\Delta x^3}{24}(6\nu^3 - 12\nu^2 + 7\nu - 1)\,u_{xxxx} + \mathcal{O}(\Delta t, \Delta x)^4. \quad (8)$$

This is the modified equation, expanded up to order three in Δt and Δx. Formally, this is the partial differential equation which is actually solved by the numerical method (1). This equation shows the different terms of the truncation error of the numerical method and their interpretation (we see in (8) the dissipative first-order term, the dispersive second-order term and the dissipative third-order term) ; in particular, the modified equation (8) shows that the scheme (1) is first-order accurate, and it gives a necessary condition $(\nu \leqslant 1)$ for the stability of the method.

If the final equation (8) is really of interest for the numerical analysis of the scheme (1), it appears however that its derivation is quite heavy and lengthy, even if the successive differentiations and eliminations can be handled using a symbolic computer algebra system as in [10]. In particular, the elimination process which leads from the *unresolved* equation (3) to the *resolved* modified equation (8) may well become much more intricate than in the above example when less simple schemes are considered, for instance in higher space dimensions or with multistep time integration methods. In such cases, even writing the difference equation (2) or the *unresolved* modified equation (3) may become a non trivial task : indeed, a spatially second order accurate scheme uses 5 points in one space dimension, but 9 points in two dimensions, and 33 points with a second-order Runge-Kutta scheme !

It is precisely the objective of this work to present a much simpler way of deriving the modified equation for a linear numerical method. Our method uses formal series expansions without any elimination step ; moreover, it has the advantage of keeping the same simplicity when multistep Runge-Kutta or predictor-corrector schemes are employed, and in any space dimensions. In [4], Chang used the same method as ours, to prove the existence of a modified equation under some restrictive asumptions, but he still used the classical algebraic elimination method for practical purposes. In our paper, we present a new derivation of the method which is constructive in the sense that it allows to actually derive modified equations.

2. THE MAIN RESULT

Our method for deriving the modified equation applies to any constant-coefficients linear numerical method. In some way it follows what Shokin does symbolically for exhibiting what he calls Γ and Π forms of a finite difference scheme in [11]. Let us consider a linear evolution partial differential equation of the following form, in one space dimension :

$$w_t = \sum_{K \geqslant 0} \gamma_K \frac{\partial^K w}{\partial x^K}, \tag{9}$$

where the right-hand-side summation is finite, and assume that the equation (9) is approximated on a uniform mesh using the explicit scheme :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sum_k A_k(\Delta_x)\, u_{j+k}^n, \tag{10}$$

(again with a finite right-hand-side summation). Then, introducing the function :

$$g_{\Delta x}(X) = \sum_k A_k(\Delta_x)\, e^{k\,\Delta x X}, \tag{11}$$

we state our main result :

PROPOSITION 1 : *Assume that the scheme* (10) *is consistent with equation* (9) *(in the classical finite-difference sense).*

Then the modified equation of the scheme (10) *writes :*

$$u_t = \sum_{k \geqslant 0} \alpha_k(\Delta t, \Delta x) \frac{\partial^k u}{\partial x^k}, \tag{12}$$

where $\sum_{k \geqslant 0} \alpha_k(\Delta t, \Delta x)\, X^k$ *is the formal series expansion of the function :*

$$\mathscr{F}(X) = \frac{\log\left(1 + \Delta t\, g_{\Delta x}(X)\right)}{\Delta t}. \quad \blacksquare \tag{13}$$

The proof starts with the following Lemma :

LEMMA 1 : *Assume that the scheme* (10) *is consistent with equation* (9), *and set :*

$$g_0(X) = \sum_{K \geqslant 0} \gamma_K X^K. \tag{14}$$

Then, the difference $g_{\Delta x}(X) - g_0(X)$ *formally tends to 0 as* Δx *tends to 0.* $\blacksquare$

Proof : The consistency of the scheme (10) implies that, formally :

$$\lim_{\Delta x \to 0} \left(\sum_k A_k(\Delta x)\, v(x + k\,\Delta x) - \sum_{K \geq 0} \gamma K \frac{d^K v}{dx^K} \right) = 0 , \tag{15}$$

for any $\mathscr{C}^\infty$ function $v(x)$. For $X \in \mathbb{R}$, we may take $v(x) = e^{Xx}$, so that (15) yields :

$$\lim_{\Delta x \to 0} \left(\sum_k A_k(\Delta x)\, e^{k\,\Delta x X} - \sum_{K \geq 0} \gamma K X^K \right) = 0 . \quad \blacksquare \tag{16}$$

Remark 1 : It is also useful to see the above proof with a slightly different point of view, using formal series expansions. The consistency of the scheme (10) is usually expressed through Taylor expansions, i.e. one writes the Taylor expansion :

$$\sum_k \sum_{p \geq 0} A_k(\Delta x) \frac{(k\,\Delta x)^p}{p\,!} \frac{d^p v}{dx^p} \tag{17}$$

of the first term in (15), and one says that the scheme is consistent if :

$$\sum_k A_k(\Delta x) \frac{(k\,\Delta x)^K}{K\,!} = \gamma_K + O(\Delta x) \quad \text{for all } K . \tag{18}$$

But obviously, (18) allows us to write :

$$\sum_k \sum_{p \geq 0} A_k(\Delta x) \frac{(k\,\Delta x X)^p}{p\,!} = \sum_{K \geq 0} \gamma_K X^K + O(\Delta x) , \tag{19}$$

for any X, which yields (16). $\blacksquare$

We can now achieve the proof of Proposition 1 ; it relies on applying the Fourier transform to the difference equation which mimics the numerical scheme.

Proof of Proposition 1 : Assume that $u(x, t)$ is bounded and satisfies the difference equation :

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = \sum_k A_k(\Delta_x)\, u(x + k\,\Delta x, t) . \tag{20}$$

We may introduce the Fourier transform $\hat{u}(\xi, \tau)$ of $u(x, t)$, and we get from (11) and (20) :

$$\left(\frac{e^{i\tau\,\Delta t} - 1}{\Delta t} \right) \hat{u}(\xi, \tau) = g_{\Delta x}(i\xi)\, \hat{u}(\xi, \tau) . \tag{21}$$

We can observe here that expanding both sides of (21) in formal series gives

$$\sum_{p \geq 1} (i\tau)^p \frac{\Delta t^{p-1}}{p!} \hat{u}(\xi . \tau) = \sum_k \sum_{p \geq 0} A_k(\Delta x) \frac{(k \Delta x)^p}{p!} (i\xi)^p \hat{u}(\xi . \tau), \quad (22)$$

whose inverse Fourier transform yields the *unresolved* modified equation

$$\sum_{p \geq 1} \frac{\Delta t^{p-1}}{p!} \frac{\partial^p u}{\partial t^p} = \sum_k \sum_{p \geq 0} A_k(\Delta x) \frac{(k \Delta x)^p}{p!} \frac{\partial^p u}{\partial x^p} \quad (23)$$

However, the *resolution* is now elementary (and writing the expansions (22) and (23) is not even useful) (21) tells us that the distribution $\hat{u}(\xi, \tau)$ vanishes except on the manifold $\mathscr{V}$ defined by the relation $\frac{e^{i\tau \Delta t} - 1}{\Delta t} = g_{\Delta x}(i\xi)$ Since we show below that the term $\Delta t g_{\Delta x}(i\xi)$ is small when Δt and Δx are small, and since the exponential function is bijective in the neighbourhood of 0, the manifold $\mathscr{V}$ is also defined for Δt and Δx small by the relation $i\tau = \frac{\log(1 + \Delta t\, g_{\Delta x}(i\xi))}{\Delta t} = \mathscr{F}(i\xi)$, where log denotes here the local inverse of the exponential function, which implies that the classical expansion $\log(1 + \delta) = \sum_{p \geq 1} (-1)^{p-1} \frac{\delta^p}{p}$ holds true We then get

$$i\tau\hat{u}(\xi, \tau) - \mathscr{F}(i\xi)\, \hat{u}(\xi, \tau) = i\tau\hat{u}(\xi, \tau) - \sum_k \alpha_k(\Delta t, \Delta x)(i\xi)^k \hat{u}(\xi, \tau) = 0, \quad (24)$$

and the inverse Fourier transform immediately gives the *resolved* modified equation (12)

It only remains to explain why the expansion in formal series is valid in (24), i e why the term $\Delta t\, g_{\Delta x}(i\xi)$ is formally small when Δt and Δx are small (notice indeed that the scheme coefficients $A_k(\Delta x)$ involve negative powers of the mesh spacing Δx, as in (1), so that difficulties may arise for small Δx) It follows indeed from (19) that $g_{\Delta x}(X)$ can be expanded in formal series under the form $g_{\Delta x}(X) = g_0(X) + \sum_{p \geq 1} \Delta x^p g_p(X)$, where the $g_p(X)$ are polynomials in X

The fraction $\mathcal{F}(X)$ then takes the form :

$$\mathcal{F}(X) = \frac{\log\left(1 + \Delta t g_0(X) + \Delta t \sum_{p \geqslant 1} \Delta x^p g_p(X)\right)}{\Delta t}, \tag{25}$$

and it is perfectly valid to expand it when Δt and Δx are small under the form :

$$\mathcal{F}(X) = g_0(X) + \sum_{\substack{p+q \geqslant 1 \\ r \geqslant 0}} \beta_{p,q,r} \Delta x^p \Delta t^q X^r \stackrel{def}{=} \sum_{k \geqslant 0} \alpha_k(\Delta t, \Delta x) X^k. \quad \blacksquare \tag{26}$$

Remark 2 : The proof of Proposition 1 is both more rigorous and more constructive than the elimination method described in Section 1. In particular, it clearly gives the truncation error of the numerical scheme (10), since the modified equation (12) finally takes the form :

$$u_t = \sum_{K \geqslant 0} \gamma_K \frac{\partial^K u}{\partial x^K} + \sum_{\substack{p+q \geqslant 1 \\ r \geqslant 0}} \beta_{p,q,r} \Delta x^p \Delta t^q \frac{\partial^r u}{\partial x^r}. \quad \blacksquare \tag{27}$$

3. SOME EXTENSIONS

The result of Proposition 1, which deals with explicit schemes in one-space dimension, using first-order accurate time integration, can be easily extended in several directions. We now describe some of them. Others may be found in [2].

3.1. Multi-dimensional schemes

Proposition 1 can be extended with no difficulty to linear constant-coefficients numerical methods in two or three space dimensions. For instance, let us consider the following partial differential equation, in two space dimensions :

$$w_t = \sum_{K, M \geqslant 0} \gamma_{K,M} \frac{\partial^{K+M} w}{\partial x^K \partial y^M}, \tag{28}$$

approximated with the explicit scheme :

$$\frac{u_{j,l}^{n+1} - u_{j,l}^n}{\Delta t} = \sum_{k,m} A_{k,m}(\Delta_x, \Delta y)\, u_{j+k,l+m}^n. \tag{29}$$

We can then state (referring to [2] for the proof) :

PROPOSITION 2 : *Assume that the scheme* (29) *is consistent with equation* (28). *Then, its modified equation writes :*

$$u_t = \sum_{k,m \geqslant 0} \alpha_{k,m}(\Delta t, \Delta x, \Delta y) \frac{\partial^{k+m} u}{\partial x^k \partial y^m}, \tag{30}$$

where $\sum_{k,m \geqslant 0} \alpha_{k,m}(\Delta t, \Delta x, \Delta y) X^k Y^m$ *is the formal series expansion of the function :*

$$\mathcal{F}(X) = \frac{\log(1 + \Delta t g_{\Delta x \Delta y}(X, Y))}{\Delta t}, \tag{31}$$

where $g_{\Delta x, \Delta y}(X, Y)$ *is the following functions of two variables :*

$$g_{\Delta x, \Delta y}(X, Y) = \sum_{k,m} A_{k,m}(\Delta x, \Delta y) e^{k \Delta x \lambda} e^{m \Delta y Y}. \quad \blacksquare \tag{32}$$

3.2. Implicit schemes

The extension of Proposition 1 to implicit schemes is also straightforward. We can state (referring again to [2] for a proof) :

PROPOSITION 3 : *Assume that the scheme* (10) *is consistent with equation* (9).

Then the modified equation of the implicit scheme :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sum_k A_k(\Delta_x) u_{j+k}^{n+1} \tag{33}$$

writes :

$$u_t = \sum_{k \geqslant 0} \alpha_k(\Delta t, \Delta x) \frac{\partial^k u}{\partial x^k}, \tag{34}$$

where $\sum_{k \geqslant 0} \alpha_k(\Delta t, \Delta x) X^k$ *is the formal series expansion of the function :*

$$\mathcal{F}(X) = \frac{\log(1 - \Delta t\, g_{\Delta x}(X))}{(-\Delta t)}. \quad \blacksquare \tag{35}$$

3.3. Multistep schemes

Let us lastly show that the method can be extended while kepping its simplicity to Runge-Kutta schemes. We state here :

PROPOSITION 4 : *Assume that the scheme* (10) *is consistent with equation* (9). *When the* N^{th}*-order Runge-Kutta method is applied to the scheme* (10), *the modified equation writes :*

$$u_t = \sum_{k \geq 0} \alpha_k(\Delta t, \Delta x) \frac{\partial^k u}{\partial x^k}, \tag{36}$$

where $\sum_{k \geq 0} \alpha_k(\Delta t, \Delta x) X^k$ *is the formal series expansion of the function :*

$$\mathscr{F}(X) = \frac{\log\left(1 + \sum_{K=1}^{N} \frac{[\Delta t g_{\Delta x}(X)]^K}{K!}\right)}{\Delta t}. \blacksquare \tag{37}$$

Proof: Let us write the scheme (10) in condensed form as $\frac{u_j^{n+1} - u_j^n}{\Delta t} = (G(u^n))_j$. For the sake of simplicity, we will only consider the second-order Runge-Kutta scheme, which writes :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = (G(u^n))_j + \frac{\Delta t}{2}(G \circ G(u^n))_j. \tag{38}$$

Proposition 4 is then a consequence of the next Lemma. ■

LEMMA 2 : *Let* G^1 *and* G^2 *be two linear schemes, defined by*

$$\begin{cases} (G^1(u))_j^n = \sum_k A_k^1(\Delta_x)\, u_{j+k}^n \\ (G^2(u))_j^n = \sum_k A_k^2(\Delta_x)\, u_{j+k}^n \end{cases},$$

and let $g_{\Delta x}^1(X)$ *and* $g_{\Delta x}^2(X)$ *be the corresponding functions associated with* G^1 *and* G^2 *respectively using* (11).

Then, the function associated by (11) *with the composed scheme* $G^1 \circ G^2$ *is simply the product* $g_{\Delta x}^1(X)\, g_{\Delta x}^2(X)$. ■

Proof: It suffices to realize that

$$(G^1 \circ G^2(u^n))_j = \sum_k \sum_m A_k^1 A_m^2\, u_{j+k+m}^n,$$

so that the associated function is $g_{\Delta x}(X) = \sum_k \sum_m A_k^1 A_m^2 \, e^{(k+m)\Delta x X}$, and Lemma 2 readily follows. ■

4. EXAMPLES

In this section, we present some derivations of modified equations, using our technique. We use a second accuracy scheme in space for the resolution of the monodimensional wave equation. We can change the right member of (1) as :

$$S(u_j) = \frac{c}{4\,\Delta x}\left(-u_{j-2} + 5\,u_{j-1} - 3\,u_j - u_{j+1}\right)$$

and obtain immediatly the $g_{\Delta x}(X)$ function :

$$g_{\Delta x}(X) = \frac{c}{4\,\Delta x}\left(-e^{-2\Delta x X} + 5\,e^{-\Delta x X} - 3 - e^{\Delta x X}\right).$$

Now we vary the time discretisation scheme.

- First order explicit scheme :
  we use the following discretisation :

$$\frac{u_j^{n+1} - u_j^n}{dt} = S(u_j^n).$$

Using relation (13), we just have to expand the following expression :

$$\mathcal{F}(X) = \frac{\log\left(1 + \Delta t g_{\Delta x}(X)\right)}{\Delta t}.$$

Up to fourth order, the Taylor expansion of $\mathcal{F}$ (which can be obtained by symbolic algebra) gives the modified equation :

$$u_t = -\,cu_x - \frac{c\,\Delta x}{2}\,\nu u_{xx} + \frac{c\,\Delta x^2}{12}\left(1 - 4\,\nu^2\right) u_{xxx} + \frac{c\,\Delta x^3}{24}\left(-3 + 2\,\nu - 6\,\nu^3\right) u_{xxxx} + O(\Delta t, \Delta x)^4.$$

- First order implicit scheme :
  we use the following discretisation :

$$\frac{u_j^{n+1} - u_j^n}{dt} = S(u_j^{n+1})$$

that becomes the relation (35) for the inverse expression that we recall :

$$\mathcal{F}(X) = \frac{\log\left(1 + \Delta t g_{\Delta x}(X)\right)}{(-\Delta t)}$$

and the Taylor expansion with symbolic algebra give till four order :

$$u_t = -cu_x + \frac{c\,\Delta x}{2}\,\nu u_{xx} + \frac{c\,\Delta x^2}{12}\,(1 - 4\,\nu^2)\,u_{xxx}$$
$$+ \frac{c\,\Delta x^3}{24}\,(-3 - 2\,\nu + 6\,\nu^3)\,u_{xxxx} + O(\Delta t, \Delta x)^4.$$

- Second order Runge-Kutta scheme :
  we use now a multistep scheme like this :

$$\frac{u_j^{n+1} - u_j^n}{dt} = S(u_j^n) + \frac{\Delta t}{2}\,S \circ S(u_j^n)$$

that becomes the relation (37) for the inverse expression that we recall :

$$\mathcal{F}(X) = \frac{\log\left(1 + \Delta t g_{\Delta x}(X) + \frac{(\Delta t g_{\Delta x}(X))^2}{2}\right)}{\Delta t}$$

and the Taylor expansion with symbolic algebra give till four order :

$$u_t = -cu_x + \frac{c\,\Delta x^2}{12}\,(1 + \nu^2)\,u_{xxx} + \frac{c\,\Delta x^3}{8}\,(-1 + \nu^3)\,u_{xxxx} + O(\Delta t, \Delta x)^4.$$

In [3], you can find two dimensionnal modified equation obtained whith this method. The more difficult work is now a good discretisation and analysis of scheme...

5. CONCLUSIONS

We have presented a very simple method for the derivation of the modified equation of any linear numerical method solving an evolution constant-coefficients linear partial differential equation. The method is much simpler than the usual technique, which derives the modified equation through Taylor expansions by a lengthy substitution and elimination process. The modified equation can be explicitly derived using our method for any linear scheme involving two time levels, in any space dimensions and for various time integration methods, either by hand or using a computer system for symbolic algebra.

As a conclusion, it is useful to summarize our method by exhibiting its relation with the Von Neumann stability analysis. The above method is indeed as simple as, and very close to the method for evaluating the amplification factor in the stability analysis : inserting $u_j^n = G(i\xi)^n e^{ij\xi \Delta x}$ in the scheme (10), one obtains with our notations $G(i\xi) = 1 + \Delta t g_{\Delta x}(i\xi)$, that is :

$$G(i\xi) = \exp(\Delta t \mathscr{F}(i\xi)) \quad \text{or} \quad \mathscr{F}(i\xi) = \frac{\log[G(i\xi)]}{\Delta t}. \tag{39}$$

These relations, which hold more generally for all linear numerical methods examined in the previous sections, summarize our « recipe » for deriving the modified equation.

REFERENCES

[1] D. A. ANDERSON, J. C. TANNEHILL and R. H. PLETCHER, *Computational fluid mechanics and heat transfer* (Hemisphere, Mc Graw-Hill, 1984).

[2] R. CARPENTIER, A. DE LA BOURDONNAYE and B. LARROUTUROU, 1994, On the derivation of the modified equation for the analysis of linear numerical methods, *CERMICS Report* No 94-26.

[3] R. CARPENTIER, 1995, *Approximation et analyse numérique d'écoulement instationnaire. Application à des instabilités tourbillonnaires,* Thesis, Université de Nice.

[4] S. C. CHANG, 1990, A critical analysis of the modified equation technique of Warming and Hyett, *J. Comput. Phys.*, **86,** pp. 107-126.

[5] D. CHARGY, 1991, *Étude numérique d'écoulements réactifs transsoniques,* Thesis, ENPC, Paris.

[6] N. GLINSKY, 1990, *Simulation numérique d'écoulements hypersoniques réactifs hors-équilibre chimique,* Thesis, Université de Nice.

[7] J. GOODMAN and A. MAJDA, 1985, The validity of the modified equation for nonlinear shock waves, *J. Comput. Phys.*, **58,** pp. 336-348.

[8] S. LANTERI, 1991, *Simulation d'écoulements aérodynamiques instationnaires sur une architecture SIMD massivement parallèle,* Thesis, Université de Nice.

[9] R. PEYRET, 1977, Résolution numérique des systèmes hyperboliques. Application à la dynamique des gaz, *ONERA Report* 1977-5.

[10] M. SPIRIDONOVA and J. A. DESIDERI, *Symbolic computations for the analysis of finite-difference schemes by the modified equation approach* (unpublished).

[11] Yu. I. SHOKIN, *The Method of Differential Approximation* (Springer Verlag, 1983).

[12] R. F. WARMING and F. HYETT, 1974, The modified equation approach to the stability and accuracy analysis of finite-difference methods, *J. Comput. Phys.*, **14,** 2, p. 159.