Monografías Matemáticas *García de Galdeano* **35**, 19–26 (2010)

A NEW MODIFIED EQUATION APPROACH FOR SOLVING THE WAVE EQUATION

C. Agut, J. Diaz and A. Ezziani

Abstract. The main topic of this work is to provide a fast and accurate solution of the wave equation. We will present new numerical schemes based on the modified equation technique using a switch between the space discretization and the time one. Numerical results illustrate the performances of these methods with respect to the accuracy and the computational burden.

Keywords: High order schemes, discontinuous Galerkin method, acoustic wave equation.

AMS classification: 65M12, 65M60, 35L05.

§1. Introduction

The solution of the full wave equation implies very high computational burdens to get high accurate results. Indeed, to improve the accuracy of the numerical solution, one must considerably reduce the space step, which is the distance between two points of the mesh representing the computational domain. Obviously this results in increasing the number of unknowns of the discrete problem. Besides, the time step, whose value fixes the number of required iterations for solving the evolution problem, is linked to the space step through the CFL (Courant-Friedrichs-Lewy) condition. The CFL number defines an upper bound for the time step in such a way that the smaller the space step is, the higher the numbers of iterations and of discrete unknowns will be. In the three-dimensional case, the problem can have more than ten millions of unknowns which must be evaluated at each time-iteration. However, high-order numerical methods can be used for computing accurate solutions with larger space and time steps. Recently, Joly and Gilbert (cf. [1]), have optimized the modified equation technique, which was proposed by Shubin and Bell (cf. [3]) for solving the wave equation and it seems to be very promising providing some improvements. In this work, we apply this technique in a original way. Indeed, most of the works devoted to the solution of the wave equation consider first the space discretization of the system before addressing the question of the time discretization. We intends here to invert the discretization process by applying first the time discretization thanks to the modified equation and after to consider the space discretization. After the time discretization an additional bilaplacian operator appears and we have therefore to consider C^1 finite elements (such as the Hermite ones) or Discontinuous Galerkin finite elements whose C^1 continuity is enforced through an appropriate penalty term. We provide a numerical comparison of the performance of the new method in order to illustrate the gains of accuracy and computational burden.

§2. Modified Equation technique

In this section, we describe the classical modified equation technique and we recall its main properties.

We consider the wave equation in a bounded domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. We impose here a Neumann boundary condition of Ω but this study can be extended to other type of boundary condition without difficulty. Similarly, for a sake of simplicity, we do not consider any source term:

$$\begin{cases} \text{Find } u : (0,T)\times\Omega \to \mathbb{R} \text{ such that} \\ \dfrac{\partial^2 u}{\partial t^2} - c^2 \Delta u = 0 & \text{in } (0,T)\times\Omega, \\ u(0,x) = u_0(x), \dfrac{\partial u}{\partial t}(0,x) = u_1(x) & \text{in } \Omega, \\ \nabla u \cdot \mathbf{n} = 0 & \text{on } \Gamma = \partial\Omega, \end{cases} \tag{1}$$

where T is the final time, c the velocity of the waves and u_0, u_1 are initial data. We assume here that the velocity is piecewise constant.

After a space discretisation (Finite Elements, Discontinous Galerkin, Finite Differences, etc.), the system can be rewritten as a linear system:

$$M\frac{d^2 U}{dt^2} + KU = 0, \tag{2}$$

where U is a vector whose components represent an approximation of u in a suitable basis of function, M is the mass matrix which is invertible and K is the stiffness matrix. To discretize (2) in time, we use Taylor expansions to obtain

$$\frac{U(t+\Delta t) - 2U(t) + U(t-\Delta t)}{\Delta t^2} = \frac{d^2 U(t)}{dt^2} + \frac{\Delta t^2}{12}\frac{d^4 U(t)}{dt^4} + O\left(\Delta t^4\right).$$

where Δt is the time step. Then, applying (2), we have that

$$\frac{d^4 U(t)}{dt^4} = M^{-1}KM^{-1}KU(t).$$

Consequently, we obtain an explicit fourth-order scheme:

$$U^{n+1} = 2U^n - U^{n-1} - \Delta t^2\left[M^{-1}K\left(U^n - \frac{\Delta t^2}{12}\left(M^{-1}KU^n\right)\right)\right], \tag{3}$$

where U^n denotes the approximation of U at time $t = n\Delta t$.

This technique is the so called modified equation technique and was introduced by Shubin and Bell ([3]). We precise that it can be applied to obtain a scheme of arbitrary even order.

This scheme is stable under the following CFL condition [1]:

$$\frac{\Delta t}{h} \leq \alpha_{LF}\sqrt{3},$$

where h is the typical space step of the mesh and α_{LF} denotes the CFL condition we would have obtained with a classical leapfrog scheme:

$$U^{n+1} = 2U^n - U^{n-1} - \Delta t^2 M^{-1} K U^n. \tag{4}$$

We remark that this scheme requires one more multiplication by $M^{-1}K$ than the classical second order leapfrog scheme, but its CFL condition is multiplied by $\sqrt{3} \simeq 1.73$, so that it increases the order of convergence by two orders, without penalizing too much the computational burden.

§3. Scheme with the bilaplacian operator

We present here the construction of a new scheme using the modified equation technique by first applying the time discretization before the space one.

3.1. Construction of the semi-discrete scheme

Using Taylor expansions on the continuous unknown, we have

$$\frac{u(t+\Delta t) - 2u(t) + u(t-\Delta t)}{\Delta t^2} = \frac{d^2 u(t)}{dt^2} + \frac{\Delta t^2}{12}\frac{d^4 u(t)}{dt^4} + O\left(\Delta t^4\right).$$

Then, applying the wave equation (1) to the second and the fourth derivative of $u(t)$ with respect to the time, we easily obtain

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} = c^2 \Delta u^n + \frac{\Delta t^2}{12} c^4 \Delta^2 u^n. \tag{5}$$

In the following, this scheme will be called "scheme with bilaplacian operator". To discretize the bilaplacian operator, we have to consider a space discretization which is able to take into account some H^2 quantities. Consequently, in this work, we have to consider C^1 finite elements (such as the Hermite ones) or Discontinuous Galerkin elements whose C^1 continuity is enforced through an appropriate penalty term.

3.2. Hermite finite elements

We first present the space discretization of (5) by Hermite elements. We restrict ourselves to the 1D-case since these elements are difficult to adapt to the higher dimensions.

Because of the bilaplacian operator, we need an additional boundary condition. Deriving two times the equation $\nabla u \cdot \mathbf{n} = 0$ with respect to the time and using the wave equation (1), we obtain

$$\frac{\partial^2 \nabla u}{\partial t^2} \cdot \mathbf{n} = \nabla \frac{\partial^2 u}{\partial t^2} \cdot \mathbf{n} = c^2 \nabla\left(\Delta u\right) \cdot \mathbf{n} = 0.$$

Consequently, we have to impose $\nabla u \cdot \mathbf{n} = 0$ and $\nabla\left(\Delta u\right) \cdot \mathbf{n} = 0$ on Γ. Similarly, for Dirichlet boundary conditions we would have $u = 0$ and $\Delta u = 0$ on Γ.

We multiply (5) by a test function $v \in H^2(\Omega)$, we integrate this equation over Ω, we apply Green formula and we use the two boundary conditions to obtain

$$\int_\Omega \left(\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} \right) v = a_1 \left(u^n, v \right) + \frac{\Delta t^2}{12} a_2 \left(u^n, v \right),$$

where

$$a_1 \left(u^n, v \right) = -c^2 \int_\Omega \nabla u^n \cdot \nabla v,$$
$$a_2 \left(u^n, v \right) = c^4 \int_\Omega \Delta u^n \Delta v - c^4 \int_\Gamma \Delta u^n \left(\nabla v \cdot \mathbf{n} \right) - c^4 \int_\Gamma \Delta v \left(\nabla u^n \cdot \mathbf{n} \right).$$

The last term of a_2 which vanishes on Γ is artificially introduced to symmetrize the bilinear form.

We consider $\Omega = [a, b] \subset \mathbb{R}$ and we introduce the following space of discretization:

$$V_h = \{ v \in C^1 \left(\Omega \right) : v_{|K} \in P^3([x_j, x_{j+1}]), \ \forall j = 1 \ldots n-1 \}.$$

where $\{x_j\}_{j=1\ldots n}$ are defined by

$$\forall j = 1 \ldots n-1, \ x_j \in [a, b] \text{ and } x_j < x_{j+1}.$$

The basis functions of Hermite's element method are defined by

$$\forall 1 \leq i, j \leq n-1, \begin{cases} \varphi_{2i-1}(x_j) = \delta_{2i-1,j}, & \varphi_{2i}(x_j) = 0, \\ \varphi'_{2i-1}(x_j) = 0, & \varphi'_{2i}(x_j) = \delta_{2i,j}. \end{cases}$$

We finally obtain the following linear system:

$$\frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} = M^{-1} K U^n, \tag{6}$$

with $M_{i,j} = \int_\Omega \varphi_i \varphi_j$, $K_{i,j} = \frac{\Delta t^2}{12} a_2(\varphi_i, \varphi_j) - a_1(\varphi_i, \varphi_j)$ and $U_i^n = u^n(x_i)$, if i is odd, or $(u^n)'(x_i)$, if i is even.

The CFL condition of this scheme is given by the following result:

Theorem 1. *A necessary and sufficient L^2-stability condition is given by*

$$c \frac{\Delta t}{h} \leq \frac{1}{\sqrt{5}},$$

where $h = \min_{j=1\ldots n-1} (x_{j+1} - x_j)$

Proof. We give the main ideas of the proof. The necessary condition is proved by a classical discrete Fourier analysis. Likewise, for the sufficient condition, we use an energy estimate to obtain

$$\lambda_{\min} \geq 0 \quad \text{and} \quad \frac{\Delta t^2}{4} \lambda_{\max} \leq 1,$$

where $\lambda_{\min} = \min\{\lambda \in \mathrm{Sp}(-M^{-1/2} K M^{-1/2})\}$ and $\lambda_{\max} = \max\{\lambda \in \mathrm{Sp}(-M^{-1/2} K M^{-1/2})\}$. □

Remark 1. The stability condition of this scheme is approximatively $\Delta t/h \leq 0.447$ and we have only one multiplication by $M^{-1}K$, whereas the stability condition of a P^3-Lagrange discretization combined with the classical modified equation technique is approximatively $\Delta t/h \leq 0.266$ and the scheme requires two multiplications by $M^{-1}K$. So the new scheme is 3.4 times faster.

In a strongly heterogeneous media, the solution is no longer C^1 because of the discontinuities of the physical parameters and Hermite elements are not adapted to this problem. Consequently, we introduce another method based on Discontinuous Galerkin method in the next section.

3.3. Discontinuous Galerkin Method

In this part, we use a Discontinuous Galerkin Method (DGM) which takes into account the discontinuities between each elements of the mesh $\mathcal{T}_h$ of Ω. More precisely, we use the Interior Penalty Discontinuous Galerkin Method [2]. First, we multiply (5) by a test function v, we integer it over each element K and we sum it over all elements of the mesh $\mathcal{T}_h$:

$$\sum_{K\in\mathcal{T}_h}\int_K \frac{u^{n+1}-2u^n-u^{n-1}}{\Delta t^2} v\,dx - \sum_{K\in\mathcal{T}_h}\int_K c^2\Delta u^n v\,dx - \frac{\Delta t^2}{12}\sum_{K\in\mathcal{T}_h}\int_K c^4\Delta^2 u^n v\,dx = 0.$$

Now, we have to introduce various notations. The set of the mesh faces are denoted $\mathcal{F}_h$ which is partitionned into two subsets $\mathcal{F}_h^i$ and $\mathcal{F}_h^b$ corresponding respectively to the interior faces and those located on the boundary. For $F \in \mathcal{F}_h^i$, we note arbitrarily K^+ and K^- the two elements sharing F and we define $\boldsymbol{\nu}$ as the unit outward normal vector pointing from K^+ to K^-.

Using a classical IPDG method, the second term of the formulation is replaced by the bilinear form a_1 defined by

$$\begin{aligned} a_1(u,v) = &\sum_{K\in\mathcal{T}_h}\int_K c^2\nabla u^n\nabla v\,dx - \sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\{\!\{c^2\nabla u^n\}\!\}\cdot\boldsymbol{\nu}\,d\sigma \\ &- \sum_{F\in\mathcal{F}_h}\int_F [\![u^n]\!]\{\!\{c^2\nabla v\}\!\}\cdot\boldsymbol{\nu}\,d\sigma + \sum_{F\in\mathcal{F}_h}\int_F \alpha_1 [\![u^n]\!]\,[\![v]\!]\,d\sigma, \end{aligned}$$

where α_1 is a well chosen penalization coefficient and $[\![\cdot]\!]$ and $\{\!\{\cdot\}\!\}$ correspond respectively to the jump and the average of a piecewise smooth function v, on an interior edge such that :

$$[\![v]\!] := v^+ - v^-, \quad \{\!\{v\}\!\} := \frac{v^+ + v^-}{2}.$$

We denote also by $v^\pm$ the restriction of v to the element $K^\pm$.

Now, we consider the third term of the formulation, denoted by Q. Using two times a Green formula, we obtain

$$Q = -\sum_{K\in\mathcal{T}_h}\int_K c^4\Delta u^n\Delta v\,dx + \sum_{K\in\mathcal{T}_h}\int_{\partial K} c^4\Delta u^n(\nabla v\cdot\mathbf{n})\,d\sigma - \sum_{K\in\mathcal{T}_h}\int_{\partial K} c^4(\nabla(\Delta u^n)\cdot\mathbf{n})\,v\,d\sigma.$$

Then, we can rewrite the second term and the third one:

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K} c^4\Delta u^n(\nabla v\cdot\mathbf{n})\,d\sigma = \sum_{F\in\mathcal{F}_h}\int_F [\![\nabla v]\!]\cdot\boldsymbol{\nu}\{\!\{c^4\Delta u^n\}\!\} + [\![c^4\Delta u^n]\!]\{\!\{\nabla v\}\!\}\cdot\boldsymbol{\nu}\,d\sigma,$$

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K} c^4(\nabla(\Delta u^n)\cdot\mathbf{n})v\,d\sigma = \sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\{\!\{c^4\nabla(\Delta u^n)\}\!\}\cdot\boldsymbol{\nu} + [\![c^4\nabla(\Delta u^n)]\!]\cdot\boldsymbol{\nu}\{\!\{v\}\!\}\,d\sigma.$$

Combining the continuity of u and $\nabla u\cdot\mathbf{n}$ across the interfaces with the wave equation (1), we deduce the continuity of Δu and $\nabla(\Delta u)\cdot\mathbf{n}$ so that

$$Q = -\sum_{K\in\mathcal{T}_h}\int_K c^4\Delta u^n\Delta v\,dx + \sum_{F\in\mathcal{F}_h}\int_F [\![\nabla v]\!]\cdot\boldsymbol{\nu}\{\!\{c^4\Delta u^n\}\!\} - \sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\{\!\{c^4\nabla(\Delta u^n)\}\!\}\cdot\boldsymbol{\nu}.$$

Since the form is not symmetric, we add the corresponding symmetric terms which vanish because of the continuity of u and $\nabla u\cdot\mathbf{n}$, and to enforce the coercivity of the form we add a suitable penalization term $\alpha_2\in\mathbb{R}$ to obtain the bilinear form

$$\begin{aligned} a_2(u,v) = Q_2 &+ \sum_{F\in\mathcal{F}_h}\int_F [\![\nabla u^n]\!]\cdot\boldsymbol{\nu}\{\!\{c^4\Delta v\}\!\} \\ &- \sum_{F\in\mathcal{F}_h}\int_F [\![u^n]\!]\{\!\{c^4\nabla(\Delta v)\}\!\}\cdot\boldsymbol{\nu} + \sum_{F\in\mathcal{F}_h}\int_F \alpha_2[\![c\nabla u\cdot\boldsymbol{\nu}]\!]\,[\![c\nabla v\cdot\boldsymbol{\nu}]\!]. \end{aligned}$$

Then, we introduce the space of discretization $V_h = \{v\in L^2(\Omega) : v_{|K}\in P^3(K), \forall K\in\mathcal{T}_h\}$ and we consider $\{\varphi_j\}_{j=1\ldots n}$, the classical discontinuous basis functions P^3 of V_h to obtain the scheme

$$U^{n+1} = 2U^n - U^{n-1} + \Delta t^2 M^{-1}\left(\frac{\Delta t^2}{12}K_2 - K_1\right),$$

where $(M)_{i,j} = \sum_{K\in\mathcal{T}_h}\int_K \varphi_i\varphi_j$, $(K_1)_{i,j} = a_1(\varphi_i,\varphi_j)$ and $(K_2)_{i,j} = a_2(\varphi_i,\varphi_j)$.

Numerical results will illustrate the fact that this scheme has the same stability condition as the classical IPDG method combined with a leapfrog scheme.

§4. Numerical Results

In this part, we present some results in the one-dimensional case. Experiments in higher dimensions are in progress and preliminary results confirms the 1D results. In all the experiments, we consider a domain $\Omega = [0, 10]$, a final time $T = 100$ and a velocity $c = 1$. We consider periodic boundary conditions, to ensure that the boundary conditions do not deteriorate the performances of the scheme. The initial conditions are

$$\begin{cases} U^0(x) = \sin(\pi x), \\ U^1(x) = \sin(\pi(x-\Delta t)), \end{cases}$$

so that the exact solution is $U(x,t) = \sin(\pi(x-t))$.

First, we compare the scheme with the bilaplacian operator to the classical P^3 FEM with the

Ndof	150	300	600	1200
P_3 FE	$\Delta x = 0.200$ $\Delta t = 0.0531$ Err = 3.39E−03	$\Delta x = 0.100$ $\Delta t = 0.0266$ Err = 2.66E−04	$\Delta x = 0.050$ $\Delta t = 0.0133$ Err = 1.75E−05	$\Delta x = 0.025$ $\Delta t = 0.0066$ Err = 1.11E−06
Δ^2 Hermite FE	$\Delta x = 0.133$ $\Delta t = 0.0584$ Err = 6.63E−03	$\Delta x = 0.067$ $\Delta t = 0.0294$ Err = 4.2E−04	$\Delta x = 0.033$ $\Delta t = 0.0147$ Err = 2.56E−05	$\Delta x = 0.017$ $\Delta t = 0.0073$ Err = 1.58E−06

Table 1: Comparison between P^3 FE and Δ^2 Hermite FE

Ndof	150	300	600	1200
DGP_3	$\Delta x = 0.256$ $\Delta t = 0.0681$ Err = 3.421E−03	$\Delta x = 0.132$ $\Delta t = 0.0349$ Err = 2.7006E−04	$\Delta x = 0.066$ $\Delta t = 0.0176$ Err = 1.809E−05	$\Delta x = 0.033$ $\Delta t = 0.0088$ Err = 1.158E−06
Δ^2 DGP_3	$\Delta x = 0.256$ $\Delta t = 0.0467$ Err = 3.297E−03	$\Delta x = 0.132$ $\Delta t = 0.0240$ Err = 1.717E−04	$\Delta x = 0.066$ $\Delta t = 0.0121$ Err = 8.088E−06	$\Delta x = 0.033$ $\Delta t = 0.0060$ Err = 4.337E−07

Table 2: Comparison between classical IPDG and Δ^2 IPDG

modified equation scheme. Table 1 presents the $L^2\,(]0,T[\,,\Omega)$-error with various choices of the number of degree of freedom (*Ndof*), the space step (Δx) and the time step (Δt).

We can easily remark that, with each method, the ratio between two consecutive errors is almost 16 that is to say the two methods are indeed fourth order methods. Furthermore, we note that the error is smaller with "P_3 FE" than with "Δ^2 Hermite FE" for a given number of degrees of freedom (i.e. for an equivalent computational burden at each time step). However, the same level of error as P^3 FE can be reached by decreasing the time step by 25%. Keeping in mind that the Δ^2 scheme requires only one multiplication by $M^{-1}K$, it is still less expensive than the classical one.
Now we present the results using a DGM with the same parameters as previously and $\alpha_1 = 8$ and $\alpha_2 = -10$ (cf. Table 2).

Once again, these results confirms that the methods are fourth order methods and we remark that the results with the scheme with the bilaplacian operator provides smaller error than the classical IPDG. Moreover, we notice that, with the bilaplacian operator, the time step is smaller than IPDG method but this problem is in balance with the fact that we have only one multiplication by $M^{-1}K$.

We now investigate the influence of the boundary conditions on the stability of the schemes. Table 3 represents the CFL conditions (numerically computed) for periodic, Neu-

	Periodic	Dirichlet	Neumann
Leapfrog scheme P_3	0.15333	0.15333	0.15333
FE P_3	0.26558	0.26558	0.26558
DG P_3	0.2655	0.2655	0.2655
Δ^2 Hermite FE	0.4471	0.4471	**0.1995**
Δ^2 DGP_3	0.1821	0.1821	0.1821

Table 3: Comparison CFL conditions

mann and Dirichlet conditions for the various schemes we have presented.

The boundary conditions do not modify the stability of the Δ^2 IPDG scheme, whereas the Neumann condition deteriorate the stability ot the Δ^2 Hermite scheme. Besides, since the IPDG scheme can be more easily extended to multidimensional cases and is more adapted to deal with heterogeneous media, we will focus on this method in future works.

§5. Conclusion

In this work, we have constructed a new scheme based on the modified equation technique and a switch between the time discretization and the space discretization. This new scheme allows to reduce the computational time and improve the accuracy of the classical methods. We are now considering the two dimensional case and heterogeneous media. Next step will be the implementation of absorbing boundary conditions.

References

[1] Gilbert, J.-C., and Joly, P. Higher order time stepping for second order hyperbolic problems and optimal CFL conditions. *Numerical Analysis and Scientific Computing for PDE's and their Challenging Applications* (2006).

[2] Grote, M. J., Schneebeli, A., and Schötzau, D. Dicontinuous galerkin finite element method for the wave equation. *SIAM J. Numer. Anal. 44* (2006), 2408–2431.

[3] Shubin, G. R., and Bell, J. B. A modified equation approach to constructing fourth-order methods for acoustic wave propagation. *SIAM J. Sci. Statist. Comput. 8* (1987), 135–151.

Cyril Agut and Julien Diaz
INRIA Research Center Bordeaux-Sud Ouest
Team-project Magique 3D
cyril.agut@inria.fr
julien.diaz@inria.fr

Abdelaaziz Ezziani
Laboratoire de Mathématiques Appliquées -
UMR CNRS 5142
Université de Pau et des Pays de l'Adour - Bt.
IPRA - BP 1155 64013 PAU Cedex
abdelaaziz.ezziani@univ-pau.fr