A generalized statement for advective-diffusive phenomena. Finite element model and applications

H. Gómez, I. Colominas*, F. Navarrina & M. Casteleiro
Group of Numerical Methods in Engineering, GMNI
Dept. of Applied Mathematics, Civil Engineering School
Universidad de A Coruña
Campus de Elviña, 15071 A Coruña, SPAIN

SUMMARY
Solving convective-diffusive transport problems is a frequent task in engineering, especially in convection dominated situations. Moreover, the standard statement for the transport problem leads to the result that mass can propagate at an infinite speed. This paradoxical result occurs as a consequence of using Fick’s law. It seems that this fact is related to the spurious oscillations that occur in the numerical solution of the standard formulation of the transport problem when the Galerkin finite element method is used for the spatial discretization.

For these reasons, we propose to use Cattaneo’s law instead of Fick’s law for the formulation of the advective-diffusive problem. Cattaneo’s law has been previously applied to pure-diffusive problems and it is a generalization of Fick’s law. The formulation of the transport problem by using Cattaneo’s law leads to a hyperbolic system of partial differential equations which can be written in conservative form. As a consequence of being a hyperbolic system, a finite diffusive velocity can be defined.

A Taylor-Galerkin procedure can be used to solve these equations. In this paper, several problems in one and two-dimensional domains have been solved to show that this new approach can be used in real engineering problems. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Convection-Diffusion, Cattaneo’s Equation, Taylor-Galerkin

1. INTRODUCTION
Solving transport problems is very common in applied science. In particular, transport problems involving convective and diffusive processes in fluid media have a great applicability in engineering (simulating the effect of pollutant spillage into the sea, rivers or the atmosphere; modelling the evolution of underground oil reserves, natural gas, etc.). This kind of phenomena can be modelled by using the so-called advective-diffusive equation. Besides its practical interest, the advection-diffusion equation also represents an adequate model for the introduction of numerical schemes for more complicated problems. For these reasons, a great

*Correspondence to: E.T.S. de Ingenieros de Caminos, Canales y Puertos, Universidad de A Coruña, Campus de Elviña, 15071 A Coruña, Spain. Email: icolominas@udc.es

Copyright © 2000 John Wiley & Sons, Ltd.
deal of effort has been devoted to the development of numerical methods for the approximation of the convective-diffusive equation. However, it is still very complicated to obtain an accurate and stable numerical solution for this equation when the convective term becomes important.

In the 1950’s, the idea rises that adding (in one way or another) numerical dissipation is the way to obtain stable solutions for this equation. In the framework of the finite difference method this idea was first proposed by von Neumann and Richtmyer [1]. It was noted early that this dissipation can be introduced by means of a non-centered difference approximation for the convective term taking into account the direction of the flow. This fact motivated the name \textit{upwind methods} for these kinds of approximations. This approach has been seen stable but overly diffusive (see for instance references [2, 3] where these methods are criticized).

In the finite element framework several different techniques can be utilized to achieve the upwind effect. The first finite element formulations were presented in [4] for ordinary differential equations and subsequently in [5] in the context of partial differential equations. These methods were based on modified weighting functions such that the element upstream of a node is weighted more heavily than the element downstream of a node. These procedures were able to deliver stable numerical results but they were subject to the same criticism as upwind differences because they also lead to overly diffusive solutions. In reference [6] it is stated that the upwind weighting can also be achieved by using a different quadrature rule for the convective term. In this paper Hughes gives the quadrature points for the one-dimensional linear element and the bilinear quadrilateral element. However, the two-dimensional procedure has been criticized for introducing excessive crosswind diffusion.

In 1979 the Winter Annual Meeting of the ASME was held in New York. In this meeting early upwind finite element formulations were reviewed [7] and other new techniques were proposed. Belytschko and Elhib [8] introduced an \textit{amplification scheme} for achieving an upwind finite element formulation. However, the behavior of the amplification scheme was shown to be similar to other upwind finite element methods. In this meeting, the idea of introducing numerical dissipation only along the streamlines is pointed out by Griffiths and Mitchell [9] and formally stated by Hughes and Brooks [10]. This is the main idea underlying the \textit{streamline upwind Petrov-Galerkin method}. Almost simultaneously Kelly et al. [11] suggested the same procedure to eliminate crosswind diffusion in multidimensional solutions. This method can be formulated by using a modified weighting function for the convective term only. This scheme was called \textit{streamline upwind} method.

Hughes and Brooks [12] proposed subsequently to apply the modified weighting function to all terms in the equation (in contrast with the application only to the convective term) in order to obtain a consistent formulation (see also [13]). This idea constitutes the last ingredient of the streamline upwind Petrov-Galerkin method also called \textit{streamline diffusion} [14, 15, 16] in mathematical circles. Later, Hughes and his co-workers proposed several techniques with a consistent stabilization. All of these methods consist of adding an extra term to the Galerkin formulations over the interior elements. This term is a function of the residual of the differential equation to ensure the consistency. The \textit{Galerkin/least-squares method} proposed in [17] can be formulated in this fashion.

In the early 1990’s the idea of enriching the Galerkin finite element method with the so-called \textit{bubbles} or \textit{wavelet functions} rose. Baiocchi, Brezzi and Franca [18] were the first to point out that the enrichment of the finite element space by summation of polynomial bubble functions results in a stabilization procedure for convection-diffusion problems. This stabilization procedure is formally similar to streamline-upwind Petrov-Galerkin and Galerkin/least-squares
methods. Indeed, in [19] it was shown that stabilized methods and bubble function methods are all approximate subgrid scale models. This result was achieved by means of the \textit{variational multiscale method} proposed by Hughes. A detailed presentation of this method can be found in [20]. Later, in an attempt to resolve thin internal layers with steep gradients or other small-scale phenomena, the concept of enriching the Galerkin finite element method with so-called \textit{residual-free bubbles} was introduced [21, 22].

An alternative approach to a stable and accurate solution of the advective-diffusive phenomena was proposed by Oñate [23] under the name of \textit{finite increment calculus}. The basic idea behind this method is that most stabilized numerical schemes can be derived by applying the standard Galerkin formulation to a so-called stabilized form of the governing differential equations of the problem. These stabilized equations are derived by establishing high-order balance statements on a finite domain (as opposed to infinitesimal).

In recent years the so-called \textit{generalized finite element method} has been introduced by Strouboulis, Babuška and Copps in [24]. Farhat, Harari and Franca [25] proposed the \textit{discontinuous enrichment method}. With the aim of obtaining a stable and accurate solution for transient advective-diffusive problems Donea, Roig and Huerta [26] introduced high order time-stepping schemes based on Padé’s approximations of exponential function. For a detailed presentation of most of these methods see the excellent book by Donea and Huerta [27].

In this paper we review the formulation of the advective-diffusive equation applied to the spillage of a pollutant into a fluid medium. In particular, we notice that Fick’s law [28] leads to the result that mass can propagate at an infinite speed. This fact is related to the appearance of spurious oscillations in the numerical solution of the advective-diffusive equation [29, 30, 31]. To overcome the infinite speed paradox we use a generalized constitutive equation proposed simultaneously by Cattaneo and Vernotte in 1958 [32, 33]. Cattaneo’s law has been used in practical applications for pure-diffusive problems [34] instead of Fick’s law, but, up to the authors’ knowledge, Cattaneo’s equation has not been used in problems with convective term. In this study the advective-diffusive problem has been formulated by using Cattaneo’s law as the constitutive equation. This approach leads to a totally hyperbolic system of conservation laws. Therefore, a finite diffusive velocity can be defined. Finally, we show that the proposed approach leads to more stable numerical solutions than the standard formulation when the Galerkin discretization is used.

The outline of this paper is as follows: In Section 2 we review the classic formulation of the advective-diffusive equation. In section 3 we study the formulation of the transport problem by using Cattaneo’s law. A numerical analysis of steady state equations (section 4) and transient equations (section 5) is performed later. Finally, section 6 is devoted to the presentation of main conclusions from this study.
2. STANDARD FORMULATION OF THE ADVECTIVE-DIFFUSIVE TRANSPORT PROBLEM

2.1. Problem statement

In this section we review the classic formulation for the advective-diffusive transport problem. The governing equations under the assumption of incompressibility are as follows:

\[
\frac{\partial u}{\partial t} + a \cdot \nabla_x (u) + \nabla_x \cdot (q) = f \tag{1.1}
\]

\[
q = -K \nabla_x (u) \tag{1.2}
\]

In the above system, (1.1) is the pollutant mass conservation equation and (1.2) is the constitutive equation known as Fick’s law. Further, \(u\) is the pollutant concentration, \(a\) is the velocity vector which satisfies the hydrodynamic equations of an incompressible fluid, \(q\) is the diffusive flux per unit fluid density, \(f\) is a source term and \(K\) is the diffusivity tensor which is assumed to be positive definite. Clearly, the system of equations (1) is fully decoupled as we can introduce (1.2) into (1.1) and solve the scalar equation

\[
\frac{\partial u}{\partial t} + a \cdot \nabla_x (u) - \nabla_x \cdot (K \nabla_x (u)) = 0 \tag{2}
\]

where source term has been removed by simplicity. It is well known that the equation (2) is a parabolic one. Therefore, boundary conditions must be imposed everywhere on the boundary of the domain. Then, let us consider the transport by convection and diffusion in a domain \(\Omega \subset \mathbb{R}^2\) with piecewise smooth boundary \(\Gamma\). The unit outward normal vector to \(\Gamma\) is denoted by \(n\). The boundary is assumed to consist of a portion \(\Gamma_D\) on which the value of \(u\) is prescribed (Dirichlet or essential conditions) and a complementary portion \(\Gamma_N\) on which flux is prescribed (Neumann or natural conditions). In addition, we know the initial distribution of the transported quantity \(u\). At this point we can state convection-diffusion initial-boundary value problem as follows: given a solenoidal velocity field \(a\), the diffusion tensor \(K\) and the necessary initial and boundary conditions, find \(u: \Omega \times [0, T] \mapsto \mathbb{R}\) such that

\[
\frac{\partial u}{\partial t} + a \cdot \nabla_x (u) - \nabla_x \cdot (K \nabla_x (u)) = 0 \quad \text{in} \quad \Omega \times [0, T] \tag{3.1}
\]

\[
u(x, 0) = u_0(x) \quad \text{on} \quad \Omega \tag{3.2}
\]

\[
u = u_D \quad \text{on} \quad \Gamma_D \times [0, T] \tag{3.3}
\]

\[(\alpha K \nabla_x (u) - \beta u a) \cdot n = h \quad \text{on} \quad \Gamma_N \times [0, T] \tag{3.4}
\]

where \(\alpha\) and \(\beta\) are real constants. Finally, it should be noted that solving (1) instead of (2) entails a greater computational cost, but it allows us to impose Neumann boundary conditions exactly. However, if we solve (2) flux boundary conditions should be imposed in a weak form.

2.2. A pure-diffusive example to analyze the infinite speed paradox

In what follows we will show that the above formulation leads to mass propagation at an infinite speed. Let us consider an (incompressible) homogeneous, isotropic (hence, if \(I\) is the identity tensor, \(K = kI\) for a certain \(k > 0\)) and one-dimensional medium. We consider a pure-diffusive situation and no source terms. We suppose the domain to be long enough to be
A generalized statement for advective-diffusive phenomena

Approximated as infinitely long, therefore no boundary conditions should be imposed. Finally, we assume that the pollutant is added to the medium as a rapid pulse. In this case we should solve the following problem: find \( u: \mathbb{R} \times [0, \infty) \mapsto \mathbb{R} \) such that

\[
\frac{\partial u}{\partial t} - k \frac{\partial^2 u}{\partial x^2} = 0 \quad \forall x \in \mathbb{R} \quad t > 0
\]

\( (4.1) \)

\[ u(x, 0) = \delta(x) \quad \forall x \in \mathbb{R} \]

\( (4.2) \)

\[ \lim_{x \to \pm \infty} u(x, t) = 0 \quad t > 0. \]

\( (4.3) \)

where \( \delta \) is the Dirac distribution. This problem can be solved by using a Fourier transform in the spatial coordinate. The solution of (4) is

\[
 u(x, t) = \frac{1}{\sqrt{4\pi kt}} e^{-\frac{x^2}{4kt}}, \quad \forall x \in \mathbb{R}, \quad t > 0.
\]

(5)

If we fix a time \( t = \tau_0 > 0 \), we can define

\[
 \tilde{u}(x) = u(x, \tau_0) = \frac{1}{\sqrt{4\pi k\tau_0}} e^{-\frac{x^2}{4k\tau_0}}
\]

(6)

which is the Gauss distribution function. Thus, \( \tilde{u}(x) > 0, \quad \forall x \in \mathbb{R} \). The previous assertion implies that polluted fluid exists in the whole domain \( \forall t > 0 \). However, at the initial time \( u(x, 0) = 0 \quad \forall x \neq 0 \), i.e., there is pollutant only in the origin of coordinates. Therefore, if we fix a generic point \( x_0 \), the following equation holds

\[
 u(x_0, \tau_0) > 0, \quad \forall \tau_0 > 0.
\]

(7)

Hence, the mean velocity of the particles in \( (x_0, \tau_0) \) is \( x_0/\tau_0 \) which is not bounded, because the above assertion holds \( \forall \tau_0 > 0 \) and \( \forall x_0 \in \mathbb{R} \). Figure 1 shows (in dashed line) the solution of (4) for \( k = 1 \) at \( t = 4 \) and \( t = 10 \).

Figure 1. Comparison at \( t = 4 \) (left) and at \( t = 10 \) (right) between the solution of (4) (dashed line) and the solution of (9) (solid line). Parameters \( k \) and \( \tau \) have a value of one.
3. FORMULATION OF THE TRANSPORT PROBLEM BY USING CATTANEO’S LAW

3.1. Problem statement

We will derive this formulation by substituting the equation (1.2) (known as Fick’s equation) by a Cattaneo-type law. Cattaneo’s equation involves a tensorial function \( \tau \). This mapping transforms each point \((x, t)\) of the domain into that point relaxation tensor. The coordinates of the relaxation tensor are specific diffusion process times. Up to the authors’ knowledge, Cattaneo’s equation has been only used in non-advective thermal problems, see for instance [34]. Thus, we had to find Cattaneo’s equation with convective term [31]. This equation has been derived by using a Lagrangian description but it will be written in this paper in Eulerian coordinates. Hence, basic equations for the transport problem described by using Cattaneo’s law are

\[
\frac{\partial u}{\partial t} + \mathbf{a} \cdot \nabla_x (u) + \nabla_x \cdot (q) = f \tag{8.1}
\]

\[
q + \tau \left( \frac{\partial q}{\partial t} + \nabla_x (q) \mathbf{a} \right) = -K\nabla_x (u) \tag{8.2}
\]

where (8.2) is Cattaneo’s law with convective term. It should be noted that we are using a generalized constitutive equation as we recover the classic formulation when \( \tau \) is the zero tensor.

3.2. A pure-diffusive example with a finite propagation velocity

In order to compare the solution of the classic formulation with the solution of the generalized formulation we now solve the Cattaneo-type counterpart of (4). In this simple case the system of equations (8) can be reduced to a second order partial differential equation (for further details, see [31]). Now we need two initial conditions because this problem involves second order derivatives with respect to the time. Then, we consider an (incompressible) homogeneous, isotropic, one-dimensional and non-convective medium. With the above assumptions we can state this problem as [31]: find \( u : \mathbb{R} \times [0, \infty) \mapsto \mathbb{R} \) such that

\[
\tau \frac{\partial^2 u}{\partial t^2} + \frac{\partial u}{\partial t} - k \frac{\partial^2 u}{\partial x^2} = 0 \quad \forall x \in \mathbb{R} \quad t > 0 \tag{9.1}
\]

\[
u(x, 0) = \delta(x) \quad \forall x \in \mathbb{R} \tag{9.2}
\]

\[
\frac{\partial u}{\partial t}(x, 0) = 0 \quad \forall x \in \mathbb{R} \tag{9.3}
\]

\[
\lim_{x \to \pm \infty} u(x, t) = 0 \quad t > 0. \tag{9.4}
\]

We may solve (9) by using subsequent Laplace and Fourier transforms (see reference [31] for a detailed resolution). In this way, we obtain

\[
u(x, t) = \begin{cases} 
\frac{1}{2} e^{-\frac{x^2}{4t}} \left[ \delta(|x| - ct) + \frac{ct}{2\pi} \left( \frac{ct}{\sqrt{\pi \sqrt{ct^2 - x^2}}} \right) + \frac{e^{-\frac{ct^2}{4}}}{\sqrt{\pi \sqrt{ct^2 - x^2}}} \right], & |x| \leq ct \\
0, & |x| > ct
\end{cases} \tag{10}
\]
where $I_0$ and $I_1$ are the modified Bessel functions of the first kind of order 0 and 1. Furthermore, $c$ is what we call mass wave celerity defined by:

$$c = \sqrt{k/\tau}. \quad (11)$$

We compare in figure 1 the solutions of (4) and (9) at $t = 4$ and $t = 10$ respectively. Clearly, if we use Cattaneo’s equation a wave front exists which advances with a celerity $c$.

3.3. Study of the proposed model as a system of conservation laws

The system (8) can not be reduced to a second order partial differential equation in multidimensional problems with a non-constant velocity field. In this case we must solve a coupled system of first order partial differential equations. This fact enlarges the computational cost of solving the transport problem but it allows us to impose Neumann boundary conditions exactly (in contrast with the imposition in a weak form). Under the assumptions that the medium is homogeneous and isotropic, the system (8) can be written as a system of conservation laws. Taking into account that $a$ is a solenoidal vector the following identities hold:

$$a \cdot \nabla_x (u) = \nabla_x \cdot (ua) \quad (12.1)$$

$$\nabla_x (q) \cdot a = \nabla_x \cdot (q \otimes a) \quad (12.2)$$

Since the medium is homogeneous and isotropic $K = kI$, $\tau = \tau I$ for certain $k, \tau \in \mathbb{R}^+$. As a consequence, we can rewrite (8) as

$$\frac{\partial u}{\partial t} + \nabla_x \cdot (ua + q) = 0 \quad (13.1)$$

$$\frac{\partial (\tau q)}{\partial t} + \nabla_x \cdot (\tau q \otimes a + kuI) + q = 0 \quad (13.2)$$

where source term has been dropped by simplicity. In what follows one and two-dimensional problems will be studied separately.

3.3.1. One-dimensional problem

In this section we will study the one-dimensional Cattaneo-type transport problem. In this simple case the governing equation is

$$\frac{\partial U}{\partial t} + \nabla_x \cdot (F) = S \quad (14)$$

where

$$U = \begin{pmatrix} u \\ \tau q \end{pmatrix}; \quad F = \begin{pmatrix} ua + q \\ \tau qa + ku \end{pmatrix}; \quad S = \begin{pmatrix} 0 \\ -q \end{pmatrix} \quad (15)$$

The system (14) can be written in non-conservative form as

$$\frac{\partial U}{\partial t} + A \frac{\partial U}{\partial x} = S \quad (16)$$

being $A$ the so-called Jacobian matrix defined by

$$A = \nabla_U (F) = \begin{pmatrix} a & 1/\tau \\ k & a \end{pmatrix} \quad (17)$$
It is apparent that the system (16) is a linear one as $A$ does not depend on $U$. On the other hand, it is well known (see for instance reference [35]) that the system (14) will be totally hyperbolic if, and only if, the matrix $A$ possesses 2 different real eigenvalues. It can be shown that

$$A = CDC^{-1}$$

where

$$C = \begin{pmatrix} \frac{1}{\tau c} & 1 \\ \tau c & -\tau c \end{pmatrix}; \quad D = \begin{pmatrix} a + c & 0 \\ 0 & a - c \end{pmatrix}$$

(18)

so (14) is, in fact, totally hyperbolic. Now we will prove that the system (16) can be diagonalized but it can not be decoupled. The so-called Riemann quasi-invariants can be defined (we call Riemann quasi-invariants those functions instead of Riemann invariants [35] because we have a source term in (14)). To prove this fact, we use (18). Hence, we can rewrite (16) as follows:

$$\frac{\partial U}{\partial t} + CDC^{-1} \frac{\partial U}{\partial x} = S$$

(19)

As a consequence of the assumption of homogeneity, (19) takes the form

$$\frac{\partial (C^{-1}U)}{\partial t} + D \frac{\partial (C^{-1}U)}{\partial x} = C^{-1}S$$

(20)

If we use the notation

$$\begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = R = C^{-1}U = \frac{1}{2} \begin{pmatrix} u + q/c \\ u - q/c \end{pmatrix}$$

(21)

for Riemann quasi-invariants the following equation holds:

$$\frac{\partial R}{\partial t} + D \frac{\partial R}{\partial x} = QR$$

(22)

In the above equation $Q$ is the matrix

$$Q = \frac{1}{2\tau} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

(23)

Therefore, as $D$ is a diagonal matrix (and $Q$ is not a diagonal one), the system of equations (22) is only coupled by the source term. The two scalar equations in (22) are two transport equations with a source term. The quantity $R_1$ is transported along the spatial coordinate with velocity $a + c$. Whereas, $R_2$ is also transported along the spatial coordinate in this case with velocity $a - c$. Hence, the direction in which each wave $R_1$ or $R_2$ is transported along the spatial coordinate depends on the sign of its corresponding propagation velocity. Therefore, depending on the values of $a$ and $c$ the solution of (22) can be the superposition of two waves traveling in the same or in the opposite direction. We will refer to this situation as **supercritical** and **subcritical flow** respectively. At this point it is very useful to introduce the following dimensionless number:

$$H = \frac{|a|}{c}$$

(24)

With the definition (24) we can characterize subcritical and supercritical flows. The first one occurs when $H < 1$ and the second one when $H > 1$. We will call **critical flow** the flow which verifies $H = 1$.

Taking into account all of this, boundary conditions which should be imposed to (22) are straightforward. Let us suppose that we have to solve (22) in a given domain $\Omega = (0, L); \ L \in$
\( \mathbb{R}^+ \) which is bounded by \( \Gamma \). We will call \textit{inflow boundary} (\( \Gamma_{\text{in}} \) in what follows) the part of the boundary in which \( \mathbf{a} \cdot \mathbf{n} < 0 \). We will call \textit{outflow boundary} (\( \Gamma_{\text{out}} \) in short) the complementary part of the boundary. Now, we define \( \Gamma_0 \) as the point \( x = 0 \) and \( \Gamma_L \) as \( x = L \). Accordingly, \( \Gamma = \Gamma_0 \cup \Gamma_L \). In supercritical flow both \( R_1 \) and \( R_2 \) should be prescribed in the inflow boundary (\( \Gamma_0 \) when \( a > 0 \) and \( \Gamma_L \) when \( a < 0 \)). If the flow is subcritical the quantity \( R_1 \) should be prescribed in \( \Gamma_0 \) and \( R_2 \) must be imposed in \( \Gamma_L \).

However, it is commonly accepted [27, 36] that a hyperbolic system of partial differential equations like (14) is well-posed when the number of imposed components of \( \mathbf{U} \) on one boundary is equal to the number of waves incoming in that boundary (number of negative eigenvalues of the jacobian matrix). Therefore, in supercritical flow both components of \( \mathbf{U} \) should be prescribed on the inflow boundary and no components of \( \mathbf{U} \) must be imposed on the outflow boundary. In subcritical flow one component of \( \mathbf{U} \) should be prescribed on the inflow boundary and another one must be imposed on the outflow boundary. In what follows, components of \( \mathbf{U} \) prescribed on the boundary will be called \textit{inflow components of \( \mathbf{U} \)}. These functions will be denoted as \( \mathbf{U}_{\text{in}} \).

Therefore, the one-dimensional Cattaneo-type transport problem can be stated as follows: given \( k, \tau > 0 \), the field velocity \( \mathbf{a} \) and the necessary initial and boundary conditions, find \( \mathbf{U}: \Omega \times [0, T] \mapsto \mathbb{R}^2 \) such that

\[
\frac{\partial \mathbf{U}}{\partial t} + \nabla \mathbf{x} \cdot (\mathbf{F}) = \mathbf{S} \quad \text{in} \quad \Omega \times [0, T] \tag{25.1}
\]

\[
\mathbf{U}(x, 0) = \mathbf{U}_0(x) \quad \text{on} \quad \Omega \tag{25.2}
\]

\[
\mathbf{U}_{\text{in}} = \mathbf{U}_{\text{in}}^n \quad \text{on} \quad \Gamma \times [0, T] \tag{25.3}
\]

being \( \mathbf{U}, \mathbf{F} \) and \( \mathbf{S} \) the vectors defined in (15).

### 3.3.2. Two-dimensional problem

Now we analyze the Cattaneo-type transport problem in a two-dimensional domain. We will use the notation \( \mathbf{q} = (q_1, q_2)^T \) and \( \mathbf{a} = (a_1, a_2)^T \). Hence (13) can be written as

\[
\frac{\partial \mathbf{U}}{\partial t} + \nabla \mathbf{x} \cdot (\mathbf{F}) = \mathbf{S} \quad \text{in} \quad \Omega \times [0, T] \tag{26}
\]

where

\[
\mathbf{U} = \begin{pmatrix} u \\ \tau q_1 \\ \tau q_2 \end{pmatrix}; \quad \mathbf{F} = \begin{pmatrix} u a_1 + q_1 & u a_2 + q_2 \\ \tau q_1 a_1 + k u & \tau q_1 a_2 \\ \tau q_2 a_1 & \tau q_2 a_2 + k u \end{pmatrix}; \quad \mathbf{S} = \begin{pmatrix} 0 \\ -q_1 \\ -q_2 \end{pmatrix} \tag{27}
\]

Note that the same notation has been used for one and two-dimensional variables, but no confusion is expected. In order to investigate the basic properties of the equation (26) it is necessary to write it in non-conservative form. In this way, if we define \( \mathbf{F}_i \) as the \( i \)-th column of the matrix \( \mathbf{F} \), the following relation holds:

\[
\nabla \mathbf{x} \cdot (\mathbf{F}) = \frac{\partial \mathbf{F}_1}{\partial x_1} + \frac{\partial \mathbf{F}_2}{\partial x_2} = A_1 \frac{\partial \mathbf{U}}{\partial x_1} + A_2 \frac{\partial \mathbf{U}}{\partial x_2} \tag{28}
\]

being \( A_1 \) and \( A_2 \) the Jacobian matrices defined by

\[
A_1 = \nabla \mathbf{U} (\mathbf{F}_1) = \begin{pmatrix} a_1 & 1/\tau & 0 \\ k & a_1 & 0 \\ 0 & 0 & a_1 \end{pmatrix}; \quad A_2 = \nabla \mathbf{U} (\mathbf{F}_2) = \begin{pmatrix} a_2 & 0 & 1/\tau \\ 0 & a_2 & 0 \\ k & 0 & a_2 \end{pmatrix} \tag{29}
\]
Now, we define the hypermatrix \( \mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)^T \). This definition allows us to use the notation \( \nabla_x \cdot (\mathbf{F}) = (\mathbf{A} \cdot \nabla_x)\mathbf{U} \) and rewrite (26) as
\[
\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{A} \cdot \nabla_x)\mathbf{U} = \mathbf{S}
\] (30)
which is a linear system of equations as \( \mathbf{A} \) does not depend on \( \mathbf{U} \). Now, we define \( \mathbf{\kappa} = (\kappa_1, \kappa_2)^T \) as an arbitrary vector with the constriction \( ||\mathbf{\kappa}|| = 1 \), where \( ||\mathbf{•}|| \) denotes the Euclidean norm of a given vector. It can be shown [35] that (30) is totally hyperbolic if the equation
\[
\det(\omega \mathbf{I} - \mathbf{A} \cdot \mathbf{\kappa}) = 0
\] (31)
possesses three different real solutions \( \omega \) for arbitrarily prescribed values of \( \mathbf{\kappa} \). It is easy to prove that (31) has the solutions
\[
\begin{align*}
\omega_1 &= \mathbf{a} \cdot \mathbf{\kappa} \\
\omega_2 &= \mathbf{a} \cdot \mathbf{\kappa} - c \\
\omega_3 &= \mathbf{a} \cdot \mathbf{\kappa} + c
\end{align*}
\] (32.1)
(32.2)
(32.3)
which are all different from each other. In addition, values written in (32) are the eigenvalues of \( \mathbf{A} \cdot \mathbf{\kappa} \) which is usually called projection matrix. Now, making some algebra we can compute the eigenvectors of \( \mathbf{A} \cdot \mathbf{\kappa} \). Then, we can define the square matrix \( \mathbf{C} \) as the matrix whose columns are the eigenvectors of the projection matrix, that is,
\[
\mathbf{C} = \begin{pmatrix} 0 & c & c \\
\kappa_2 & -k\kappa_1 & k\kappa_1 \\
-k\kappa_1 & -k\kappa_2 & k\kappa_2 \end{pmatrix}
\] (33)
Let \( \mathbf{D} \) be the diagonal matrix such that the elements placed in the main diagonal are the eigenvalues \( \omega_1, \omega_2, \omega_3 \), then, the following relation holds
\[
\mathbf{D} = \mathbf{C}^{-1} (\mathbf{A} \cdot \mathbf{\kappa}) \mathbf{C}
\] (34)
It is important to note that even though matrices \( \mathbf{A}_1 \) and \( \mathbf{A}_2 \) are diagonalizable, the system (30) can not be diagonalized because \( \mathbf{A}_1 \) and \( \mathbf{A}_2 \) are diagonalizable in two different bases. As a consequence, Riemann quasi-invariants can not be defined in 2D. At this point, it is very useful to introduce the following dimensionless number
\[
H = \frac{||\mathbf{a}||}{c}
\] (35)
which is the two-dimensional counterpart of (24). This number plays a similar role to Mach number [37] in gas dynamics or Froude number [37] in shallow water problems. Then, we define supercritical flow as one characterized for a local \( H \) number such that \( H > 1 \). In the same way, we define subcritical flow as one characterized for \( H < 1 \) and critical flow as the flow which satisfies \( H = 1 \). It should be noted that in supercritical flow it is not possible a pollutant transport towards upstream. This fact occurs because the circular (diffusive) mode of propagation travels with a velocity lesser than the convective one.

Imposing boundary conditions for the system (30) is not trivial. Further, an inadequate choice could affect the existence and uniqueness of solutions. A very frequent situation is to have to impose a boundary condition at a fixed solid wall. In this situation only the normal
A component of flux $q$ must be prescribed. Then, the boundary condition in this case will be $q \cdot n = 0$. However, computational domains are usually limited by other kinds of boundaries, as well. These are the so-called inflow and outflow boundaries. In these types of boundaries the issue of boundary conditions is more complicated. It is well known that boundary conditions for scalar hyperbolic problems are only imposed on the inflow portion of the boundary (see, for instance [38]). For systems of hyperbolic equations only inflow components can be prescribed. As a consequence of that, a Riemann analysis in the direction of the outward normal is necessary. It has been shown that Cattaneo-type transport equations can be diagonalized (however they remain coupled by the source term) in 1D which is the case of the outward normal to the boundary. In this way, three different propagation velocities are found:

$$\omega_1 = a \cdot n$$
$$\omega_2 = a \cdot n - c$$
$$\omega_3 = a \cdot n + c$$

In order to obtain a well-posed problem so many components of $U$ must be imposed as negative velocities (incoming information) there are in (36). Therefore, in a supercritical inflow boundary all components of $U$ must be prescribed while no components should be imposed in a supercritical outflow boundary. In contrast, in a subcritical outflow boundary only one component of $U$ will be prescribed while 2 components of $U$ must be imposed in a subcritical inflow boundary. Independently of boundary conditions, an initial condition must be imposed to (26). Consequently, the two-dimensional Cattaneo-type transport problem can be stated as follows: given $k, \tau > 0$, a solenoidal velocity field $a$ and the necessary initial and boundary conditions, find $U: \Omega \times [0, T] \rightarrow \mathbb{R}^3$ such that

$$\frac{\partial U}{\partial t} + \nabla_x \cdot (F) = S$$
in $\Omega \times [0, T]$ (37.1)
$$U(x, 0) = U_0(x)$$
on $\Omega$ (37.2)
$$U^{in} = U^{in}_0$$
on $\Gamma_D \times [0, T]$ (37.3)
$$F^{in} n = F_N$$
on $\Gamma_N \times [0, T]$ (37.4)

being $F$, $S$ the vectors defined in (27), $U^{in}$ the inflow components of $U$ and $F^{in}$ the inflow components of the flux matrix.

4. NUMERICAL ANALYSIS OF THE STEADY STATE CATTANEO-TYPE TRANSPORT EQUATIONS

In this section we analyze the behavior of the numerical solution of the steady-state Cattaneo-type transport equation. We will refer to the one-dimensional problem in this part of the paper (two-dimensional problems will be studied as transient problems). The governing equations in this case are:

$$\nabla_x \cdot (F) = S$$

being $F$ and $S$ the vectors defined in (15). The above equation can be written in non-conservative form as follows:

$$A \frac{dU}{dx} = S$$

Copyright © 2000 John Wiley & Sons, Ltd.
Int. J. Numer. Meth. Engng 2000; 00:1–6
Prepared using nmeauth.cls
where $U$ is the vector defined in (15) and $A$ the Jacobian matrix defined in (17). Let us consider the domain $\Omega = (0, L)$, $L \in \mathbb{R}^+$ bounded by $\Gamma$. Clearly, $\Gamma = \Gamma_0 \cup \Gamma_L$ if we define $\Gamma_0$ as the point $x = 0$ and $\Gamma_L$ as $x = L$. By using this notation the inflow boundary (which will be denoted by $\Gamma_{\text{in}}$) will be $\Gamma_0$ when $a > 0$ and $\Gamma_L$ in the opposite case. The outflow boundary will be the complementary of the inflow one. Therefore, this problem can be stated as follows:

\[
A \frac{dU}{dx} = S \quad \text{in } \Omega \quad \text{(40.1)}
\]
\[
U_{\text{in}} = U^D_{\text{in}} \quad \text{on } \Gamma \quad \text{(40.2)}
\]

where $U_{\text{in}}$ represents the inflow components of $U$. It is straightforward that (40.1) can be written as:

\[
\frac{dq}{dx} = -\frac{du}{dx} \quad \text{(41.1)}
\]
\[
k \frac{du}{dx} + \tau a \frac{dq}{dx} = -q \quad \text{(41.2)}
\]

By introducing (41.1) into (41.2) it can be shown that

\[
q = -(k - \tau a^2) \frac{du}{dx} \quad \text{(42)}
\]

which is just true in this simple case.

Now, in order to compare the proposed model (Cattaneo-type) with the standard one (Fick-type) we will write (41) as an equivalent second order differential equation, namely

\[
a \frac{du}{dx} - (k - \tau a^2) \frac{d^2u}{dx^2} = 0 \quad \text{(43)}
\]

To derive equation (43) has been necessary to derive (41.2), so spurious solutions could have been introduced in (43). However, we can avoid to obtain spurious solutions by imposing to (43) the same boundary conditions which have been imposed to (40). In supercritical flow we should impose $u$ and $q$ on the inflow boundary. The boundary condition regarding $q$ can be imposed by using the relation (42). We must not impose boundary conditions on the supercritical outflow boundary. In a subcritical problem one component of $U$ ($u$ for instance) should be prescribed on the inflow boundary and the another one ($q$ in this case) should be imposed on the outflow boundary. However, there is no problem in prescribing concentration on $\Gamma_0$ and $\Gamma_L$ in subcritical flow because we are imposing one condition at each boundary.

In what follows we will compare the standard formulation with the proposed one. This matter will be easier if we use the expression (43) to describe the Cattaneo-type transport. In this way, we can recover the standard formulation simply by setting $\tau = 0$ in (43). We will begin this study by analyzing an homogeneous problem with Dirichlet boundary conditions. In this example Fick’s law will be used. Therefore, we state the following problem: find a function $u: [0, L] \mapsto \mathbb{R}$ such that

\[
a \frac{du}{dx} - k \frac{d^2u}{dx^2} = 0; \quad x \in (0, L) \quad \text{(44.1)}
\]
\[
u(0) = u_0 \quad \text{(44.2)}
\]
\[
u(L) = u_L \quad \text{(44.3)}
\]
Let $0 = x_0 < x_1 < \cdots < x_N = L$ be a uniform partition of the interval $[0, L]$. We call $h$ the distance between two consecutive nodes. Let us call

$$P_e = \frac{ah}{k}$$

the mesh Péclet number which expresses the ratio of convective to diffusive transport. If we solve (44) by using the standard Galerkin method and linear finite elements we obtain the following discrete equation at an interior node $j$ [39]:

$$(1 - P_e)u_{j+1} - 2u_j + (1 + P_e)u_{j-1} = 0$$

(46)

In the above equation $u_j$ is the finite element approximation of $u(x_j)$ and $u_0, u_N$ are the values given by the boundary conditions of (44). In addition, difference equations (46) can be solved exactly (see, for instance, reference [40]) as they are linear equations. The exact solution of (46) (subject to boundary conditions (44.2) and (44.3)) is

$$u_j = \frac{1}{1 - \left(\frac{1 + P_e}{1 - P_e}\right)^j} \left\{ u_0 \left[ \left(\frac{1 + P_e}{1 - P_e}\right)^j - \frac{1}{N} \left(\frac{1 + P_e}{1 - P_e}\right)^j \right] + u_L \left[ 1 - \left(\frac{1 + P_e}{1 - P_e}\right)^j \right] \right\}$$

(47)

so oscillations will occur when $|P_e| > 1$. On the other hand, the exact solution of (44) is

$$u(x_j) = \frac{1}{1 - e^{\frac{-ah}{k}}} \left[ u_0 \left( e^{\frac{ah}{k}j} - e^{\frac{ah}{k}N} \right) + u_L \left( 1 - e^{\frac{ah}{k}j} \right) \right]$$

(48)

A simple comparison between (47) and (48) shows that the approximate solution will be equal to the exact one if the following relation holds

$$e^{2P_ej} = \left(\frac{1 + P_e}{1 - P_e}\right)^j \quad \forall j = 0, \ldots, N$$

(49)

Relation (49) is only satisfied for $P_e = 0$ (pure-diffusive problem). However, it can be shown that when $|P_e| \leq 1$ the approximate solution (47) is, in fact, the exact solution of the problem

$$a \frac{du}{dx} - k^* \frac{d^2u}{dx^2} = 0; \quad x \in (0, L)$$

(50.1)

$$u(0) = u_0$$

(50.2)

$$u(L) = u_L$$

(50.3)

for a certain $k^* \leq k$. To show this assertion we will use (47) and (48). The relation between $k$ and $k^*$ is governed by

$$e^{\frac{ah}{k^*}} = \left(\frac{1 + P_e}{1 - P_e}\right)^j \quad \forall j = 0, \ldots, N$$

(51)

We want to obtain $k^*$ such that (51) holds. If we admit complex solutions, then $k^*$ can always be determined. If we exclusively admit real numbers, then (51) has a solution only when $|P_e| \leq 1$. The solution of (51) is

$$k^* = k \frac{2P_e}{\ln \left(\frac{1 + P_e}{1 - P_e}\right)}$$

(52)
By means of (52) we notice that \( k^* \to 0 \) as \(|P_e| \to 1\) and \( k^* \to k \) as \(|P_e| \to 0\). See figure 2 where \( k^*/k \) is represented for \( P_e \in [-1,1] \). Therefore, the standard Galerkin method applied to (44) solves exactly an underdiffusive equation. On the other hand, the equation (52) can be rearranged as

\[
k^* = k - k \left( 1 - \frac{2P_e}{\ln \left( \frac{1+P_e}{1-P_e} \right)} \right) < k
\]  

(53)

If we compare the diffusive coefficient \( k^* \) with the coefficient which results of using Cattaneo’s law (see equation (43)) the following result is achieved: when we solve (44) by using the standard Galerkin method we obtain the solution of a Cattaneo-type transport problem defined by the following relaxation time:

\[
\tau_{G-FEM} = \frac{h}{a} \left( \frac{1}{2P_e} - \frac{1}{\ln \left( \frac{1+P_e}{1-P_e} \right)} \right)
\]  

(54)

Therefore, an “artificial” relaxation time has been introduced by the Galerkin formulation. As a result, a finite velocity of propagation can be defined in the discrete equation (46):

\[
c_{G-FEM} = \frac{a}{\left( 1 - \frac{2P_e}{\ln \left( \frac{1+P_e}{1-P_e} \right)} \right)^{1/2}}
\]  

(55)

By using the relation (55) it is easy to compute the value of “artificial” \( H \) (the dimensionless number defined in (24)) for a certain \( P_e \). In figure 2 has been represented the “artificial” \( H \) as a function of \( P_e \). As a result, when we solve the problem (44) for \(|P_e| < 1\) by using the standard Galerkin method we are really solving a Cattaneo-type transport problem in subcritical flow. Therefore, we solve a well posed problem because boundary conditions (44.2), (44.3) can be imposed in subcritical flow. However, as \(|P_e| \to 1\) the problem which is really solved tends to an ill-posed problem.

In what follows we will solve the steady state Cattaneo-type transport equation in subcritical and supercritical flows. In order to make easier the comparison between the Cattaneo-type
transport and the standard formulation of the transport problem we will use the equation (43) to describe the proposed model. However, we must take into account that boundary conditions which should be imposed are (40.2).

4.1. Numerical examples in subcritical flow

In this section we will analyze the problem
\[ \frac{\partial^2 u}{\partial x^2} - (k - \tau a^2) \frac{\partial^2 u}{\partial x^2} = 0; \quad x \in (0, L) \]  
(56.1)
\[ u(0) = u_0 \]  
(56.2)
\[ u(L) = u_L \]  
(56.3)
which represents a Cattaneo-type transport problem only in subcritical flow. Let us consider once more the partition of \([0, L]\) defined by the nodes \(0 = x_0 < x_1 < \cdots < x_N = L\). We call \(h = L/N\). At this point it is very useful to define the dimensionless number
\[ H_e = \frac{ah}{2(k - \tau a^2)} \]  
(57)
which plays a similar role to \(P_e\) in the standard description of the transport problem. If we solve (56) by using the standard Galerkin method and linear finite elements, the following difference equations are found:
\[ (1 - H_e)u_{j+1} - 2u_j + (1 + H_e)u_{j-1} = 0; \quad \forall j = 1, \ldots, N - 1 \]  
(58)
where \(u_0\) and \(u_N\) are given by the boundary conditions (56.2) and (56.3). In the same way as (46), difference equations (58) can be solved exactly and the stability condition
\[ |H_e| \leq 1 \]  
(59)
can be found. If we take \(\tau = 0\) in (59) we obtain
\[ |P_e| \leq 1 \]  
(60)
which constitutes a stability condition for the standard formulation. Relations (59) and (60) seem to be useless because they can only be applied to (56.1) and (44.1). Indeed, the above assertion is true in the case of (60). However, the asymptotic behavior of (59) is equivalent (except for a scale factor) to impose that the grid step size is smaller than typical sizes related to the waves which give the solution of the Cattaneo-type transport problem. As we said before, the waves which determine the solution propagate with celerities \(a - c\) and \(a + c\). Thus, typical sizes upstream and downstream are \(\tau(c - a)\) and \(\tau(a + c)\), respectively. Hence, it is possible to show [31] that
\[ h < \min (\tau(c - a), \tau(a + c)) \]  
(61)
tends to (59) as \(a\) tends to the mass wave celerity \(c\), except for a scale factor. Hence, stability condition (61) could be used to develop stability conditions for more complicated problems by using its physical meaning.

Now we represent the approximate solution and the exact solution of the Cattaneo-type transport problem. We use several sets of values for the parameters of the problem. Two groups

\[ \text{Copyright } \odot 2000 \text{ John Wiley & Sons, Ltd.} \]  
\[ \text{Int. J. Numer. Meth. Engng 2000; 00:1–6} \]  
\[ \text{Prepared using nmeauth.cls} \]
Figure 3. Three subcritical transport problems by using Cattaneo’s law. These problems are defined by $k = 1$, $\tau = 0.01$ and three different $H$ values: $H = 0.7$ (left), $H = 0.88278221857319$ (center), $H = 0.975$ (right). Numerical solutions are obtained by using the Galerkin FEM with linear shape functions. A 20 element grid has been used.

of numerical examples will be presented. At each group the relaxation time is a constant. In the same way, the grid step size, the diffusivity, the domain length and the boundary values are the same for all the numerical examples. However, at each group we will show three results defined by the fluid velocity $a$. For all examples in section 4 we use a 20 element discretization, $L = 1$ (thus, $h = 0.05$) and $k = 1$.

- **Group 1: small relaxation time**
  This first group of results is defined by $\tau = 0.01$. This is a small value for the relaxation time $\tau$, and hence this example is near Fick’s law. By using the above values for $k$ and $\tau$ we obtain the diffusive wave celerity $c = \sqrt{k/\tau} = 10$. Thus, if $|a| \geq 10$, then (56) does not represent a Cattaneo-type problem. Our next step will be calculate the maximum $a$ value to obtain a stable solution of (56) by using the stability condition $|H_e| \leq 1$. If we do that we will see that the numerical scheme used will give unstable solutions when $|a| > 8.8278221857319$. Therefore, we can say that the numerical solution of (56) is stable for almost all possible values of $a$, because (56) does not represent a Cattaneo-type transport problem when $|a| \geq 10$.

In figure 3 we show the numerical (dashed line) and the exact (solid line) solutions for three $a$ values. On the left, solutions for $a = 7$ are plotted. The middle graphic shows solutions for $a = 8.8278221857319$ which is the greatest $a$ value that gives a stable solution. Finally, we plot solutions for $a = 9.75$ on the right.

- **Group 2: medium relaxation time**
  This group of problems is defined by $\tau = 1$. Therefore, the mass wave celerity is $c = \sqrt{k/\tau} = 1$. In addition, according to the stability condition (59), the largest velocity that gives a stable solution is $a = 0.9875812194848$. Hence, we will obtain stable solutions if $|a| \leq 0.9875812194848$. We show three numerical tests for this relaxation time in figure 4. Solutions for $a = 0.97$ (left), $a = 0.9875812194848$ (center) and $a = 0.995$ (right) are plotted.
4.2. Numerical examples in supercritical flow

In this section, the numerical resolution of the Cattaneo-type transport problem in supercritical flow will be performed. We will use linear finite elements and the standard Galerkin method. As we said above, in supercritical flow $u$ and $q$ must be prescribed on the inflow boundary. Boundary conditions involving $q$ will be imposed by using (42). There is no restriction in supposing $a > 0$. In this way, we can determine the inflow boundary which is $\Gamma_0$. Accordingly, we will analyze the problem

$$\frac{d}{dx} - (k - \tau a^2) \frac{d^2u}{dx^2} = 0; \quad x \in (0, L) \quad (62.1)$$

$$u(0) = u_0 \quad (62.2)$$

$$\frac{d}{dx}(0) = -\frac{q_0}{k - \tau a^2} \quad (62.3)$$

which could also be analyzed by solving the system of equations (41) provided with boundary conditions $u(0) = u_0$, $q(0) = q_0$. Further, by solving (41) we could impose flux boundary conditions in a strong form instead of imposing it in a weak form. However, we analyze the Cattaneo-type transport problem by solving (62) in order to be consistent with the previous section. Once more, two groups of numerical examples will be presented. At each group the relaxation time is a constant. The values of $L$, $k$ and $h$ are the same as in the previous section. The boundary values ($u_0 = 1$ and $q_0 = -1$) are the same for all the numerical tests.

- Group 1: small relaxation time

  This first group is defined by $\tau = 0.01$. This is a small value for the relaxation time $\tau$, and hence this example is near Fick’s law. It is straightforward to compute $c = \sqrt{k/\tau} = 10$. It has already been said that (62) represents a Cattaneo-type transport problem when $|a| > 10$. The minimum $a$ value to obtain a stable solution is given by the condition $|H_e| \leq 1$. Making some algebra in the above relation we conclude that stable solutions will be obtained when $|a| \geq 11.32782218537319$. Therefore, we can say that the numerical solution of (62) is stable for almost all possible $a$ values.
Figure 5. Three supercritical transport problems by using Cattaneo’s law. These problems are defined by $k = 1, \tau = 0.01$ and three different $H$ values: $H = 1.2$ (left), $H = 1.05$ (right). Numerical solutions are obtained by using the Galerkin FEM with linear shape functions. A 20 element grid has been used.

Figure 6. Three supercritical transport problems by using Cattaneo’s law. These problems are defined by $k = 1, \tau = 1$ and three different $H$ values: $H = 1.05$ (left), $H = 1.005$ (right). Numerical solutions are obtained by using the Galerkin FEM with linear shape functions. A 20 element grid has been used.

In figure 5 we show the numerical (dashed line) and exact (solid line) solutions for three $a$ values. On the left, solutions for $a = 12$ are plotted. The middle graphic shows solutions for $a = 1.25781218537319$ which is the minimum $a$ value that gives a stable solution. Finally, we plot solutions for $a = 1.05$ on the right.

- **Group 2: medium relaxation time**
  This group is defined by the relaxation time $\tau = 1$. Therefore, the mass wave celerity is $c = 1$. We show three numerical tests for this relaxation time in figure 6. Solutions for $a = 1.05$ are plotted on the left. The middle graphic shows solutions for $a = 1.01257812194848$ which is the minimum $a$ value that gives a stable solution. Finally, we plot solutions for $a = 1.005$ on the right.
4.3. Conclusions from the numerical examples

By means of the above numerical results we have shown that if we use Cattaneo’s law we obtain stable solutions in a very significative part of the domain of $a$. This is true even if we use the standard Galerkin method. Hence, from a practical point of view, we can say that the transport equation by using Cattaneo’s law is a stable equation because the values of $a$ that make its solution unstable are negligible even for small relaxation times. Further, we conclude that as $\tau$ increases transport problem becomes more stable. Indeed, it can be proved that the size (in the velocity domain) of the interval that leads to unstable solutions is

\[ I = \frac{h}{\tau} \quad (63) \]

which decreases as $\tau$ increases. In order to prove the above assertion we will find the $a$ values that make

\[ |H_e| = 1 \quad (64) \]

It is easy to prove that (64) possesses 4 real solutions, namely

\[ a_1 = -\frac{h}{4\tau} - \sqrt{\left(\frac{h}{4\tau}\right)^2 + c^2} \quad (65.1) \]
\[ a_2 = -\frac{h}{4\tau} + \sqrt{\left(\frac{h}{4\tau}\right)^2 + c^2} \quad (65.2) \]
\[ a_3 = -a_2 \quad (65.3) \]
\[ a_4 = -a_1 \quad (65.4) \]

Further, it is straightforward that $a_1 < 0$, $a_1 < -c$, $a_2 > 0$, $a_2 < c$. Taking into account all of this, the interval of velocities that makes the numerical solution unstable has a size of

\[ I = a_4 - a_2 + a_3 - a_1 = -2(a_1 + a_2) = \frac{h}{\tau} \quad (66) \]

as we said above.

5. NUMERICAL ANALYSIS OF THE TRANSIENT CATTANEO-TYPE TRANSPORT EQUATIONS

In this section we will mainly refer to two-dimensional Cattaneo-type transport equations. Three-dimensional equations could be discretized in the same way as we will discretize two-dimensional ones. We will use Taylor-Galerkin methods proposed by Donea in [41] (see also [42, 43]). Such methods represent an attempt to take into account (by a Taylor series in time extended to second, third or fourth order) the directional character of propagation of information in convective transport. In this study second and third order Taylor-Galerkin methods will be used.

5.1. Second order Taylor-Galerkin method

In this section the numerical discretization of the Cattaneo-type transport equations will be performed. Governing equations written in conservative form are (26) which we rewrite here.
for completeness

\[ \frac{\partial U}{\partial t} + \nabla_x \cdot (F) = S \]  

(67)

In the above

\[ U = \begin{pmatrix} u \\ \tau q_1 \\ \tau q_2 \end{pmatrix}; \quad F = \begin{pmatrix} u a_1 + q_1 \\ \tau q_1 a_1 + k u \\ \tau q_2 a_2 + k u \end{pmatrix}; \quad S = \begin{pmatrix} 0 \\ -q_1 \\ -q_2 \end{pmatrix} \]  

(68)

The relation (28) and the notation (29) will also be used. We will solve the equation (67) in a domain \( \Omega \) with a piecewise smooth boundary \( \Gamma \). The unit outward normal vector to \( \Gamma \) is denoted by \( n = (n_1, n_2)^T \). In addition, one initial condition and the necessary boundary conditions (see section 3.3.2) should be imposed to this equation.

The numerical algorithm proposed by Donea can be described as follows: Let \( U^n \) be the finite element approximation at time \( t^n = n\Delta t \), being \( \Delta t \) the time increment. The solution \( U^{n+1} \) at the next time level \( t^{n+1} = t^n + \Delta t \) is obtained by performing a second order Taylor series expansion about \( t = t^n \), as

\[ U^{n+1} = U^n + \Delta t \left( \frac{\partial U}{\partial t} \right)^n + \frac{\Delta t^2}{2} \left( \frac{\partial^2 U}{\partial t^2} \right)^n \]  

(69)

Then, time derivatives of \( U \) are replaced by spatial derivatives using the original equation (67):

\[ \frac{\partial U}{\partial t} = S - \nabla_x \cdot (F) \]  

(70.1)

\[ \frac{\partial^2 U}{\partial t^2} = B(S - \nabla_x \cdot (F)) + \nabla_x \cdot (P) \]  

(70.2)

where \( B \) is the Jacobian matrix associated to the source term, namely

\[ B = \nabla_U (S) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1/\tau & 0 \\ 0 & 0 & -1/\tau \end{pmatrix} \]  

(71)

and \( P \) is the matrix which results of assembling the column vectors

\[ P_i = -A_i(S - \nabla_x \cdot (F)); \quad i = 1, 2 \]  

(72)

If we define \( \Delta U = U^{n+1} - U^n \) and we take into account that the medium is homogeneous, the following relation holds:

\[ \Delta U = b - \nabla_x \cdot (G) \]  

(73)

being

\[ b = \Delta t S + \frac{\Delta t^2}{2} BS \]  

(74.1)

\[ G = \Delta t F + \frac{\Delta t^2}{2} BF - \frac{\Delta t^2}{2} P \]  

(74.2)

The weighted residual formulation of (73) by using the standard Galerkin method leads to the following vectorial equation at an interior node \( B \)

\[ \sum_A M_{BA} \Delta U_A = f_B \]  

(75)
A GENERALIZED STATEMENT FOR ADVECTIVE-DIFFUSIVE PHENOMENA

where

\[ M_{BA} = \int_\Omega N_B N_A d\Omega \quad (76.1) \]

\[ f_B = \int_\Omega N_B b d\Omega + \int_\Omega G \nabla_x (N_B) d\Omega - \int_\Gamma N_B (G_1 n_1 + G_2 n_2) d\Gamma. \quad (76.2) \]

\( N_A, N_B \) are the interpolating functions which possess \( C^0 \) continuity and \( G_i, \ i = 1, 2 \) is the \( i \)-th column of the matrix \( G \).

Several numerical tests performed in one-dimensional problems have been shown that the algorithm (75) can be implemented by using the so-called lumped mass matrix (obtained by nodal integration of (76.1)) without losing excessive accuracy.

5.2. Third order Taylor-Galerkin method

The third order Taylor-Galerkin method is based on a Taylor series expansion up to third order in time. In order to make possible an implementation with \( C^0 \)-continuous elements, the third order derivative of \( U \) is not totally replaced by spatial derivatives but it is written in a mixed spatial-temporal form. This mixed form leads to a modification of the usual mass matrix when a forward Euler approximation is used for the temporal part of the mixed derivative. Another procedure to achieve third order of accuracy in time which allows to use \( C^0 \) interpolating functions is to perform a two-step implementation to march in time. We will use the scheme proposed by Selmin in [44], namely

\[ \tilde{U}^n = U^n + \frac{1}{3} \Delta t \left( \frac{\partial U}{\partial t} \right)^n + \alpha \Delta t^2 \left( \frac{\partial^2 U}{\partial t^2} \right)^n \quad (77.1) \]

\[ U^{n+1} = U^n + \Delta t \left( \frac{\partial U}{\partial t} \right)^n + \frac{1}{2} \Delta t^2 \left( \frac{\partial^2 U}{\partial t^2} \right)^n \quad (77.2) \]

where \( \alpha \) is a parameter which only influences the coefficient of the fourth order term. Selmin proposed to determine \( \alpha \) in order to obtain (in the one-dimensional linear pure-advective problem) a scheme with the same phase response as the one-step third order Taylor-Galerkin. This procedure leads to \( \alpha = 1/9 \). In addition, in two-dimensional problems the two-step algorithm has a greater stability domain than the one-step one (see [44]). In the same way as the second order Taylor-Galerkin method, the third order algorithm will be implemented by using a lumped mass matrix, as well.

5.3. Numerical examples

This section is devoted to the presentation of the results obtained by using the numerical schemes presented in sections 5.1 and 5.2. Two practical examples will be presented. The first one consists of a pollutant transport in a rectangular channel. We will solve this problem by using two different velocity fields. The second example consists of the convection and diffusion of a Gaussian hill in a pure rotation velocity field. The unstructured meshes were generated by using the code GEN4U based on the algorithm by Sarrate and Huerta [45].

5.3.1. Subcritical flow in a rectangular channel

We consider a channel of 10 m. length and 2 m. width over the rectangle \([0, 10] \times [0, 2]\). A typical mesh of quadrilaterals has been used.
Grid steps size are $\Delta x_1 = \Delta x_2 = 0.25$ m. The values of the parameters are $k = 1$ and $\tau = 1$ which implies $c = 1$. The velocity field employed in the calculations is

$$a(x_1, x_2) = (0.8x_2(2 - x_2), 0)^T$$

(78)

which is a divergence free field. By using relations (35) and (78) it is straightforward that this flow is subcritical at each point of the domain. As initial conditions we consider

$$u(x_1, x_2) = e^{-(x_1^2 + x_2^2)}$$

(79.1)

$$q(x_1, x_2) = 0$$

(79.2)

The initial condition involving concentration has been plotted in figure 7. As boundary conditions, we prescribe concentration and $q_2$ on the inflow boundary; $q_1$ on the outflow boundary and $q_2 = 0$ at the solid wall (in all cases we impose the values which result of initial conditions). In what follows we will present the results obtained by using second and third order Taylor-Galerkin methods. A time step of $\Delta t = 0.1$ s. has been used in the calculations. We step in time until the steady state has been reached. The solution is considered to be steady when

$$\text{Residual} = \frac{||U^{n+1} - U^n||}{||U^{n+1}||} \leq 10^{-5}$$

(80)

- Second order Taylor-Galerkin method

We show concentration solutions at 4 different times: At $t = 3$ s. (figure 8), at $t = 6$ s. (figure 9), at $t = 9$ s. (figure 10) and at $t = 12$ s. (figure 11). Further, we plot the steady state solution in figure 12.

- Third order Taylor-Galerkin method

We show concentration solutions at 4 different times: At $t = 3$ s. (figure 13), at $t = 6$ s. (figure 14), at $t = 9$ s. (figure 15) and at $t = 12$ s. (figure 16). Further, we plot the steady state solution in figure 17.
Figure 9. Subcritical channel flow. Solution at $t = 6$ s. by using second order Taylor-Galerkin method.

Figure 10. Subcritical channel flow. Solution at $t = 9$ s. by using second order Taylor-Galerkin method.

Figure 11. Subcritical channel flow. Solution at $t = 12$ s. by using second order Taylor-Galerkin method.

Figure 12. Subcritical channel flow. Steady state solution by using second order Taylor-Galerkin method.

Figure 13. Subcritical channel flow. Solution at $t = 3$ s. by using third order Taylor-Galerkin method.
5.3.2. Transcritical flow in a rectangular channel

We consider a channel of 10 m. length and 2 m. width over the rectangle $[0, 10] \times [0, 2]$. The computational mesh (1375 elements) is plotted in figure 18. The values of the parameters are $k = 10^{-2}$ and $\tau = 1$. Therefore, the mass wave celerity is $c = 0.1$. We will solve this problem by using the velocity field

$$a(x_1, x_2) = (2x_2(2 - x_2), 0)^T$$  \hspace{1cm} (81)

which is a solenoidal one. By using (35) we can find the flow to be supercritical except near the walls of the channel. Indeed, this is a high velocity problem as it satisfies $H = 20$ at the straight line $x_2 = 1$. We use (79) as initial conditions. The one involving concentration has been depicted in figure 19. As boundary conditions, we prescribe all unknowns on the
A GENERALIZED STATEMENT FOR ADVECTIVE-DIFFUSIVE PHENOMENA

Figure 18. Transcritical channel flow. Computational mesh (1375 elements).

Figure 19. Transcritical channel flow. Initial condition involving concentration.

Figure 20. Transcritical channel flow. Solution at $t = 1.5$ s. by using second order Taylor-Galerkin method.

supercritical inflow boundary (we impose the values which result of the initial conditions). At the solid wall we impose $q_2 = 0$ which is also consistent with the initial conditions. In what follows we will present the results obtained by using second and third order Taylor-Galerkin methods. A time step of $\Delta t = 2 \times 10^{-2}$ s. has been used in the calculations. We march in time until the relation (80) is satisfied.

- Second order Taylor-Galerkin method
  We show concentration solutions at 6 different times: At $t = 1.5$ s. (figure 20), at $t = 3$ s. (figure 21), at $t = 4.5$ s. (figure 22), at $t = 6$ s. (figure 23) and at $t = 12$ s. (figure 24). Further, we plot the steady state solution in figure 25.

- Third order Taylor-Galerkin method
  We show concentration solutions at 6 different times: At $t = 1.5$ s. (figure 26), at $t = 3$ s. (figure 27), at $t = 4.5$ s. (figure 28), at $t = 6$ s. (figure 29) and at $t = 12$ s. (figure 30). Further, we plot the steady state solution in figure 31.

5.3.3. The rotating cone problem In this example we study the convection and diffusion of a Gaussian hill in a pure rotation. A uniform mesh of $30 \times 30$ four-node elements over the unit square $[-1/2, 1/2] \times [-1/2, 1/2]$ is employed in the calculations. The values of the parameters
Figure 21. Transcritical channel flow. Solution at $t = 3$ s. by using second order Taylor-Galerkin method.

Figure 22. Transcritical channel flow. Solution at $t = 4.5$ s. by using second order Taylor-Galerkin method.

Figure 23. Transcritical channel flow. Solution at $t = 6$ s. by using second order Taylor-Galerkin method.

Figure 24. Transcritical channel flow. Solution at $t = 12$ s. by using second order Taylor-Galerkin method.

Figure 25. Transcritical channel flow. Steady state solution by using second order Taylor-Galerkin method.
Figure 26. Transcritical channel flow. Solution at $t = 1.5$ s. by using third order Taylor-Galerkin method.

Figure 27. Transcritical channel flow. Solution at $t = 3$ s. by using third order Taylor-Galerkin method.

Figure 28. Transcritical channel flow. Solution at $t = 4.5$ s. by using third order Taylor-Galerkin method.

Figure 29. Transcritical channel flow. Solution at $t = 6$ s. by using third order Taylor-Galerkin method.

Figure 30. Transcritical channel flow. Solution at $t = 12$ s. by using third order Taylor-Galerkin method.
Figure 31. Transcritical channel flow. Steady state solution by using third order Taylor-Galerkin method.

Figure 32. The rotating cone problem. Initial condition involving concentration.

are $k = 10^{-2}$ and $\tau = 1$. The velocity field is $\mathbf{u}(x_1, x_2) = (-x_2, x_1)^T$ which is a solenoidal one. Therefore, by using (35) it can be shown that into the open circle $x_1^2 + x_2^2 < 0.1^2$ the flow is subcritical, on the circumference $x_1^2 + x_2^2 = 0.1^2$ the flow is critical and in the complementary part of the domain the flow is supercritical. As initial conditions we use

$$u(x_1, x_2) = e^{-60(x_1^2+x_2^2)}$$

$$q(x_1, x_2) = 0$$

The initial condition involving concentration has been depicted in figure 32. We impose boundary conditions consistent with initial conditions on the inflow boundary. We have solved this problem by using second and third order Taylor-Galerkin methods. A time step of $\Delta t = 0.01$ s. has been used in the calculations. We have marched in time until a complete revolution has been finished. Solutions after a quarter of a revolution, two quarters of a revolution, three quarters of a revolution, and after a complete revolution have been plotted in figures 33–36.
Figure 33. The rotating cone problem. Solution after a quarter of a revolution (left) and two quarters of a revolution (right) by using the second order Taylor-Galerkin method.

Figure 34. The rotating cone problem. Solution after three quarters of a revolution (left) and after a complete revolution (right) by using the second order Taylor-Galerkin method.
Figure 35. The rotating cone problem. Solution after a quarter of a revolution (left) and two quarters of a revolution (right) by using the third order Taylor-Galerkin method.

Figure 36. The rotating cone problem. Solution after three quarters of a revolution (left) and after a complete revolution (right) by using the third order Taylor-Galerkin method.
6. CONCLUSIONS

In this paper we propose to use Cattaneo’s law as the constitutive equation of the advective-diffusive transport problem. This approach leads to a wave-like solution. The Cattaneo-type formulation avoids the infinite speed paradox which is reached by using the standard formulation. The proposed formulation constitutes a generalized approach for advective-diffusive phenomena because the standard formulation can be considered as a subcase of the proposed one. When Cattaneo’s law is used two kinds of flow can occur: subcritical flow (pollutant can propagate towards any direction) and supercritical flow (propagation towards upstream is not possible).

From a numerical point of view we have shown that the discrete equations of the standard steady model represent, in fact, a Cattaneo-type transport when the standard Galerkin formulation is employed. Therefore, an ill-posed discrete problem can be obtained as boundary conditions could be incompatible with discrete equations. We have shown that the proposed model leads to more stable numerical solutions than the classic one when the standard Galerkin method is used for the spatial discretization. Finally, several numerical tests have been performed to show that the Cattaneo-type model can be used in real engineering problems.

ACKNOWLEDGEMENTS

This work has been partially supported by Grant # PGDIT01PXI11802PR of the SXID of the “Xunta de Galicia”, Grant # PGDIT03PXIC11802PN of the SXID of the “Xunta de Galicia”, Grant # DPI200200297 of the SGPICt of the “Ministerio de Ciencia y Tecnología” of the Spanish Government, research fellowships of the “Universidad de A Coruña” and research fellowships of the “Fundación de la Ingeniería Civil de Galicia”.

REFERENCES


32. Cattaneo, M.C., Sur une forme de l’équation de la chaleur liminant le paradoxe d’une propagation.