

RESEARCH ARTICLE

Open Access



# On the issue that Finite Element discretizations violate, nodally, Clausius's postulate of the second law of thermodynamics

Alejandro Limache<sup>1†</sup> and Sergio Idelsohn<sup>1\*,2†</sup>

\* Correspondence:

sergio@cimne.upc.edu

<sup>†</sup>Both authors contributed equally to this work.

<sup>1</sup>CIMEC-Research Center of Computational Methods (UNL/CONICET), Ruta 168 s/n, Predio Conicet "Dr A. Cassano", 3000 Santa Fe, Argentina  
Full list of author information is available at the end of the article

## Abstract

Discretization processes leading to numerical schemes sometimes produce undesirable effects. One potentially serious problem is that a discretization may produce the loss of validity of some of the physical principles or mathematical properties originally present in the continuous equation. Such loss may lead to uncertain results such as numerical instabilities or unexpected non-physical solutions. As a consequence, the compatibility of a discrete formulation with respect to intrinsic physical principles might be essential for the success of a numerical scheme. This paper addresses such type of issue. Its main objective is to demonstrate that standard Finite Element discretizations of the heat conduction equation violate Clausius's postulate of the second law of thermodynamics, at nodal level. The problem occurs because non-physical, reversed nodal heat-fluxes arise in such discretizations. Conditions for compatibility of discrete nodal heat-fluxes with respect to Clausius's postulate are derived here and named discrete thermodynamic compatibility conditions (DTCC). Simple numerical examples are presented to show the undesirable consequences of such failure. It must be pointed out that such DTCCs have previously appeared in the context of the study of the conditions that make discrete solutions to satisfy the discrete maximum principle (DMP). However, the present article does not put attention on such mathematical principle but on the satisfaction of a fundamental physical one: the second law of thermodynamics. Of course, from the presented point of view, it is clear that the violation of such fundamental law will cause, among different problems, the violation of the DMP.

**Keywords:** Finite Element discretization, Violation of the second law of thermodynamics, Heat equation, Clausius's postulate

## Background

Numerical methods intend to solve, in a discrete approximation, physical phenomena described by continuous differential equations. However, it is important to be aware that, due to the discretization procedure, a physical principle originally present in the continuous equation could no longer remain valid in the corresponding numerical scheme. This can be the cause of severe failures of numerical methods, including the fact that smooth

but non-physical solutions can be obtained. Different types of physical compatibility problems have been studied in the literature. For example, Limache et al. [1] have studied the consequences of using Finite Element discretizations that violate the physical principle of objectivity. Some authors have studied the issue of developing time-integration methods that restore energy and momentum conservation [2]. Many authors have also study discretization methods that preserve the satisfaction of the Maximum Principle. The Maximum Principle states maximal properties that characterize solutions of certain second order PDEs (like the heat equation to be discussed here). The discrete maximum principle (DMP) [3–7], then refers to the satisfaction of these maximal properties by the corresponding discrete solutions.

Quite similarly to the examples mentioned above, this informative research paper deals with the issue of the compatibility of spatial discretizations with respect to Clausius's postulate of second law of thermodynamics, at nodal level. The issue is revealed by studying the general structure of spatial discretizations of the heat equation. From the resulting semi-discrete equations it is seen that their discrete operators must satisfy certain algebraic conditions in order to guarantee that only thermodynamically compatible nodal heat-fluxes exist. If these conditions, named here discrete thermodynamic compatibility conditions (DTCC), are not satisfied non-physical reversed heat fluxes will appear between nodes, violating Clausius's postulate. Other types of DTCC related to other thermodynamic aspects, like energy conservation, may exist and will not be considered here.

The present article is organized as follows. In "Space-discretizations of the unsteady heat conduction equation" section, the unsteady heat conduction equation is introduced together with a general expression of its corresponding (spatially) discrete equations. FE spatial discretizations are presented as particular cases of the above general expression. Also for completeness, time-discretization is briefly discussed. In "Discrete thermodynamic compatibility conditions" section, Clausius's postulate is presented and the DTCC are derived. In "Thermodynamic incompatibility of Finite Element spatial discretizations" section, FE spatial discretizations are analyzed and it is seen that they do not satisfy the DTCC. Simple examples are given to show the effects of the violation of Clausius's postulate. In "On the issue of generating thermodynamically compatible Finite-Element spatial discretizations" section, it is discussed what alternative non-consistent Finite Element formulations can be used in order to recover discrete thermodynamic compatibility. In "Related final comments" section, final comments and open issues are presented and discussed.

**Notation** Given an arbitrary field  $f(\mathbf{x}, t)$ , function of position  $\mathbf{x}$  and time  $t$ , the partial derivative with respect to time will be denoted as  $\dot{f}(\mathbf{x}, t)$ , so:

$$\dot{f}(\mathbf{x}, t) \doteq \frac{\partial f(\mathbf{x}, t)}{\partial t}$$

Similarly, given an arbitrary function  $g = g(t)$  of time  $t$ , the time-derivative will be denoted as  $\dot{g}(t)$ , so:

$$\dot{g}(t) \doteq \frac{dg(t)}{dt}$$

Whenever there is no risk of confusion, the explicit dependence on  $\mathbf{x}$  and  $t$  will be dropped so,  $\dot{f}(\mathbf{x}, t)$  and  $\dot{g}(t)$ , will be simply written as  $\dot{f}$  and  $\dot{g}$ , respectively.

## Space-discretizations of the unsteady heat conduction equation

### The heat conduction equation

The heat conduction equation:

$$\rho c_v \dot{T} = \nabla \cdot (\kappa \nabla T) \quad (1)$$

defines the evolution of physical temperature  $T(\mathbf{x}, t)$  as a function of time  $t$  and position  $\mathbf{x}$  in a body occupying a domain  $\Omega$  in space. In the above equation,  $\rho$  is the material density,  $c_v$  is the specific heat and  $\kappa$  the conductivity. In Eq. (1), it has been assumed that there are not external heat sources inside the domain. The heat conduction equation is based on the physical assumption that the heat flux  $\mathbf{q}$  is connected to the temperature gradient through Fourier's law of heat-flux:

$$\mathbf{q} = -\kappa \nabla T \quad (2)$$

Equations (1), (2) imply that the increase in temperature in an domain is directly proportional to the net heat-fluxes  $q$  entering the domain:

$$\dot{T} \propto \sum_{\text{fluxes}} q \quad (3)$$

To solve Eq. (1), initial conditions:

$$T(\mathbf{x}, 0) = T_0(\mathbf{x}) \quad (4)$$

and boundary conditions must be provided. Here, only homogeneous Neumann conditions will be used:

$$\mathbf{q} \cdot \mathbf{n} = 0 \quad \text{at} \quad \partial\Omega \quad (5)$$

The above condition and the absence of heat sources guarantee that the body is fully isolated from the exterior.

*Remark 1* The consideration of fully isolated bodies allows to study the mechanics of heat conduction in pure form, without the interference of external perturbations. The addition of heat sources or the use of other boundary conditions do not affect the results presented in this paper and their addendum would only mean an unnecessary complication.

### Spatial discretizations of the heat equation

Most commonly used discretization methods are based on the reduction of the infinite dimensional representation of the temperature field  $T(\mathbf{x}, t)$  to a finite-dimensional representation  $T^h(\mathbf{x}, t)$  in terms of values of temperature  $T_j(t) = T(\mathbf{x}_j, t)$  at certain points  $\mathbf{x}_j$  in the domain. These discrete points are called nodes. Following the adopted notation convention, nodal temperature-rates will be denoted by  $\dot{T}_j(t)$ . Assuming a spatial discretization in terms of  $n$  nodal points, the exact temperature distribution  $T(\mathbf{x}, t)$  is approximated by a discrete approximation  $T^h(\mathbf{x}, t)$ :

$$T(\mathbf{x}, t) \simeq T^h(\mathbf{x}, t) = \sum_{j=1}^n \varphi_j(\mathbf{x}) T_j(t) \quad (6)$$

where  $\varphi_j(\mathbf{x})$  are basis functions whose explicit form depend on the particular method being used. Time-differentiation of Eq. (6) leads to the following approximation of temperature-rates:

$$\dot{T}(\mathbf{x}, t) \simeq \dot{T}^h(\mathbf{x}, t) = \sum_{j=1}^n \varphi_j(\mathbf{x}) \dot{T}_j(t) \quad (7)$$

Whenever possible the explicit dependence on  $t$  will be dropped, so  $T_j(t)$  and  $\dot{T}_j(t)$  will be simply denoted by  $T_j$  and  $\dot{T}_j$ , respectively. When Eqs. (6), (7) are replaced into a differential or integral form of Eq. (1), the following general form of semi-discrete equations is obtained:

$$\mathbb{M}\dot{\mathbf{T}} = -\mathbb{K}\mathbf{T} \quad (8)$$

In the equation above  $\mathbf{T} = [T_j]$  is the column vector of nodal temperatures and  $\dot{\mathbf{T}} = [\dot{T}_j]$  is the column vector of nodal temperature-rates.  $\mathbb{M} = [\mathbb{M}_{ij}]$  and  $\mathbb{K} = [\mathbb{K}_{ij}]$  are matrix-operators known as the mass matrix and the diffusion matrix, respectively. The mass matrix is associated to the discretization of the LHS of Eq. (1) and the diffusion matrix to the discretization of the RHS of Eq. (1). Discrete methods like for example: finite volume, finite element and smoothed particle hydrodynamics present such form of semi-discrete equations. Note that Eq. (8) can be written as:

$$\dot{\mathbf{T}} = -\mathbb{H}\mathbf{T} \quad (9)$$

where

$$\mathbb{H} = \mathbb{M}^{-1}\mathbb{K} \quad (10)$$

$\mathbb{H}$  is called the *effective diffusion matrix* of the system. Equations (8) and (9) can be written in index notation as:

$$\sum_{j=1}^n \mathbb{M}_{ij} \dot{T}_j = - \sum_{j=1}^n \mathbb{K}_{ij} T_j, \quad \dot{T}_i = - \sum_{j=1}^n \mathbb{H}_{ij} T_j \quad (11)$$

### Galerkin Finite Element spatial discretizations

In the particular case of FE discretizations, the domain  $\Omega$  is partitioned into  $m$  non-overlapping elemental subdomains  $\Omega^e$ . The matrices  $\mathbb{M}$  and  $\mathbb{K}$  are computed as the following assembly of elemental matrices:

$$\mathbb{M}_{ij} = \sum_e^m \mathbb{M}_{ij}^{(e)}, \quad \mathbb{K}_{ij} = \sum_e^m \mathbb{K}_{ij}^{(e)}, \quad (12)$$

where in the Galerkin approach they are given by

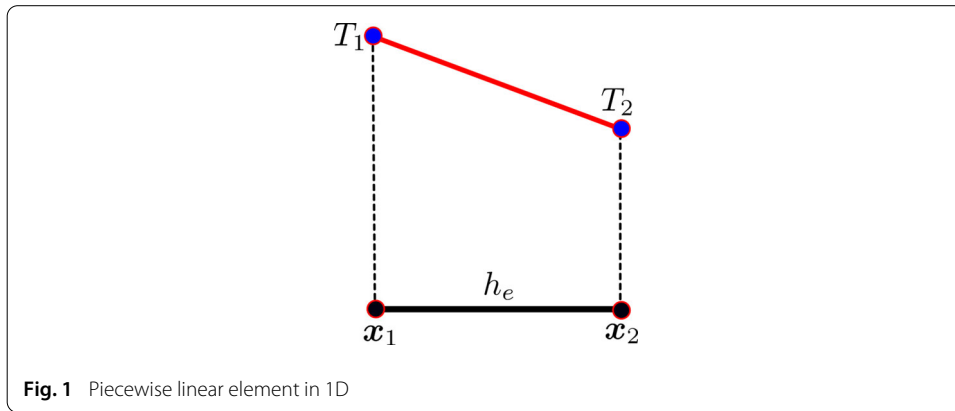
$$\mathbb{M}_{ij}^{(e)} = \rho c_v \int_{\Omega^e} \varphi_i^{(e)} \varphi_j^{(e)} dv, \quad \mathbb{K}_{ij}^{(e)} = \kappa \int_{\Omega^e} \nabla \varphi_i^{(e)} \cdot \nabla \varphi_j^{(e)} dv \quad (13)$$

In the equations above,  $\varphi_i^{(e)}$  denotes the value of the basis function  $\varphi_i(\mathbf{x})$  corresponding to node  $\mathbf{x}_i$  inside the elemental domain  $\Omega^e$ . Matrices  $\mathbb{M}^{(e)} = [\mathbb{M}_{ij}^{(e)}]$  and  $\mathbb{K}^{(e)} = [\mathbb{K}_{ij}^{(e)}]$  are usually called the elemental mass and diffusion matrices, respectively. In the 1D case with piecewise linear elements, as the one shown in Fig. 1, the elemental matrices are given by (see [8]):

$$\mathbb{M}^{(e)} = \frac{\rho c_v h_e}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \mathbb{K}^{(e)} = \frac{\kappa}{h_e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (14)$$

where  $h_e$  is the length of the element. In the 2D case with linear triangular elements, the elemental mass matrix turns out to be independent of the triangle's shape, being only proportional to the triangle's area  $\sigma^{(e)}$ . In this case, the general expression for  $\mathbb{M}^{(e)}$  is ([9, pp. 473]):

$$\mathbb{M}^{(e)} = \rho c_v \frac{\sigma^{(e)}}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (15)$$



On the other hand, following the DMP works [10,11], the diffusion elemental matrices for linear triangular elements are given by the following expression:

$$\mathbb{K}^{(e)} = \frac{\kappa}{2} \begin{bmatrix} \cot(\alpha_2) + \cot(\alpha_3) & -\cot(\alpha_3) & -\cot(\alpha_2) \\ -\cot(\alpha_3) & \cot(\alpha_1) + \cot(\alpha_3) & -\cot(\alpha_1) \\ -\cot(\alpha_2) & -\cot(\alpha_1) & \cot(\alpha_1) + \cot(\alpha_2) \end{bmatrix} \quad (16)$$

where  $\alpha_i$  denote the triangle's inner angle associate to node  $i$ .

It is worth to mention here that some FE approaches modify the consistency of the discretization by replacing  $\mathbb{M}$  by a diagonal lumped matrix  $\tilde{\mathbb{M}}$ . In this case, the effective diffusion matrix is  $\tilde{\mathbb{H}} = \tilde{\mathbb{M}}^{-1} \mathbb{K}$ . Using that the diagonal elements of  $\tilde{\mathbb{M}}$  are given by  $\tilde{\mathbb{M}}_{ii} = m_i = \sum_j \mathbb{M}_{ij}$ , one gets that:

$$\tilde{\mathbb{H}}_{ij} = \frac{1}{m_i} \mathbb{K}_{ij} \quad (17)$$

Note that, in this case, every row of  $\tilde{\mathbb{H}}$  is directly proportional to the corresponding row of elements of  $\mathbb{K}$ .

#### Exact solutions of spatial discretizations of the heat conduction equation

A general exact solution of Eq. (9) exists (for constant mass and diffusion matrices):

$$\mathbf{T}(t) = e^{-\mathbb{H}t} \mathbf{T}(0) \quad (18)$$

#### Time and space discretizations of the heat conduction equation

Usually, Eq. (9) is further discretized in time so as to arrive to fully discrete schemes. Different time-discretization methods can be used for this purpose. For example, if the Euler explicit method were used, the fully discrete scheme will look like:

$$\mathbf{T}^{(n+1)} = \mathbf{T}^{(n)} - \Delta t \mathbb{H} \mathbf{T}^{(n)} \quad (19)$$

where  $\mathbf{T}^{(n)}$  denotes the vector of nodal temperature values at discrete time  $t^n$ . The stability of the above scheme is guaranteed if the time-step  $\Delta t$  is chosen such that:

$$\Delta t \leq \Delta t_{sta} = \frac{2}{\lambda_{\max}} \quad (20)$$

where  $\lambda_{\max}$  is the maximum eigenvalue of matrix  $\mathbb{H}$ . Note that condition (20) is a generalized version of the standard Fourier stability condition:

$$\Delta t \leq \Delta t_{Fourier} = \frac{1}{2} \frac{h^2}{\kappa} \quad (21)$$

which appears when finite-differences are used.

### Discrete thermodynamic compatibility conditions

Thermodynamics is the part of physics that studies the relationships between heat and other energy forms. The first law of thermodynamics expresses conservation of energy. The second law expresses an evolutionary property of all physical processes. Different equivalent statements of the second law exist. Here, the oldest one will be used. It was given around 150 years ago by Rudolf Clausius, it is known as Clausius's postulate, and according to his own words is [12]:

**Clausius's postulate:** "Heat can never pass from a colder to a warmer body (without some other change, connected therewith, occurring at the same time). Everything we know concerning the interchange of heat between to bodies of different temperatures confirm this, for heat everywhere manifests a tendency to equalize existing differences of temperature, and therefore to pass in contrary direction, i.e. from warmer to colder bodies. Without further, explanation, therefore, the truth of the principle will be granted".<sup>1</sup>

Mathematically Clausius's postulate can be expressed as follows.

**Clausius's postulate, bis:** Consider two regions  $\mathcal{R}_i$  and  $\mathcal{R}_j$  with different temperatures  $T_i$  and  $T_j$ , respectively. Assume  $T_j > T_i$ . Let us use the convention that heat  $q$  entering a system is positive. If the two regions above enter in thermal contact then: Heat can only flow from  $\mathcal{R}_j$  towards  $\mathcal{R}_i$ , so:

$$\text{if } T_j > T_i \text{ then } q_{j \rightarrow i} \geq 0 \quad (22)$$

In turn, according to Eq. (3), this positive heat supply  $q_{j \rightarrow i}$  towards  $i$  will necessarily produce an increase of temperature in  $\mathcal{R}_i$ , i.e. a positive contribution to temperature-rates in  $\mathcal{R}_i$ , i.e.:

$$\text{if } T_j > T_i \text{ then } \dot{T}_i|_{\text{contribution}} \geq 0 \quad (23)$$

Of course, the heat equation (1) satisfies Clausius's postulate at any point.<sup>2</sup> However, not necessarily, a consistent discretization of such equation will fully satisfy such postulate. Next some algebraic DTCC will be deduced. These DTCC have to be satisfied by standard discretizations in order to avoid violations of Clausius's postulate, at nodal level:

#### DTCC (Clausius's postulate—DTCC)

To be thermodynamically compatible with Clausius's postulate, a spatial discretization of the heat equation must always generate effective-diffusion matrices  $\mathbb{H}$  with non-positive off-diagonal elements:

$$\mathbb{H}_{ij} \leq 0 \quad \forall j \neq i \quad (24)$$

*Proof* Consider the general form of heat equation's semi-discrete equations derived in "Space-discretizations of the unsteady heat conduction equation" section:

<sup>1</sup> At that time, Clausius used such postulate to derive an expression of a new state variable he called entropy [13]. In terms of this variable, a second equivalent statement of the second law was given:

**Entropy statement:** In any physical process, the change of entropy of an isolated system can only be greater or equal than zero.

<sup>2</sup> This is true because the PDE (1) imposes the satisfaction of the two laws of thermodynamics by construction. In particular, note that the presence of Fourier's law of heat flux ( $\mathbf{q} = -\kappa \nabla T$ ) inside Eq. (1), forces the satisfaction of the second law in the form of Clausius's postulate.

$$\dot{\mathbf{T}} = -\mathbb{H} \mathbf{T} \quad (25)$$

Expanding Eq. (25) in components, one has:

$$\dot{T}_i = - \sum_j \mathbb{H}_{ij} T_j \quad (26)$$

Separating diagonal terms ( $j = i$ ) of the off-diagonal terms ( $j \neq i$ ), Eq. (26) can be rewritten as:

$$\dot{T}_i = -\mathbb{H}_{ii} T_i - \sum_{j \neq i} \mathbb{H}_{ij} T_j = -(\mathbb{H}_{ii} + \sum_{j \neq i} \mathbb{H}_{ij}) T_i - \sum_{j \neq i} \mathbb{H}_{ij} (T_j - T_i) \quad (27)$$

So in general, one has:

$$\dot{T}_i = -\mathbb{h}_i T_i - \sum_{j \neq i} \mathbb{H}_{ij} (T_j - T_i) \quad (28)$$

where the coefficient  $\mathbb{h}_i = \sum_j \mathbb{H}_{ij}$ , equal to the row sum of the effective diffusion matrix, has been defined.

Now, in Eq. (28), let us isolate the  $j$ th-contribution of an arbitrary node  $j \neq i$  to the temperature rate of node  $i$ :

$$\dot{T}_i|_{j \text{ contribution}} = -\mathbb{H}_{ij} (T_j - T_i) \quad (29)$$

The procedure leading to Eq. (29) from Eq. (28) is physically equivalent to completely isolate the small region  $\mathcal{R}_i$  associated to node  $\mathbf{x}_i$  from all other regions except from the small region  $\mathcal{R}_j$  associated to node  $\mathbf{x}_j$ . Since Clausius's postulate should be valid over these two nodal regions in thermal contact. Application of condition Eq. (23) into Eq. (29) implies that:

$$\mathbb{H}_{ij} \leq 0 \quad \forall j \neq i \quad (30)$$

As a consequence, a discrete formulation will be thermodynamically compatible with Clausius's postulate, at nodal level, only if all off-diagonal coefficients of  $\mathbb{H}$  are nonpositive. This completes the proof.

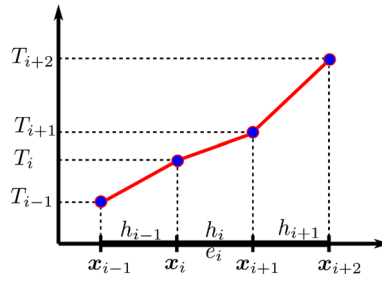
The DTCC given in Eq. (24) define necessary conditions that must be satisfied by the effective diffusion matrix  $\mathbb{H}$  of any spatial discretization in order to preserve thermodynamic compatibility of numerical solutions. If such discrete conditions are not satisfied, the second law of thermodynamics will be violated at nodal level because of the emergence of reversed, non-physical heat flows going from colder temperature regions to warmer temperature regions.

### Thermodynamic incompatibility of Finite Element spatial discretizations

This section is devoted to study the compatibility of FE discretizations with respect to the above presented DTCC. Only the case of linear elements in 1D and 2D will be considered.

#### Thermodynamic incompatibility of 1D Finite Element discretizations

Consider the problem of solving the heat equation in a 1D body using FE discretizations. The body is completely isolated from the exterior [there are not external heat sources and there is not heat flow through its two boundaries ( $\mathbf{x} = 0$  and  $\mathbf{x} = L$ )]. In this case, given any mesh made of  $m$  consecutive non-overlapping segments as the ones shown in Fig. 2, the FE matrices  $\mathbb{M}$  and  $\mathbb{K}$  can be computed using Eqs. (12) and (14).



**Fig. 2** 1D mesh discretization of space

For the sake of easy reading, assume the body has unit length  $L = 1$  and unit physical coefficients  $\kappa, \rho, c_v = 1$ . First, let us consider the elemental case where the mesh is made of only one single element (i.e.  $m = 1$ ) so the single element occupies the total length of the body. In this case one has that:  $\mathbb{M} = \mathbb{M}^{(e)}$  and  $\mathbb{K} = \mathbb{K}^{(e)}$  so the effective diffusion matrix turns out to be:

$$\mathbb{H} = \mathbb{M}^{(e)-1} \mathbb{K}^{(e)} = \frac{6\kappa}{\rho c_v L^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (31)$$

Since all off-diagonal coefficients are non-positive ( $\mathbb{H}_{12} \leq 0$  and  $\mathbb{H}_{21} \leq 0$ ), Eq. (24) is satisfied, so one gets the following result:

**Result 1** Trivial ( $m = 1$ ) 1D FE spatial discretizations based on piecewise-linear elements and consistent mass matrices always satisfy the DTCC.

Likewise, one would expect that the same result holds if the mesh has more elements ( $m > 1$ ). The following simple example will demonstrate that this is not so. Consider the 1D-body of length  $L = 1$  and assume that it is discretized by a mesh of four nodes and 3 segments of equal length  $L/3$ . The nodes are located at positions  $x_1 = 0, x_2 = 1/3L, x_3 = 2/3L, x_4 = L$ . Using Eq. (14), the FE mass and diffusion matrices turn out to be:

$$\mathbb{M} = \begin{bmatrix} 2/18 & 1/18 & 0 & 0 \\ 1/18 & 4/18 & 1/18 & 0 \\ 0 & 1/18 & 4/18 & 1/18 \\ 0 & 0 & 1/18 & 2/18 \end{bmatrix}, \quad \mathbb{K} = \begin{bmatrix} 3 & -3 & 0 & 0 \\ -3 & 6 & -3 & 0 \\ 0 & -3 & 6 & -3 \\ 0 & 0 & -3 & 3 \end{bmatrix} \quad (32)$$

Then, in this case, the effective diffusion matrix turns out to be:

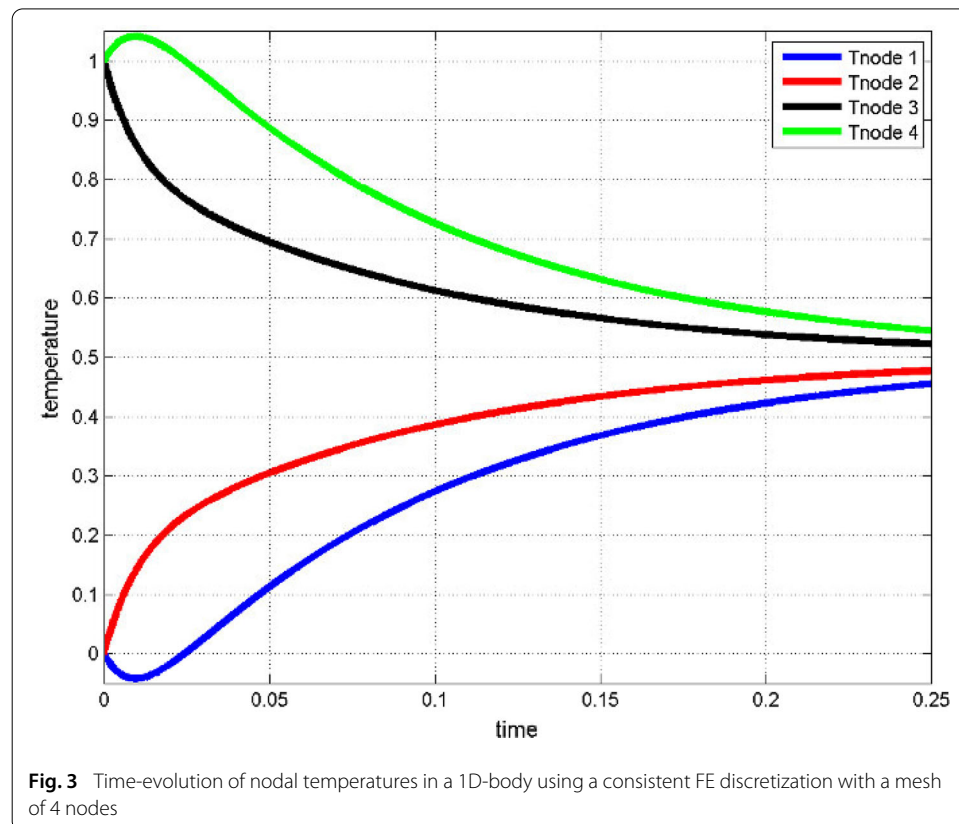
$$\mathbb{H} = \begin{bmatrix} 39.6 & -50.4 & 14.4 & -3.6 \\ -25.2 & 46.8 & -28.8 & 7.2 \\ 7.2 & -28.8 & 46.8 & -25.2 \\ -3.6 & 14.4 & -50.4 & 39.6 \end{bmatrix} \quad (33)$$

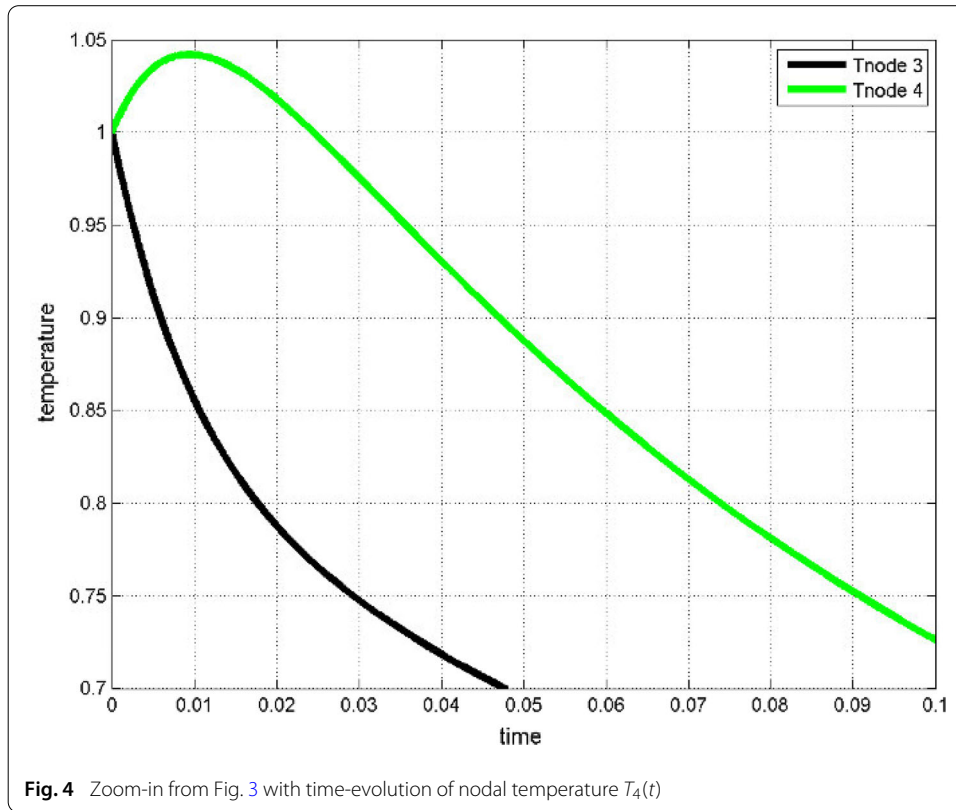
From Eq. (33) one sees that some off-diagonal coefficients of  $\mathbb{H}$  do not satisfy condition (24), having positive values (14.4 and 7.2). So, this FE discretization is nodally thermodynamically incompatible. Actually, it can be shown that, for any non-trivial mesh  $m > 1$ , FE effective diffusion matrices will have some positive off-diagonal elements violating DTCC conditions. Therefore, one has that:

**Result 2** Non-trivial ( $m > 1$ ) 1D FE discretizations based on piecewise-linear elements and consistent mass matrices always violate the DTCC.



The above result implies that 1D FE discretizations are, nodally, thermodynamically incompatible with Clausius's postulate. Then, their use should produce visible non-physical results. This is what actually occurs as shown in the following simple numerical experiment. Consider the isolated 1D-body being used, discretized with the 3-element mesh defined above [the corresponding effective diffusion matrix is given by Eq. (33)]. Assume the body has the following initial nodal temperature distribution:  $T(0) = [T_1(0) = 0.0, T_2(0) = 0.0, T_3(0) = 1.0, T_4(0) = 1.0]^T$ . According to this distribution the initial maximum and minimum temperatures in the body are  $T_{\max} = 1.0$  and  $T_{\min} = 0.0$ , respectively. Once the heat conduction process starts, regions with higher temperatures should immediately start decreasing their temperature because they should be transmitting heat to all surrounding regions which are at lower temperatures. On the contrary, as a consequence of the received heat, lower temperature regions should start increasing their temperature. Of course, these natural evolution should continue until a uniform steady-temperature is reached everywhere in the body. Now, let us determine the nodal temperatures predicted by the FE discretization. For this purpose Eq. (18) is used. The obtained nodal temperature evolution is shown in Fig. 3. Since at location  $x_4$  the body has initially the maximum body-temperature ( $T_4(0) = T_{\max} = 1.0$ ), heat should flow from this region towards colder regions, so body-temperature  $T(x_4, t) = T_4(t)$  at  $x_4$  should decrease continuously from  $T_{\max}$ . However, this is not what the FE solution predicts. On the contrary, as shown by the yellow line in Figs. 3 and 4, FE solution predicts that at initial times ( $t \leq 0.01$  s apprx.), the body-temperature  $T_4(t)$  at such location will increase instead of decrease (going above the maximum temperature  $T_{\max}$ ). Of course, this





is a completely non-physical behavior. A similar but reversed situation occurs at location  $\mathbf{x}_1$ . There, the body has initially the minimum temperature  $T_1(0) = T_{\min} = 0.0$  so heat should flow from warmer regions towards this region, this positive flow of heat should cause a continuous increase of temperature at such location. Instead, the numerical FE solution predicts (blue line of Fig. 3), that at initial times ( $t \leq 0.01$  apprx.), the temperature  $T_1(t)$  in  $\mathbf{x}_1$  will become negative instead of increase its value. Again, this prediction is thermodynamically incompatible. Both observed phenomena at locations  $\mathbf{x}_1$  and  $\mathbf{x}_4$  are caused by the non-physical reversed heat-fluxes produced by the thermodynamic incompatibility of the FE discretization, at nodal level.

#### Thermodynamic incompatibility of 2D Finite Element discretizations

In this section, the thermodynamic compatibility of 2D FE spatial discretizations based on linear triangular elements is investigated. Simple counter-examples are presented to show that, contrary to what one may have expected, such discretizations are nodally incompatible with Clausius's postulate.

For simplicity, consider fully isolated 2D bodies with unit physical coefficients  $\kappa$ ,  $\rho$ ,  $c_v = 1$ . First, let us consider a body formed by a single triangular element  $e$  whose nodes are defined by coordinates:

$$\mathbf{x}^{(e)} = \begin{bmatrix} 0.0 & 0.0 \\ 1.0 & 0.0 \\ 1.5 & 1.0 \end{bmatrix}$$

From Eqs. (15) and (16), it turns out that the elemental matrices are given by:

$$\mathbb{M}^{(e)} = \begin{bmatrix} 1/12 & 1/24 & 1/24 \\ 1/24 & 1/12 & 1/24 \\ 1/24 & 1/24 & 1/12 \end{bmatrix} \quad (34)$$

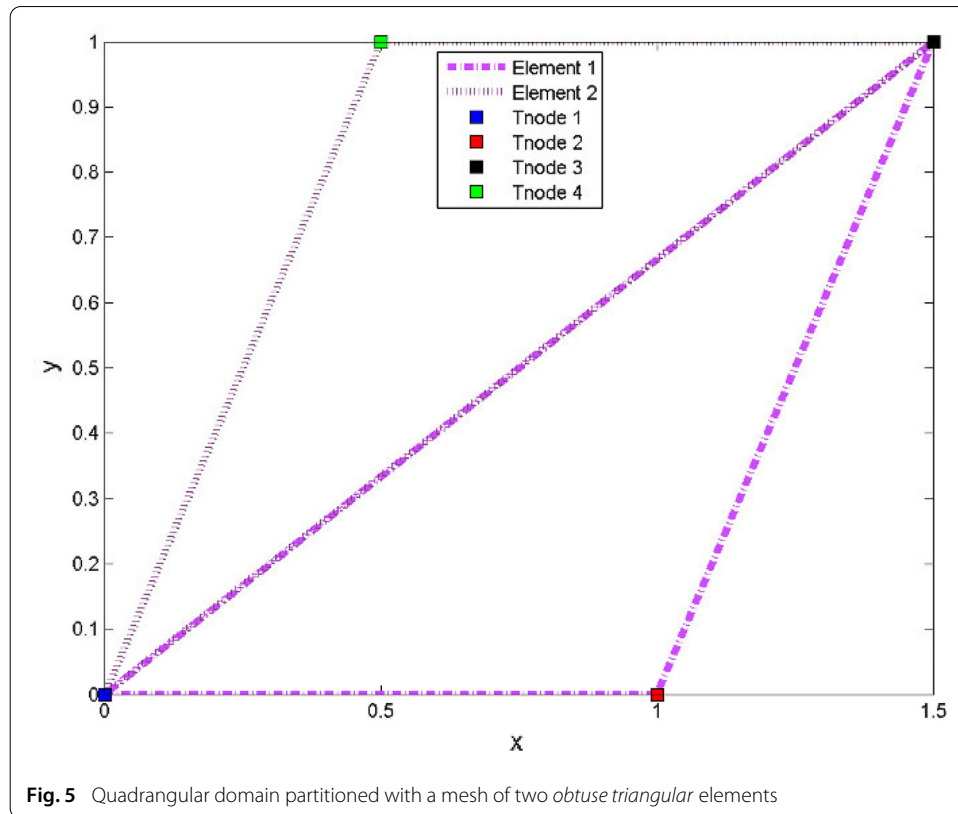
$$\mathbb{K}^{(e)} = \begin{bmatrix} 0.625 & -0.875 & 0.250 \\ -0.875 & 1.625 & -0.750 \\ 0.250 & -0.750 & 0.500 \end{bmatrix} \quad (35)$$

As a consequence, the effective diffusion matrix of this elemental discretization turns out to be:

$$\mathbb{H} = \mathbb{M}^{(e)-1} \mathbb{K}^{(e)} = \begin{bmatrix} 15 & -21 & 6 \\ -21 & 39 & -18 \\ 6 & -18 & 12 \end{bmatrix} \quad (36)$$

Since some off-diagonal elements are positive ( $\mathbb{H}_{13}, \mathbb{H}_{31} = 6 > 0$ ), the above triangular element is nodally incompatible with Clausius's postulate.

Another simple numerical example is presented to show that consistent 2D-FE spatial discretizations are also nodally thermodynamically incompatible for non-trivial meshes ( $m > 1$ ). Consider the problem of heat conduction in a 2D-body with quadrangular shape, as the one shown in Fig. 5. Assume the vertices of the quadrangular-body are defined by the following coordinates:



$$\mathbf{x}^{(e)} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{bmatrix} = \begin{bmatrix} 0.0 & 0.0 \\ 1.0 & 0.0 \\ 1.5 & 1.0 \\ 0.5 & 1.0 \end{bmatrix}$$

Assume the body is discretized in space by a mesh formed by two equal triangular elements  $e_1$  and  $e_2$ , as shown in Fig. 5. The elements are defined by node indices [1, 2, 3] and [3, 4, 1], respectively. Using Eq. (12), the global effective diffusion matrix  $\mathbb{H}$  for this case is:

$$\mathbb{H} = \begin{bmatrix} 13.5 & -10.5 & 6.0 & -9.0 \\ -20.25 & 29.25 & -18.75 & 9.75 \\ 6.0 & -9.0 & 13.5 & -10.5 \\ -18.75 & 9.75 & -20.25 & 29.25 \end{bmatrix} \quad (37)$$

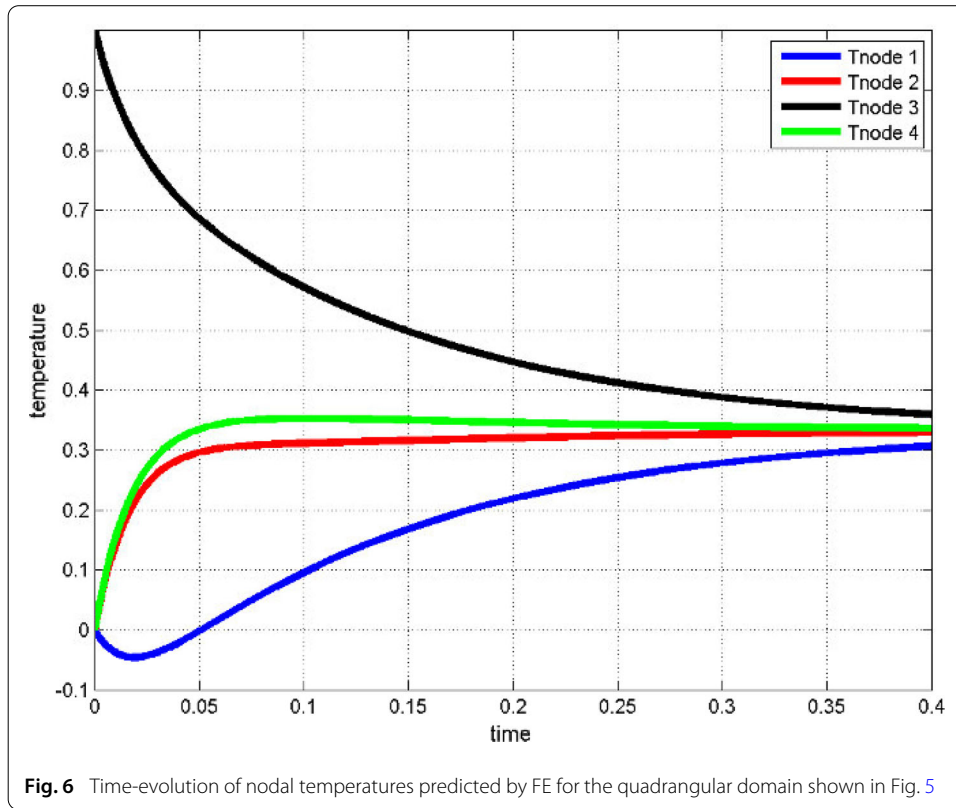
The existence of positive off-diagonal coefficients (6.0 and 9.75) in the effective diffusion matrix  $\mathbb{H}$  imply that condition (24) is violated. Actually, Thomée and Wahlbin [14, pp. 15–16] have proved in the context of the DMP that, for meshes (triangulations) with more than one element, the FE method always produces effective diffusion matrices  $\mathbb{H} = \mathbb{M}^{-1}\mathbb{K}$  that violate condition (24). Then, the following result follows:

**Result 3** 2D FE discretizations with piecewise-linear triangular elements and consistent mass matrices always violate the DTCC.

Such violation of Clausius's postulate is observable in numerical experiments. As an example, consider the 2D FE discretization of the quadrangular-body mentioned above and assume that the following initial temperature distribution is given to the body:  $T(0) = [T_1(0) = 0.0, T_2(0) = 0.0, T_3(0) = 1.0, T_4(0) = 0.0]$ . Note that the initial temperatures given to the body are between  $T_{\min} = 0.0$  and  $T_{\max} = 1.0$ , in particular note that, initially, node 1 has the lowest temperature value. Recall that the body is fully isolated from the exterior. The time-evolution of body temperatures predicted by the FE discretization can be computed using Eq. (18) and it is shown in Fig. 6. There one can see that the FE discrete solution predicts a non-physical evolution of the temperature at node 1 (blue line): such region has the minimum value of temperature, then by Clausius's postulate it should always received heat and increase its temperature. However, this not what it is predicted by the numerical solution: at initial times ( $t: 0.0 \leq t \leq 0.05$ ), it predicts that the nodal temperature will further decrease (as if it were giving heat instead of receiving it). This is a completely non-physical result.

**Remark 2** Although, in the example above, the non-physical effects are clearly seen for  $t: 0.0 \leq t \leq 0.05$  sec, it is important to note that the existence of reversed nodal heat-fluxes due to the incompatible discretization occurs not only at those times but along the whole simulation.

A natural question to be made is whether the issue of nodal thermodynamic incompatibility of consistent FE Discretizations, discussed here for the 1D and 2D cases, also happens in the 3D case. Recent Finite-Element formulas presented in DMP works [15–18]) indicate that the problem is also present in 3D FE discretizations.



### On the issue of generating thermodynamically compatible Finite-Element spatial discretizations

In “Thermodynamic incompatibility of Finite Element spatial discretizations” section, it has been shown that 1D and 2D consistent FE spatial discretizations (with piecewise linear interpolating functions) always generate effective diffusion matrices  $\tilde{\mathbb{H}}$  that violate the DTCC. Then, such discretizations are physically incompatible with Clausius’s postulate, at nodal level. This negative result opens the question if FE discretizations can be modified in order to recover nodal thermodynamic compatibility. A partially positive answer to this question can be given if lumped mass matrices are used instead of using the consistent mass matrix. In this case, as mentioned in “Spatial discretizations of the heat Equation” section, the effective-diffusion matrix of the discretization is  $\tilde{\mathbb{H}} = \tilde{\mathbb{M}}^{-1}\mathbb{K}$ . Since each of row  $\tilde{\mathbb{H}}$  is proportional to the corresponding row of  $\mathbb{K}$  [see Eq. (17)]. The following result (see [14]) is valid:

**Result 4** Any FE effective-diffusion matrix  $\tilde{\mathbb{H}}$  is DTCC if and only if the corresponding  $\mathbb{K}$  is DTCC.

The above result indicates the FE discretizations with lumped mass are thermodynamically compatible if their respective diffusion matrices satisfy the DTCC. Now, due to the relationship (12) between  $\mathbb{K}$  and its conforming elemental diffusion matrices  $\mathbb{K}^{(e)}$ , the following result holds:

**Result 5** Any FE diffusion matrix  $\mathbb{K}$  is DTCC if all of its conforming elemental matrices  $\mathbb{K}^{(e)}$  are DTCC.

From Eq. (14) and Results 4 and 5, one gets that:

**Result 6** 1D FE spatial discretizations with piecewise-linear elements and lumped mass matrices always satisfy the DTCC.

Regrettably, the same result is not valid for lumped 2D FE discretizations. The problem in this case is that not all elemental diffusion matrices  $\mathbb{K}^{(e)}$  are DTCC, so Result 5 can not be used. More specifically, only elemental matrices formed by acute or right angles are DTCC (see Eq. 16). Elements formed by obtuse triangles generate diffusion matrices  $\mathbb{K}^{(e)}$  that violate the DTCC and then produce non-physical reversed nodal fluxes. Note that the satisfaction of the DTCC by elemental diffusion matrices  $\mathbb{K}^{(e)}$  is only a sufficient but not a necessary condition for satisfaction of the DTCC by  $\mathbb{K}$ . In particular, note that a 2D Delaunay triangulation can have some obtuse triangular-elements but still generate a global matrix  $\mathbb{K}$  satisfying DTCC (see [14, 19, p. 78]). Although from a purely algebraic point of view, thermodynamic compatibility of matrix  $\mathbb{K}$  can be satisfied without the strict need that every elemental matrix  $\mathbb{K}^{(e)}$  be thermodynamically compatible, from a physical point of view, it does not seem right to allow having such thermodynamically incompatible elemental contributions. As a summary of the above discussion, the following result can be stated:

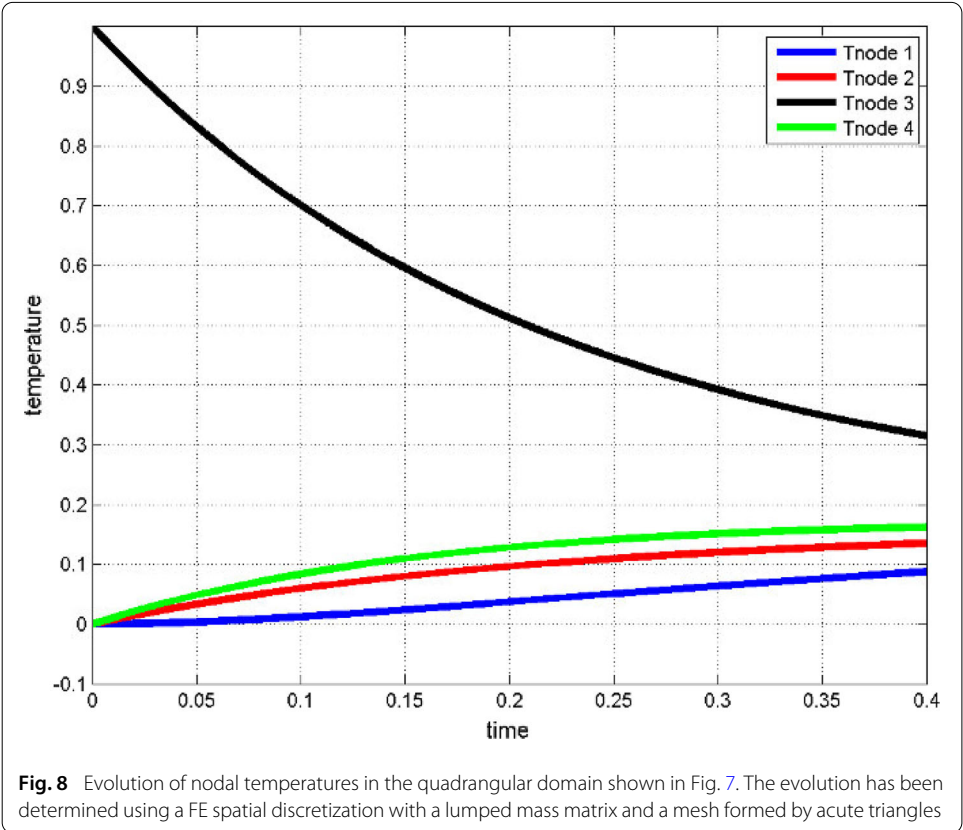
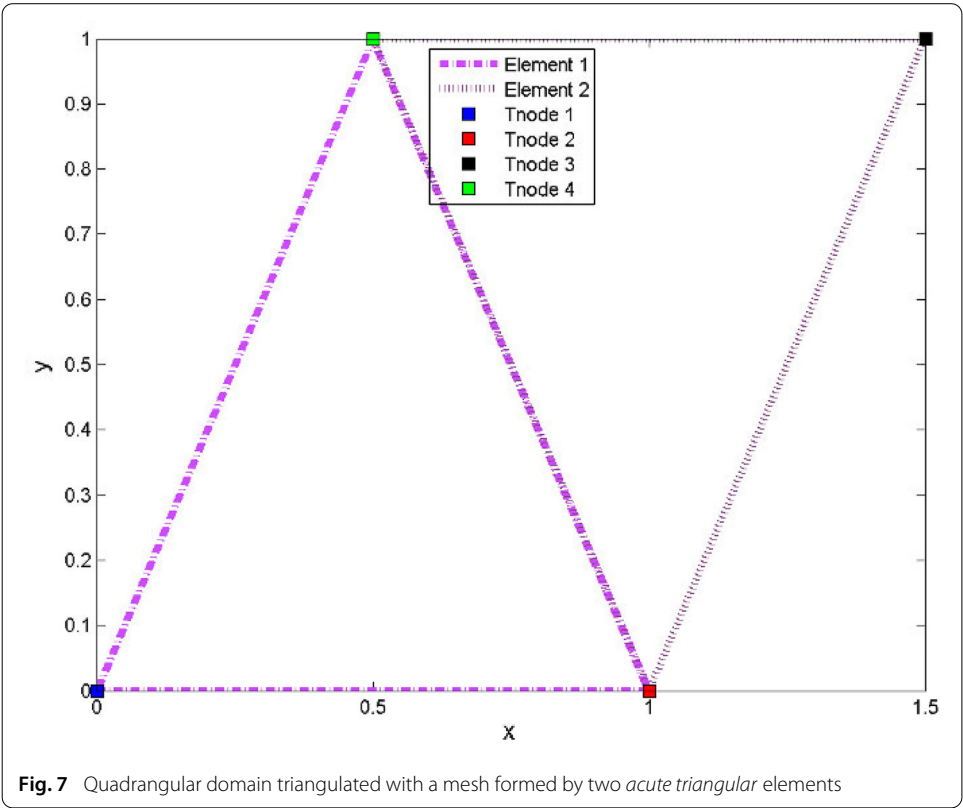
**Result 7** 2D FE spatial discretizations with piecewise-linear triangular elements and lumped mass matrices will be DTCC, both, at elemental and global level, only if the mesh is formed by acute and right triangles. If the mesh is a Delaunay Triangulation, such lumped FE discretization will still be DTCC, globally, but each of their elemental contributions will not.

Result 7 states that arbitrary meshes can not be used if one wants to have 2D FE formulations that are DTCC.

All the above results can be verified in the context of the 1D and 2D examples given in “Thermodynamic incompatibility of 1D Finite Element discretizations” and “Thermodynamic incompatibility of 2D Finite Element discretizations” sections, by using of the lumped mass matrix instead the consistent mass matrix. In particular, the reader can verify that in the case of the quadrangular body, the triangular mesh contains obtuse triangles, so the lumped discretization approach still fails. However, if the mesh is re-triangulated, with two triangular elements formed by the node indices [1, 2, 4] and [2, 3, 4] as shown Fig. 7. Then, in this case, since the mesh is formed only by acute triangles, the effective-diffusion matrix for the lumped FE discretization takes the form:

$$\tilde{\mathbb{H}} = \tilde{\mathbb{M}}^{-1} \mathbb{K} = \begin{bmatrix} 3.75 & -2.25 & 0 & -1.5 \\ -1.125 & 3.375 & -0.75 & -1.5 \\ 0.0 & -1.5 & 3.75 & -2.25 \\ -0.75 & -1.5 & -1.125 & 3.375 \end{bmatrix} \quad (38)$$

which agrees with Result 7. The time-evolution of temperatures in the quadrangular-body predicted by this discretization is shown in Fig. 8. Note that, this time, the nodal temperatures are physically consistent and stay within the real physical limits imposed by the initial values of maximum and minimum temperatures in the body. In particular, node 1 which has the lowest temperature in the body correctly starts to increase its temperature. Compare this result with the non-physical solution shown in Fig. 6.





### Related final comments

In the derivation of the DTCC only spatial discretizations have been considered, time  $t$  has been left as a continuous variable. From this, it follows that the results discussed in this paper will affect any numerical scheme having any type of time-discretization. This is so, because by means of numerical consistency any time-discretization will tend to the time-continuous solution, as  $\Delta t \rightarrow 0$ . As a consequence, for example not only the Euler explicit method (presented in “Exact solutions of spatial discretizations of the heat conduction equation” section but all time-integration numerical schemes are affected by the presented results. Even implicit schemes are affected.

Note that the violation of the DTCC does not imply that the discrete solution will not converge to the exact solution as the mesh is refined. However, one might expect that thermodynamically compatible algorithms should be more stable or converge faster than the ones that are not. As mentioned in Eq. (20), the stability condition depends directly on the maximum eigenvalue  $\lambda_{\max}$  of the discretization matrix  $\mathbb{H}$ . So stability changes will depend on how the satisfaction or the non-satisfaction of the DTCC will alter the value of  $\lambda_{\max}$ . The issue of improving convergence if thermodynamically compatible formulations are used, could be related to the work of Ciarlet and Raviart [20] who proved uniform convergence of FE solutions. DTCC and non DTCC solutions could also be compared using error estimation procedures [21,22]

Similarly, it must be pointed that a scheme that violates the DTCC is not necessarily unstable. In particular, note that all non-DTCC examples presented in “Thermodynamic incompatibility of 1D Finite Element discretizations” and “Thermodynamic incompatibility of 2D Finite Element discretizations” sections are stable and can be solved without problems. Note also that those cases could also be solved using fully-discrete time-integration algorithms, like for example, Euler’s method (19) and they will be stable as long as the used time-steps  $\Delta t$  are below the stability limit of the time-integration algorithm. For the case of Euler method this stability limit is given by Eq. (20).

Note that numerical solutions obtained in “Thermodynamic incompatibility of 1D Finite Element discretizations” and “Thermodynamic incompatibility of 2D Finite Element discretizations” sections prove that the use of thermodynamically incompatible formulations can produce non-physical solutions which are not only stable but also smooth and non-oscillatory. Then, the decision to use a thermodynamically incompatible discretization anyway, can be very dangerous because not necessarily the obtained solutions will show signs that something is evidently wrong.

Another well-known source of spurious spatial numerical oscillations are the ones produced in thermal shocks. It is known that consistent 1D FE formulations can generate non-physical thermal shock oscillations if [23]:

$$\Delta t \leq \Delta t_{\text{shock}} \quad (39)$$

This means that, for any given mesh size, as  $\Delta t \rightarrow 0$  consistent FE formulations will generate non-physical oscillations in thermal shocks. One open question is if these non-physical oscillations can be fully attributed to the nodal thermodynamical incompatibility of consistent FE discretizations (see Result 2). The fact that condition (39) disappears if lumped mass (thermodynamically consistent) FE discretizations are used ([23,24]), supports this conjecture.



Another point to be analyzed in the future is if popular stabilization techniques [25,26] used to avoid the above mentioned spurious numerical oscillations [24] are somehow connected to the DTCCs presented in this article. One may intuitively think that, by means of these techniques, diffusion is added to compensate for unphysical heat flow. However, to the authors is uncertain if stabilization techniques are smart enough to detect (and correct) the thermodynamic incompatibility of any FE discretization. Of course these topics need further research.

The matrix property of non-positive off-diagonal elements [such as condition (24)] emerging here in the context of the satisfaction of the second law of thermodynamics seems to appear in different types of previously studied matrices, such as Metzler Matrices in dynamical systems [27], non-negative matrices, Z-matrices and M-matrices in economics, discretization methods of PDEs, Markov processes, stability, etc. (see [28], [29]). In particular, it can be shown that  $-\mathbb{H}$  must be a Metzler matrix and  $\mathbb{H}$  must be a Z-matrix. Also it can be shown that  $\mathbb{B} = h_{mm}\mathbb{I} - \mathbb{H}$  must be a nonnegative matrix (where  $h_{mm} = \max(\mathbb{H}_{ii})$  and  $\mathbb{I}$  is the identity matrix).

Finally note that, the thermodynamical incompatibility phenomenon discussed here applies not only to heat transfer problems but to any type of diffusive process. For example, it is completely equivalent to phenomena of negative viscosity associated with the diffusion equation for momentum

$$\frac{\partial \mathbf{v}}{\partial t} = \nu \nabla^2 \mathbf{v} \quad (40)$$

involving kinematic viscosity  $\nu = \rho/\mu$ . Such phenomena correspond to diffusion of momentum from low velocity to high velocity regions of the flow and could be shown to also constitute a violation of the second law of thermodynamics. Finally, another parallel problem (that could also be called negative diffusion) can occur with the diffusion equation for partial densities

$$\frac{\partial \rho_i}{\partial t} = D_i \nabla^2 \rho_i \quad (41)$$

involving mass diffusivity  $D_i$ . Now, counter-gradient diffusion of mass such as a sweet cup of coffee becoming bitter spontaneously with time would also indicate the same type of violation of the second law of thermodynamics.

## Conclusions

This paper describes the issue that consistent Finite Element spatial discretizations with linear elements lead to numerical schemes that are nodally thermodynamically incompatible. These incompatible schemes violate Clausius's postulate of the second law of thermodynamics at nodal level producing reversed non-physical heat fluxes. Discrete mathematical conditions for nodal thermodynamic compatibility are presented. These conditions have appeared before in the context of the satisfaction of the discrete maximum principle. However, the purpose of this article is to emphasize that the root of such mathematical problem may lie in the violation of such fundamental physical principle. Simple numerical examples in 1D and 2D are used to demonstrate that the use of thermodynamically incompatible discretizations can produce non-physical numerical solutions.

**Requirements to restore nodal thermodynamic compatibility of FE formulations are presented.**

#### Author details

<sup>1</sup>CIMEC-Research Center of Computational Methods (UNL/CONICET), Ruta 168 s/n, Predio Conicet "Dr A. Cassano", 3000 Santa Fe, Argentina, <sup>2</sup>International Center for Numerical Methods in Engineering (CIMNE), Institutió Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

Received: 12 February 2016 Accepted: 1 April 2016

Published online: 22 April 2016

#### References

1. Limache A, Sanchez P, Dalcin L, Idelsohn S. Objectivity tests for navier-stokes simulations: the revealing of non-physical solutions produced by laplace formulations. *Comput Methods Appl Mech Eng.* 2008;197:1703–59.
2. Bathe K-J. Conserving energy and momentum in nonlinear dynamics: a simple implicit time integration scheme. *Comput Struct.* 2007;85:437–45.
3. Varga R. On a discrete maximum principle. *J SIAM Numer Anal.* 1966;3(2):355–9.
4. Fujii H. Some remarks on finite element analysis of time-dependent field problems. In: Yamada Y, Gallagher RH, editors. *In theory and practice in finite element analysis*. Tokyo: University of Tokyo Press; 1973. p. 91–106.
5. Brandt A. Generalized local maximum principles for finite-difference operators. *Math Comput.* 1973;27(124):685–718.
6. Rank E, Katz C, Werner H. On the importance of the discrete maximum principle in transient analysis using finite element methods. *Int J Numer Methods Eng.* 1983;19:1771–82.
7. Bertolazzi E, Manzini G. A second-order maximum principle preserving finite volume method for steady convection-diffusion problems. *SIAM J Numer Anal.* 2005;43(5):2172–99.
8. Larson M, Bengzon F. *The finite element method: theory, implementation and applications*. Texts in computational science and engineering, vol 10. Berlin: Springer; 2013.
9. Zienkiewicz O, Taylor R. *The finite element method*. 5th ed. Oxford: Butterworth-Heinemann; 2000.
10. Draganescu A, Dupont TF, Scott LR. Generalized local maximum principles for finite-difference operators. *Math Comp.* 2004;27(124):685–718.
11. Faragó I, Horváth R, Korotov S. Discrete maximum principle for linear parabolic problems solved on hybrid meshes. *Appl Numer Math.* 2005;53:249–64.
12. Clausius R. On a modified form of the second fundamental theorem in the mechanical theory of heat. *Phil Mag J Sci.* 1856;12(77):81–98.
13. Clausius R. *The mechanical theory of heat*. London: McMillan and Co.; 1879.
14. Thomée V, Lars B, Wahlbin L. On the existence of maximum principles in parabolic finite element equations. *Math Comp.* 2008;77(261):11–9.
15. Korotov S, Neittaanmaki P. Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle. *Math Comp.* 1998;70:107–19.
16. Vejchodský T, Korotov S, Hannukainen A. Discrete maximum principle for parabolic problems solved by prismatic finite elements. *Math Comput Simul.* 2010;80:1758–70.
17. Kosik R, Fleischmann P, Haindl B, Pietra P, Selberherr S. On the interplay between meshing and discretization in three-dimensional diffusion simulation. *IEEE Trans Comput-Aided Design Integr Circuits Syst.* 2000;19(11):1233–40.
18. Cordes C, Putti M. Accuracy of Galerkin finite elements for groundwater flow simulations in two and three-dimensional triangulations. *Int J Numer Meth Eng.* 2001;52:371–87.
19. Strang G, Fix G. *An analysis of the finite element method*. Upper Saddle River: Prentice-Hall; 1973.
20. Ciarlet P, Raviart P. Maximum principle and uniform convergence for the finite element method. *Comput Methods Appl Mech Eng.* 1973;2:17–31.
21. Ladeveze P, Rougeot P, Blanchard P, Moreau J. Local error estimators for finite element linear analysis. *Comput Methods Appl Mech Eng.* 1999;176(1):231–46.
22. Ladeveze P, Chouaki A. Application of a posteriori error estimation for structural model updating. *Inverse Prob.* 1999;15(1):49.
23. Bergheau J-M, Fortunier R. *Finite element simulation of heat transfer*. Hoboken: Wiley; 2008.
24. Wood WL, Lewis RW. A comparison of time marching schemes for the transient heat conduction equation. *Int J Numer Methods Eng.* 1975;9:679–89.
25. Idelsohn SR, Heinrich JC, Oñate E. Petrov-galerkin methods for the transient advective—diffusive equations with sharp gradients. *Int J Numer Methods Eng.* 1996;39:1455–73.
26. Idelsohn S, Nigro N, Storti M, Buscaglia G. A petrov-galerkin formulation for advection-reaction-diffusion problems. *Comput Methods Appl Mech Eng.* 1996;136:27–46.
27. Kaczorek T. *Positive 1D and 2D systems*. Berlin: Springer-Verlag; 2002.
28. Berman A, Plemmons RJ. *Non-negative matrices in the mathematical sciences*, SIAM; 1994.
29. Minc H. *Non-negative matrices*. Hoboken: Wiley; 1988.