Indicator of space error in residue for transitory thermics

Summary

During digital simulations by finite elements, obtaining a gross profit is not sufficient any more. The user is increasingly petitioning of **space error analysis compared to its grid**. **It needs for support methodological and tools “numériquo” - “data-processing” pointed to measure the quality of its studies and to improve them.**
To this end, the indicators of space error a posteriori make it possible to locate, on each element, a cartography of error on which the tools of mending of meshes will be able to rest: the first calculation on a coarse grid makes it possible to exhume the map of error starting from the data and of the solution discretized (from where the term “a posteriori”), refinement is carried out then locally by treating on a hierarchical basis this information.
new indicator a posteriori (known as “in pure residue”) who has just been established **for post” - “to treat the thermal solveurs of *Code_Aster*** is based on their local residues extracted semi” - “discretizations in time. Via the option ‘ERTH_ELEM’ of CALC_ERREUR, it uses the thermal fields (EVOL_THER) emanating from THER_LINEAIRE and of THER_NON_LINE. **It thus supplements the offer of the code in term of advanced tools making it possible to improve quality of the studies, their mutualisations and their comparisons**.
The goal of this note is to detail theoretical, digital work and data processing which governed its establishment. With regard to the theoretical study we, initially, limited ourselves to the linear thermics of a motionless structure discretized by the finite elements isoparametric standards. But, in practice**, the perimeter of use of this option was partially extended to nonlinear thermics**.
One gives to the reader the properties and the theoretical and practical limitations of the exhumed indicator, while connecting these considerations, which can sometimes appear a little “ethereal”, to a precise parameter setting of the accused operators and to the choices of modeling of the code. One tried constantly to bind different the items approached, while detailing, has minimum, of a little technical demonstrations seldom clarified in the specialized literature.

Contents

1 Problems – Description of the document

During digital simulations by finite elements obtaining a gross profit is not sufficient any more. **The user is increasingly petitioning of space error analysis compared to his grid**. **It needs for support methodological and pointed digital tools to measure the quality of its studies and to improve them.**
For example, the precision of the results is often degraded by local singularities (corners, heterogeneities…). One then seeks the good strategy to identify these critical areas and them to refine/déraffiner in order to optimize the compromise site/total error. And this, with the highhest possible degree of accuracy, in an automatic, reliable way (the error analysis must be itself the least approximate possible!) robust and with the lower costs.

For each type of finite elements, one in general has**estimates a priori of the space error** [bib1], [bib3]. But those are checked only asymptotically (when the size h elements tends towards zero) and they require a certain level of regularity which is precisely not reached in the zones with problem. Moreover, these increases underlie two types of strategies to improve calculation:

- "methods" -" p " which consists in locally increasing the order of the finite elements,
- "methods" -" h " which locally refines in order to decrease the geometrical characteristics of the elements.

We are interested here in the second strategy, but through another class of indicators: **indicators of errors a posteriori** . Since work founders of I. BABUSKA and W. RHEINBOLDT [bib18], the importance of this kind of indicator is well established and they arouse a growing interest, as well in pure digital analysis [bib5], [bib6], [bib7] as in the field of the applications [bib4], [bib16]. They were in particular established and used in N3S, TRIFOU and it *Code_Aster* (for linear mechanics cf [R4.10.01], [R4.10.02]). For a panorama of the existing indicators, one will be able to consult the reference book of R. VERFURTH [bib7] or, the report of X. DESROCHES [bib16].

To take again a sales leaflet of M.FORTIN (cf [bib17] pp468" - "469), the development of the estimate a posteriori is justified mainly by three reasons:

- the first is the need for establishing the precision of the results got by a calculation finite elements: which credit to grant to them? Are all the phenomena and all the data which intervene" - "they well taken into account in modeling?
- the second objective is to make it possible whoever to use a computer code without having to intervene in the construction of the grid in order to obtain the necessary total precision,
- finally, the third direction of study is more particularly directed towards the three-dimensional problems for which the size of the grids is limited by the place memory available and the cost of the resolution.

These specifications reveal two duaux problems : to estimate the precision of the solution obtained on the principal parameters of simulation and to propose means of calculating a new solution which respects a minimal precision. The first problem is truly that of the estimate of error whereas the second relates to the associated adaptive methods (refinement/déraffinement, mending of meshes, displacement of points, follow-up of border…).
Thus these indicators make it possible to locate on each element a cartography of error on which the tools of mending of meshes will be able to rest.

Note:

> *One prefers to him the denomination of “indicator” to the usual terminology of “estimator” (translation literal of English “the error estimator “). Account” - “held owing to the fact that it has the same theoretical limitations that those of the solvor finite elements (that it “post - milked “), that he is him even often sullied with digital approximations and that he is exhumed via relations of equivalence utilizing many constants dependent on the problem… the information which it under” - “tightens truly gives “only one order of magnitude” of the required space error. In spite of these restrictions, these cartographies of error a posteriori do not remain less important about it, and in any case, they constitute the only type of accessible information in this field.*

The first calculation on a coarse grid makes it possible to associate, with each element of the triangulation, an indicator calculated starting from the discretized data and first discrete solution. Refinement is carried out then locally by treating on a hierarchical basis this information.

In short, and in a nonexhaustive way, **the use of an indicator possibly coupled with a remaillor:**

- provides a certain estimate of the error of space discretization,
- get a better frequency of errors due to the local singularities,
- allows to improve modeling of the facts of the case (materials, loadings, sources…),
- allows to optimize (even precision with the lower costs) and to make reliable the process of convergence of the grid,
- to estimate and qualify a calculation for a class of grid given.

These considerations show clearly that the calculation of these estimators (which is finally only a post” - “treatment of the problem considered) must:

- to be much less expensive than that of the solution,
- to require only the discretized data and the calculated solution,
- to be able to be localised,
- to be equivalent (in a particular form) to the exact error.

We will see, that with the indicators in residue, one can obtain only one total increase of the exact error united with a local decrease of this same error. But these hight delimiters and lower of the error are complementary because, the first ensures us to have obtained a solution with a certain tolerance, while the second enables us to locally optimize the number of points to respect this precision and not to over-estimate it. They utilize constants which not depending on the discretizations space and temporal.

The goal of this note is to detail theoretical, digital work and data processing which governed the indicator installation of error a posteriori making it possible “post-to treat” the thermal solveurs of *Code_Aster* . It is about an indicator in pure residue initiated by the option ‘ `ERTH_ELEM` ’ of `CALC_ERREUR` .
With regard to the theoretical study we, initially, limited ourselves to the linear thermics of a motionless structure discretized by the finite elements isoparametric standards. But, in practice, the perimeter of use of this option was partially extended to nonlinear thermics. For more details on the perimeter of use and functional of the thermal indicator and an example of use, one will be able to refer to [§6].

The indicator a posteriori that we propose is an indicator in pure residue based on the local residues of the semi strong equation" - "discretized in time. For certain elements of the theoretical study (and in particular its groundwork) we took as a starting point the innovative work by C. BERNARDI and B. METIVET [bib6]. They have extended they-even, of elliptic with parabolic, the results of R. VERFURTH [bib7]. They in particular were interested in calculations of indicators on the model case of the equation of heat with homogeneous condition of Dirichlet, semi-discretized in time by a diagram of implicit Euler. We extended these results to with the problems really dealt by the linear operator of thermics of the code, `THER_LINEAIRE`. They are mixed problems in extreme cases (Cauchy-Dirichlet-Neumann-Robin) inhomogenous, linear, with coefficients variable and discretized by one θ- method.

A basic work was thus undertaken for encircling the theoretical springs of the subjacent thermal problem well and to extrapolate the results of the problem models preceding. This in order to try to approach to modelings and the perimeter of the code while detailing often induced mathematical subtleties in the articles of Article a special effort was brought to put in prospect the choices led in *Code_Aster* compared to research, last and current, like clarifying the general philosophy of these indicators.
One gives to the reader the properties and the theoretical and practical limitations of the indicators released while connecting these considerations to a precise parameter setting of the operator `CALC_ERREUR` accused in this postprocessing. One tried constantly to bind different the items approached, to limit the recourse to long mathematical digressions, while detailing has minimum many "technical" demonstrations seldom clarified a little in the specialized literature.

This document is articulated around the following parts:

- In **the first time, a theoretical study is led** in order to underline holding them and outcomes of the subjacent thermal problem, and, their possible links with the choices of modeling of the code. First of all, one determines the Abstracted Variational Framework (CVA) minimum (cf [§2.1]) on which one will be able to rest to show the existence and the unicity of a field of temperature solution (cf [§2.2]). By recutting these pre-necessary theoretical with the practical constraints of the users, one from of deduced from the limitations as for the types of geometry and the licit loadings. Then one studies the evolution of the properties of stability of the problem (cf [§3]) during the process of semi-discretization in time and space.
- These results of controllability are very useful to create the standards, the techniques and the inequalities which intervene in the genesis of the indicator in residue. After having introduced the usual notations of this kind of problems (cf [§4.1]), **one exhumes a possible formulation of the indicator** as well as the increase of the total error (cf [§4.2]) and the decrease of the site error associated (cf [§4.4]). Various types of space indicators (cf [§4.3]) are evoked and one details several used strategies of construction of indicators into parabolic (cf [§4.5]). In this same paragraph, the temporal aspect of the problem is also examined through the contingencies of management of the space error with respect to that of the step of time.
- In a third part (cf [§5]), the principal contributions of these theoretical chapters and their links with the thermal solveurs of the code are summarized.
- Finally, one concludes by approaching the practical difficulties from implementation (cf [§6.1]), the environment necessary (cf [§6.2]), the parameter setting (cf [§6.3]) and it **perimeter of use (cf [§6.4]) of the indicator actually established in the operator of postprocessing `CALC_ERREUR`**. An example of use extracted from a case official test (`TPLL01J`) is also detailed (cf [§6.5]).

Warning:

> *The reader in a hurry and/or not very interested by the theoretical springs genesis indicator error and subjacent thermal problem can, from the start, to jump to [§5] which recapitulates the principal theoretical contributions of the preceding chapters.*

2 The problem in extreme cases

2.1 Context

One is considered **limited open motionless body occupying related** Ω of R^q (q =2 or 3) of **border** $\partial\Omega:=\Gamma:=\bigcup_{i=1}^{3}\Gamma_i$ **regular** characterized by its voluminal heat with constant pressure $\rho C_p(\mathbf{x})$ (the vectorial variable $\mathbf{x}$ symbolize the couple here (x,y) (resp. (x,y,z)) for q =2 (resp. q =3))*)* and its coefficient of isotropic thermal conductivity $\lambda(\mathbf{x})$.

Note:

> *One will thus not take account of a possible displacement of the structure (cf.* `THER_NON_LINE_MO` *[R5.02.04]).*

These data materials are supposed to be independent of time (modeling `THER` of *Code_Aster*) and constants by element (discretization P_0).

Note:

> *With modeling* `THER_FO` *these characteristics can depend on time. As of the first versions of the code and before the installation of* `THER_NON_LINE` *, it made it possible to simulate "pseudonym" non-linearities. Taking into account its rather marginal use, we will not be interested in this modeling.*

One is interested in the changes of the temperature in any point $\mathbf{x}$ opened and at any moment $t\in[0,\tau[\ (\tau>0)$, when the body is subjected to limiting conditions and loadings independent of the temperature but being able to depend on time. It is about voluminal source $s(\mathbf{x},t)$, boundary conditions of type imposed temperature $f(\mathbf{x},t)$ (on the external portion of surface Γ_1), imposed normal flow $g(\mathbf{x},t)$ (on Γ_2) and exchanges convectif $h(\mathbf{x},t)$ and $T_{ext}(\mathbf{x},t)$ (on Γ_3).
One places oneself thus within the framework of application of the operator `THER_LINEAIRE` [R5.02.01] of *Code_Aster* by retaining only the conductive aspects of this linear thermal problem.

Note:

> *Non-linearities pose serious theoretical problems [bib2] to show the existence, the unicity and the stability of the possible solution. Some are still completely open… But in practice, that by no means prevents from "stretching" the perimeter of use of the estimator of error which will be exhumed rigorously for linear thermics, to nonlinear thermics (operator* `THER_NON_LINE` *[R5.02.02]).*

This problem in extreme cases **mixed** (of type **Cauchy-Dirichlet-Neumann-Robin** (also called condition of Fourier) inhomogenous, linear and with variable coefficients) is formulated

$$(P_0)\begin{cases}\rho\, C_p\dfrac{\partial T}{\partial t}-\mathrm{div}(\lambda\,\nabla T)=s & \Omega\times]0,\tau[\\ T=f & \Gamma_1\times]0,\tau[\\ \lambda\dfrac{\partial T}{\partial n}=g & \Gamma_2\times]0,\tau[\\ \lambda\dfrac{\partial T}{\partial n}+hT=hT_{ext} & \Gamma_3\times]0,\tau[\\ T(\mathbf{x},0)=T_0(\mathbf{x}) & \Omega\end{cases}$$ éq 2.1-1

Note:

- *In this theoretical study of the mixed problem (P_0) , it is supposed that the border dissociates in portions on which acts a condition inevitably limits nonhomogeneous. This assumption is not in fact not paramount and one can suppose the existence of a portion Γ_4 , such as $\Gamma_4 := \Gamma - \bigcup_{i=1}^{3} \Gamma_i \neq \emptyset$, on which intervenes a homogeneous condition of Neumann (thus, when one builds the variational formulation associated with the strong formulation (P_0) , the terms of edges related to this zone disappear. The problem remains well posed then since it is thermically unconstrained in this zone. By means of computer, it is besides well what occurs, since the terms of edges are initialized to zero). In practice, it is often the case besides.*
- *It will be supposed that the coefficient of exchange $h(t, \mathbf{x})$ is positive what is the case in Code_Aster (cf [U4.44.02 §4.7.3]). And that will facilitate a little the things to us in the demonstrations to come (cf for example property 5).*
- *The condition of Robin modelling the convectif exchange (keyword `EXCHANGE`) on a portion of edge of the field, can duplicate itself to take account of exchanges between two under-parts of the border in opposite (keyword `ECHANGE_PAROI`). This limiting condition models a thermal resistance of interface*

$$\text{Avec } \Gamma_3 = \Gamma_{12} \cup \Gamma_{21} \text{ , } T_i = T_{|\Gamma_{ij}} \text{ on a } \begin{cases} \lambda \dfrac{\partial T_1}{\partial n} + hT_1 = hT_2 & \Gamma_{12} \times]0, \tau[\\ \lambda \dfrac{\partial T_2}{\partial n} + hT_2 = hT_1 & \Gamma_{21} \times]0, \tau[\end{cases} \quad \text{éq 2.1-2}$$

Not to weigh down the writing of the problem and insofar as this option is similar to the condition of Robin with the external medium, we will not specifically mention it in calculations which will follow.

- *The condition of Dirichlet can spread in the form of linear relations between the degrees of freedom (keyword `LIAISON_*`) to simulate, in particular, of geometrical symmetries of the structure.*

$$\text{Avec } \Gamma_1 = \Gamma_{12} \cup \Gamma_{21} \text{ , } T_i = T_{|\Gamma_{ij}} \text{ on a} (\text{LIAISON\_GROUP})$$

$$\sum_i \beta_{1i} T_1^i(\mathbf{x}, t) + \sum_j \beta_{2j} T_2^j(\mathbf{x}, t) = \gamma(\mathbf{x}, t) \text{ sur } \Gamma_1 \times]0, \tau[$$

$$\text{ou plus simplement } \sum_i \beta_i T_i(\mathbf{x}, t) = \gamma(\mathbf{x}, t) \text{ sur } \Gamma_1 \times]0, \tau[\ (\text{LIAISON\_DDL}) \quad \text{éq 2.1-3}$$

In the same way features `LIAISON_UNIF` and `LIAISON_CHAMNO` allow to impose the same temperature (unknown) on a set of nodes. They constitute a surcouche of the preceding conditions by imposing couples (β, γ) individuals. Not to weigh down the writing of the problem and insofar as these options are only typical cases of the generic condition of Dirichlet, we will not specifically mention them in calculations which will follow.

- *When the material is **anisotropic** conductivity is modelled by a diagonal matrix expressed in the reference mark of orthotropism of material. That basically does not change following calculations which take account only isotropic case. It is just necessary to take care not to commutate more, under the limiting conditions of Neumann and of Robin, the scalar product with the normal and multiplication by conductivity.*
- *For one **transitory calculation** , the initial temperature can be selected in three different ways: by carrying out a stationary calculation over the first moment, by fixing it at a uniform or unspecified value created by one `CREA_CHAMP` and by carrying out a recovery starting from a preceding*

transitory calculation. This choice of the condition of Cauchy does not have any incidence on the theoretical study which will follow.

- *We will not treat the case where (almost) all the loadings are multiplied by the same function dependent on time (option `FONC_MULT` , this functionality adapted well for certain mechanical problems is disadvised in thermics, because it can return in conflict with the temporal dependence of the loadings and, in addition, it applies selectively with each one of them. It was not taken again besides in `THER_NON_LINE`).*

It is shown that the functional framework more the most convenient general and for "the catch in hand" of this parabolic problem is the following.

For the geometry:

Ω opened locally limited the only one with dimensions one of its border, (H1)

Γ variety of dimension $q-1$, lipschitzienne or C^1 by piece (H2)

For the data:

$$s\in L^2\left(0,\tau\,;H^{-1}(\Omega)\right)\quad T_0\in L^2(\Omega)$$
$$f\in L^2\left(0,\tau\,;H^{1/2}(\Gamma_1)\right),\; g\in L^2\left(0,\tau\,;H^{-1/2}(\Gamma_2)\right),\;\; T_{ext}\in L^2\left(0,\tau\,;H^{-1/2}(\Gamma_3)\right) \qquad \text{(H3)}$$
$$\rho\;,\;C_p\;,\;\lambda\in L^{\infty}(\Omega)\quad h\in L^2\left(0,\tau\,;L^{\infty}(\Gamma_3)\right)$$

who allows us to obtain a solution in the following intersection

$$T\in L^2\left(0,\tau\,;H^1(\Omega)\right)\cap C^0\left(0,\tau\,;L^2(\Omega)\right) \qquad \text{éq 2.1-4}$$

Note:

That is to say $(X,\|\|_X)$ *Banach, one notes* $L^p(0,\tau\,;X)$ *the space of the functions* $t\rightarrow v(t)$ *strongly measurable for measurement* dt *such as* $\|v\|_{0,\tau\,;\,p,\,X}=\left(\int_0^{\tau}\|v(t)\|_X^p\,dt\right)^{\frac{1}{p}}<+\infty$ *. It is Banach, therefore a space of Hilbert for the associated standard.*

The introduction as of these spaces of Hilbert particular "space times" comes from **need for separating the variables** $\mathbf{x}$ and t. Any function $u:(\mathbf{x},t)\in Q_{\tau}:=\Omega\times]0,\tau[\rightarrow u(\mathbf{x},t)\in\Re$ can in fact of being identified (by using the theorem of Fubini) with another function $\tilde{u}:t\in]0,\tau[\rightarrow\{\tilde{u}(t):\mathbf{x}\in\Omega\rightarrow\tilde{u}(t)(\mathbf{x})=u(\mathbf{x},t)\}$. The transformation $u\rightarrow\tilde{u}$ constituting an isomorphism, one will simplify the expressions thereafter while noting *U* what should have been meant $\tilde{u}$.

Note:

- *The fact of separating, in first, the time of the variable of space makes it possible to be strongly inspired by the conceptual tools developed for the elliptic problems. It is completely coherent besides with the sequence "semi-discretization in time/total discretization in space" which usually chairs the determination of a formulation usable in practice.*
- *The assumptions on the geometry ensure us of the property of 1-prolongation of the open one* Ω *. Thus one will be able to confuse the space of Hilbert*

$$B^1(\Omega):=\left\{u\in L^2(\Omega)/\nabla\mathbf{u}\in\left(L^2(\Omega)\right)^q\right\}$$

on which it is convenient to work, with space

$$H^1(\Omega):=\left\{u\in D'(\Omega)/\exists\,U\in H^1(\Re^q)\quad\text{avec }u=U_{|\Omega}\right\}$$

for which the standard theoretical results on the traces, the densities of space and the *standards equivalent are licit.*

- *Taking into account the character lipschitzien of the border the theoretical results which will follow will be able to apply to **structures comprising of the corners** (outgoing or returning). On the other hand treatment of **points** or of **points of graining** fate of this theoretical framework general. In the same way, the fact that the open one must locally be located the same with dimensions one of its border, prevent (theoretically) the treatment of **crack** . To deal with this kind of problem rigorously, an approach consists in correcting the basic functions of the finite elements by a function suitable centered on the internal end of the crack (cf P. GRISVARD. School of Analysis Digital CEA-EDF-INRIA on the breaking process, pp183-192, 1982).*
- ***The indicator in residue using the solution of the problem in temperature, its theoretical limitations are thus, at best, identical to those of the aforesaid problem.***

Taking into account the formulation [éq 2.1-1] one will thus be interested in a solution belonging to following functional space:

Note:

This space comprises also the possible conditions of "generalized" Dirichlet of linear relations type between ddls.

$$T \in W := \left\{ u \in H^1(\Omega) / \gamma_{0,1}\, \mathrm{u} := u_{|\Gamma_1} = f \right\} \quad \text{éq 2-1-5}$$

Moreover, thanks to the geometrical assumptions (H_1) and (H_2), there exists an operator of raising (compound of the operator of usual raising and the operator of prolongation by zero apart from Γ_1) $R : H^{\frac{1}{2}}(\Gamma_1) \rightarrow H^1(\Omega)$ linear, continuous and surjective such as:

$$\gamma_{0,1} Rf = f \quad \forall f \in H^{\frac{1}{2}}(\Gamma_1) \quad \text{éq 2-1-6}$$

One will thus be able **to make the problem initial homogeneous in Dirichlet** while being interested more but in the solution

$$u \in V := \left\{ u \in H^1(\Omega) / \gamma_{0,1} u := u_{|\Gamma_1} = 0 \right\} \quad \text{éq 2-1-7}$$

resulting from the decomposition

$$T := u + Rf \quad \text{éq 2-1-8}$$

Note:

That is to say $\left(X, \|\ \|_X\right)$ Banach, one notes $L^p(0,\tau ; X)$ the space of the functions $t \rightarrow \nu(t)$ strongly measurable for measurement dt such as $\|\nu\|_{0,\tau\, ;\, p,\, X} = \left(\int_0^\tau \|\nu(t)\|_X^p dt \right)^{\frac{1}{p}} < +\infty$. It is Banach, therefore a space of Hilbert for the associated standard.

This change of variable produces the problem simplified in u

$$(P_1)\begin{cases} \rho\, C_p \dfrac{\partial u}{\partial t} - div(\lambda\, \nabla u) = \hat{s} & \Omega\times]0,\tau[\\ u=0 & \Gamma_1\times]0,\tau[\\ \lambda \dfrac{\partial u}{\partial n} = \hat{g} & \Gamma_2\times]0,\tau[\\ \lambda \dfrac{\partial u}{\partial n} + hu = \hat{h} & \Gamma_3\times]0,\tau[\\ u(0)=u_0 & \Omega \end{cases}$$ éq 2-1-9

with the new second member

$$\hat{s} := s - \rho\ C_p \frac{\partial Rf}{\partial t} + div(\lambda\ \nabla Rf)\ \in L^2(0,\tau\,;H^{-1}(\Omega))\ ,\qquad \text{éq 2-1-10}$$

new loadings

$$\hat{g} := g - \lambda\ \frac{\partial Rf}{\partial n}\ \in L^2\left(0,\tau\,,H^{-1/2}(\Gamma_2)\right)\qquad \text{éq 2-1-11}$$

$$\hat{h} := h(T_{ext} - Rf) - \lambda\ \frac{\partial Rf}{\partial n}\ \in L^2\left(0,\tau\,,H^{-1/2}(\Gamma_3)\right)\qquad \text{éq 2-1-12}$$

and the new initial condition

$$u_0(.) := T_0(.) - Rf(.,0)\ \in L^2(\Omega)\qquad \text{éq 2-1-13}$$

Note:

- *It **theoretical raising** , which can appear a little “ethereal”, a completely concrete anchoring in the digital techniques has put in work to solve this kind of problem. It corresponds to one **substitution (this technique is not used in Code_Aster, one prefers to him the technique of double dualisation via ddls of Lagrange [R3.03.01]) of the limiting conditions of Dirichlet** . By renumbering the unknown factors so that these conditions appear in the last, the comparison can be schematized in the following matric form*

$$\begin{bmatrix} A & 0 \\ 0 & Id \end{bmatrix} \begin{bmatrix} T \\ T_{\Gamma_1} := Rf_{\Gamma_1} \end{bmatrix} = \begin{bmatrix} \hat{s} := s - \sum_{j>J} a_{ji} f_j \\ f \end{bmatrix}$$

The assumptions of regularity on the border also ensure us of the good following properties for the workspaces. One then will be able to place itself within the usual abstracted variational framework.

Lemma 1

Under the assumptions (H_1) and (H_2) workspaces W and V with Hilberts are provided with the standard induced by $H^1(\Omega)$.

Proof:

The result comes simply owing to the fact that the application traces $\gamma_{0,1} : H^1(\Omega) \to L^2(\Gamma_1)$ is the made up one of the application traces usual $\gamma_0 : H^1(\Omega) \to H^{\frac{1}{2}}(\Gamma) \subset L^2(\Gamma)$ linear, continuous and surjective (taking into account as of assumptions selected) and of the operator of restriction on Γ_1 also linear, continuous and surjective. Of share their definition, one from of deduced that W and V are sev closed of $H^1(\Omega)$. It of Hilberts is thus provided with the standard $\|\ \|_{1,\Omega}$.

Lemma 2

Under the assumptions (H_1) and (H_2), the standard and the pseudo norm induced by $H^1(\Omega)$ are equivalent on functional space V. One will note $P(\Omega) > 0$ the constant of Poincaré relaying this equivalence

$$\forall\, v \in V\ \ \|v\|_{1,\Omega} \le P(\Omega)\ |v|_{1,\Omega}$$

Note:

One will note thereafter $\|u\|_{\infty,\Omega}:=\underset{pp.\,t\,\in\,\Omega}{\text{supess}}|u(t)|$ and $\forall(u,v)\in\left(H^m(\Omega)\right)^2 (u,v)_{m,\Omega}:=\sum_{|\alpha|\le m}\left(\partial^\alpha u,\partial^\alpha v\right)_{L^2(\Omega)}$, $\|u\|^2_{m,\Omega}:=\sum_{|\alpha|\le m}\|\partial^\alpha u\|^2_{L^2(\Omega)}$ et $|u|^2_{m,\Omega}:=\sum_{|\alpha|=m}\|\partial^\alpha u\|^2_{L^2(\Omega)}$.

Proof:

This result is a corollary of the inequality of Poincaré checked by the open ones called of "Nikodym" of which been part Ω taking into account the assumptions selected. There are however two cases:

- that is to say the problem is really mixed and comprises limiting conditions others that those of Dirichlet, $\text{mes}(\Gamma-\Gamma_1)\neq 0$ (see the demonstration [bib1] §III.7.2 pp922-925),
- either one takes into account only conditions of type imposed temperature, $\text{mes}(\Gamma-\Gamma_1)=0$, $V=H_0^1(\Omega)$ and one finds the standard result of equivalence of the standard and the pseudo norm on this space (see for example the demonstration [bib3] pp18-19).

The compilation of the preceding results makes it possible to encircle it **Tally Variational Abstract** (CVA) on which will rest the weak formulation:

- $H_0^1(\Omega)\subset V\subset H^1(\Omega)$,
- $V\subset H:=L^2(\Omega)=H'\subset V'\subset H^{-1}(\Omega)$ while identifying H and its dual,
- there is a linear canonical injection continues V in H,
- V is dense in H and the injection is compact (it inherits in that the properties $H^1(\Omega)$ with respect to H),
- V is provided with the pseudo norm induced by $H^1(\Omega)$ and H of its usual standard.

Note:

According to a formulation of the theorem of compactness of Rellich adapted to spaces of Sobolev on open (for example, theorem 1.5.2 [bib3] pp29-30).

2.2 Strong formulation with weak

By multiplying the principal equation of the problem in extreme cases [éq 2.1-1] by a function test $v\in V$ and by using the theorems of Green and Reynolds (to commutate the integral in space and derivation in time, with Ω fixed and from characteristic materials independent of time), one obtains:

$$\frac{d}{dt}\int_\Omega \rho\, C_p u(t) v\, dx+\int_\Omega \lambda\,\nabla u(t).\nabla v\, dx=\int_\Omega \hat{s}(t) v\, dx+\int_\Gamma \lambda\,\frac{\partial u(t)}{\partial n} v\, d\sigma \qquad \text{éq 2.2-1}$$

By introducing the limiting conditions into [éq 2.2-1], it occurs **weak formulation** (with the direction **distributions (within this framework general, the temporal derivative is thus to take with the weak direction) temporal** of $D'(]0,\tau[)$) following:
The solution is sought

$$u\in L^2(0,\tau\,;V)\cap C^0(0,\tau\,;H) \qquad \text{éq 2.2-2}$$

checking the problem

$$(P_2)\begin{cases} trouver\, u : t\in]0,\tau[\,\rightarrow u(t)\in V\; tel\, que \\ \forall v\in V\, \dfrac{d}{dt}(\rho\, C_p u(t), v)_{0,\Omega} + a(t\,;u(t),v) = (b(t),v) \\ u(0)=u^0 \end{cases} \qquad \text{éq 2.2-3}$$

with

$$a(t;u(t),v) := \int_\Omega \lambda\, \nabla u(t).\nabla v\; dx + \int_{\Gamma_3} h(t)\,\gamma_{0,3} u(t)\,\gamma_{0,3} v\; d\sigma$$
$$(b(t),v) := \langle \hat{s}(t), v\rangle_{-1\times 1,\Omega} + \langle \hat{g}(t), \gamma_{0,2} v\rangle_{-\frac{1}{2}\times\frac{1}{2},\Gamma_2} + \langle \hat{h}(t), \gamma_{0,3} v\rangle_{-\frac{1}{2}\times\frac{1}{2},\Gamma_3} \qquad \text{éq 2.2-4}$$

while noting $\langle\, ,\,\rangle_{p\times q,\Theta}$ the hook of duality enters spaces $H^p(\Theta)$ and $H^q(\Theta)$.

Note:

- *The unknown field and the function test belong to the same functional space, which is more comfortable from a digital and theoretical point of view.*
- *The hooks of duality will not be able to be transformed into integrals with the classical direction (as for the surface term of (T has;. .)) that if one restricts the space of membership of the new source and the new loadings with*

$$\hat{s}\in L^2(0,\tau\,;L^2(\Omega)),\;\; \hat{g}\in L^2(0,\tau\,;L^2(\Gamma_2))\text{ et }\hat{h}\in L^2(0,\tau\,;L^2(\Gamma_3)) \qquad \textit{éq 2.2-5}$$

According to [éq 2-1-10] [éq 2-1-12] this restriction can be translated on the initial loadings in the form

$$f\in L^2\left(0,\tau\,;H^{\frac{3}{2}}(\Gamma_1)\right),\; s\in L^2(0,\tau\,;L^2(\Omega)),\; g\in L^2(0,\tau\,;L^2(\Gamma_2))\text{ et }T_{ext}\in L^2(0,\tau\,;L^2(\Gamma_3))$$
éq 2.2-6

- *The formulation (P_2) a direction has well, because it is shown that*

$$t\rightarrow a(t\,;u(t),v)\in L^2(]0,\tau[)\subset D'(]0,\tau[)$$
$$t\rightarrow \rho\; C_p u(t)\in L^2(0,\tau\,;V)\;\; et\;\; v\in V \Rightarrow t\rightarrow(\rho\; C_p u(t),v)_{0,\Omega}\in L^2(]0,\tau[)\subset D'(]0,\tau$$
$$t\rightarrow \hat{s}(t)\in L^2(0,\tau\,;H^{-1}(\Omega))\;\; \textbf{et}\;\; v\in H^1(\Omega)\subset H^{-1}(\Omega)$$
$$\Rightarrow t\rightarrow\langle \hat{s}(t),v\rangle_{-1\times 1,\Omega}\in L^2(]0,\tau[)\subset D'(]0,\tau[)$$
$$t\rightarrow \hat{g}(t)\in L^2\left(0,\tau\,;H^{-\frac{1}{2}}(\Gamma_2)\right)\;\; \textbf{et}\;\; \gamma_{0,2}v\in H^{\frac{1}{2}}(\Gamma_2)\subset H^{-\frac{1}{2}}(\Gamma_2)$$
$$\Rightarrow t\rightarrow\langle \hat{g}(t),\gamma_{0,2}v\rangle_{-\frac{1}{2}\times\frac{1}{2},\Gamma_2}\in L^2(]0,\tau[)\subset D'(]0,\tau[)$$

and one finds obviously the same thing for the term of exchange on Γ_3 .

- *In the surface integrals one will note henceforth $u(t)$ and v what should be noted (in any rigour) $\gamma_{0,i}u(t)\,\text{et}\,\gamma_{0,i}v$.*
- *Membership of the solution with $L^2(0,\tau\,;V)$ rise from the assumptions on the data and properties of the differential operators and trace. The fact that it*

must also belong with $C^0(0,\tau\,;H)$ comes just from the necessary one justification of the condition of Cauchy.

One can then be interested in **the existence and with the unicity of the solution of the initial problem** (P_0) while showing **its equivalence with** (P_2) and by applying to this last a parabolic alternative of the theorem of Lax-Milgram.

> **Theorem 3**
> Within the abstracted variational framework (CVA) definite previously and by supposing that the assumptions (H$_1$), (H$_2$) and (H$_3$) are checked, then the problem (P_2) admits a solution and only one $u\in L^2(0,\tau\,;V)\cap C^0(0,\tau\,;H)$.

Proof:

This result comes from theorems 1 & 2 of the "Dautray-Lions" (cf [bib3], §XVIII pp615-627). To use them it is necessary nevertheless to check

- Mesurability of the bilinear form $\forall(u(t),v)\in V^2 \quad t\rightarrow a(t\,;u(t),v)\ \text{sur}\ ]0,\tau[$
- Its continuity on $V\times V$

$$pp\ t\in]0,\tau[\quad |a(t\,;u(t),v)|\leq\|\lambda\|_{\infty,\Omega}|u(t)|_{1,\Omega}|v|_{1,\Omega}+\|h(t)\|_{\infty,\Gamma_3}\|u(t)\|_{\frac{1}{2},\Gamma_3}\|v\|_{\frac{1}{2},\Gamma_3}$$

$$\forall(u(t),v)\in V^2 \qquad \leq\max\left(\|\lambda\|_{\infty,\Omega},\|h(t)\|_{\infty,\Gamma_3}C_3^2P^2(\Omega)\right)|u(t)|_{1,\Omega}|v|_{1,\Omega}$$

with C_3 the constant of continuity of the operator of trace on Γ_3 and $P(\Omega)$ the constant of Poincaré.

- Its V- ellipticity compared to H

$$pp\ t\in]0,\tau[\quad a(t\,;v,v)+\frac{\beta}{2}\|v\|_{0,\Omega}^2\geq C_0^{-2}\left(\|\lambda\|_{\infty,\Omega}-\|h(t)\|_{\infty,\Gamma_3}C_3^2\right)\|v\|_{0,\Omega}^2$$

$$\forall v\in V \ \Rightarrow\ a(t\,;v,v)+\underbrace{\beta}_{>0}\|v\|_{0,\Omega}^2\geq\underbrace{\left\{\frac{\beta}{2}+C_0^{-2}\left(\|\lambda\|_{\infty,\Omega}-\|h(t)\|_{\infty,\Gamma_3}C_3^2\right)\right\}}_{\alpha>0}\|v\|_{0,\Omega}^2$$

with C_0 the constant of continuity of the canonical injection of $H^1(\Omega)$ in $L^2(\Omega)$.

- The continuity of the linear form $b(t)$ on V

$$pp\ t\in]0,\tau[\quad |(b(t),v)|\leq\|\hat{s}(t)\|_{-1,\Omega}\|v\|_{1,\Omega_3}+\|\hat{g}(t)\|_{-\frac{1}{2},\Gamma_2}\|v\|_{\frac{1}{2},\Gamma_2}+\|\hat{h}(t)\|_{-\frac{1}{2},\Gamma_3}\|v\|_{\frac{1}{2},\Gamma_3}$$

$$\forall v\in V \qquad \leq P(\Omega)\max\left(\|\hat{s}(t)\|_{-1,\Omega},\|\hat{g}(t)\|_{-\frac{1}{2},\Gamma_2}C_2,\|\hat{h}(t)\|_{-\frac{1}{2},\Gamma_3}C_3\right)|v|_{1,\Omega}$$

with C_2 the constant of continuity of the operator of trace on Γ_2 .

> **Theorem 4**
> Problems (P_0) and (P_2) are equivalent and thus the initial problem admits a solution and only one $u\in L^2(0,\tau\,;V)\cap C^0(0,\tau\,;H)$.

Proof:

The existence and the unicity of the solution of the problem (P_0) of course result from the preceding theorem, once the equivalence of the two problems was shown. It thus remains to prove the opposite implication $(P_2) \Rightarrow (P_0)$ who is very hard to exhume "not formally". In particular the limiting conditions of Neumann, Robin and the condition of Cauchy are difficult to obtain rigorously. The "Dautray - Lions" proposes a very technical demonstration ([bib1] §XVIII pp637-641). By adapting his results one shows that in our case, the limiting conditions on Γ_i in fact are checked, not on $L^2\left(0,\tau, H^{-\frac{1}{2}}(\Gamma_i)\right)$, but on space $(B_i)' \supset H_{00}^{-\frac{1}{2}}(\Gamma_\tau^i)$ (while noting $\Gamma_\tau^i := \Gamma_i \times]0,\tau[$) defined as being the dual topological one of

$$B_i := \left\{ w \in H^{\frac{1}{2}}(\partial\Omega_\tau) \cap L^2(\Gamma_\tau^i) / \exists v \in L^2(0,\tau; V) \text{ avec } v_{|\Omega\times\{0\}} = v_{|\Omega\times\{\tau\}} = 0 \text{ et } v_{|\Gamma_i} = w \right\}$$

Note:

- *Because of **low regularity** imposed on **thermal conductivity** , $\lambda \in L^\infty(\Omega)$, one cannot claim with **standard regularity** " " $u \in H^2(\Omega)$. Indeed in the case, for example, of one **bi--material** (with $\Omega = \Omega_1 \cup \Sigma \cup \Omega_2$) from which the characteristics are distinct on both sides of the border Ω , [éq 2-1-9] and the theorem of the divergence imposes*

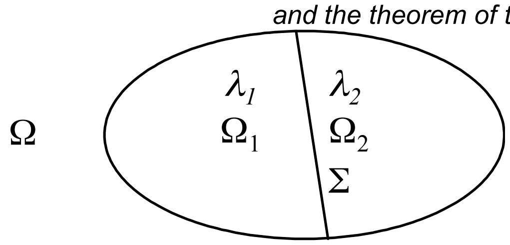


Figure 2.2-a: Example of bi--material

$$\lambda_1 \frac{\partial u(t)}{\partial n}\Big|_{\Omega_1} = \lambda_2 \frac{\partial u(t)}{\partial n}\Big|_{\Omega_2} \textbf{ dans } H_{00}^{-\frac{1}{2}}(\Sigma) \quad pp\ t \in]0,\tau[$$

However $\lambda_1 \neq \lambda_2$, therefore the condition of transmission cannot be carried out on the internal border Σ

$$\frac{\partial u(t)}{\partial n}\Big|_{\Omega_1} \neq \frac{\partial u(t)}{\partial n}\Big|_{\Omega_2} \quad pp\ \Sigma \quad pp\ t \in]0,\tau[$$

*Thus $u(t) \in H^2(\Omega_1) \cup H^2(\Omega_2)$ do not involve $u(t) \in H^2(\Omega)$. This restriction will not enable us to exhume, as in [bib6], of **increases of the "strong" type** total space error and site indicator of error. Within our framework of work plus general one will have to be satisfied **estimates of the "weak" type.***

- *This kind of problem also meets when open polyhedric the nonconvex ones are treated (for example comprising a returning corner). Open polyhedric (known as polygonal into two-dimensional) is a finished meeting of polyhedrons. A polyhedron is a finished intersection of closed half spaces.*
- *To obtain estimates of the "strong" type, it is necessary to concede more regularity on the geometry and the loadings*

Γ *variety of dimension $q-1$, C_2 by piece (property of 2-prolongation)* *(H 4)*

$$s \in L^2(0,\tau\,;L^2(\Omega)) \quad T_0 \in H^1(\Omega)$$
$$f \in L^2\left(0,\tau\,;H^{\frac{3}{2}}(\Gamma_1)\right),\ g \in L^2\left(0,\tau\,;H^{\frac{1}{2}}(\Gamma_2)\right),\ \ T_{ext} \in L^2\left(0,\tau\,;H^{\frac{1}{2}}(\Gamma_3)\right) \qquad (H_5)$$
$$\rho\,,\ C_p\,,\ \lambda \in L^\infty(\Omega) \quad h \in L^2(0,\tau\,;L^\infty(\Gamma_3))$$

What allows obtaining a solution in the following intersection

$$u \in L^2(0,\tau\,;H^2(\Omega)) \cap C^0(0,\tau\,;H^1(\Omega)) \qquad \text{éq 2.2-7}$$

Now that we made sure of the existence and the unicity of the solution within the functional framework required by the operators of *Code_Aster*, we go **semi-to discretize in time (P_0)** then **to spatially discretize the whole by a method finite elements**. In parallel, we will study its properties of stability. They we will be very useful to create the standards, the techniques and the inequalities which will intervene in the genesis of the indicator of error in residue.

3 Discretization and controllability

3.1 Controllability of the continuous problem

By not making no concession on the assumptions of regularity seen in the preceding paragraph, there is increase known as "weak" (to take again a terminology in force in the article which was used as a basis for our study [bib6]) following.

Property 5

Within the abstracted variational framework (CVA) definite previously and by supposing that the assumptions (H_1), (H_2) and (H_3) are checked, one has **"weak" controllability of the continuous problem** (with $K_1\left(\|\lambda\|_{\infty,\Omega}, mes(\Gamma_i), \gamma_{0,i}, P(\Omega)\right)>0$)

$$pp\ t \quad \|\sqrt{\rho C_p}u(t)\|^2_{0,\Omega}+\int_0^t \|\sqrt{\lambda}\nabla u(\xi)\|^2_{0,\Omega}d\xi \leq \|\sqrt{\rho C_p}u_0\|^2_{0,\Omega}+ K_1\left\{\int_0^t \|\hat{s}(\xi)\|^2_{-1,\Omega}+\|\hat{g}(\xi)\|^2_{-\frac{1}{2},\Gamma_2}+\|\hat{h}(\xi)\|^2_{-\frac{1}{2},\Gamma_3}d\xi\right\} \qquad \text{éq 3.1-1}$$

Proof:

One here will detail this a little technical demonstration because, on the one hand, the specialized literature seldom returns in this level of details and, on the other hand, one will re-use same methodology to exhume all increases which will follow one another in this theoretical part of the document. First of all, by multiplying the equation of [éq 2.1-1] by $u(t)$, while integrating spatially on Ω, then temporally on $[0,t]$ avec $t\in[0,\tau[$ one obtains, like characteristic materials are supposed to be independent of time,

$$\frac{1}{2}\int_0^t \frac{\partial}{\partial t}\left(\rho C_p u(\xi), u(\xi)\right)_{0,\Omega}d\xi-\int_0^t\left(div\left(\lambda\nabla u(\xi)\right), u(\xi)\right)_{0,\Omega}d\xi=\int_0^t\langle\hat{s}(\xi), u(\xi)\rangle_{-1\times 1,\Omega}d\xi \qquad \text{éq 3.1-2}$$

By using the formula of Green and the conditions limiting of [éq 2.1-1] one obtains

$$\frac{1}{2}\left(\|\sqrt{\rho C_p}u(t)\|^2_{0,\Omega}-\|\sqrt{\rho C_p}u_0\|^2_{0,\Omega}\right)+\int_0^t\left(\lambda\nabla u(\xi), \nabla u(\xi)\right)_{0,\Omega}d\xi+\int_0^t h(\xi)u^2(\xi)d\xi= \int_0^t\left\{\langle\hat{s}(\xi), u(\xi)\rangle_{-1\times 1,\Omega}+\langle\hat{g}(\xi), u(\xi)\rangle_{-\frac{1}{2}\times\frac{1}{2},\Gamma_2}+\langle\hat{h}(\xi), u(\xi)\rangle_{-\frac{1}{2}\times\frac{1}{2},\Gamma_3}\right\}d\xi \qquad \text{éq 3.1-3}$$

One can oust the term of exchange of [éq 3.1-3] because it is supposed that $h(t)\geq 0$ pp t . By using an argument of duality, the inequality of Cauchy-Schwartz, lemma 2 and the relation $2ab\leq\left(\frac{a}{\alpha}\right)^2+(b\alpha)^2\ (\alpha>0)$, one obtains

$$\int_0^t\langle\hat{s}(\xi), u(\xi)\rangle_{-1\times 1,\Omega}d\xi\leq\frac{1}{2}\left(\frac{1}{\alpha^2}\int_0^t\|\hat{s}(\xi)\|^2_{-1,\Omega}d\xi+\frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\alpha^2\int_0^t\|\sqrt{\lambda}\nabla u(\xi)\|^2_{0,\Omega}d\xi\right) \qquad \text{éq 3.1-4}$$

One carries out same work on the loadings, thus defining the parameters β and γ by taking again the notations of theorem 3 (for C_i ...), then one inserts these inequalities in [éq 3.1-3]

$$\|\sqrt{\overline{\rho C_p}}\, u(t)\|_{0,\Omega}^2+\left(2-\frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\left(\alpha^2+C_2^2\beta^2+C_3^2\gamma^2\right)\right)\int_0^t \|\sqrt{\lambda}\,\nabla u(\xi)\|_{0,\Omega}^2\, d\xi \leq$$

$$\|\sqrt{\overline{\rho C_p}}\, u_0\|_{0,\Omega}^2+\int_0^t\left\{\frac{\|\hat{s}(\xi)\|_{-1\times 1,\Omega}^2}{\alpha^2}+\frac{\|\hat{g}(\xi)\|_{-\frac{1}{2}\times\frac{1}{2},\Gamma_2}^2}{\beta^2}+\frac{\|\hat{h}(\xi)\|_{-\frac{1}{2}\times\frac{1}{2},\Gamma_3}^2}{\gamma^2}\right\}d\xi \qquad \text{éq 3.1-5}$$

It now remains to seek a triplet of strictly positive realities (α,β,γ), not privileging any particular term, in order to reveal a constant independent of the solution and parameter setting in front of the term in gradient. One arbitrarily chooses to pose

$$2-\frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\left(\alpha^2+C_2^2\beta^2+C_3^2\gamma^2\right)=1 \qquad \text{éq 3.1-6}$$

Maybe, for example,

$$\begin{cases}\alpha^2=\dfrac{\|\lambda\|_{\infty,\Omega}\left(\text{mes}(\Gamma_1)+1\right)}{P^2(\Omega)\left(\text{mes}(\Gamma)+3\right)}\\[2ex] \beta^2=\dfrac{\|\lambda\|_{\infty,\Omega}\left(\text{mes}(\Gamma_2)+1\right)}{C_2^2P^2(\Omega)\left(\text{mes}(\Gamma)+3\right)}\\[2ex] \gamma^2=\dfrac{\|\lambda\|_{\infty,\Omega}\left(\text{mes}(\Gamma_3)+1\right)}{C_3^2P^2(\Omega)\left(\text{mes}(\Gamma)+3\right)}\end{cases} \qquad \text{éq 3.1-7}$$

From where increase [éq 3.1-1] while taking

$$K_1=\frac{P^2(\Omega)\left(\text{mes}(\Gamma)+3\right)}{\|\lambda\|_{\infty,\Omega}}\max\left(\frac{1}{\left(\text{mes}(\Gamma_1)+1\right)},\frac{C_2^2}{\left(\text{mes}(\Gamma_2)+1\right)},\frac{C_3^2}{\left(\text{mes}(\Gamma_3)+1\right)}\right) \qquad \text{éq 3.1-8}$$

Note:

- *The recourse to measurements of the external borders is a trick making it possible the inequality to support the passage to limit ($\Gamma_i \rightarrow 0$) when one or more limiting conditions have suddenly missed in this mixed problem.*
- *While placing itself within the particular framework of a homogeneous problem of Cauchy-Dirichlet with characteristic materials constants equal to the unit*

$$\lambda=\rho\, C_p=1\ ,\quad \Gamma_2=\Gamma_3=\emptyset \quad \textbf{et} \quad \hat{s}=s \qquad \text{éq 3.1-9}$$

and by introducing particular standards on $V=H_0^1(\Omega)$ and its dual

$$\|\hat{s}(t)\|^*_{-1,\Omega}=\sup_{v\in V,\ v\neq 0}\frac{\langle \hat{s}(t), v\rangle_{-1\times 1,\Omega}}{\|v\|^*_{1,\Omega}} \text{ avec } \|v\|^*_{1,\Omega}=\frac{mes(\Gamma)+1}{(mes(\Gamma)+3)\,P^2(\Omega)}\|v\|_{1,\Omega} \qquad \text{éq 3.1-10}$$

one finds well the inequality (2) pp427 of [bib6].

- *If one allows more regularity on the geometry (H_4) and on the data (H_5), one can exhume it **during, known as "extremely",** preceding property. The control of the solutions which it operates is of course more precise than with [éq 3.1-1] because it is carried out via stronger standards. Contrary to "weak" increase, it makes too **to intervene directly the infinite standard of the coefficient of convectif exchange** . His obtaining here will not be detailed because this family of increase is not essential for the calculation of the required indicator.*

3.2 Semi-discretization in time

A step of time is fixed Δt such as $\frac{\tau}{\Delta t}$ that is to say an entirety N and such as the temporal discretization is regular: $t_0=0,\ t_1=\Delta t,\ t_2=2\Delta t \cdots t_n=n\Delta t$.

Note:

This assumption of regularity does not have really importance, it just makes it possible to simplify the writing of the semi-discretized problem. To model an unspecified transient at the moment t_n , it is just enough to replace Δt by $\Delta t_n=t_{n+1}-t_n$.

semi - discretization in time of [éq 2.1-1] by θ **- method** lead to the following problem:
The continuation is sought

$$(u^n)_{0\leq n\leq N}\in V \qquad \text{éq 3.2-1}$$

such as

$$(P_1^{n+1})\begin{cases}\rho C_p\dfrac{u^{n+1}-u^n}{\Delta t}-\theta\,\mathrm{div}(\lambda\,\nabla u^{n+1})-(1-\theta)\,\mathrm{div}(\lambda\,\nabla u^n)=\theta\ \hat{s}^{n+1}+(1-\theta)\,\hat{s}^n & \Omega\quad 0\leq n\leq N-1\\ u^{n+1}=0 & \Gamma_1\quad 0\leq n\leq N-1\\ \lambda\,\dfrac{\partial u^{n+1}}{\partial n}=\hat{g}^{n+1} & \Gamma_2\quad 0\leq n\leq N-1\\ \lambda\,\dfrac{\partial u^{n+1}}{\partial n}+h^{n+1}\,u^{n+1}=\hat{h}^{n+1} & \Gamma_3\quad 0\leq n\leq N-1\\ u^0(.)=u_0 & \Omega\end{cases}$$

éq 3.2-2

while posing

$$\Xi^n = \Xi\left(\mathbf{x}, n\frac{\tau}{\Delta t}\right) \text{ avec } \Xi \in \{u, \hat{s}, \hat{h}, h, \hat{g}\} \text{ et } 0 \leq n \leq N$$

While multiplying [éq 3.2-2] by a function test v and while integrating on Ω , one of course finds (via the formula of Green) the variational formulation [éq 2.2-3] semi-discretized in time

$$\left(P_2^{n+1}\right)\begin{cases}\text{Etant donnés } u^n, \hat{s}^n, \hat{s}^{n+1}, \hat{g}^n, \hat{g}^{n+1}, \hat{h}^n, \hat{h}^{n+1}, h^n, h^{n+1} \\ \text{Calculer } u^{n+1} \in V \text{ tel que} \\ \left(\rho\, C_p u^{n+1}, \nu\right)_{0,\Omega} + \Delta t\ a\left(n \Delta t \theta\ ;\ u_\theta^{n+1}, \nu\right) = \left(\rho\, C_p u^n, \nu\right)_{0,\Omega} + \Delta t\left(b_\theta^n, \nu\right) \quad \left(\forall \nu \in V\right)\end{cases} \quad \text{éq 3.2-3}$$

with

$$\begin{aligned} &\Xi_\theta^{n+1} := \theta\ \Xi^{n+1} + (1-\theta)\,\Xi^n \quad \text{où} \quad \Xi \in \{u, hu, b, \hat{s}, \hat{g}, \hat{h}\} \\ &a\left(n \Delta t \theta\ ; u_\theta^{n+1}, \nu\right) := \int_\Omega \lambda\ \nabla u_\theta^{n+1} . \nabla \nu\ dx + \int_{\Gamma_3} (hu)_\theta^{n+1} \nu\ d\sigma \\ &\left(b_\theta^{n+1}, \nu\right) := \langle \hat{s}_\theta^{n+1}, \nu \rangle_{-1\times 1, \Omega} + \langle \hat{g}_\theta^{n+1}, \gamma_{0,2} \nu \rangle_{-\frac{1}{2}\times\frac{1}{2}, \Gamma_2} + \langle \hat{h}_\theta^{n+1}, \gamma_{0,3} \nu \rangle_{-\frac{1}{2}\times\frac{1}{2}, \Gamma_3}\end{aligned} \quad \text{éq 3.2-4}$$

This semi-discretization in time made it possible to transform our parabolic problem into an elliptic problem to which one can apply the theorem of standard Lax-Milgram. The assumptions of this theorem are checked easily thanks to the results of continuity and ellipticity of the demonstration of theorem 3. From where the existence and the unicity of the continuation $\left(u^n\right)_{0 \leq n \leq N} \in V$ searched.

Note:

- *While posing $Rf = 0$ one finds well the semi-discretized variational formulation of Code_Aster (cf [R5.02.01 §5.1.3]). (Or them) the condition (S) of Dirichlet (generalized or not) are checked within the space of work W which the solution must belong. Moreover, while implicitant completely θ - method (Euler retrogresses) one finds the formulation of code SYRTHES [bib9].*
- *To be able semi-to discretize by θ - method one needs to restrict the membership of the new source with $\hat{s} \in C^0\left(0, \tau\, ; H^{-1}(\Omega)\right)$ (to be able to take a value in a given moment). In addition, the initialization of the iterative process [éq 3.2-3] necessarily involves $u_0 \in H^1(\Omega)$.*
- *To simplify the expressions one will not mention any more the temporal dependence of the bilinear form (T has;. .) (for the implicitation of the term of exchange), it will remain implied by that of the solution.*

As for the continuous problem, by not making no concession on the assumptions of regularity, there is "weak" increase following:

Property 6

By supposing that the assumptions of property 5 are checked, that it θ -diagram is unconditionally stable ($\theta \geq \frac{1}{2}$), that $\hat{s} \in C^0\left(0,\tau ; H^{-1}(\Omega)\right)$ and $u_0 \in H^1(\Omega)$, one has **"weak" controllability of the problem semi-discretized in time** (with $K_1\left(\|\lambda\|_{\infty,\Omega}, mes(\Gamma_i), \gamma_{0,i}, P(\Omega)\right) > 0$)

$$\begin{aligned} &\|\sqrt{\rho C_p} u^{n+1}\|^2_{0,\Omega} + \Delta t \|\sqrt{\lambda} \nabla u_\theta^{n+1}\|^2_{0,\Omega} \leq \frac{1}{2}\|\sqrt{\rho C_p} u^{n+1}\|^2_{0,\Omega} + \frac{4\theta - 3}{2}\|\sqrt{\rho C_p} u^n\|^2_{0,\Omega} \\ &\forall 0 \leq n \leq N-1 \qquad + \frac{\Delta t}{2}\|\sqrt{\lambda} \nabla u_\theta^{n+1}\|^2_{0,\Omega} + \frac{K_1 \Delta t}{2}\left(\|\hat{s}_\theta^{n+1}\|^2_{-1,\Omega} + \|\hat{g}_\theta^{n+1}\|^2_{-\frac{1}{2},\Gamma_2} + \|\hat{h}_\theta^{n+1}\|^2_{-\frac{1}{2},\Gamma_3}\right) \end{aligned}$$

éq 3.2-5

Proof:

This inequality is obtained easily by taking again the stages described in the demonstration of property 5. It is necessary, on the other hand, to multiply [éq 3.2-2] by the particular function test

$$u_\theta^{n+1} := \theta\, u^{n+1} + (1-\theta) u^n \in V \qquad \text{éq 3.2-6}$$

and to oust the term of exchange by the argument

$$0 < \min\left(h^n, h^{n+1}\right) \|u_\theta^{n+1}\|^2_{0,\Gamma_3} \leq \int_{\Gamma_3} (hu)_\theta^{n+1} u_\theta^{n+1} dx \leq \max\left(h^n, h^{n+1}\right) \|u_\theta^{n+1}\|^2_{0,\Gamma_3} \qquad \text{éq 3.2-7}$$

In addition there is not that time the term source and the loadings which require the trick [éq 3.1-4], it should also be set up on the cross term $(2\theta - 1)\int_\Omega \rho\, C_p u^{n+1} u^n dx$. From where a fourth parameter η checking a system of the type [éq 3.1-6]

$$\begin{aligned} &|2 - \frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\left(\alpha^2 + C_2^2 \beta^2 + C_3^2 \gamma^2\right) = 1 \\ &|2\theta - \eta^2|1 - 2\theta| = 1 \end{aligned} \qquad \text{éq 3.2-8}$$

Note:

- *If one in the case of is not placed **conditionally stable diagram** , in addition to the digital problems which are likely to occur at the time of implementation the effective of the operator, one will not be able to determine the parameters* $(\alpha, \beta, \gamma, \eta)$ *governing the equation [éq 3.2-8].*
- *While placing themselves within the particular framework [éq 3.1-9] of the article [bib6] and by taking again the equivalent standards [éq 3.1-10], like* $\frac{4\theta - 3}{2} < \frac{1}{2}$ *, one finds well the inequality (5) pp428.*

While stating [éq 3.2-5] for the values of $m \in \{0,1\ldots,n\}$ and by summoning these increases until n, one obtains the "weak" increase following which takes account of the history of the solutions and the data.

Corollary 7

Under the assumptions of property 6, there is increase

$$\|\sqrt{\rho\, C_p} u^n\|^2_{0,\Omega} + \Delta t \sum_{m=0}^{n-1} \|\sqrt{\lambda}\, \nabla u_\theta^{m+1}\|^2_{0,\Omega} + 4(1-\theta) \sum_{m=0}^{n-1} \|\sqrt{\rho\, C_p} u^m\|^2_{0,\Omega} \leq (4\theta - 3) \|\sqrt{\rho\, C_p} u_0\|^2_{0,\Omega}$$

$$\forall 0 \leq n \leq N \qquad + K_1 \Delta t \sum_{m=0}^{n-1} \left(\|\hat{s}_\theta^{m+1}\|^2_{-1,\Omega} + \|\hat{g}_\theta^{m+1}\|^2_{-\frac{1}{2},\Gamma_2} + \|\hat{h}_\theta^{m+1}\|^2_{-\frac{1}{2},\Gamma_3} \right)$$

éq 3.2-9

or more simply

$$\|\sqrt{\rho\, C_p} u^n\|^2_{0,\Omega} + \Delta t \sum_{m=0}^{n-1} \|\sqrt{\lambda}\, \nabla u_\theta^{m+1}\|^2_{0,\Omega} \leq \|\sqrt{\rho\, C_p} u_0\|^2_{0,\Omega}$$

$$\forall 0 \leq n \leq N \qquad + K_1 \Delta t \sum_{m=0}^{n-1} \left(\|\hat{s}_\theta^{m+1}\|^2_{-1,\Omega} + \|\hat{g}_\theta^{m+1}\|^2_{-\frac{1}{2},\Gamma_2} + \|\hat{h}_\theta^{m+1}\|^2_{-\frac{1}{2},\Gamma_3} \right)$$

éq 3.2-10

Proof:

Obtaining [éq 3.2-9] being already explained, it remains to be shown [éq 3.2-10]. This "coarse" inequality more comes simply owing to the fact that

$$4(1-\theta) \sum_{m=0}^{n-1} \|\sqrt{\rho\, C_p} u^m\|^2_{0,\Omega} \geq 0$$

$$(4\theta - 3) \|\sqrt{\rho\, C_p} u_0\|^2_{0,\Omega} \leq \|\sqrt{\rho\, C_p} u_0\|^2_{0,\Omega}$$

éq 3.2-11

Note:

- *One can obviously pass the same remark as [bib6] by noting that the last term of [éq 3.2-9] is a sum of Riemann which tends towards the last term of [éq 3.1-1] when the step of time tends towards zero. In addition, if one introduces the function (*with $\chi_{[n\Delta t,(n+1)\Delta t]}$ the temporal function characteristic of the interval $[n\Delta t,(n+1)\Delta t]$) $u(t) = u_\theta^{n+1} \chi_{[n\Delta t,(n+1)\Delta t]}(t)$ *closely connected per pieces in [éq 3.1-1], one finds exactly [éq 3.2-9].*
- *As for [éq 3.1-1], by adopting the less restrictive approaches (H_4) and (H_5), one finds a version "strong" of properties 6 and 7.*

3.3 Error of temporal discretization

The preceding results on the continuous problem and its form semi-discretized in time are re-used jointly to study the controllability of the error of temporal discretization

$$\forall\, 0 \le n \le N \quad e^n := u^n - u(n\Delta t)$$
$$e^0 = 0 \qquad \text{éq 3.3-1}$$

One starts by revealing this error by withdrawing from the equation [éq 3.2-2] the relations

$$\frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} \frac{\partial u(\xi)}{\partial t} d\xi = \frac{u((n+1)\Delta t) - u(n\Delta t)}{\Delta t}$$
$$\theta\ \rho\ C_p \frac{\partial u((n+1)\Delta t)}{\partial t} = \theta\ \mathrm{div}(\lambda\ \nabla u((n+1)\Delta t)) + \theta\ \hat{s}((n+1)\Delta t) \qquad \text{éq 3.3-2}$$
$$(1-\theta)\ \rho\ C_p \frac{\partial u(n\Delta t)}{\partial t} = (1-\theta)\ \mathrm{div}(\lambda\ \nabla u(n\Delta t)) + (1-\theta)\ \hat{s}(n\Delta t)$$

that is to say

$$\rho\ C_p \frac{e^{n+1} - e^n}{\Delta t} - \mathrm{div}(\ \lambda\ \nabla e_\theta^{n+1}) = \frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} \frac{\partial u(\xi)}{\partial t} d\xi + \rho\ C_p \left(\frac{\partial u}{\partial t}\right)_\theta \qquad \text{éq 3.3-3}$$

while noting

$$e_\theta^{n+1} := \theta\ e^{n+1} + (1-\theta)\, e^n$$
$$\left(\frac{\partial u}{\partial t}\right)_\theta := \theta\ \frac{\partial u}{\partial t}((n+1)\Delta t) + (1-\theta)\frac{\partial u}{\partial t}(n\Delta t) \qquad \text{éq 3.3-4}$$

From this expression one can describe, via the recourse to the formula of Taylor, the "weak" controllability of the error of temporal discretization. But to be able to use the derivative temporal of the solution one continues needs a minimum of regularity in t, for example by conceding that

$$u \in H^1(0,\tau\,;V) \cap H^2(0,\tau\,;H^{-1}(\Omega)) \qquad \text{éq 3.3-5}$$

Property 8

By supposing that the solution checks the additional assumption of temporal regularity [éq 3.3-5], one has **"weak" controllability of the error of temporal discretization**

$$\forall\, 0 \le n \le N \quad \|\sqrt{\rho\ C_p}\, e^n\|_{0,\Omega}^2 + \Delta t \sum_{m=0}^{n-1} \|\sqrt{\lambda}\ \nabla e_\theta^{m+1}\|_{0,\Omega}^2 \le$$
$$\frac{K_1 (\Delta t)^3 (\rho\ C_p)^2}{4} \sum_{m=0}^{n-1} \left((1-\theta)\frac{\partial^2 u}{\partial t^2}(m\Delta t) - \theta\ \frac{\partial^2 u}{\partial t^2}((m+1)\Delta t)\right)$$

éq 3.3-6

Proof:

While evaluating [éq 3.3-3] by a formula of Taylor to order 2, one utilizes the derivative second temporal of the solution and one shows that the continuation of error $\left(e^n\right)_{0\le n\le N}\in V$ check a problem similar to [éq 3.2-2] (by supposing that the temporal discretization of the limiting conditions are exact)

$$\left(P_3^{n+1}\right)\begin{cases}\rho C_p\dfrac{e^{n+1}-e^n}{\Delta t}-\mathrm{div}\left(\lambda\,\nabla e_\theta^{n+1}\right)= \\ \qquad \dfrac{\rho C_p\Delta t}{2}\left(\left(1-\theta\right)\dfrac{\partial^2 u}{\partial t^2}\left(n\Delta t\right)-\theta\,\dfrac{\partial^2 u}{\partial t^2}\left(\left(n+1\right)\Delta t\right)\right) & \Omega\quad 0\le n\le N-1\\ e^{n+1}=0 & \Gamma_1\quad 0\le n\le N-1\\ \lambda\,\dfrac{\partial e^{n+1}}{\partial n}=0 & \Gamma_2\quad 0\le n\le N-1\\ \lambda\,\dfrac{\partial e^{n+1}}{\partial n}+h^{n+1}\,e^{n+1}=0 & \Gamma_3\quad 0\le n\le N-1\\ e^0(.)=0 & \Omega\end{cases}$$

éq 3.3-7

One can then apply the second result of corollary 7 to him from where [éq 3.3-6] (one could, of course, just as easily to apply the gross profit of this corollary or that of property 6 from which it rises).

Note:

- *While placing itself within the particular framework [éq 3.1-9] of the article [bib6] with an implicit scheme* $\left(\theta=1\right)$ *and by taking again the equivalent standards [éq 3.1-10] one finds well the inequality (8) pp429. It is enough to make tend* $\Delta t\to 0$ *and to approximate the integral by the sum of Riemann which constitutes the second member of [éq 3.3-6].*
- *The existence and the unicity of the continuation* $\left(e^n\right)$ *rise of course from that of* $\left(u^n\right)$ *but one can also the redémontrer by applying the theorem of Lax-Milgram to the weak formulation rising from [éq 3.3-7].*

3.4 Total discretization in time and space

It is supposed that the field Ω is **polyhedric or not** and that it is discretized spatially by one **regular family** $\left(T_h\right)_h$ **triangulations**. Because of this regularity finite element method applied to $\left(P_2^{n+1}\right)$ converge when the largeest diameter of the elements K of $\left(T_h\right)_h$ tends towards zero

$$h:=\max_{K\in T_h} h_K\to 0 \qquad \text{éq 3.4-1}$$

Note:

- *Finite elements* $\left(K, P_K, \Sigma_K\right)$ *equivalents with same elements of reference are closely connected, they check relations of compatibility on their common borders and the constraints geometrical [éq 3.4-1] and [éq 3.4-2].*
- *It is pointed out that the diameter of K is reality* $h_K:=\max_{\left(\mathbf{x},\mathbf{y}\right)\in K^2}\left|\mathbf{x}-\mathbf{y}\right|$.

While noting ρ_K the roundness (one recalls that the roundness of K is reality $\rho_K := \max\{\text{diamètre des sphères} \subset K\}$) associated with K, finite elements of $(T_h)_h$ satisfy also the constraint

$$\exists \sigma > 0 / \ \frac{h_K}{\rho_K} \leq \sigma \qquad \text{éq 3.4-2}$$

In the usual triplet (K, P_K, Σ_K) one defines polynomial space as being that of the polynomials of total degree lower or equal to k on K

$$P_K := P_k(K) \qquad \text{éq 3.4-3}$$

and approximation spaces it (with the "weak" direction) associated

$$V_h := \{ v_h \in V / \forall K \in T_h \ \ v_{h_{|K}} \in P_k(K) \} \subset V \qquad \text{éq 3.4-4}$$

To conclude, one will note Π_h, the operator of projection which associates with the solution continues its V_h – interpolated

$$\begin{aligned} \Pi_h : V &\rightarrow V_h \\ v &\rightarrow v_h \end{aligned} \qquad \text{éq 3.4-5}$$

Note:

> *With a regular family of triangulations, this operator of interpolation is continuous and it can be written $\Pi_h v = \sum_i v(\mathbf{x_i}) N_i$ while noting $\mathbf{x_i}$ tops of the grid and N_i their function of associated form.*

He will be of a very particular importance when it is necessary to describe the increase which will exhume the indicator of error.

Note:

- ***In practice the grids are often polygonal, the approximation Ω_h of Ω becomes more rudimentary then than in the polyhedric case. To preserve the convergence of the method it is then necessary to resort to isoparametric elements** (cf [bib3] pp113-123 or P. GRISVARD. Behavior of the solutions of year elliptic boundary problem in has polygonal gold polyhedral domain. Numerical solution of PDE, ED. Academic Near, 1976).*
- *The indicator in residue was established in Code_Aster only for the isoparametric elements (triangle, quadrangle, face, tetrahedron, pentahedron and hexahedron). Moreover, as it is **simplexes** or of **parallélotopes** , **associated triangulation is regular** (cf [bib3] pp108-112).*
- *For the simplexes the relation [éq 3.4-2] results by the existence of a lower limit on the angles and, for the parallélotopes, in the existence of an upper limit controlling the relationship between the height, the width and the length.*
- *In the definition [éq 3.4-4] of V_h , they are the intrinsic relations of compatibility to the family of elements which assures us*

$$\forall h, K \ \ v_{h_{|K}} \in P_k(K) \subset H^1(K) \Rightarrow \ v_h \in H^1(\Omega := \cup \bar{K}) \qquad \textit{éq 3.4-6}$$

In the literature one often prefers the more regular definition to him

$$V_h^* := V_h \cap C^0(\Omega) \qquad \textbf{\textit{éq 3.4-7}}$$

By regaining the semi-discretized shape $\left(P_2^{n+1}\right)$ with functions tests in V_h one obtains the problem completely discretized in time and space (for one h fixed) according to:
The continuation is sought

$$\left(u_h^n\right)_{0\leq n\leq N}\in V_h \qquad \text{éq 3.4-8}$$

initialized by

$$u_h^0 := \Pi_h u_0 \qquad \text{éq 3.4-9}$$

checking the following problem

$$\left(P_2^{h,n+1}\right)\begin{cases}\text{Etant donnés } u_h^n, \hat{s}^n, \hat{s}^{n+1}, \hat{g}^n, \hat{g}^{n+1}, \hat{h}^n, \hat{h}^{n+1}, h^n, h^{n+1}\\ \text{Calculer } u_h^{n+1}\in V_h \text{ tel que}\\ \left(\rho\, C_p u_h^{n+1}, v_h\right)_{0,\Omega}+\Delta t\; a\left(u_{\theta,h}^{n+1}, v_h\right)=\left(\rho\, C_p u_h^n, v_h\right)_{0,\Omega}+\Delta t\left(b_\theta^{n+1}, v_h\right)\quad\left(\forall v_h\in V_h\right)\end{cases}$$

éq 3.4-10

Just as one supposed in the preceding paragraph as **temporal discretization of the loadings was exact**

$$e_\chi^n := \Xi^n - \Xi\left(n v\, t\right)=0 \text{ avec } \Xi\in\left\{\hat{s}, \hat{h}, h, \hat{g}\right\} \text{ et } 0\leq n\leq N \qquad (H_6)$$

, one supposes here moreover than them **space discretization is too**

$$\forall h\;\; \Xi_h^n := \Pi_h \Xi^n = \Xi^n \text{ avec } \Xi\in\left\{\hat{s}, \hat{h}, h, \hat{g}\right\} \text{ et } 0\leq n\leq N \qquad (H_7)$$

In ***Code_Aster*, these assumptions can not be checked** and it will be seen that they impact **quality of the indicator in residue** and its relations between equivalence and the exact error (cf [§4.3]). In practice, even if one is obliged to compose with this approximation, it is not truly problematic as long as the loadings "are not kicked up a rumpus too much" in time and space.

By applying the theorem of standard Lax-Milgram following the groundwork developed in the demonstration of theorem 3, one shows the existence and the unicity of the continuation $\left(u_h^n\right)_n$ in the closed sev (it is thus Hilbert, pre-necessary essential for the use of the famous theorem) V_h of Hilbert V. Moreover, while applying second result of corollary 7 (one could, of course, just as easily have applied the gross profit of this corollary or that of property 6 from which it rises), **"weak" controllability of the completely discretized problem** takes the following shape:

Property 9

While being based on the triangulation defined previously and by supposing that assumptions (H_6) and (H_7) are checked, one has increase

$$\begin{aligned}&\|\sqrt{\rho\, C_p}u_h^n\|_{0,\Omega}^2+\Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\,\nabla u_{\theta,h}^{m+1}\|_{0,\Omega}^2\leq\|\sqrt{\rho C_p}\,\Pi_h u_0\|_{0,\Omega}^2\\ &\forall\, 0\leq n\leq N \qquad +K_1\Delta t\sum_{m=0}^{n-1}\left(\|\hat{s}_\theta^{m+1}\|_{-1,\Omega}^2+\|\hat{g}_\theta^{m+1}\|_{-\frac{1}{2},\Gamma_2}^2+\|\hat{h}_\theta^{m+1}\|_{-\frac{1}{2},\Gamma_3}^2\right)\end{aligned} \qquad \textbf{éq 3.4-11}$$

while noting $u_{\theta,h}^{m+1} := \theta\; u_h^{m+1}+\left(1-\theta\right)u_h^m$.

Note:

- *While placing itself within the particular framework [éq 3.1-9] of the article [bib6] with an implicit scheme $(\theta=1)$ and by taking again the equivalent standards [éq 3.1-10] one finds well the inequality (14) pp430.*
- *By adopting the less restrictive approaches (H_4) and (H_5), one finds a version "strong" of this increase utilizing the standard H^1 field result.*

Now that we determined the functional framework ensuring us of the existence and the unicity of the continuation discrete solution and to study the evolution of the controllability of the problem during discretizations, we will pool these "ethereal" results to release increase a little where the indicator will intervene.

4 Indicator in pure residue

4.1 Notations

To build the site indicator of error one will require **following notations** :

- The whole of the faces (resp. nodes) of the element K is indicated by $S(K)$ (resp. $N(K)$).
- The whole of the nodes associated with one with its faces F (pertaining to $S(K)$) is noted $N(F)$.

Note:

To make simple, one will indicate under the term "face", with dimensions one of a finite element in 2D or one of his faces in 3D.

- The diameter of the element K (resp. of one of its faces F) is noted h_K (resp. h_F).
- The whole of the triangulation (T_h) breaks up in the form

$$T_h:=T_{h,\Omega}\cup T_{h,1}\cup T_{h,2}\cup T_{h,3}$$

while noting ($T_{h,i}$) the whole of the finite elements having a face contained in the border Γ_i .
- With same logic, the whole of the faces of the triangulation (T_h) breaks up in the form

$$S_h:=S_{h,\Omega}\cup S_{h,1}\cup S_{h,2}\cup S_{h,3}$$

with

$$\forall i\in\{1,2,3\} \quad S_{h,i}:=\{\partial K/K\in T_h \quad \partial K\subset\Gamma_i\}=\bigcup_{K\in T_{h,i}} S(K)$$

- In the same way, the whole of the nodes of the triangulation (□$_H$) breaks up in the form

$$N_h:=N_{h,\Omega}\cup N_{h,1}\cup N_{h,2}\cup N_{h,3}$$

- The function "bubble" associated with K (resp. F) is noted ψ_K (resp. ψ_F).

Note:

It is the function of $D(\Omega)$ (together of the indefinitely derivable functions and with compact support) resulting from the theorem of truncation on compact: its support is limited to compact in question (here K or F) and it is worth between 0 and 1 on sound interior (with the topological direction of the term). It is thus worthless on the border of compact and outside that - Ci.

- One notes P_F the operator of raising on K traces on F, built starting from an operator of raising fixed on the element of reference.
- Union of the finite elements of the triangulation dividing at least a face with K is noted
$$\Delta_K := \bigcup_{S(K)\cap S(K')\neq\emptyset} K'$$
- Union of the finite elements of the triangulation containing *F* in their border is noted
$$\Delta_F := \bigcup_{F\in S(K')} K'$$
- Union of the finite elements of the triangulation which share at least a node with K (resp. with F) is noted
$$\omega_K := \bigcup_{N(K)\cap N(K')\neq\emptyset} K' \quad \text{(resp. } \omega_F := \bigcup_{N(F)\cap N(K')\neq\emptyset} K' \text{).}$$

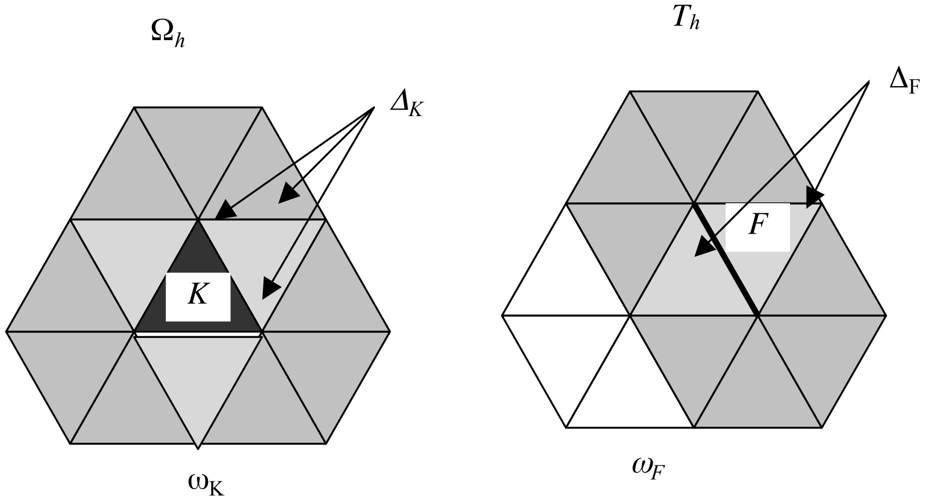


Figure 4.1-a: Designation of the types of vicinities for K and F .

4.2 Increase of the total space error

We will thus see how to obtain **a local indicator of calculable error starting from the data and discrete solution** $\left(u_h^n\right)_n$. As the discretized workspace is included in continuous space $V_h\subset V$, one can re-use [éq 3.2-3] with v_h. While withdrawing to him [éq 3.4-10] it occurs (with n and h fixed and while supposing (H_6) and (H_7))

$$\left(\rho\ C_p\left(u^{n+1}-u_h^{n+1}\right),\nu_h\right)_{0,\Omega}+\Delta t\ a\left(\left(u_\theta^{n+1}-u_{\theta,h}^{n+1}\right),\nu_h\right)=\left(\rho\ C_p\left(u^n-u_h^n\right),\nu_h\right)_{0,\Omega}\quad\left(\forall\nu_h\in V_h\right)\qquad \text{éq 4.2-1}$$

Note:

- *This relation states the orthogonal character of the space error with respect to the elements of V $_H$.*
- *It supposes in addition that the discretization is " **consistent** " i.e. there is not **additional errors introduced by the digital integration of the integrals** . In practice it is of course not the case!*

Let us consider the following linear form

$$A(v):=\left(\rho\, C_p\left(u^{n+1}-u_h^{n+1}\right),v\right)_{0,\Omega}+\Delta t\; a\left(u_\theta^{n+1}-u_{\theta,h}^{n+1},v\right)\quad(\forall v\in V)\qquad \text{éq 4.2-2}$$

who will be used to us as main idea during this demonstration. By packing it via [éq 4.2-1], one obtains

$$\begin{aligned}A(v)=&\left(\rho\, C_p\left(u^n-u_h^n\right),v\right)_{0,\Omega}+\left(\rho\, C_p\left(u^n-u_h^n\right),\left(v_h-v\right)\right)_{0,\Omega}+\\(\forall v\in V)\qquad&\left(\rho\, C_p\left(u^{n+1}-u_h^{n+1}\right),\left(v-v_h\right)\right)_{0,\Omega}+\Delta t\; a\left(\left(u_\theta^{n+1}-u_{\theta,h}^{n+1}\right),v-v_h\right)\end{aligned}\qquad \text{éq 4.2-3}$$

While taking [éq 3.2-3] after having replaced v_h by $v-v_h\in V$, one can build

$$\begin{aligned}&\left(\rho\, C_p\left(u^{n+1}-u_h^{n+1}-u^n+u_h^n\right),v-v_h\right)_{0,\Omega}+\Delta t\; a\left(u_\theta^{n+1}-u_{\theta,h}^{n+1},v-v_h\right)_{0,\Omega}=\\(\forall v\in V)\qquad&\Delta t\left(b_\theta^{n+1},v-v_h\right)_{0,\Omega}-\Delta t\; a\left(u_{\theta,h}^{n+1},v-v_h\right)-\left(\rho\, C_p\left(u_h^{n+1}-u_h^n\right),v-v_h\right)_{0,\Omega}\end{aligned}\qquad \text{éq 4.2-4}$$

Then $A(v)$ becomes

$$\begin{aligned}A(v)=&\left(\rho\, C_p\left(u^n-u_h^n\right),v\right)_{0,\Omega}+\Delta t\left(b_\theta^{n+1},v-v_h\right)-\\(\forall v\in V)\qquad&\left(\rho\, C_p\left(u_h^{n+1}-u_h^n\right),v-v_h\right)_{0,\Omega}-\Delta t\; a\left(u_{\theta,h}^{n+1},v-v_h\right)\end{aligned}\qquad \text{éq 4.2-5}$$

Then one breaks up the last three terms on each element K triangulation and one applies, with the last, the formula of Green

$$\begin{aligned}A(v)=&\left(\rho\, C_p\left(u^n-u_h^n\right),v\right)_{0,\Omega}+\Delta t\sum_{K\in T_h}\int_K\left(\hat{s}_\theta^{n+1}-\rho\, C_p\frac{u_h^{n+1}-u_h^n}{\Delta t}+\operatorname{div}\left(\lambda\,\nabla u_{\theta,h}^{n+1}\right)\right)\left(v-v_h\right)\,dx\\&-\frac{\Delta t}{2}\sum_{F\in S_{h,\Omega}}\int_F\left[\lambda\,\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\left(v-v_h\right)\,d\sigma\\\forall\; v-v_h\in V\qquad&+\Delta t\sum_{F\in S_{h,2}}\int_F\left(\hat{g}_\theta^{n+1}-\lambda\,\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right)\left(v-v_h\right)\,d\sigma\\&+\Delta t\sum_{F\in S_{h,3}}\int_F\left(\hat{h}_\theta^{n+1}-\lambda\,\frac{\partial u_{h,\theta}^{n+1}}{\partial n}-\left(hu_h\right)_\theta^{n+1}\right)\left(v-v_h\right)\,d\sigma\end{aligned}$$

éq 4.2-6

Note:

- *One allowed oneself to replace the hooks of duality of [éq 3.2-4] by integrals and one can apply the formula of Green because to the compact one K assumptions (H_4) and (H_5) are checked (while replacing Ω by K and Γ_i by $\partial K\cap\Gamma_i$). One thus has*

$$v-v_h\in H^1(K),\; u_h\in H^2(K),\; \hat{s}\in L^2(K),\; \hat{g}\in L^2\left(\partial K\cap\Gamma_2\right)\;\textbf{et}\;\hat{h}\in L^2\left(\partial K\cap\Gamma_3\right)\qquad \textit{éq 4.2-7}$$

Let us point out some properties of the operator Π_h of projection L^2- local introduces by P. CLEMENT [bib8]

$$\begin{aligned}\Pi_h : V \subset L^2(\Omega) \rightarrow V_h \\ \nu \rightarrow \nu_h\end{aligned}$$ éq 4.2-8

It checks in particular increases of errors of projection

$$\begin{aligned}&\forall \nu \in V \qquad \|\nu-\Pi_h \nu\|_{0,K} := \|\nu-\nu_h\|_{0,K} \leq C_4\, h_K \|\nu\|_{1,\omega_K} \\ &\forall K \in T_h,\ \forall F \in S(K) \qquad \|\nu-\Pi_h \nu\|_{0,F} := \|\nu-\nu_h\|_{0,F} \leq C_5 \sqrt{h_F} \|\nu\|_{1,\omega_F}\end{aligned}$$ éq 4.2-9

where constants C_4 and C_5 depend on the smallest angles of the triangulation. By taking this operator of space projection and by applying the inequality of Cauchy-Schwartz to [éq 4.2-6] it thus occurs:

$$\begin{aligned}A(\nu)-\left(\rho\, C_p\left(u^n-u_h^n\right),\nu\right)_{0,\Omega} \leq &\Delta t C_4 \sum_{K\in T_h} h_K \left\|\hat{s}_\theta^{n+1}-\rho\, C_p \frac{u_h^{n+1}-u_h^n}{\Delta t}+\mathrm{div}\left(\lambda \nabla u_{\theta,h}^{n+1}\right)\right\|_{0,K} \|\nu\|_{1,\omega_K} \\ &+\frac{\Delta t}{2} C_5 \sum_{F\in S_{h,\Omega}} \sqrt{h_F} \left\|\left[\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\right\|_{0,F} \|\nu\|_{1,\omega_F} \\ \forall \nu \in V \qquad &+\Delta t C_5 \sum_{F\in S_{h,2}} \sqrt{h_F} \left\|\hat{g}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right\|_{0,F} \|\nu\|_{1,\omega_F} \\ &+\Delta t C_5 \sum_{F\in S_{h,3}} \sqrt{h_F} \left\|\hat{h}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}-\left(h u_h\right)_\theta^{n+1}\right\|_{0,F} \|\nu\|_{1,\omega_F}\end{aligned}$$

éq 4.2-10

This inequality clearly lets show through a possible formulation of the indicator in pure residue:

Definition 10

Within the framework of the operator of transitory thermics linear of *Code_Aster*, **continuation** $\left(\eta^n(K)\right)_{0\leq n\leq N}^{K\in T_h}$ **local indicators theoretical** can be written in the form

$$\begin{aligned}\eta^{n+1}(K) := &h_K \left\|\hat{s}_\theta^{n+1}-\rho\, C_p \frac{u_h^{n+1}-u_h^n}{\Delta t}+div\left(\lambda \nabla u_{h,\theta}^{n+1}\right)\right\|_{0,K}+\frac{1}{2}\sum_{F\in S_\Omega(K)} \sqrt{h_F} \left\|\left[\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\right\|_{0,F}+ \\ &\sum_{F\in S_2(K)} \sqrt{h_F} \left\|\hat{g}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right\|_{0,F}+\sum_{F\in S_3(K)} \sqrt{h_F} \left\|\hat{h}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}-\left(h u_h\right)_\theta^{n+1}\right\|_{0,F}\end{aligned}$$

éq 4.2-11

It is initialized by

$$\begin{aligned}\eta^0(K) := &h_K \left\|\hat{s}^0+div\left(\lambda\ \nabla u_h^0\right)\right\|_{0,K}+\frac{1}{2}\sum_{F\in S_\Omega(K)} \sqrt{h_F} \left\|\left[\lambda \frac{\partial u_h^0}{\partial n}\right]\right\|_{0,F}+ \\ &\sum_{F\in S_2(K)} \sqrt{h_F} \left\|\hat{g}^0-\lambda \frac{\partial u_h^0}{\partial n}\right\|_{0,F}+\sum_{F\in S_3(K)} \sqrt{h_F} \left\|\hat{h}^0-\lambda \frac{\partial u_h^0}{\partial n}-\ h^0 u_h^0\right\|_{0,F}\end{aligned}$$

éq 4.2-12

continuation $\left(\eta^n(\Omega)\right)_{0\leq n\leq N}$ **total indicators theoretical** is defined as being

$$\forall\, 0\leq n\leq N \quad \eta^n(\Omega) := \left(\sum_{K\in T_h} \eta^n(K)^2\right)^{\frac{1}{2}}$$ **éq 4.2-13**

Note:

- *While placing itself within the particular framework [éq 3.1-9] of the article [bib6] with an implicit scheme ($\theta=1$) one finds well the definition (24) pp432.*
- *Whatever the initialization retained for thermal calculation, one starts the temporal continuation of cartography of indicators of error as if one were in hover: no the term of temporal finished difference, $n+1=0$ (in Code_Aster a transitory field of temperature is initialized with index 0) and $\theta=1$.*
- *It should be stressed that this indicator is composed of four terms: **term principal** , called **voluminal term of error** , controlling the good checking of the equation of heat, to which are added **three secondary terms** checking the good behaviour of the limiting conditions (**terms of jump, flow and exchange**). In `2D-PLAN` or in `3D` (resp. in `2D-AXI`), if the unit of the geometry is the meter, the unit of the first is it W.m (resp. $W.m.rad^{-1}$) and that of the other terms is it $W.m^{\frac{1}{2}}$ (resp. $W.m^{\frac{1}{2}}.rad^{-1}$). Attention thus with the units taken into account for the geometry when one is interested in the gross amount of the indicator and not in his relative value!*
- *While taking as a starting point the increases developed by R. VERFURTH (cf [bib7] pp84-94) for the Poisson's equation one could have taken as indicator the root of the sum of the squares of the terms quoted above.*

$$\tilde{\eta}^{n+1}(K):=\left\{\begin{array}{l} h_K^2 \|\hat{s}_\theta^{n+1}-\rho C_p \frac{u_h^{n+1}-u_h^n}{\Delta t}+div\left(\lambda \nabla u_{h,\theta}^{n+1}\right)\|_{0,K}^2+\frac{1}{2}\sum_{F\in S_\Omega(K)} h_F \|\left[\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\|_{0,F}^2+ \\ \sum_{F\in S_2(K)} h_F \|\hat{g}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\|_{0,F}^2+\sum_{F\in S_3(K)} h_F \|\hat{h}_\theta^{n+1}-\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}-\left(hu_h\right)_\theta^{n+1}\|_{0,F}^2 \end{array}\right\}^{\frac{1}{2}}$$

éq 4.2-14

This definition leads to an increase of the total error which is more optimal than that which will be released thereafter. But we preferred, to remain homogeneous with the writings of B. METIVET [bib6] and with the estimator in linear mechanics already installation in the code, to hold us with the version of definition 10.

While resting on [éq 4.2-10] and definition 10 one can then exhume the increase of the following total error:

Property 11

Under the assumptions of properties 6, of (H_6) and by using definition 10, one has, **at the total level**, **"weak" increase of the error** (with $K_2\left(\|\lambda\|_{\infty,\Omega}, P(\Omega), C_4, C_5\right)>0$) **via the history of the indicators**

$$\begin{aligned}&\|\sqrt{\rho\, C_p}\left(u^n - u_h^n\right)\|_{0,\Omega}^2 + 4(1-\theta)\sum_{m=0}^{n-1}\|\sqrt{\rho\, C_p}\left(u^m - u_h^m\right)\|_{0,\Omega}^2 + \Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\,\nabla\left(u_\theta^{m+1} - u_{\theta,h}^{m+1}\right)\|_{0,\Omega}^2\\ \forall\, 0\le n\le N \quad &\le (4\theta-3)\|\sqrt{\rho\, C_p}\left(u_0 - u_h^0\right)\|_{0,\Omega}^2 + K_2\Delta t\sum_{m=0}^{n}\left(\eta^m(\Omega)\right)^2\end{aligned}$$

éq 4.2-15

or more simply

$$\begin{aligned}&\|\sqrt{\rho\, C_p}\left(u^n - u_h^n\right)\|_{0,\Omega}^2 + \Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\,\nabla\left(u_\theta^{m+1} - u_{\theta,h}^{m+1}\right)\|_{0,\Omega}^2\\ \forall\, 0\le n\le N \quad &\le \|\sqrt{\rho\, C_p}\left(u_0 - u_h^0\right)\|_{0,\Omega}^2 + K_2\Delta t\sum_{m=0}^{n}\left(\eta^m(\Omega)\right)^2\end{aligned}$$

éq 4.2-16

Proof:

The estimates [éq 4.2-15] [éq 4.2-16] are obtained by reiterating the same process as for properties 5.6 and 7. One takes in [éq 4.2-10] the particular function test

$$\nu := u_\theta^{n+1} - u_{\theta,h}^{n+1} \quad \text{éq 4.2-17}$$

One ousts the term of exchange by the usual argument

$$\int_{\Gamma_3}\left(h\left(u-u_h\right)\right)_\theta^{n+1}\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}\right)dx \; > 0 \quad \text{éq 4.2-18}$$

It is necessary to apply the trick [éq 3.1-4] to the cross term $(2\theta-1)\int_\Omega \rho\, C_p\left(u^{n+1} - u_h^{n+1}\right)\left(u^n - u_h^n\right)dx$ and on the product utilizing the indicator. One has then to find the parameters α and β checking a system of the type [éq 3.2-8]

$$\begin{aligned}&|2 - \frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\alpha^2 = 1\\ &|2\theta - \beta^2|1-2\theta| = 1\end{aligned} \quad \text{éq 4.2-19}$$

who admits solution only if the diagram is unconditionally stable ($\theta \ge \frac{1}{2}$). From where increase [éq 4.2-15] [éq 4.2-16] while taking

$$K_2 = \frac{P^2(\Omega)}{\|\lambda\|_{\infty,\Omega}}\max\left(C_4^2, C_5^2\right) \quad \text{éq 4.2-20}$$

The inequality [éq 4.2-16] more "coarse" results from the same sales leaflet as for corollary 7.

Note:

- *While placing itself within the particular framework [éq 3.1-9] of the article [bib6] with an implicit scheme ($\theta=1$) one finds well the inequality (25) pp432 (with $c=max(1,K2)$).*
- *By adopting the less restrictive approaches (H_4) and (H_5), one finds a version "strong" of this property.*
- *This property can be shown more quickly while noticing than the inequation [éq 4.2-10] is similar to the equation of the problem semi-discretized in time [éq 3.2-3]: except for the inequality, while changing u by $u-u_h$ and while taking as term $\left(b_\theta^n, v\right)$ the second member of [éq 4.2-10]. One can then directly apply the corollary 7 to him who is during required estimate!*
- ***Of [éq 4.2-15] [éq 4.2-16] it appears that, at one moment given, the error on the approximation of the condition of Cauchy and the history of the total indicators intervene on the total quality of the solution obtained. One will be able to thus minimize overall the error of approximation due to the finite elements in the course of time while re-meshing "advisedly", via the continuation of indicators, the structure. Because, in practice, one realizes that the refinement of the meshs makes it possible to decrease their error and thus cause a drop in the temporal sum of the indicators. The total error will butt (and it is moral) against the value floor of the error of approximation of the initial condition (which will tend it-also to drop of course!). The indicator "over-estimates" the space error overall.***
- *With the other alternative of indicator [éq 4.2-14] one finds the same type of increase. However the constant K_2 change. It is multiplied by the constant C_6 checking (cf [bib7] pp90)*

$$\sum_{K\in T_h} \|v\|^2_{1,\omega_K} + \sum_{F\in S_h} \|v\|^2_{1,\omega_F} \leq C_6 \|v\|^2_{1,\Omega} \qquad \text{éq 4.2-21}$$

$$\tilde{K}_2 := C_6 K_2 \qquad \text{éq 4.2-22}$$

According to the definitions [éq 2-1-8], [éq 2-1-10] with [éq 2-1-13] if the taking into account of the limiting conditions of Dirichlet (generalized or not), via the ddls of Lagrange, is exact (what is the case in *Code_Aster*)

$$\forall h \quad Rf_h^n := \Pi_h Rf^n = Rf^n = Rf\left(n\Delta t\right) \quad 0\leq n\leq N \qquad (H_8)$$

the preceding property produces the following corollary then:

Corollary 11bis

Under the assumptions of property 11 by supposing (H_8), there is the increase of **the total space error expressed in temperature**

$$\begin{aligned} &\|\sqrt{\rho C_p}\left(T^n - T_h^n\right)\|^2_{0,\Omega} + \Delta t \sum_{m=0}^{n-1} \|\sqrt{\lambda}\nabla\left(T_\theta^{m+1} - T_{\theta,h}^{m+1}\right)\|^2_{0,\Omega} \\ \forall 0\leq n\leq N \qquad &\leq \|\sqrt{\rho C_p}\left(T_0 - T_h^0\right)\|^2_{0,\Omega} + K_2 \Delta t \sum_{m=0}^{n} \left(\eta^m(\Omega)\right)^2 \end{aligned} \qquad \textbf{éq 4.2-23}$$

by using definition 10 of the indicator also expressed in temperature

$$u\Rightarrow T,\ \hat{s}\Rightarrow s,\ \hat{g}\Rightarrow g \text{ et } \hat{h}\Rightarrow hT_{ext} \qquad \textbf{éq 4.2-24}$$

4.3 Various types of possible indicators

By extrapolating a remark of [bib5] (pp194-195) it appears that increases of property 11 can be exhumed while taking as indicator

$$\eta_{p,t}^{n+1}(K):=h_K^r\|\hat{s}_\theta^{n+1}-\rho\ C_p\frac{u_h^{n+1}-u_h^n}{\Delta t}+div(\lambda\ \nabla u_{h,\theta}^{n+1})\|_{L^p(K)}+\frac{1}{2}\sum_{F\in S_\Omega(K)}h_F^s\left\|\left[\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\right\|_{L^t(K)}+\sum_{F\in S_2(K)}h_F^s\|\hat{g}_\theta^{n+1}-\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\|_{L^t(K)}+\sum_{F\in S_3(K)}h_F^s\|\hat{h}_\theta^{n+1}-\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}-(hu_h)_\theta^{n+1}\|_{L^t(K)}$$ éq 4.3-1

where constants r and s are worth

$$t\geq 1,\ \ p>1\ (2D\ \ q=2)\ \text{ou}\ \ p\geq\frac{6}{5}\ (3D\ \ q=3)$$

$$\begin{cases} r(q,p):=q+1-\frac{q}{2}-\frac{q}{p} \\ s(q,t):=q-\frac{1}{2}-\frac{q-1}{2}-\frac{q-1}{t} \end{cases}$$ éq 4.3-2

Note:

> *Just to introduce this generic shape of indicators, one passes from the notation hilbertienne of the standards of spaces to the notation of Lebesgue*

It is parameterized by the types of standards voluminal and surface which intervene for its obtaining. Contrary to the indicator which we chose ($\eta_{2,2}^{n+1}(K)$ who corresponds to $p=t=2$), some use the voluminal standard L^1 ($p=t=2$) or on the contrary the infinite standard.

This last formulation, just like its simplified form of definition 10 (or [éq 4.2-14]), **constitute an indicator of error well a posteriori** because its calculation requires only the knowledge of materials, the loadings, the geometrical data, of θ and of the approximate field solution u_h accused thermal problem. **However the exact estimate of the indicator is not always possible when one has complicated loadings. Two approaches are then possible:**

- That is to say one **approximate the integrals** who return in the composition of definition 10 by one **formula of squaring.**
- That is to say **the loadings are approximated** by a linear combination of simpler functions which will be able to allow an exact integration. Generally one uses same architecture as that which was installation for the finite elements modelling the field of temperature.

Note:

> - *In both cases the loadings are "prisoners of the selected vision finite elements" to model the field solution.*

These two strategies are equivalent and in *Code_Aster* it is the first which was retained : the voluminal integral is calculated by a formula of Gauss, those surface by a formula of Newton-Dimensions.
Both introduce one **skew in the calculation of the estimator** who can be represented by introducing the approximate versions of the loadings and the source (into the initial problem in T and in the problem transformed into u)

$$s_{\theta,h}^{n+1},\ g_{\theta,h}^{n+1},\ T_{ext,\theta,h}^{n+1} \text{ et } h_{\theta,h}^{n+1} \qquad \text{éq 4.3-3}$$

$$\hat{s}_{\theta,h}^{n+1},\ \hat{g}_{\theta,h}^{n+1},\ \hat{h}_{\theta,h}^{n+1} \text{ et } h_{\theta,h}^{n+1} \qquad \text{éq 4.3-4}$$

in spaces of voluminal approximation (for the source) and surface (for the loadings)

$$\begin{aligned} X_h(\Omega) &:= \left\{ \nu_h \in L^2(\Omega) / \forall K \in T_h \ \ \nu_{h_{|K}} \in P_{l_1}(K) \right\} \\ X_h(\Gamma_i) &:= \left\{ \nu_h \in L^2(\Gamma_i) / \forall F \in S_{h,i} \ \ \nu_{h_{|F\cap\Gamma_i}} \in P_{l_i}(F \cap \Gamma_i) \right\} \end{aligned} \qquad \text{éq 4.3-5}$$

In fact, one introduces **two types of digital errors during the calculation of the indicator** : that inherent in **formulas of squaring** (for polynomial loadings of a high nature) and that due to **voluminal term**. Indeed, this last requires a double derivation which one carries out in three stages because in *Code_Aster* one does not recommend the use of the derivative second of the functions of forms.

Note:

> *They were recently introduced to treat the derivation of the rate of refund of energy (cf [R7.02.01 § Annexe 1]).*

On the one hand, one calculates (in the thermal operator) the heat flux at the points of gauss, then one extrapolates the values with the nodes corresponding by smoothing local (cf [R3.06.03] `CALC_CHAMP` with `THERMICS='FLUX_ELNO'`) before calculating the divergence of the vectors flow at the points of Gauss. With finite elements quadratic the intermediate operation is only approximate (one assigns like value to the median nodes the half the sum of their values to the extreme nodes). However digital tests (limited) showed that, even in P_2 , this approach does not provide results very different from those obtained by a direct calculation via good the derivative second.

Note:

- *Indices l_1 , l_2 , l_3 of these polynomial spaces can be unspecified and different from that of the approximate solution: k . However, to prevent that these terms do not become prevalent (**it is a question of rather estimating the error on the solution than that on the modeling of the loadings**) one will tend to take $l_i \geq k-2 \quad (i=1,2,3)$.*

Definition 10 and the weak estimate 11 partner are rewritten then in the following form. This new definition, $\eta_R^{n+1}(K)$, is subscripted by one R (one takes again in that the usual notations of [bib6] and [bib7]) (for "reality") in order to notify well that it corresponds better to the values which are calculated indeed in the code.

Definition 12

Within the framework of the operator of transitory thermics linear of *Code_Aster*, **continuation** $\left(\eta_R^n(K)\right)_{0\le n\le N}^{K\in T_h}$ **local indicators realities** can be written in the form

$$\eta_R^{n+1}(K):=h_K\left\|\hat{s}_{\theta,h}^{n+1}-\rho\, C_p\frac{u_h^{n+1}-u_h^n}{\Delta t}+div\left(\lambda\,\nabla u_{h,\theta}^{n+1}\right)\right\|_{0,K}+\frac{1}{2}\sum_{F\in S_\Omega(K)}\sqrt{h_F}\left\|\left[\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right]\right\|_{0,F}+ \sum_{F\in S_2(K)}\sqrt{h_F}\left\|\hat{g}_{\theta,h}^{n+1}-\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right\|_{0,F}+\sum_{F\in S_3(K)}\sqrt{h_F}\left\|\hat{h}_{\theta,h}^{n+1}-\lambda\frac{\partial u_{h,\theta}^{n+1}}{\partial n}-\left(h_h u_h\right)_\theta^{n+1}\right\|_{0,F}$$

éq 4.3-6

It is initialized by

$$\eta_R^0(K):=h_K\left\|\hat{s}_h^0+div\left(\lambda\,\nabla u_h^0\right)\right\|_{0,K}+\frac{1}{2}\sum_{F\in S_\Omega(K)}\sqrt{h_F}\left\|\left[\lambda\frac{\partial u_h^0}{\partial n}\right]\right\|_{0,F}+ \sum_{F\in S_2(K)}\sqrt{h_F}\left\|\hat{g}_h^0-\lambda\frac{\partial u_h^0}{\partial n}\right\|_{0,F}+\sum_{F\in S_3(K)}\sqrt{h_F}\left\|\hat{h}_h^0-\lambda\frac{\partial u_h^0}{\partial n}-h_h^0u_h^0\right\|_{0,F}$$

éq 4.3-7

continuation $\left(\eta^n(\Omega)\right)_{0\le n\le N}$ **total indicators realities** is defined as being

$$\forall\, 0\le n\le N\quad \eta_R^n(\Omega):=\left(\sum_{K\in T_h}\eta_R^n(K)^2\right)^{\frac{1}{2}}$$

éq 4.3-8

Note:

- *One can pass the same remarks as for his "theoretical" alter ego. They is also declined according to the formulations [éq 4.2-14]* $\tilde{\eta}_R^n(K)$ *and [éq 4.3-1], [éq 4.3-2]* $\eta_{R,p,t}^n(K)$ *.*

While being based on the results of property 11, definition 12 and the triangular inequality one can then exhume the increase of the following real total error (one began again that the simplified version):

Property 13

Under the assumptions of properties 6, (H_6) and by using definition 12, one has, at the total level, **"weak" increase of the error** (with $K_2\left(\|\lambda\|_{\infty,\Omega}, P(\Omega), C_4, C_5\right)>0$) via **the history of the real indicators**

$$\begin{aligned}&\|\sqrt{\rho\, C_p}\left(u^n-u_h^n\right)\|_{0,\Omega}^2+\Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\,\nabla\left(u_\theta^{m+1}-u_{\theta,h}^{m+1}\right)\|_{0,\Omega}^2\\ \forall\, 0\le n\le N\quad &\le\|\sqrt{\rho\, C_p}\left(u_0-u_h^0\right)\|_{0,\Omega}^2+K_2\Delta t\sum_{K\in T_h}\left(\eta_R^0(K)\right)^2+\sum_{m=0}^{n-1}\left\{\left(\eta_R^{m+1}(K)\right)^2+h_K^2\|\hat{s}_{\theta,h}^{m+1}-\hat{s}_\theta^{m+1}\|_{0,K}^2\right\}+\\ &K_2\Delta t\sum_{K\in T_h}\sum_{m=0}^{n-1}\left\{\sum_{F\in S_2(K)}h_F\|\hat{g}_{\theta,h}^{m+1}-\hat{g}_\theta^{m+1}\|_{0,F}^2+\sum_{F\in S_3(K)}h_F\|\hat{h}_{\theta,h}^{m+1}-\hat{h}_\theta^{m+1}-\left(h_h u_h\right)_\theta^{m+1}+\left(hu_h\right)_\theta^{m+1}\|_{0,F}^2\right\}\end{aligned}$$

éq 4.3-9

Under (H_8), there is the same expression in **temperature**

$$\begin{aligned}&\|\sqrt{\rho\, C_p}\left(T^n-T_h^n\right)\|_{0,\Omega}^2+\Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\,\nabla\left(T_\theta^{m+1}-T_{\theta,h}^{m+1}\right)\|_{0,\Omega}^2\\ \forall\, 0\le n\le N\quad &\le\|\sqrt{\rho\, C_p}\left(T_0-T_h^0\right)\|_{0,\Omega}^2+K_2\Delta t\sum_{K\in T_h}\left(\eta_R^0(K)\right)^2+\sum_{m=0}^{n-1}\left\{\left(\eta_R^{m+1}(K)\right)^2+h_K^2\|s_{\theta,h}^{m+1}-s_\theta^{m+1}\|_{0,K}^2\right\}+\\ &K_2\Delta t\sum_{K\in T_h}\sum_{m=0}^{n-1}\left\{\sum_{F\in S_2(K)}h_F\|g_{\theta,h}^{m+1}-g_\theta^{m+1}\|_{0,F}^2+\sum_{F\in S_3(K)}h_F\|\left(h_h\left(T_{ext,h}-T_h\right)\right)_\theta^{m+1}-\left(h\left(T_{ext}-T_h\right)\right)_\theta^{m+1}\|_{0,F}^2\right\}\end{aligned}$$

éq 4.3-10

by using definition 12 of **the indicator also expressed in temperature**

$$u\Rightarrow T,\ \hat{s}\Rightarrow s,\ \hat{g}\Rightarrow g\ \text{et}\ \hat{h}\Rightarrow hT_{ext}$$ **éq 4.3-11**

Note:

- *As for the theoretical value there is a morals with the history because, when one will refine, the total error will butt against the value floor due to the approximations of the initial condition, the limiting conditions and the source. One cannot get results of better quality that the data input of the problem!*

4.4 Decrease of the local space error

Before exhuming the decrease of the space error, one will have to introduce some complementary results:

Lemma 14

It is shown that there exist strictly positive constants C_l (l=6… 11) checking

$$\begin{aligned}
&\forall v \in P_{\sup\{k,l_1,l_2,l_3\}}(K) \quad C_6 \|\psi_K v\|_{0,K} \leq \|v\|_{0,K} \leq C_7 \|\psi_{K^{\frac{1}{2}}} v\|_{0,K} \\
&\qquad \|\nabla \psi_K v\|_{0,K} \leq C_8 h_K^{-1} \|\psi_K v\|_{0,K} \\
&\forall v \in P_{\sup\{k,l_1,l_2,l_3\}}(F) \quad C_9 h_F^{-\frac{1}{2}} \|\psi_K P_F v\|_{0,\Delta_F} \leq \|v\|_{0,F} \leq C_{10} \|\psi_{F^{\frac{1}{2}}} v\|_{0,F} \\
&\qquad \|\nabla \psi_K v\|_{0,\Delta_F} \leq C_{11} h_F^{-1} \|\psi_K v\|_{0,\Delta_F}
\end{aligned} \qquad \text{éq 4.4-1}$$

Proof:

One passes to the element of reference then one uses the fact that the standards are equivalent on polynomial spaces considered, because they are of finished size (cf [bib5] pp196-98, [bib7] [§1]).

These preliminary results are crucial to determine a decrease of the site error by the real indicator. But one will see that one will be able to obtain only one opposite room of [éq 4.3-9], [éq 4.3-10].

Property 15

Under the assumptions of property 6, of (H_6) and while being based on definition 12 and the lemma 14, one has, **at the local level**, **the "weak" decrease of the error** (with $K_3(C_i\,, i=6\cdots 11)>0$) **via the real indicator**

$$\eta_R^{n+1}(K)\le K_3\left\{\begin{array}{r} h_K\|\sqrt{\rho\ C_p}\dfrac{u^{n+1}-u_h^{n+1}-u^n-u_h^n}{\Delta t}\|_{0,\Delta_K}+\|\sqrt{\lambda}\,\nabla\left(u_\theta^{n+1}-u_{\theta,h}^{n+1}\right)\|_{0,\Delta_K}+\\ h_K\|\hat{s}_\theta^{n+1}-\hat{s}_{\theta,h}^{n+1}\|_{0,\Delta_K}+h_F^{\frac{1}{2}}\|\hat{g}_\theta^{n+1}-\hat{g}_{\theta,h}^{n+1}\|_{0,\Delta_K\cap\Gamma_2}+\\ h_F^{\frac{1}{2}}\|\hat{h}_\theta^{n+1}-\hat{h}_{\theta,h}^{n+1}-\left(hu_h\right)_\theta^{n+1}+\left(h_h u_h\right)_\theta^{n+1}\|_{0,\Delta_K\cap\Gamma_3}\end{array}\right\}$$

$$\forall\, 0\le n\le N-1$$

éq 4.4-2

Under (H_8), there is the same expression in **temperature**

$$\eta_R^{n+1}(K)\le K_3\left\{\begin{array}{r} h_K\|\sqrt{\rho\ C_p}\dfrac{T^{n+1}-T_h^{n+1}-T^n-T_h^n}{\Delta t}\|_{0,\ \Delta_K}+\|\sqrt{\lambda}\,\nabla\left(T_\theta^{n+1}-T_{\theta,h}^{n+1}\right)\|_{0,\Delta_K}+\\ h_K\|s_\theta^{n+1}-s_{\theta,h}^{n+1}\|_{0,\Delta_K}+h_F^{\frac{1}{2}}\|g_\theta^{n+1}-g_{\theta,h}^{n+1}\|_{0,\Delta_K\cap\Gamma_2}+\\ h_F^{\frac{1}{2}}\|h_\theta^{n+1}\left(T_{ext,\theta}^{n+1}-T_\theta^{n+1}\right)-h_{\theta,h}^{n+1}\left(T_{ext,\theta,h}^{n+1}-T_{\theta,h}^{n+1}\right)\|_{0,\Delta_K\cap\Gamma_3}\end{array}\right\}$$

$$\forall\, 0\le n\le N-1$$

éq 4.4-3

by using definition 12 of **the indicator also expressed in temperature**

$$u\Rightarrow T,\ \hat{s}\Rightarrow s,\ \hat{g}\Rightarrow g\ \text{et}\ \hat{h}\Rightarrow hT_{ext}$$ **éq 4.4-4**

Proof:

This a little technical demonstration comprises three stages which will consist in successively raising each term of the indicator [éq 4.3-6] (by using the inequalities of the property 14) and to gather increases obtained:

Firstly, one will replace in the equation [éq 4.2-6] the term in $\nu-\nu_h$ by the product w_K utilizing the function "bubble" of K

$$\forall K\in T_h\quad \nu_K:=\hat{s}_{\theta,h}^{n+1}-\rho\ C_p\frac{u_h^{n+1}-u_h^n}{\Delta t}+div\left(\lambda\ \nabla u_{\theta,h}^{n+1}\right)$$

$$w_K:=\psi_K\nu_K$$ éq 4.4-5

From where succession of increases, via [éq 4.4-1] and the inequality of Cauchy-Schwartz,

$$\|\nu_K\|_{0,K}^2 \le C_7^2 \int_K w_K \nu_K \, dx \le C_7^2 \left\{ \left(\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}, w_K \right)_{0,\Omega} + a\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}, w_K\right) - \left(\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}, w_K\right) \right\}$$

$$\le C_7^2 \max\left(1, C_8\right) \left\{ \left\|\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}\right\|_{0,K} + h_K^{-1} \|\sqrt{\lambda}\,\nabla\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}\right)\|_{0,K} + \|\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}\|_{0,K} \right\} \|w_K\|_{0,K}$$

$$\Rightarrow \|\nu_K\|_{0,K} \le \frac{C_7^2}{C_6} \max\left(1, C_8\right) \left\{ \left\|\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}\right\|_{0,K} + \|\sqrt{\lambda}\,\nabla\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}\right)\|_{0,K} + \|\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}\|_{0,K} \right\}$$

éq 4.4-6

Then, one reiterates the same process for the surface terms $w_{F,i}$

$$\forall F \in S(K) \cap S_{h,\Omega} \quad \nu_{F,1} := \left[\lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n}\right] \qquad \text{éq 4.4-7}$$

$$w_{F,1} := \psi_K P_F \nu_{K,1}$$

$$\forall F \in S(K) \cap S_{h,2} \quad \nu_{F,2} := \hat{g}_{h,\theta}^{n+1} - \lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n} \qquad \text{éq 4.4-8}$$

$$w_{F,2} := \psi_F P_F \nu_{F,2}$$

$$\forall F \in S(K) \cap S_{h,3} \quad \nu_{F,3} := \hat{h}_{h,\theta}^{n+1} - \lambda \frac{\partial u_{h,\theta}^{n+1}}{\partial n} - \left(h_h u_h\right)_\theta^{n+1} \qquad \text{éq 4.4-9}$$

$$w_{F,3} := \psi_F P_F \nu_{F,3}$$

Maybe, for example, for $i=1$ succession of increases, via [éq 4.4-1] and the inequality of Cauchy - Schwartz,

$$\|\nu_{F,1}\|_{0,F}^2 \le C_{10}^2 \int_F w_{F,1} \nu_{F,1} \, d\sigma \le C_{10}^2 \left\{ \begin{array}{l} \left(\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}, w_{F,1} \right)_{0,\Omega} + a\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}, w_{F,1}\right) \\ \quad - \left(\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}, w_{F,1}\right)_{0,\Omega} - \left(\nu_K, w_{F,1}\right)_{0,\Omega} \end{array} \right\}$$

$$\le C_{10}^2 \max\left(1, C_{11}\right) \left\{ \begin{array}{l} \left\|\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}\right\|_{0,\Delta_F} + h_F^{-1} \|\sqrt{\lambda}\,\nabla\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}\right)\|_{0,\Delta_F} \\ \quad + \|\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}\|_{0,\Delta_F} + \|\nu_K\|_{0,\Delta_F} \end{array} \right\} \|w_{F,1}\|_{0,\Delta_F}$$

$$\Rightarrow \|\nu_{F,1}\|_{0,F} \le \frac{C_{10}^2}{C_9} \max\left(1, C_{11}\right) \left\{ \begin{array}{l} h_F^{\frac{1}{2}} \left\|\frac{u^{n+1} - u_h^{n+1} - u^n - u_h^n}{\Delta t}\right\|_{0,\Delta_F} + h_F^{-\frac{1}{2}} \|\sqrt{\lambda}\,\nabla\left(u_\theta^{n+1} - u_{\theta,h}^{n+1}\right)\|_{0,\Delta_F} \\ \quad + h_F^{\frac{1}{2}} \|\hat{s}_\theta^{n+1} - \hat{s}_{\theta,h}^{n+1}\|_{0,\Delta_F} + h_F^{\frac{1}{2}} \|\nu_K\|_{0,\Delta_F} \end{array} \right\}$$

éq 4.4-10

Finally it is enough to carry out the linear combination implying [éq 4.4-9] and [éq 4.4-10] to conclude (because $h_F \le h_K$ et $\forall \nu \; \|\nu\|_{0,\Delta_F} \le \|\nu\|_{0,\Delta_K}$ avec $F \in S(K)$).

Note:

- *This local decrease of the error is also declined according to the formulations [éq 4.2-14] $\tilde{\eta}_R^n(K)$ and [éq 4.3-1], [éq 4.3-2] $\eta_{R,p,t}^n(K)$.*
- *While placing itself within the particular framework [éq 3.1-9] of the article [bib6] with an implicit scheme ($\theta=1$) one finds well the inequality (29) pp432.*
- *By adopting the less restrictive approaches (H_4) and (H_5), one finds a version "strong" of this property.*
- ***This result provides only one opposite room of total increase [éq 4.3-9], [éq 4.3-10] but within the framework of this kind of indicator one will not be able to obtain better compromise. These estimates are optimal within the meaning of [bib5]** . They show the equivalence of the sum hilbertienne indicators with the space part of the total exact error. The constants of equivalence are independent of the parameters of discretizations in space and in time, they depend only on the smallest angle of the triangulation.*
- ***This increase of the real indicator of error shows, which if one very locally refines (around K) in order to decrease $\eta_R^n(K)$, one is not ensured of a reduction in the error in an immediate vicinity of the zone concerned (in Δ_K). The indicator "underestimates" the space error locally and only a more macroscopic refinement carries out theoretically a reduction in the error (cf property 13).***

4.5 Complements

The constant K_3 just like its preceding alter ego, K_2, **depends intrinsically on the type of limiting conditions enriching the equation by initial heat as well as type of temporal and space discretization**. To try to free itself from this last constraint, SR. GAGO [bib10] proposes (on a problem 2D models) a dependence of the constant K_2 according to **type of finite elements used**. She is written

$$K_2 := \frac{\tilde{K}_2}{\sqrt{24\,p^2}} \qquad \text{éq 4.5-1}$$

where p is the degree of the polynomial of interpolation used ($p=1$ for TRIA3 and QUAD4, $p=2$ for TRIA6 and QUAD8/9). From where the idea, once the indicator of total error calculated, to multiply it by this "corrective" constant $\frac{1}{\sqrt{24\,p^2}}$. This strategy was implicitly retained for the calculation of the indicator of error in mechanics (option 'ERME_ELEM' of CALC_ERREUR, cf [R4.10.02 §3]). We however did not adopt it for thermics because this constant was given only empirically on the equation of Laplace 2D. We do not want to thus skew the values of the indicators.

It was question, until now, only of cards of indicators of space errors calculated at a given moment of the transient of calculation. But, in fact, **there exist several ways for to build indicators of error on a parabolic problem** :

- one can very well, first of all, semi-to discretize the strong formulation space some and to control its space error by indicator of error adapted a posteriori to the stationary case (in our elliptic case). Then one applies a solvor, of step and order variables, treating the ordinary differential equations (for example [bib10] [bib11] [bib12]),
- a second strategy consists in semi-discretizing in time then in space and determining the indicator of one moment space error given (for example [bib4] [bib6] [bib13]) starting from the local residues of the semi-discretized form. One applies a linear solvor to the variational form allowing to repeatedly build the solution at one moment given starting from the solution of the previous moment,
- another possibility consists in discretizing simultaneously in time and space on suitable finite elements and controlling their "space-time" errors in a coupled way (for example [bib14] [bib15]).

It **last scenario is most tempting** from a theoretical point of view because he proposes a complete control of the error and he makes it possible to avoid unfortunate decouplings as for the possible refinements/déraffinements controlled by a criterion with respect to the other (cf following paragraph). He is however very heavy to set up in a large industrial code such as *Code_Aster*. It supposes indeed, to be optimal, nothing less than one separate management step of time by finite elements. What from the point of view of architecture supporting the finite elements of the code is a true challenge!
The second scenario is thus preferred to him who has the large advantage of being able to be established directly in a code D finite elements because this it is based above all on the resolution of the completely discretized system. These is the indicators which was set up in N3S, TRIFOU and *Code_Aster*.

Within the framework of one **true "space-time" discretization of the problem** (scenario 3), in any rigour, one is obtained **"space-time" indicator for each element of discretization** $K\times\left[t_n, t_{n+1}\right]$ who is the balanced sum of three terms:

1) the residue of the calculated solution and the data discretized compared to the strong formulation of the problem $\left(P_0\right)$ evaluated on $K\times\left[t_n, t_{n+1}\right]$,
2) the jump space through $\partial K\times\left[t_n, t_{n+1}\right]$ of the operator traces associated (who naturally connects the formulations weak and strong via the formula of Green),
3) the temporal jump through $K\times\partial\left[t_n, t_{n+1}\right]$ calculated solution.

The solution which was installation does not make it possible obviously to reveal explicitly it **term of temporal jump**. **It re-appears** however **implicitly**, because of method of particular temporal semi-discretization, in **all terms in** θ **definitions 10 and 12**.

On the other hand, the fact of being interested mainly only in the space discretization and its possible refinement/déraffinement should not occult certain contigences with respect to the management of the step of time. Indeed, at the time of **transitory calculations comprising of abrupt variations of loadings and/or sources in the course of time**, for example of the thermal shocks, the fields of calculated temperatures $T^n\left(0<n\leq N\right)$ can **to oscillate spatially and temporally**. Moreover, they can violate it " **principle of the maximum** " by taking values apart from the terminals imposed by the condition of Cauchy and the conditions limiting. To overcome this digital phenomenon one parasitizes shows, on a canonical case without condition of exchange (cf [R3.06.07 §2]), that the step of time must remain between two terminals:

$$\Delta t_{\min}\left(h\right)<\Delta t<\Delta t_{\max}\left(\theta\right) \qquad \text{éq 4.5-2}$$

In practice, it is difficult to have an order of magnitude of these terminals, one has thus difficulty, if oscillations are detected, modifying the step of time in order to respect [éq 4.5-2]. In addition, this kind of operation is not always possible sometimes because it is necessary to precisely take into account the abrupt variations of loadings (in particular when Δt is too small).

When Δt **is too large** one can function in **Implicit Euler** $\left(\theta=1\right)$ what will cause to gum the upper limit.

On the other hand **when it is too weak**, two palliative strategies are offered to the user:

- **diagonaliser the matrix of mass via the lumpés elements** (cf [R3.06.07 §4] [§5]) proposed in the code (that requires installation to treat the elements P_2 or modeling `2D_AXI`),
- **to decrease the size of the meshs** (that increases complexities necessary calculation and memory).

It is from this point of view that them **refinements/déraffinements practised on the faith of our indicator can have an incidence**. The fact of refining will not pose any problem on the other hand while déraffinant one can deteriorate very well the decrease of [éq 4.5-2]. It is necessary thus to be very circumspect if one uses the option déraffinement software LOBSTER (encapsulated for *Code_Aster* in `MACR_ADAP_MAIL` option '`DERAFFINEMENT`' [U7.03.01]) on case test comprising a thermal shock.

We now will summarize the principal contributions of the preceding and their holding and bordering theoretical chapters with respect to the thermal calculation set up in *Code_Aster*.

5 Summary of the theoretical study

That is to say **(P_0) the problem in extreme cases mixed** (of linear type inhomogenous Cauchy-Dirichlet-Neumann-Robin and with variable coefficients) solved by **the operator `THER_LINEAIRE`**

$$(P_0)\begin{cases} \rho\, C_p \dfrac{\partial T}{\partial t} - \mathrm{div}(\lambda\, \nabla T) = s & \Omega\times]0,\tau[\\ T = f & \Gamma_1\times]0,\tau[\\ \lambda \dfrac{\partial T}{\partial n} = g & \Gamma_2\times]0,\tau[\\ \lambda \dfrac{\partial T}{\partial n} + hT = hT_{ext} & \Gamma_3\times]0,\tau[\\ T(\mathbf{x},0) = T^0(\mathbf{x}) & \Omega \end{cases} \quad \text{éq 5-1}$$

Taking into account as of choice modelings operated in *Code_Aster* (by `AFFE_MATERIAU`, `AFFE_CHAR_THER`...) it is determined **Tally Variational Abstract** (CVA cf [§2]) minimal on which one will be able to rest to show **the existence and the unicity of a field of temperature solution** (cf [§2]). By recutting these pre-necessary theoretical a little "ethereal" with the practical constraints of the users, one from of deduced from the limitations as for the types of geometry and the licit loadings. Then, **while semi-discretizing in time and space by the usual methods of the code** (which one makes sure of course of the cogency and owing to the fact that they preserve the existence and the unicity of the solution), one studies **evolution of the properties of stability of the problem** (cf [§3]). These results of controllability are very useful for us to create the standards, the techniques and the inequalities which intervene in the genesis of the indicator in residue. In these stages of discretization we also briefly approach the influence of such or such theoretical assumption on **functional perimeter of the operators of the code**.

Before summarizing the principal theoretical results concerning the indicator of error, we go repréciser some notations:

- a step of time is fixed Δt such as $\dfrac{\tau}{\Delta t}$ that is to say an entirety *NR* and that the temporal discretization is regular: $t_0=0,\ t_1=\Delta t,\ t_2=2\Delta t \cdots t_n=n\Delta t$,

  Note:

  > This assumption of regularity does not have really importance, it just makes it possible to simplify the writing of the semi-discretized problem. To model an unspecified transient at the moment t_n , it is just enough to replace Δt by $\Delta t_n = t_n + 1 - t_{n.}$

- that is to say θ the parameter of θ - method semi-discretizing temporally (P_0) ,
- are T^n and T_h^n fields of temperatures at the moment t_n $(0 \le n \le N)$, exact solutions of the initial problem (P_0) , respectively semi-discretized in time and completely discretized in time and space.

Taking into account as of modelings installation in the code, we can suppose that **temporal discretization of the loadings and the source is exact** and that **taking into account, via Lagranges, of the limiting conditions (generalized or not) of Dirichlet is too**. On the other hand, one of the approaches to model the digital approximations carried out during the integral calculus of the indicator of error, consists in supposing **inaccurate space discretization of the loadings and the source**. Their approximate values are noted

$$s^{n+1}_{\theta,h},\ g^{n+1}_{\theta,h},\ T^{n+1}_{ext,\theta,h} \text{ et } h^{n+1}_{\theta,h} \qquad \text{éq 5-2}$$

while posing

$$\chi^{n+1}_{\theta}=\theta\chi\left(\mathbf{x},(n+1)\frac{\tau}{\Delta t}\right)+(1-\theta)\chi\left(\mathbf{x},n\frac{\tau}{\Delta t}\right) \text{ avec}\chi\in\{T,s,T_{ext},g,h\} \text{ et } 0\leq n\leq N-1 \qquad \text{éq 5-3}$$

Note:

This kind of indicator installation of (in mechanics as in thermics) is also sullied with another type of digital approximations related to calculations of the derivative second of the voluminal term (cf [§4.3]). Its effect can possibly feel when one is interested in the intrinsic value of the voluminal error for sources very kicked up a rumpus on a coarse grid.

They exist two constants then K_2 and K_3 independent of the parameters of discretization in time and space, depending only on the smallest angle of the triangulation and the type of problem, which make it possible to build:

- One **increase of the total space error** (the history of the total real indicator **"on - estimates"** the total space error)

$$\|\sqrt{\rho C_p}(T^n-T^n_h)\|^2_{0,\Omega}+\Delta t\sum_{m=0}^{n-1}\|\sqrt{\lambda}\nabla(T^{m+1}_{\theta}-T^{m+1}_{\theta,h})\|^2_{0,\Omega}$$
$$\forall 0\leq n\leq N \quad \leq\|\sqrt{\rho C_p}(T_0-T^h_0)\|^2_{0,\Omega}+K_2\Delta t\sum_{K\in T_h}(\eta^0_R(K))^2+\sum_{m=0}^{n-1}\left\{(\eta^{m+1}_R(K))^2+h^2_K\|s^{m+1}_{\theta,h}-s^{m+1}_{\theta}\|^2_{0,K}\right\}+$$
$$K_2\Delta t\sum_{K\in T_h}\sum_{m=0}^{n-1}\left\{\sum_{F\in S_2(K)}h_F\|g^{m+1}_{\theta,h}-g^{m+1}_{\theta}\|^2_{0,F}+\sum_{F\in S_3(K)}h_F\|(h_h(T_{ext,h}-T_h))^{m+1}_{\theta}-(h(T_{ext}-T_h))^{m+1}_{\theta}\|^2_{0,F}\right\}$$

éq 5-4

- One **decrease of the local space error** (it " **underestimate** " the local space error)

$$\eta^{n+1}_R(K)\leq K_3\left\{\begin{array}{c} h_K\|\sqrt{\rho C_p}\frac{T^{n+1}-T^{n+1}_h-T^n-T^n_h}{\Delta t}\|_{0,\Delta_K}+\|\sqrt{\lambda}\nabla(T^{n+1}_{\theta}-T^{n+1}_{\theta,h})\|_{0,\Delta_K}+\\ h_K\|s^{n+1}_{\theta}-s^{n+1}_{\theta,h}\|_{0,\Delta_K}+h^{\frac{1}{2}}_F\|g^{n+1}_{\theta}-g^{n+1}_{\theta,h}\|_{0,\Delta_K\cap\Gamma_2}+\\ h^{\frac{1}{2}}_F\|h^{n+1}_{\theta}(T^{n+1}_{ext,\theta}-T^{n+1}_{\theta})-h^{n+1}_{\theta,h}(T^{n+1}_{ext,\theta,h}-T^{n+1}_{\theta,h})\|_{0,\Delta_K\cap\Gamma_3}\end{array}\right\}$$
$$\forall 0\leq n\leq N-1$$

éq 5-5

- With **continuation** $\left(\eta^n_R(K)\right)^{K\in T_h}_{0\leq n\leq N}$ **local real indicators** (by using the notations of [§4.1])

$$\begin{aligned}\eta_R^{n+1}(K) &:= \eta_{R,vol}^{n+1}(K)+\eta_{R,saut}^{n+1}(K)+\eta_{R,flux}^{n+1}(K)+\eta_{R,éch}^{n+1}(K)\\ &:= h_K \| s_{\theta,h}^{n+1}-\rho C_p \frac{T_h^{n+1}-T_h^n}{\Delta t}+div(\lambda \nabla T_{h,\theta}^{n+1})\|_{0,K} + \frac{1}{2}\sum_{F\in S_\Omega(K)} \sqrt{h_F} \| \left[\lambda \frac{\partial T_{h,\theta}^{n+1}}{\partial n}\right] \|_{0,F} + \\ &\sum_{F\in S_2(K)} \sqrt{h_F} \| g_{\theta,h}^{n+1} - \lambda \frac{\partial T_{h,\theta}^{n+1}}{\partial n} \|_{0,F} + \sum_{F\in S_3(K)} \sqrt{h_F} \| (h(T_{ext}-T))_{\theta,h}^{n+1} - \lambda \frac{\partial T_{h,\theta}^{n+1}}{\partial n} \|_{0,F}\end{aligned}$$ éq 5-6

who is initialized by

$$\begin{aligned}\eta_R^0(K) &:= h_K \| s_h^0 + div(\lambda \nabla T_h^0) \|_{0,K} + \frac{1}{2}\sum_{F\in S_\Omega(K)} \sqrt{h_F} \| \left[\lambda \frac{\partial T_h^0}{\partial n}\right] \|_{0,F} + \\ &\sum_{F\in S_2(K)} \sqrt{h_F} \| g_h^0 - \lambda \frac{\partial T_h^0}{\partial n} \|_{0,F} + \sum_{F\in S_3(K)} \sqrt{h_F} \| (h(T_{ext}-T))_h^0 - \lambda \frac{\partial T_h^0}{\partial n} \|_{0,F}\end{aligned}$$

éq 5-7

This local continuation makes it possible to build **continuation** $(\eta^n(\Omega))_{0\le n\le N}$ **total real indicators**

$$\forall 0 \le n \le N \quad \eta_R^n(\Omega) := \left(\sum_{K\in T_h} \eta_R^n(K)^2\right)^{\frac{1}{2}}$$ éq 5-8

Of [éq 5-4] (cf [§4.2]) it appears that, at one moment given, the error on the approximation of the condition of Cauchy and the history of the total indicators intervene on the total quality of the solution obtained. **One will be able to thus minimize overall the error of approximation due to the finite elements in the course of time while re-meshing "advisedly", via the continuation of indicators, the structure**. Because, in practice, one realizes that the refinement of the meshs makes it possible to decrease their error and thus cause a drop in the temporal sum of the indicators. **The total error will butt (and it is moral) against the value floor due to the approximations of the initial condition, the limiting conditions and the source** (which will tend it-also to drop of course!)**. One cannot get results of better quality that the data input of the problem!**

The result [éq 5-5] (cf [§4.4]) provides only one **opposite room of total increase** [éq 5-4] ("must" would have been to reveal also an increase at the local level) but, within the framework of this kind of indicator, one will not be able to obtain better compromise. **These estimates are optimal within the meaning of [bib5]**. They illustrate the equivalence of the sum hilbertienne indicators with the space part of the total exact error. The constants of equivalence are independent of the parameters of discretizations in space and in time, they depend only on the smallest angle of the triangulation and the type of with dealt problem.

According to this increase of the indicator [éq 5-6], if one refines very locally (around the element K) in order to decrease $\eta_R^n(K)$, one is not ensured of a reduction in the error in an immediate vicinity of the zone concerned (in Δ_K). **The indicator "underestimates" the space error locally and only a more macroscopic refinement carries out a reduction in the error theoretically**.

Only in pure residue, all one **"zoology" of indicators of space error are permissible** (cf [§4.3]), we retained a type similar of it to that already set up for the mechanics of *Code_Aster*. Being based on the solutions and the discrete loadings of the moment running and the previous moment (except with the first step of time), **its theoretical limitations are thus, at best, those inherent in the resolution of the problem in temperature**: no zones comprising of crack, point or points of reflection not, problem to the multi-material interfaces, θ - unconditionally stable diagram, regular family of triangulation, polygonal grid discretized by isoparametric finite elements, oscillations and violation of the principle of the maximum (cf [§4.5]). Of course, in practice, one very often passes in addition to, and this without encumbers, this perimeter of "theoretical" use.

But it is necessary well to keep in mind, that as a "simple postprocessing" of $(\mathbf{P_0})$, the indicator cannot unfortunately provide more reliable diagnosis in the zones where the resolution of the initial problem stumbles (close to crack, shock…). Its denomination prudently reserved of indicator (instead of the usual terminology of estimator) is in these typical cases more than ever of setting! But if, in these extreme cases, its gross amount perhaps prone to guarantee, **its utility as an effective and convenient supplier of cards of error for a mending of meshes or a refinement/déraffinement remains completely justified**.

In the same vein, even if the formulation [éq 5-6] were established only in the transitory linear case, isotropic or not, defines by (P_0), one could also stretch his perimeter of use to non-linear (operator `THER_NON_LINE`), in different limiting conditions (`ECHANGE_PAROI` for example) or with other types of finite elements (lumpés isoparametric elements, elements of structure…) (cf [§2.1]). For more information on the "data-processing" perimeter corresponding to his effective establishment in the code, one can refer to [§6.2] or the user's documentation of `CALC_ERREUR` [U4.81.06].

It was question, until now, only of cards of indicators of space errors calculated at a given moment of the transient of calculation. But, in fact, **there exist several ways for to build indicators of error on a parabolic problem** (cf [§4.5]). That that we retained does not allow a complete control of the error and it always requires one **certain vigilance when one deals with problems of the shocks type** (the same one as for the problem post-treaty!). It reveals only implicitly the term of jump temporal in all the terms in θ of [éq 5-6].

To finish, it should be stressed that this indicator is thus composed of four terms:

- **principal term**, called **voluminal term of error**, controlling the good checking of the equation of heat,
- to which are added **three secondary terms** checking the good behaviour of **space jumps** and of the limiting conditions: **terms of flow and exchange**.

In `2D-PLAN` or in `3D` (resp. in `2D-AXI`), if the unit of the geometry is the meter, the unit of the first is it *W.m* (resp. $W.m.rad^{-}$) and that of the other terms is it $W.m^{\frac{}{}}$ (resp. $W.m^{\frac{1}{2}}.rad^{-}$). **Attention thus with the units taken into account for the geometry when one is interested in the gross amount of the indicator and not in his relative value!**

We now will approach, after the practical difficulties of implementation in the code, the environment necessary and its perimeter of use. One will conclude for an example of use drawn from a case official test.

6 Implementation in *Code_Aster*

6.1 Particular difficulties

To calculate this kind of indicator it is necessary to compose with the vision "elementary calculation + assembly" generally deployed in all the codes finite elements. However the estimate, at the local level, of $\eta(K)$ requires, not only the knowledge of its local fields, but also that of its close meshs. **One thus needs to carry out a "total calculation" at the level of Δ_K, in local calculation!** A copied strategy on what had been set up for the estimator in mechanics consists in transmitting this kind of information in **components of wide cards** who they will be transmitted in argument of entry of `CALCULATION`. It is this kind of contingency which explains the heterogeneity of treatment at the time of the overloads of loadings between the thermal solveurs and the calculation of our indicator (cf [§6.2]).

Another type of difficulty, digital this time, relates to the calculation of the voluminal term. Indeed, it requires a double derivation which one carries out in three stages, because in *Code_Aster* one does not recommend the use of the derivative second of the functions of forms.

Note:

> *They were recently introduced to treat the derivation of the rate of refund of energy (cf [R7.02.01 § Annexe 1]).*

On the one hand, one calculates (in the thermal operator) the vector flow at the points of gauss, then one extrapolates the corresponding values with the nodes by local smoothing (cf [R3.06.03] `CALC_CHAMP` with `THERMICS='FLUX_ELNO'` and [§6.2]) in order to calculate its divergence at the points of Gauss. With quadratic finite elements the intermediate operation is only approximate (one assigns like value to the median nodes the half the sum of their values to the extreme nodes). However digital tests (limited) showed that this approach does not provide results very different from those obtained by a direct calculation via good the derivative second.

Lastly, it was necessary to determine various geometrical characteristics (diameters, normals, jacobiens…), connector industries of the elements in opposite and to reach the data which they underlie in all the cases envisaged by the code (started from grid symmetrized and/or heterogeneous, loading function or reality, non-linear material, all isoparametric elements 2D/3D and all the thermal loadings).
Beyond these fastidious developments, **a large effort of data-processing validation "géométrico -" was made** to try to track possible bugs in this entrelac of small formulas. These hard tests on small cases model tests (`TPLL01A/H` for `2D_PLAN/3D` and `TPNA01A` for `2D_AXI`) appeared profitable (including for the indicator in mechanics and the elements lumpés!) and essential. Because one does not lay out, to my knowledge, of theoretical values allowing to validate in certain situations these indicators: **"nothing resembles any more one value of indicator… than another value of indicator!". For another thing and, although in a process of validation that is not the panacea, it is thus necessary to try to release a maximum of confidence in all these components**.

6.2 Environment necessary/parameter setting

The calculation of this indicator is carried out, via the option '`ERTH_ELEM`'L' operator of postprocessing `CALC_ERREUR`, on one `EVOL_THER` (provides to the keyword `RESULT)` coming from a former thermal calculation (linear or not, transient or stationary, isotropic or orthotropic, via `THER_LINEAIRE` or `THER_NON_LINE`, cf more precise perimeter [§6.4]).
As one already underlined, it **the recourse to the option requires as a preliminary '`FLUX_ELNO`'** of `CALC_CHAMP` who determines the values of the vector heat flux to the nodes (cf example of use [§6.5]).
This indicator consists of fifteen components by elements and for a given moment. In order to be able post-to treat them via `POST_RELEVE` or GIBI one needs to extrapolate these fields by element into cubes fields with the nodes by element. The addition of the option '**`ERTH_ELNO`**'(after L' call to '`ERTH_ELEM'`) allows to carry out this purely data-processing transformation. For one moment and a given finite element, it does nothing but duplicate the fifteen components of the indicator on each nodes of the element.

For carrying out the integral postprocessing of desired thermal calculation well, it is necessary:

- **To carry out on all the geometry, TOUT=' OUI'** (value by default, if not calculation stops in ERREUR_FATALE). This provisional choice was led by data-processing and functional contingencies, because thus all the finite elements are seen affecting a homogeneous indicator calculated with the same number of terms (if not quid of the concept of term of jump and term of CL at the edge of the field considered?). In addition the tool of refinement/déraffinement of the code (software LOBSTER encapsulated in MACR_ADAP_MAIL), emerged natural of our cartographies of error, does not allow to treat only parts of grids.

  Note:

  > *That poses problems of propagation of subdivisions to preserve the conformity of the triangulation. In fact, to divert this kind of contingency, it would be necessary, either to define a buffer zone making the junction between the “dead” zone of the grid and the zone “activates” to treat, or in manner more optimal but also much more difficult from a point of view structures, to reduce it to a layer of joined elements.*

- **To provide the same temporal parameter setting** : value of θ (value by default equalizes to 0.57) provided to the keyword PARM_THETA ; if necessary if with a transitory problem are dealt, should be informed the usual fields TOUT/NUMÉRIQUE/LISTE_ORDRE with licit values with respect to thermal calculation. The calculation of the history of the indicator can then be carried out as from any moment of a transient, knowing that with the first increment one carries out calculation like in hover ($\theta=1$, $n+1=0$ and not of term of finished difference cf [éq 5-7]).
  Moreover, in hover, if the user provides a value of θ different from 1, one imposes this last value to him after having informed some.
  In a related way, one detects the request for supply of cards of errors between noncontiguous sequence numbers (one has one ALARM) or the data of one EVOL_THER not comprising a field of temperature and vector flow to the nodes (calculation stops in ERREUR_FATALE). The value of θ and the number of sequence number taken into account are traced in the file message [§6.3]. The sequence number and the corresponding moment accompany also each occurrence by indicator of error in the file result ([§6.3]).
- **To use the same loadings and by complying with the rules of particular overloads** with the options of error analyses of this operator. Thus, in the thermal solveurs (and mechanics) one incorporates the limiting conditions of the same type, whereas in error analyses of CALC_ERREUR (and thus also with our indicator) one can take into account, for a kind of limiting condition given, only the last provided to the keyword EXCIT. **The order of these loadings thus is a crucial importance !**

  Note:

  > *This restriction finds its base in the first remark of the preceding paragraph. For making well it would be necessary, either concaténer on the elements of skin concerned all the limiting conditions, or to provide to elementary calculations of the variable cards of sizes containing all the loadings exhaustively. The first solution seems by far most optimal but also hardest to implement. It would then be necessary also to make the same thing for the indicator in residue of mechanics (OPTION=' ERRE_ELGA_NORE').*

However, in the event of conflict between loadings of the same type, one often can and easily to find a solution palliative via AFFE_CHAR_THER adequate. The user is informed presence of several occurrence of the same type of loading by a message of ALARM and lists it loadings actually taken into account is traced in the file message ([§6.3]).
The code stops on the other hand in ERREUR_FATALE if the provided loadings pose certain problems (interpolation of loadings function, access to the components, presence of CHAMPGD coefficient of exchange and absence of CHAMPGD outside temperature or vice versa….),

- **Within the same framework general** : value of the model (parameter MODEL), necessary materials (CHAM_MATER), structure EVOL_THER data (RESULT) and result (assignment of CALC_ERREUR with possibly one " reuse " réentrant). They are traced in the file message ([§6.3]).

If the user does not respect this necessary homogeneity of parameter setting (with the rules of overload near) between the thermal solvor and the tool for postprocessing, it being exposed to skewed results even completely false (without inevitably a message of ALARM or one ERREUR_FATALE stops, one cannot all control and/or prohibit!). There then remains only judge of the relevance of his results.

Let us recapitulate all this parameter setting of the operator CALC_ERREUR directly impacting the calculation of the indicator of space error in thermics.

Keyword factor	*Keyword*	*Value by default*	*Value obligatory (O) or advised (C)*
	MODEL		Idem thermal calculation (O)
	CHAM_MATER		Idem thermal calculation (O)
	ALL	'YES'	'YES' (O)
	TOUT/NUMÉRIQUE/LISTE_ORDRE	'YES'	'YES' (C)
	PARM_THETA	0.57	Idem thermal calculation (O)
	RESULT		EVOL_THER thermal calculation (O)
	reuse		EVOL_THER thermal calculation (C)
EXCIT	LOAD		Idem thermal calculation + rule of overload (O)
	OPTION		**'ERTH_ELEM'** 'ERTH_ELNO'
	INFORMATION	1	1 (C)

Table 6.2-1: Summary of the parameter setting of CALC_ERREUR impacting the calculation of the indicator

Note:

- *In transient, it (strongly) is advised to calculate the history of the indicator over moments of calculations contiguous. If not, the postprocessing of the temporal semi-discretization will be distorted, and according to the devoted formula… the user will become only judge of the relevance of his results.*

6.3 Presentation/analysis of the results of the error analysis

The option '`ERTH_ELEM`' provides in fact, not one, but **fifteen components by finite elements** K **and by step of time** t_{n+1} . Indeed, **for each of the four terms** of [éq 5-6], _the voluminal principal term and the three surface secondary terms _, one not only calculates **the absolute error**, but too **a term of standardisation** (the theoretical value of the discretized loadings that one would have had to find) and **the associated relative error**. By summoning these three families of four contributions, one establishes also the total absolute error, the total term of standardisation and the total relative error. What makes the account well!

The fact of dissociating the contributions of each component of this indicator makes it possible to compare their relative importances and to target strategies of refinement/déraffinement on a certain kind of error. Even if the voluminal term (representing the good checking of the equation of heat) and the term of jump (related to modeling finite elements) remain the dominating terms, it can prove to be useful to measure the errors due to certain type of loading in order to refine their modeling or to re-mesh the accused frontier zones.

Moreover this kind of strategy can be easily diverted of its primary goal in order to make refinement/déraffinement by zone: it is enough to impose, only in this zone, a kind of limiting condition fictitious (with very bad value in order to cause a large error).
Way of calculating of these components and the name of their component "of reception" in the field symbolic system '`ERTH_ELEM_TEMP`' of `the EVOL_THER` are recapitulated in the table below (while being based on the nomenclature of [éq 5-6]).

	Absolute error	Relative error (in %)	Term of standardisation
Voluminal term	$\eta^{n+1}_{R,vol}(K)$ **TERMVO**	$\frac{\eta^{n+1}_{R,vol}(K)}{N^{n+1}_{R,vol}(K)}\times 100.$ **TERMV2**	$N^{n+1}_{R,vol}(K):=h_K\|s^{n+1}_{\theta,h}\|_{0,K}$ **TERMV1**
Term of jump	$\eta^{n+1}_{R,saut}(K)$ TERMSA	$\frac{\eta^{n+1}_{R,saut}(K)}{N^{n+1}_{R,saut}(K)}\times 100.$ TERMS2	$N^{n+1}_{R,saut}(K):=\frac{h_F^{\frac{1}{2}}}{2}\left\|\left[\lambda\ \frac{\partial T^{n+1}_{\theta,h}}{\partial n}\right]\right\|_{0,F}$ TERMS1
Term of flow	$\eta^{n+1}_{R,flux}(K)$ TERMFL	$\frac{\eta^{n+1}_{R,flux}(K)}{N^{n+1}_{R,flux}(K)}\times 100.$ TERMF2	$N^{n+1}_{R,flux}(K):=h_F^{\frac{1}{2}}\|g^{n+1}_{\theta,h}\|_{0,F}$ TERMF1
Term of exchange	$\eta^{n+1}_{R,éch}(K)$ TERMEC	$\frac{\eta^{n+1}_{R,éch}(K)}{N^{n+1}_{R,éch}(K)}\times 100.$ TERME2	$N^{n+1}_{R,éch}(K):=h_F^{\frac{1}{2}}\|(h(T_{ext}-T))^{n+1}_{\theta,h}\|_{0,F}$ TERME1
Total	$\eta^{n+1}_{R}(K):=\sum_i \eta^{n+1}_{R,i}(K$ **ERTABS**	$\frac{\eta^{n+1}_{R}(K)}{N^{n+1}_{R}(K)}\times 100.$ **ERTREL**	$N^{n+1}_{R}(K):=\sum_i N^{n+1}_{R,i}(K)$ **TERMNO**

Table 6.3-1: Components of the indicator of error.

For the absolute error and the term of standardisation, in 2D-PLAN or in 3D (resp. in 2D-AXI), if the unit of the geometry is the meter, the unit of the first term is it W.m (resp. $W.m.rad^{-}$) and that of the other terms is it $W.m^{\frac{1}{2}}$ (resp. $W.m^{\frac{1}{2}}.rad^{-1}$).

Attention thus with the units taken into account for the geometry when one is interested in the gross amount of the indicator and not in his relative value!

This information is accessible in three forms:

- For each moment of the transient, these fifteen values are summoned on all the grid (one makes the same thing as in the table [Table 6.3-1] while replacing K by Ω) and traced in a table of the file result (.RESU).

```
**********************************************
THERMICS: ESTIMATOR OF ERROR IN RESIDUE
**********************************************

IMPRESSION OF THE TOTAL STANDARDS:

SD EVOL_THER     RESU_1
SEQUENCE NUMBER       3
MOMENT            5.0000E+00
ERROR              ABSOLUTE   /  RELATIVE/STANDARDISATION
TOTAL              0.5863E-05   0.2005E- 04%  0.2923E+02
VOLUMINAL TERM    0.3539E-05   0.0000E+ 00%  0.0000E+00
JUMP TERM         0.2217E-05   0.1006E- 04%  0.2205E+02
FLOW TERM         0.4384E-06   0.3886E- 05%  0.1128E+02
TERM EXCHANGES      0.4591E-06   0.5755E- 05%  0.7977E+01

**********************************************
```

Example 6.3-1: Layout of the option ' ERTH_ELEM_TEMP ' in the file result

- It is stored by means of computer in the fifteen components of the field symbolic system 'ERTH_ELEM_TEMP' SD_RESULTAT thermics. The variables of access of this field are, for each mesh (in our example M1), the sequence number (NUME_ORDRE) and the moment (INST). With the option 'ERTH_ELNO_ELEM'there is the same thing for each node of L' element considered.

```
FIELD BY ELEMENT AT THE POINTS OF GAUSS OF REFERENCE SYMBOL
ERTH_ELEM_TEMP
 SEQUENCE NUMBER: 3 INST:  5.00000E+00
 M1         ERTABS        ERTREL        TERMNO
            TERMVO        TERMV2        TERMV1
            TERMSA        TERMS2        TERMS1
            TERMFL        TERMF2        TERMF1
            TERMEC        TERME2        TERME1
        1  0.5863E-05   0.2005E-04  0.2923E+02
           0.3539E-05   0.0000E+00  0.0000E+00
           0.2217E-05   0.1006E-04  0.2205E+02
           0.4384E-06   0.3886E-05  0.1128E+02
           0.4591E-06   0.5755E-05  0.7977E+01
                                 ........

FIELD BY ELEMENT AT THE POINTS OF GAUSS OF REFERENCE SYMBOL
ERTH_ELNO_ELEM
 SEQUENCE NUMBER: 3 INST:  5.00000E+00
 M1         ERTABS        ERTREL        TERMNO
            TERMVO        TERMV2        TERMV1
            TERMSA        TERMS2        TERMS1
            TERMFL        TERMF2        TERMF1
            TERMEC        TERME2        TERME1
 N1        0.5863E-05   0.2005E-04  0.2923E+02
           0.3539E-05   0.0000E+00  0.0000E+00
           0.2217E-05   0.1006E-04  0.2205E+02
           0.4384E-06   0.3886E-05  0.1128E+02
           0.4591E-06   0.5755E-05  0.7977E+01
 N3        0.5863E-05   0.2005E-04  0.2923E+02
                                 ........
```

Example 6.3-2: Layouts, via `IMPR_RESU` , components of the field symbolic system ' `ERTH_ELEM_TEMP` '/' `ERTH_ELNO_ELEM'` in the file result

- One can also trace the values of each one of these components in the file message (.MESS) by initializing the keyword INFORMATION to 2. However this functionality rather reserved for the developers (for maintenance or of the pointed diagnoses) also revealed complementary impressions (documented but too exhaustive) on the elements constituting the indicator and the characteristics of the finite elements and their vicinities.

```
TE0003 **********
NOMTE/L2D THPLTR3          /   T
 RHOCP     2.0000000000000
 ORIENTATION NETS     1.0000000000000
…
 ---> TERMVO/TERMV1    1.2499997764824    1.2499997764826
CURRENT MESH <<<  3 TRIA3
 DIAMETER     3.5355335898314D-02
  EDGES OF THE TYPE SEG2


 NUMBER OF ARETE/HF   1    2.4999997764826D-02
 MANY TOPS   2
 CONNECTOR INDUSTRY   1  2
 XN     0.59999992847442   0.59999992847442
 YN    -0.80000005364418  -0.80000005364418
 JAC     1.2499998882413D-02    1.2499998882413D-02
 <<< CLOSE MESH   2  QUAD4
 IGREL/IEL  1  2
 INOV LOCAL/GLOBAL  2  5
….
*********************************************
      TOTAL ON THE MESH   2
 ERROR             ABSOLUTE   /  RELATIVE    /  MAGNITUDE
 TOTAL              0.5900D-03   0.1079D- 03%   0.5466D-03
 VOLUMINAL TERM   0.1768D-01   0.1000D- 03%   0.1768D-01
 JUMP TERM        0.5882D-03   0.1080D- 03%   0.5448D-03
 FLOW TERM        0.0000D+00   0.0000D+ 00%   0.0000D+00
 TERM EXCHANGES     0.0000D+00   0.0000D+ 00%   0.0000D+00
 *********************************************
```

Example 6.3-3: Layout, via INFO=2 , in the file message

Note:

- *When the term of standardisation is null (a certain kind of loading or of source is null, as it is the case in the examples [Example 6.3-1] and [Example 6.3-2] above with the voluminal term), one does not calculate the term of relative error associated. There remains initialized to zero.*
- *Moreover, to calculate indeed the absolute error relative to a worthless limiting condition (a flow or a condition of exchange) it should be forced as a function via `AFFE_CHAR_THER_F` adhoc. And this for simple data-processing contingencies, which make that with a constant loading, one cannot make the distinction between:*

1) *worthless limiting condition : the user imposes zero on this portion of border and he wants to test the associated absolute error,*
2) *worthless limiting condition : there are no limiting conditions on this edges,*

- *Tests of not-regression "numérico-data processing" showed that the manner of modelling the loadings and the source, as constants or functions, could notably influence the values of very small terms of error (especially in relative error of course) and worry the user unnecessarily. This phenomenon is explained by differences in codings of the discretized loadings [éq 5-2]. This kind of behavior is also found as soon as one changes linear solvor, preconditionnor, method of renumerotation, of platform…*
- *In hover, when one uses a nonworthless source with linear finite elements, the term principal is very badly estimated since it requires a double derivation of the field of temperature. One `ALARM` thus prevents the user and enjoins it to pass into quadratic.*

6.4 Perimeter of use

This indicator was developed, for the moment, only on the isoparametric elements (`TRIA3/6`, `QUAD4/8/9`, `TETRA4/10`, `PENTA6/13/15` and `HEXA8/20/27`) and for modelings `PLAN`, `PLAN_DIAG`, `AXIS`, `AXIS_DIAG`, `3D` and `3D_DIAG`. It thus does not calculate the contributions of the elements of structure of type hull (`COQUE_PLAN`, `COQUE_AXIS`, `HULL`), pyramids (`PYRAM5` and `PYRAM13`) and of the modeling of Fourier (`AXIS_FOURIER`). It does not make it possible either to calculate the terms of jumps of these elements with the authorized elements. **However, if a grid comprises licit and illicit elements, calculation does not stop** and, via`OPTION` – 2 in the suitable catalogues of elements, **one warns the user of not taken into account of the aforesaid elements**. Indeed to carry out this postprocessing, it is necessary as a preliminary to call, explicitly, the option '`FLUX_ELNO`' (calculation of the vector heat flux to the nodes) and, implicitly, '`INIT_MAIL_VOIS`' (determination of the characteristics of the vicinity Δ_K of an element K). **One is thus tributary of their respective perimeters of use**.

It is also necessary to keep in mind **some more minor rules** but which can be of a very particular importance for very precise studies:

- The calculation of the indicator treats only the elements of the grid pertaining to the model indicated by the keyword `MODEL` order `CALC_ERREUR`. One can thus work with grids (not cleaned) comprising "meshs of outline" to which one allots a different model.
- In one **grid in dimension** q, one calculates the terms of jump and loading, only on **elements of skin of dimension** $q-1$. Therefore, one treats the relations of `TRIA/QUAD` with `SEG` and relations `TETRA/PENTA/HEXA` with `FACE`. For example, in the event of presence of segments in a three-dimensional grid, the option will not stop but she will not take into account their (possible) contributions.
- The option '`ERTH_ELEM_TEMP`' and its preliminary options do not know them **`PYRAM`**. Their contributions will be ignored. This gap comes from their introduction into *Code_Aster* more recent than those of the already quoted preliminary options.

Note:

In any event these elements are minority in a grid 3D and are generated only by the voluminal free maillor of GIBI, which creates some locally to supplement portions of grids hexahedral.

- In 2D, **one should not accidentally intercalate a segment between two triangles or quadrangles**, if not the term of jump of these elements will not be calculated and one will enquérira oneself wrongly of the existence of a possible limiting condition. Calculation will not stop but with this interface, the value of the indicator will be incomplete. However, for special needs (charging density internal and localised in a structure, fissures…), one can of course allow this kind of situation. In 3D, the problem arises of course also when one intercalates quadrangles or triangles between two `FACE` contiguous.
- the same type of imbroglio occurs when **two points of the grid are superimposed** geometrically. There still, calculation should not stop, but the value of the indicator will be incomplete on the level of this zone of covering,
- **If one works with a grid which results from operations of symmetrization**, it is necessary to try not to be in the two preceding cases. Moreover, on both sides of the axis of symmetry, **the close meshs do not have inevitably (with in particular maillor GIBI) of the orientations which meet the standard of *Code_Aster*** (they should be reversed). The calculation of the indicator, for which this information is crucial (to define the external normals in each mesh and the connector industries in opposite), detects the problem by calculating the jacobien of each mesh. In 2D, an algorithm of substitution makes it possible to circumvent the problem and to rebuild the tables of connector industry "nodes of the element running nodes of its neighbors". In 3D, the problem is much more difficult and individual with each element, the code thus stops in `ERREUR_FATALE` in the event of problem.
- **If one wants to refine or déraffiner its grid with `MACR_ADAP_MAIL` [U7.03.01]**, the grid should comprise only triangles or tetrahedrons. Concerning the loadings surface or voluminal, the

"good practice" consists in using only groups of meshs. If groups of nodes are used, one must expect distorted calculations, because after some refinements, other points will have probably formed part geometrically of the zone concerned with GROUP_NO without seeing itself affecting any loading (one cannot modify the composition of one GROUP_NO in the course of session!).

For specific loadings or points of statement (on which go, for example, to rest POST_RELEVE_T) GROUP_NO is licit. On the other hand, it is not advised to use meshs directly (MY) or of the nodes (NO) (apart from group), because in this case, with the liking of the renumérotations, LOBSTER probably will lose their trace. It can preserve the memory of the meshs or the nodes only through one name of GROUP_MA or of GROUP_NO. Thanks to this mechanism, it can adopt a Lagrangian vision of becoming of these meshs or these points!

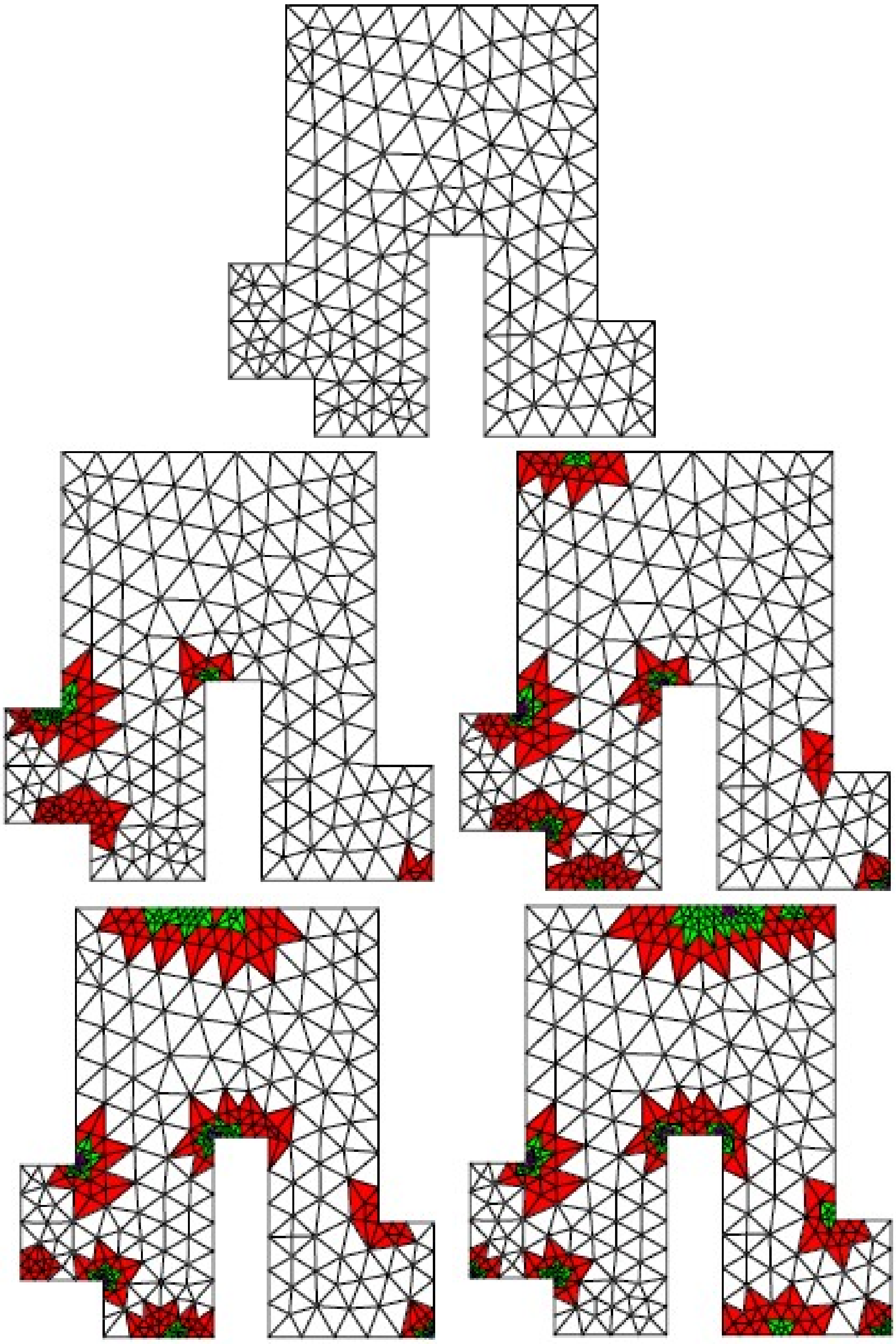
The calculation of the indicator takes place indifferently on one EVOL_THER coming from THER_LINEAIRE or of THER_NON_LINE , stationary or transitory, isotropic or orthotropic, and, on a motionless structure with a grid by elements answering the preceding criteria.

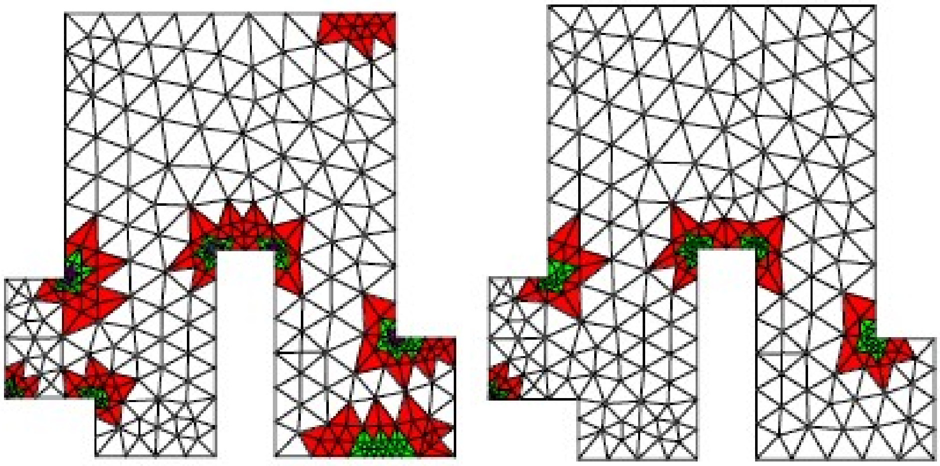
Into non-linear one takes into account non-linearities of materials and the modification of the problem in enthalpy. However one does not treat the possible contributions of non-linear loadings (FLUX_NL and RADIATION). The user is informed by it by one ALARM, just like he is informed not taken into account of a limiting condition of type ECHANGE_PAROI. Indeed, into linear one recognizes, for the moment, only the contributions of the loadings SOURCE, FLUX_REP and EXCHANGE. For the taking into account of these loadings, particular rules of overload are applied (cf [§6.2]).

6.5 Example of use

To familiarize itself with the use of this indicator in thermics and its possible coupling with the LOBSTER encapsulation (for more information, one will be able to consult the site http://www.code - aster.com/outils/homard) via MACR_ADAP_MAIL [U7.03.01] one can take as a starting point the case test TPLL01J [V4.02.001].

In this other example extracted the LOBSTER website, coupling ERTH_ELEM/MACR_ADAP_MAIL [U7.03.01] simulates the circulation of a "hot" fluid on both sides of a metal part bent (in top and bottom, via a condition ofEXCHANGE depending on time in AFFE_CHAR_THER_F). The circulation of the fluid is carried out left towards the line.
The precision is especially necessary at the ends of the structure, on the level of the propagation of the fluid: thanks to the indicating coupling of error/tool of refinement-déraffinement, the grid thus remains fine in edge of part, in the zone where concentrates the "hot" fluid. Finally it is déraffiné in the back, once the fluid passed.
It is also noted that, as envisaged by the theory (cf remarks [§2.2]), the resolution of the thermal problem "is blunted" in the returning corners and that the indicator of error (although it is also penalized in these zones) announces this irrefutable fact (even when the part cooled).





Example 6.5-2: Use of the option ' ERTH_ELEM' coupled with LOBSTER

7 Conclusion – Outline

During digital simulations by finite elements, obtaining a gross profit is not sufficient any more. **The user is increasingly petitioning of space error analysis compared to the grid. He needs for support methodological and pointed tools "numériquo-data processing" to measure the quality of his studies and to improve them.**
To this end, the indicators of space error a posteriori make it possible to locate, on each element, a cartography of error on which the tools of mending of meshes will be able to rest: the first calculation on a coarse grid makes it possible to exhume the map of error starting from the data and of the solution discretized (from where the term "a posteriori"), refinement is carried out then locally by treating on a hierarchical basis this information.
The new indicator a posteriori which has just been established post-to treat the thermal problems of *Code_Aster* is based on their local residues extracted the semi-discretizations in time. Via the option 'ERTH_ELEM' of CALC_ERREUR, it uses the thermal fields (EVOL_THER) emanating from THER_LINEAIRE and of THER_NON_LINE.
This new indicator supplements the offer of the code in term of advanced tools making it possible to improve quality of the studies, their mutualisations and their comparisons. Indeed, indicators of error in mechanics and macro of refinement/déraffinement MACR_ADAP_MAIL [U7.03.02] are already available. It remains to supplement the perimeter of use of these tools and, to pack them, in particular for better managing non-linearities and the interactions space error/temporal error.

Note:

> *Estimator by smoothing of constraints of Zhu & Zienkiewicz (CALC_ERREUR + OPTION 'ERZ1/ERZ2_ELEM' [R4.10.01]) and indicator in pure residue ('ERME_ELEM' [R4.10.02]).*

Thereafter, the prospects for this work are several orders:

- **From a functional point of view,** the completeness of this indicator could also improve by taking into account possible nonlinear limiting conditions (FLUX_NL and RADIATION) and of the exchanges between walls (ECHANGE_PAROI). In the long term, it would also be necessary to be able to be pressed on finite elements of structure (hull…), pyramids and power to deal with problems of convection-diffusion (operator THER_NON_LINE_MO [R5.02.04]).
- **From a theoretical point of view**, when new limiting conditions are used and/or when one is based on new modelings (hull, beam…), a study "numériquo - functional" similar to that of this document, should be carried out to consider limitations theoretical and practical (with respect to *Code_Aster*) of such an indicator and to exhume its adhoc formulation.
- Let us recall finally that one **string of indicators of error a posteriori are available,** and, that enough little was tested and validated on industrial cases. In order to refine diagnoses, to establish comparisons and to set up strategies of mending of meshes per class of problem, it would be interesting to pack the list of the indicators available. Various indicators in residue plus local problem thus appeared more effective (but also more expensive) during digital tests (into elliptic) in N3S [bib5].

  Note:

  > *The indicator is the standard of the solution of a local, of the same problem standard than the initial problem, but discretized on spaces moreover high degree and whose second member is the residue. According to the limiting conditions affixed with this local problem, one distinguishes some from various types. They thus mix the vision "bases hierarchical" and the aspects "residue" with the indicators of error a posteriori.*

- The ideal consists in discretizing simultaneously in time and space on suitable finite elements and controlling their "space-time" errors in a coupled way. **This "space-time" indicator** give access to a complete control of the error and it makes it possible to avoid unfortunate decouplings as for the possible refinements/déraffinements controlled by a criterion with respect to the other (cf discussion [§4.5]). It is however very heavy to set up in a large industrial code such as *Code_Aster*. It supposes indeed, to be optimal, nothing less than one separate management step of time by finite elements. What from the point of view of architecture supporting the finite elements of the code is a true challenge!

8 Bibliography

- R. DAUTRAY & J. - L. LIONS and al. mathematical Analysis and digital calculation for sciences and technology. ED. Masson, 1985.
- J. - L. LIONS. Some methods of resolution of the problems in extreme cases non-linear. ED. Dunod, 1969.
- P.A. RAVIART & J.M. THOMAS. Introduction to the digital analysis of the partial derivative equations. ED. Masson, 1983.
- C. BERNARDI, O. BONNIN, C. LANGOUET & B. METIVET. Residual error indicators for linear problems. Extension to the Navier-Stokes equations. Proc. Int. Conf. Finite Elements in Fluids, Venezia 95, pp 347-356. Note HI72/95/018,1995.
- C. BERNARDI, B. METIVET & R. VERFURTH. Working group "adaptive Grid": digital analysis of indicators of errors. Note HI73/93/062, 1993.
- C. BERNARDI & B. METIVET. Indicator of error for the equation of heat. European review of the finite elements, flight n°9, n°4, pp425-438, 2000.
- R. VERFURTH. With review of a posteriori error and estimate adaptive mesh-refinement technical. ED. Wiley & Teubner, 1996.
- P. CLEMENT. Approximation by finite local element functions using regularization. RAIRO Analyzes digital, flight n°9, pp77-84, 1975.
- I. RUUP & PENIGUEL. Code SYRTHES: conduction and radiation. Theoretical handbook of V3.1. Note HE41/98/048, 1998.
- S. ADJERID & J.E. FLATHERTY. With room refinement finite element method for 2D parabolic systems. SIAM J.Sci.Stat.Comput., 9, pp795-811, 1988.
- MR. BIETERMAN & I. BABUSKA. The finite element method for parabolic equations, a posteriori error estimate. Numer. Maths. 40, pp339-371, 1982.
- R.E. Adaptive BIENNER & al. Year finite element method for steady and transient problems. SIAM J.Sci.Stat.Comput., 8, pp529-549, 1987.
- F. BORNEMANN. Year adaptive multilevel approach to parabolic equations. 3 shares in IMPACT of comp. In Sci. And Engrg. 2, pp279-317, 1990. 3, pp93-122, 1991. 4, pp1-45, 1992.
- K. ERIKSSON & C. JOHNSON. Adaptive finite element methods for parabolic problems. SIAM J.Nume.Anal., 28, pp43-77, 1991.
- C. JOHNSON & V. THOMEE. Year a posteriori error estimate and adaptive timestep control for has backward Euler discretization of has parabolic problem. SIAM J.Nume.Anal, 27, pp277-291, 1990.
- X. DESROCHES. Estimators of errors in linear elasticity. Note HI75/93/118, 1993.
- Mr. FORT and al. Estimate a posteriori and adaptation of grids. European review of the finite elements. Vol. 9,4,2000.
- I. BABUSKA & W. RHEINBOLT. A posteriori error estimates for the finite element method. International Newspaper for Numerical Methods in Engineering, vol. 12, pp.1597-1615, 1978.

9 Description of the versions of the document

Version Aster	Author (S) Organization (S)	Description of the modifications
01/06/00	***O. BOITEAU*** *(EDF/SINETICS)*	Initial text