Записки научных
семинаров ПОМИ
Том 508, 2021 г.

S. I. Repin

ERROR IDENTITIES FOR PARABOLIC INITIAL BOUNDARY VALUE PROBLEMS

Abstract. The paper is concerned with error identities for a class of parabolic equations. One side of such an identity is a natural measure of the distance between a function in the corresponding energy class and the exact solution of the problem in question. Another side is either directly computable or serves as a source of fully computable error bounds. Particular forms of the identities can be viewed as analogs of the hypercircle identity well known for elliptic problems. It is shown that identities possess an important consistency property. Therefore, the identities and the corresponding error estimates can be used in quantitative analysis of direct and inverse problems associated with parabolic equations. The first part of the paper deals with linear parabolic equations. A class of nonlinear problems is considered in the second part. In particular, this class includes problems, whose spatial parts are presented by the α-Laplacian operator.

§1. Introduction

Quantitative analysis of differential equations usually operates with approximations instead of exact the solutions (which are unknown except rather special cases). Therefore, it is important to understand how accurately an approximate solution found by some method represents the exact one. For this purpose, we need special mathematical tools known as estimates of deviations from the exact solution (or a posteriori estimates of the functional type). In the context of elliptic type problems, they have beed derived and comprehensively studied over the past 20 years (see [17] and many other references cited therein). First a posteriori estimates of the functional type were derived for the evolutionary heat equation in [16] (the simplest form of such an estimate is (2.7)). In [7–9] and some other publications, analogous estimates were obtained and numerically tested for more general parabolic equations. In this paper, we deduce error identities,

---

Key words and phrases: parabolic equations, deviations from exact solution, error identities, hypercircle error estimates, a posteriori estimates of the functional type.

The research was partially supported by RFBR grant No. 20-01-00397.

which imply a posteriori error estimates in terms of norms stronger than those used in the above mentioned publications.

In general terms, the problem is as follows. Consider an abstract boundary value problem $\mathcal{A}u = f$ generated by an operator $\mathcal{A} : V \to V^*$, whoose exact solution is u. Assume that $v \in V$ is a function (approximation) compared with u, so that $e := v - u$ is the error. The ultimate goal of error analysis (which is not always achievable) is to obtain the identity

$$\boldsymbol{\mu}(e) = \mathbb{F}(v, \mathcal{D}), \tag{1.1}$$

where $\boldsymbol{\mu}(e)$ is a suitable measure of the error and the functional $\mathbb{F}$ is computable and depends only on known function v and problem data $\mathcal{D}$ (domain, coefficients of the differential operator, boundary conditions, etc.). If $\mathcal{A}$ is a bounded linear operator associated with Banach spaces V and V^*, then the simplest form of (1.1) is obvious:

$$\|\mathcal{A}e\|_V = \|\mathrm{R}(v)\|_{V^*}, \tag{1.2}$$

where $\mathrm{R}(v) := \mathcal{A}v - f$ is the equation *residual*.

Practical applicability of (1.2) depends on the definitions of V and V^*. Regrettably, in the majority of cases the identity (1.2) does not generate efficient error control tools because it is unable to simultaneously satisfy two important conditions: *computability* and *consistency*. Computability means that one part of the identity (or of an error estimate generated by the identity) does not contain unknown functions (such as the exact solution u) and can be directly computed. Consistency is usually understood in the sense that an error measure (and the respective error estimate) must tend to zero for any sequence of approximations that converges to the exact solution in the natural energy space, in which this (generalized) solution is uniquely defined. Getting a fully computable and consistent error estimate for a class of problems may be a challenging task.

It is easy to illustrate difficulties related to (1.2) with the paradigm of the basic elliptic problem

$$\Delta u = f \quad \text{in } \Omega, \quad u = 0 \text{ on } \Gamma := \partial\Omega, \tag{1.3}$$

where Ω is a bounded Lipschitz domain in $\mathbb{R}^d$. If we operate with generalized solutions and define V as the energy space $\overset{\circ}{W}{}^1_2(\Omega)$, then V^* is the corresponding dual space H^{-1} supplied with a supremum type norm. Unlike integral type norms, this norm is incomputable because the supremum is taken over an infinite amount of functions. If the amount of functions is reduced to some finite dimensional subspace, then the equality in (1.2) is lost and we have some lower error bound only. An attempt to overcome this difficulty and get a computable majorant of $\|\mathrm{R}(v)\|_{V^*}$ using special properties of v (Galerkin orthogonality) is known in the literature as the *explicit residual method* (e.g., see [22]). In practice, this way leads to error indicators rather than to efficient and guaranteed error bounds.

Another form of the identity (1.2) for the problem (1.3) arises if it possesses a classical solution and v has an extra regularity, so that we can select V^* with a computable (integral type) norm. For the case $V^* = L^2(\Omega)$, such an identity has the form

$$\|\Delta e\|_\Omega = \|\Delta v - f\|_\Omega, \tag{1.4}$$

where $\|\cdot\|_\Omega$ denotes the L^2-norm. It is not difficult to deduce a similar relation for the parabolic problem

$$u_t - \Delta u + f = 0 \quad \text{in } Q_T = \Omega \times (0,T), \tag{1.5}$$

$$u(x,0) = \phi(x), \tag{1.6}$$

$$u(x,t) = 0 \text{ on } S_T := \Gamma \times (0,T). \tag{1.7}$$

Assume that a function v satisfies the initial and boundary conditions and f, u, and v are sufficiently regular, so that the residual $\mathrm{R}v := v_t - \Delta v + f$ belongs to the space $L^2(Q_T)$ and is easily computable. Then, we obtain the following error identity:

$$\|e_t - \Delta e\|_{Q_T} = \|\mathrm{R}v\|_{Q_T}.$$

The left hand side has a more transparent representation provided that u and v admit formal transformations below, namely

$$\int\limits_{Q_T} |e_t - \Delta e|^2 dx = \|e_t\|^2_{Q_T} + \|\Delta e\|^2_{Q_T} - 2\int\limits_0^T \int\limits_\Omega e_t \Delta e dx dt$$

and

$$\int\limits_0^T \int\limits_\Omega e_t \Delta e dx dt = -\int\limits_0^T \int\limits_\Omega \nabla e \cdot \nabla e_t dx dt = -\frac{1}{2}\|\nabla e(\cdot,T)\|^2_\Omega.$$

In this case, we arrive at a particular form of (1.2), where the error measure $\boldsymbol{\mu}(e)$ is decomposed into norms associated with different error components, i.e.,

$$\|e_t\|^2_{Q_T} + \|\Delta e\|^2_{Q_T} + \|\nabla e(\cdot,T)\|^2_\Omega = \|\mathrm{R}v\|^2_{Q_T}. \tag{1.8}$$

At first glance (1.8) looks attractive. However, the identities (1.4) and (1.8) have a serious drawback: they exploit strong form of the equation residual. Approximations generated by the majority of numerical methods are not adapted to minimise such norms (e.g., for the most popular finite element approximations of the Courant type the residual norm in (1.4) may be too strong and, therefore, the identity may be principally invalid). Almost all numerical methods are aimed to minimise weaker norms of residuals associated with functional spaces containing the corresponding generalised solutions (e.g., see [1, 2, 20]). Hence if the residual norm in the space V^* is selected as in (1.4), then the important *consistency* property may be lost. In such a case, the identity becomes practically useless because a sequence v_k may converge to u in the energy norm but the error measure and the residual norm may not tend to zero (they may grow or even be equal to $+\infty$).

Error identities that are both computable and consistent are derived by more sophisticated methods and have more complicated forms (see a consequent exposition in [17, 18]). For elliptic problems, they also include the error $e^* = y^* - p^*$ associated with a function y^* (approximation of the exact flux p^*). There common form is

$$\boldsymbol{\mu}(e, e^*) = \mathbb{F}(v, y^*, \mathcal{D}), \tag{1.9}$$

where $\boldsymbol{\mu}(e, e^*)$ is a combined measure of the errors e and e^* and the functional $\mathbb{F}$ depends only on approximations v and y^* (which are supposed to be known) and problem data (domain, coefficients of the differential operator, boundary conditions, etc.). The measure is a nonnegative functional that must satisfy natural conditions: $\boldsymbol{\mu}(e, e^*) \geqslant 0$ and $\boldsymbol{\mu}(e, e^*) = 0$ if and only if $v = u$ and $y^* = p^*$.

In [18], error identities of the type (1.9) (and more complicated estimates) are studied for a vide class of variational problems. However, in many cases, getting error identities in the form (1.9) is impossible because the right hand side includes additional terms that depend on e and/or e^*. In particular, the problem (1.5)–(1.7) belongs to this class. The respective identity (2.3) cannot be used directly, but implies a computable error majorant (2.7). These results (obtained in earlier publications) are briefly discussed in Section 2.

In Section 3, we deduce an advanced error identity (3.3) which operates with norm stronger than in (2.3). It yields a hypercircle type estimate for the problem (1.5)–(1.7) and two-sided bounds of errors (4.1) and (4.2).

Unlike (2.3) and (2.7), they bound a combined (primal–dual) norm of the error.

In Section 4, error identities are derived using stronger assumptions on the exact solution u and approximation v. The corresponding identity (3.15) is related to a norm stronger than in (3.3) (it additionally includes the norms $\|e_t\|^2_{Q_T}$ and $[\![\, \|e^*\|^2_\Omega \,]\!]_0^T$).

Finally, in Section 5 we consider a class of nonlinear diffusion equations, where the spatial parts are associated with monotone elliptic operators. For them, we derive the general error identity and discuss its particular form generated by the problems which spatial parts are defined by the α–Laplacian operator.

§2. Notation and background

For the scalar and vector valued functions in Ω, we use the standard Lebesgue and Sobolev spaces $L^p(\Omega)$ and $W^l_p(\Omega)$ (where $l, p \geqslant 1$) and mark them above by $\circ$ if the respective functions vanish on S_T, L^2 norms of the functions in Ω and Q_T are denoted by $\|\cdot\|_\Omega$ and $\|\cdot\|_{Q_T}$, respectively.

We use standard notation for the Bochner spaces. For a separable Banach space X endowed with the norm $\|\cdot\|_X$, the space $L^2(0,T;X)$ contains functions with the norm $\|v\|^2_{L^2(0,T;X)} := \int_0^T \|\nabla v\|^2_X \; dt < \infty$.

By $\{\!| \, g \, |\!\}_\omega$ we denote the mean value of g in $\omega \subset \Omega$ and use the notation

$$[\![g(t)]\!]_0^T := g(T) - g(0),$$

e.g., for $v = v(x,t)$ we write $[\![\|v\|_\Omega]\!]_0^T$ instead of $\|v(x,T)\|_\Omega - \|v(x,0)\|_\Omega$.

Spatial derivative of v with respect to x_i is denoted by $v_{,i}$ and time derivative by v_t or $\partial_t v$. Spatial gradient and divergence are denoted by ∇ and div, respectively.

In what follows, we use the spaces

$$W_2^{1,0}(Q_T) := L^2(0,T,W_2^1(\Omega)) \text{ and } \overset{\circ}{W}{}_2^{1,0}(Q_T) := L^2(0,T,\overset{\circ}{W}{}_2^1(\Omega))$$

supplied with the norm

$$\|w\|_{1,Q_T} := \left(\int\limits_0^T (\|\nabla w\|^2_\Omega + \|w\|^2_\Omega) \, dt \right)^{1/2}.$$

For the functions in $\overset{\circ}{W}{}_2^{1,0}(Q_T, \mathbb{R}^d)$, the norm $\|\nabla w\|_{Q_T}$ is equivalent to $\|w\|_{1,Q_T}$. Next, let

$$\overset{\circ}{W}{}_2^{1,1}(Q_T) := \left\{ w \in W_2^1(Q_T),\ w = 0 \text{ on } S_T \right\}.$$

For functions in this space we introduce weighted norms (μ, $\nu > 0$)

$$[\![\, w \,]\!]_{\mu,\nu,Q_T} := \left(\mu \|\nabla w\|_{Q_T}^2 + \nu \|w_t\|_{Q_T}^2 \right)^{1/2}. \tag{2.1}$$

Following [6], by $\overset{\circ}{W}{}_2^{\Delta,1}(Q_T)$ we denote the subspace of $\overset{\circ}{W}{}_2^{1,0}(Q_T, \mathbb{R}^d)$ that consists of the functions such that

$$\|w\|_{\Delta,1,Q_T} := \int\limits_{Q_T} (w^2 + w_t^2 + |\nabla w|^2 + (\Delta w)^2)\, dx\, dt < +\infty.$$

Also, we use functional spaces associated with vector valued functions (fluxes). They are $Y^*(Q_T) := L^2(Q_T, \mathbb{R}^d)$ and the space

$$Y^*_{\mathrm{div}}(Q_T) := \{ y^* \in Y^*(Q_T) \mid \mathrm{div} y^* \in L^2(Q_T) \}$$

supplied with the norm $\|y^*\|_{\mathrm{div},Q_T} := \left(\|y^*\|_{Q_T}^2 + \|\mathrm{div} y^*\|_{Q_T}^2 \right)^{1/2}$.

First, we consider the linear problem (1.5)–(1.7) with $f \in L^2(Q_T)$ and $\phi \in \overset{\circ}{W}{}_2^1(\Omega)$. Under these assumptions, the problem is uniquely solvable in the space $\overset{\circ}{W}{}_2^{\Delta,1}(Q_T)$ (see [5, 6]). The function u satisfies the integral identity

$$\int\limits_{Q_T} \nabla u \cdot \nabla w\, dx\, dt - \int\limits_{Q_T} u w_t\, dx\, dt + \int\limits_{\Omega} (u(x,T) w(x,T) - u(x,0) w(x,0)) dx$$
$$= \int\limits_{Q_T} f w\, dx\, dt \qquad \forall w \in V_0 := \overset{\circ}{W}{}_2^{1,1}(Q_T). \tag{2.2}$$

For all $t \in [0, T]$, the solution u has traces from $L^2(\Omega)$ on cross-sections of Q_T that continuously change with respect to t.

Let $v \in V_0$. We rewrite (2.2) in the form

$$\int\limits_{Q_T} \nabla(u - v) \cdot \nabla w\, dx dt + \int\limits_{Q_T} (u_t - v_t) w\, dx dt = \int\limits_{Q_T} (f w - \nabla v \cdot \nabla w - v_t w)\, dx dt,$$

set $w = -e = u - v$, and arrive at the identity

$$\|\nabla e\|^2_{Q_T} + \frac{1}{2}\left[\!\left[\, \|e\|^2_\Omega \right]\!\right]_0^T = \int\limits_{Q_T} (\nabla v \cdot \nabla e + v_t e - fe)\, dxdt. \tag{2.3}$$

This identity meets natural consistency requirements because both its parts tend to zero for any sequence v_k that converges to u in $\overset{\circ}{W}{}^{1,1}_2(Q_T)$. However, the right hand side of (2.3) contains unknown e, so that it is not applicable for direct measurement of errors. To get a fully computable estimate, in [16] it was suggest to split the right hand side using an additional vector valued function $y^* \in Y^*_{\rm div}(Q_T)$ (which can be viewed as an approximation of the exact flux $p^* = \nabla u$). Then (2.3) comes in the form

$$\|\nabla e\|^2_{Q_T} + \frac{1}{2}\left[\!\left[\, \|e\|^2_\Omega \right]\!\right]_0^T = \int\limits_{Q_T} (\nabla v - y^*) \cdot \nabla e\, dxdt - \int\limits_{Q_T} \mathrm{R}(v, y^*) e\, dxdt,$$

where div denotes the spatial divergence and

$$\mathrm{R}(v, y^*) := \mathrm{div} y^* + f - v_t.$$

The integrals can be estimated from above by different methods. In the simplest case, we apply the estimates

$$\Big| \int\limits_{Q_T} (\nabla v - y^*) \cdot \nabla e\, dxdt \Big| \leqslant \|\nabla v - y^*\|_{Q_T} \|\nabla e\|_{Q_T}, \tag{2.4}$$

$$\Big| \int\limits_{Q_T} \mathrm{R}(v, y^*) e\, dxdt \Big| \leqslant C(\Omega) \|\mathrm{R}(v, y^*)\|_{Q_T} \|\nabla e\|_{Q_T}, \tag{2.5}$$

where $C(\Omega)$ is a constant in the inequality[1]

$$\|w\|_\Omega \leqslant C(\Omega) \|\nabla w\|_\Omega \qquad \forall w \in \overset{\circ}{W}{}^1_2(\Omega). \tag{2.6}$$

As a result, we obtain the estimate (see [16, 17])

$$(2 - \gamma)\|\nabla e\|^2_{Q_T} + \left[\!\left[\, \|e\|^2_\Omega \right]\!\right]_0^T \leqslant \frac{1}{\gamma}\Big(\|\nabla v - y^*\|_{Q_T} + C(\Omega) \|\mathrm{R}(v, y^*)\|_{Q_T} \Big)^2, \tag{2.7}$$

[1]An easily computable bound of $C(\Omega)$ is known: if $\Omega \subset \{x \in \mathbb{R}^d \mid a_i < x_i < b_i,\ b_i - a_i = l_i\}$ then $C(\Omega) \leqslant \pi \left(\sqrt{\sum\limits_{i=1}^d \frac{1}{l_i^2}} \right)^{-1}$.

in which $\gamma \in (0, 2]$ and y^* is any vector valued function in $Y^*_{\rm div}$. Changing γ within the admissible limits generates a collection of norms with different weights.

The efficiency of (2.7) and other similar estimates in practical computations has been tested in [3, 7–9]. It was shown that the estimates provide guaranteed and realistic error bounds for different classes of numerical approximations (finite element, incremental, spectral, Iga). Certainly, the quality of (2.7) depends on the choice of y^*, which should be either defined by a proper reconstruction of the numerical flux or by minimization of the right hand side of (2.7) over a certain finite dimensional subspace of $Y^*_{\rm div}$. Several more sophisticated estimates has been also derived from (2.3). They are sharper than (2.7), but nevertheless, may also overestimate the error. In part, this is because the identity (2.3) and all the estimates that follow from it are related to only one part of the error (associated with e). Below we deduce more general error identities that also include the error e^* and show that they imply the estimates related to stronger error norms.

§3. Error identities for the problem (1.5)–(1.7)

In this section, we deduce error identities (3.3), (3.15), and (3.18) for the classical linear parabolic problem. They are based on different regularity assumptions and contain different error measures $\boldsymbol{\mu}(e, e^*)$. We show that the identities are consistent with respect to sequences of approximations that satisfy natural converging properties.

3.1. The basic identity. For $v \in V_0$ and $y^* \in Y^*_{\rm div}(Q_T)$ we rewrite (2.3) in the form

$$2\|\nabla e\|^2_{Q_T} + [\![\, \|e\|^2_\Omega \,]\!]_0^T = 2\int\limits_{Q_T} (\nabla v - y^*) \cdot \nabla e \, dxdt - 2\int\limits_{Q_T} \mathrm{R}(v, y^*) e \, dxdt \tag{3.1}$$

and notice that

$$\begin{aligned}\|e^*\|^2_{Q_T} &= \|y^* - \nabla u\|^2_{Q_T} \\ &= \|\nabla e\|^2_{Q_T} + 2\int\limits_{Q_T} \nabla e \cdot (y^* - \nabla v) dxdt + \|y^* - \nabla v\|^2_{Q_T}.\end{aligned} \tag{3.2}$$

Summation of (3.1) and (3.2) yields the identity

$$\boxed{\||(e,e^*)\||^2_{Q_T} + [\![\, \|e\|^2_\Omega \,]\!]_0^T = \|\nabla v - y^*\|^2_{Q_T} - 2\int_{Q_T} \mathrm{R}(v,y^*)e\; dxdt} \tag{3.3}$$

for the combined error norm

$$\||(e,e^*)\||_{Q_T} := \left(\|\nabla e\|^2_{Q_T} + \|e^*\|^2_{Q_T}\right)^{1/2},$$

which is stronger than in (2.3). The identity (3.3) satisfies the required convergence properties and is consistent. Indeed, let v_k and y_k^* $(k = 1, 2, ...)$ be two sequences of approximations such that $v_k(x, 0) = \phi(x)$ and

$$v_k \to u \text{ in } \overset{\circ}{W}{}_2^{1,1}(Q_T) \quad \text{and } y_k^* \to p^* \text{ in } Y^*(Q_T). \tag{3.4}$$

Then the errors $e_k := v_k - u$ and $e_k^* := y_k^* - p^*$ tend to zero in the corresponding spaces. Moreover, $e_k(\cdot, T)$ tends to zero in $L^2(\Omega)$ so that the left hand side of (3.3) tends to zero. Consider the right hand side of (3.3). Notice that

$$\begin{aligned}
\int\limits_{Q_T} \mathrm{R}(v_k, y_k^*)e\; dxdt &= \int\limits_{Q_T} (fe_k - (v_k)_t e_k - y_k^* \cdot \nabla e_k)dxdt \\
&= \int\limits_{Q_T} (fe_k - \frac{1}{2}\frac{d}{dt}e_k^2 - u_t e_k - y_k^* \cdot \nabla e_k)dxdt \\
&= \int\limits_{Q_T} (fe_k - u_t e_k - p^* \cdot \nabla e_k - \frac{1}{2}\frac{d}{dt}e_k^2 + e_k^* \cdot \nabla e_k)dxdt \\
&= \int\limits_{Q_T} (e_k^* \cdot \nabla e_k - \frac{1}{2}\frac{d}{dt}e_k^2)dxdt.
\end{aligned}$$

Hence the right hand side of (3.3) is equal to

$$\|\nabla v_k - y_k^*\|^2_{Q_T} + \|e_k(\cdot, T)\|^2_\Omega - 2\int\limits_{Q_T} e_k^* \cdot \nabla e_k dxdt.$$

In view of (3.4) $\nabla v_k \to \nabla u = p^*$ in $Y^*(Q_T)$, $\|e_k^*\|_{Q_T} \to 0$ and the quantity tends to zero as $k \to +\infty$.

3.2. The hypercircle identity. A particular form of (3.3) can be viewed as an analog of the hypercircle error identity known for elliptic problems (see [10, 11])[2]. Define the set

$$\mathbb{Q}_f := \Big\{ (v, y^*) \in V_0 \times Y^*_{\mathrm{div}}(Q_T) \;\Big|\; \int\limits_{Q_T} (y^* \cdot \nabla w - f w + v_t w) dx dt = 0 \; \forall\, w \in V_0 \Big\}. \tag{3.5}$$

Using (3.3), we conclude that for any pair $(v, y^*) \in \mathbb{Q}_f$ it holds

$$|\!|\!|(e, e^*)|\!|\!|^2_{Q_T} + [\![\, \|e\|^2_\Omega \,]\!]_0^T = \|\nabla v - y^*\|^2_{Q_T}. \tag{3.6}$$

We see that in the parabolic case the error measure additionally includes the term $[\![\, \|e\|^2_\Omega \,]\!]_0^T$. The right hand side of (3.6) contains only known functions v and y^* and is fully computable. Obviously, (3.6) is also consistent. However, from the practical point of view this identity has the same drawback as the one for the elliptic case: $\mathbb{Q}_f$ contains a differential condition, which must be satisfied a.e. in Q_T. In Section 4, we will show how to overcome this difficulty and deduce computable error bounds out of the identity (3.3) and other more general identities discussed below.

3.3. Identities using additional regularity. Now we consider the case, where the solution and its approximations possess additional differentiability in time. Let u and v belong to the set

$$\overset{\circ}{W}{}^{1,1+}_2(Q_T) := \{ w \in \overset{\circ}{W}{}^{1,1}_2(Q_T) \;|\; \partial_t w_{,i} \in L^2(\Omega) \quad i = 1, 2, ..., d \}.$$

The exact flux and its approximations are also assumed to be more regular, so that

$$p^*, y^* \in Y^{*+}_{\mathrm{div}} := \{ y^* \in Y^*_{\mathrm{div}} \;|\; y^*_t \in L^2(Q_T, \mathbb{R}^d) \}.$$

[2]For the problem (1.3) the hypercircle identity reads $\|\nabla e\|^2_\Omega + \|e^*\|^2_\Omega = \|\nabla v - y^*\|^2_\Omega$ for any $v \in \overset{\circ}{W}{}^1_2(\Omega)$ and $y^* \in H(\Omega, \mathrm{div})$ such that $\mathrm{div} y^* + f = 0$.

In view of (1.5) and (1.6), for any test function $w \in \overset{\circ}{W}{}_2^{1,1+}(Q_T)$ and $v \in \overset{\circ}{W}{}_2^{1,1+}(Q_T)$ we have the integral relation

$$\int\limits_{Q_T} \Big((u_t - v_t)\, w + \nabla(u-v)\cdot\nabla w \Big) w dxdt$$
$$= \int\limits_{Q_T} (fw - v_t w - \nabla v\cdot\nabla w)\, dxdt. \quad (3.7)$$

Let us set here $w = -e_t = u_t - v_t$. Then (3.7) implies

$$\int\limits_{Q_T} \Big(e_t^2 + \nabla e\cdot\nabla e_t \Big) dxdt = \int\limits_{Q_T} (v_t e_t + \nabla v\cdot\nabla e_t - f e_t)\, dxdt. \quad (3.8)$$

Since $e_t = 0$ on S_T, for any $y^* \in Y^{*+}_{\rm div}$ it holds

$$\int\limits_{Q_T} (e_t {\rm div} y^* + y^*\cdot\nabla e_t) dxdt = \int\limits_0^T \int\limits_\Gamma (y^*\cdot n) e_t\, dxdt = 0. \quad (3.9)$$

From (3.8) and (3.9), we obtain

$$\int\limits_{Q_T} \Big(e_t^2 + \nabla e\cdot\nabla e_t \Big) dxdt = \int\limits_{Q_T} (\nabla v - y^*)\cdot\nabla e_t\, dxdt - \int\limits_{Q_T} {\rm R}(v,y^*) e_t\, dxdt. \quad (3.10)$$

Consider the first term in the right hand side of (3.10). It is easy to see that

$$\int\limits_{Q_T} (\nabla v - y^*)\cdot\nabla e_t\, dxdt = \int\limits_{Q_T} (\nabla v - y^*)\cdot\nabla(v_t - u_t)\, dxdt$$
$$= {\rm H}(\nabla v - y^*) + \int\limits_{Q_T} (\nabla v - y^*)\cdot(y_t^* - p_t^*)\, dxdt$$
$$= {\rm H}(\nabla v - y^*) + \int\limits_{Q_T} (\nabla v - p^*)\cdot(y_t^* - p_t^*)\, dxdt + \int\limits_{Q_T} (p^* - y^*)\cdot(y_t^* - p_t^*)\, dxdt, \quad (3.11)$$

where

$${\rm H}(\nabla v - y^*) := \int\limits_{Q_T} (\nabla v - y^*)\cdot(\nabla v_t - y_t^*)\, dxdt = \frac{1}{2} \big[\!\big[\, \|\nabla v - y^*\|_\Omega^2 \,\big]\!\big]_0^T .$$

Notice that

$$\int\limits_{Q_T} (y^* - p^*) \cdot (y^*_t - p^*_t)\, dxdt = \int\limits_{Q_T} e^* \cdot e^*_t dxdt = \frac{1}{2} \left[\!\left[\, \|e^*\|^2_\Omega \right]\!\right]_0^T \tag{3.12}$$

and

$$\int\limits_{Q_T} (\nabla v - p^*) \cdot (y^*_t - p^*_t)\, dxdt = \int\limits_{Q_T} \nabla e \cdot (y^*_t - \nabla v_t) dxdt + \int\limits_{Q_T} \nabla e \cdot (\nabla v_t - p^*_t)\, dxdt$$
$$= \int\limits_{Q_T} \nabla e \cdot (y^*_t - \nabla v_t) dxdt + \int\limits_{Q_T} \nabla e \cdot \nabla e_t\, dxdt. \tag{3.13}$$

By (3.10)–(3.13), we obtain

$$2\|e_t^2\|_{Q_T} + \left[\!\left[\, \|e^*\|^2_\Omega \right]\!\right]_0^T = 2\mathrm{H}(\nabla v - y^*)$$
$$+ 2\int\limits_{Q_T} \nabla e \cdot (y^*_t - \nabla v_t) dxdt - 2\int\limits_{Q_T} \mathrm{R}(v, y^*) e_t\, dxdt. \tag{3.14}$$

Summation of (3.3) and (3.14) yields an advanced error identity (cf. (2.1))

$$[\![\, e \,]\!]_{1,2,Q_T} + \|e^*\|^2_{Q_T} + \left[\!\left[\, \|e\|^2_\Omega \right]\!\right]_0^T + \left[\!\left[\, \|e^*\|^2_\Omega \right]\!\right]_0^T = \|\nabla v - y^*\|^2_{Q_T} + 2\mathrm{H}(\nabla v - y^*)$$
$$+ 2\int\limits_{Q_T} (y^* - \nabla v)_t \cdot \nabla e\, dxdt - 2\int\limits_{Q_T} \mathrm{R}(v, y^*)(e + e_t)\, dxdt \tag{3.15}$$

whose left hand side is a combined norm generated by different norms of e and e^*. This norm is stronger than the norm in (3.3). Therefore, analysing its properties with respect to sequences of approximations in addition to (3.4) we need to impose one more condition:

$$\|(\nabla v_k)_t\|_{Q_T} \text{ and } \|(y^*_k)_t\|_{Q_T} \quad \text{are bounded in } L^2(Q_T, \mathbb{R}^d). \tag{3.16}$$

This condition is related to the behaviour of approximations. It is not very demanding and can be easily verified in practical computations. With such an additional requirement imposed on the approximation sequences both sides of the identity tend to zero as $k \to +\infty$ (for the last norm in the left hand side it follows from (3.12)).

At the end of this section, we deduce an error identity related to a norm even stronger than in (3.15). Notice that

$$\|\mathrm{R}(v,y^*)\|^2_{Q_T}+\|e_t\|^2_{Q_T}=\|\mathrm{div} y^*+f\|^2_{Q_T}-2\int\limits_{Q_T}(\mathrm{div} y^*+f)v_t dxdt$$

$$+2\|v_t\|^2_{Q_T}+\|u_t\|^2_{Q_T}-2\int\limits_{Q_T}v_tu_tdxdt=\|\mathrm{div} y^*+f\|^2_{Q_T}-2\int\limits_{Q_T}(\mathrm{div} y^*+f)u_tdxdt$$

$$+2\int\limits_{Q_T}(\mathrm{div} y^*+f)(u_t-v_t)dxdt+\|u_t\|^2_{Q_T}+2\int\limits_{Q_T}v_t(v_t-u_t)dxdt$$

$$=\|\mathrm{R}(u,y^*)\|^2_{Q_T}-2\int\limits_{Q_T}(\mathrm{div} y^*+f-v_t)e_tdxdt$$

$$=\|\mathrm{div} e^*\|^2_{Q_T}-2\int\limits_{Q_T}\mathrm{R}(v,y^*)e_tdxdt. \quad (3.17)$$

We withdraw (3.17) from (3.15) and obtain

$$[\![\, e \,]\!]^2_{1,1,Q_T}+\|e^*\|^2_{\mathrm{div},Q_T}+\left[\!\left[\,\|e\|^2_\Omega\,\right]\!\right]_0^T+\left[\!\left[\,\|e^*\|^2_\Omega\,\right]\!\right]_0^T$$
$$=\|\nabla v-y^*\|^2_{Q_T}+2\mathrm{H}(\nabla v-y^*)+\|\mathrm{R}(v,y^*)\|^2_{Q_T}$$
$$+2\int\limits_{Q_T}(y^*-\nabla v)_t\cdot\nabla e\,dxdt-2\int\limits_{Q_T}\mathrm{R}(v,y^*)e\,dxdt. \quad (3.18)$$

This identity operates with the full primal–dual error norm in $\overset{\circ}{W}{}^{1,1+}_2(Q_T)\times Y^*_{\mathrm{div}}(Q_T)$. Now the condition (3.16) is not sufficient to guarantee that the right hand side of (3.18) tend to zero for approximation sequences (3.4). It is additionally required that

$$y^*_k\to p^* \text{ in } Y^*_{\mathrm{div}}(Q_T). \quad (3.19)$$

This condition is more demanding. In the next section, we discuss a way to overcome difficulties generated by (3.19).

§4. Error estimates for the problem (1.5)–(1.7)

Error identities (3.3), (3.15), and (3.18) contain unknown function e in their right hand sides. Therefore, they cannot be directly applied for analysis of approximation errors. However, they serve as the basis for deriving fully computable error estimates, which we deduce below.

4.1. Estimates generated by (3.3). First, we discuss computable bounds of errors that follow from (3.3). The last term in the right hand side of (3.3) is the only one that contains unknown function u. We use (2.5) and Young's inequality with $\beta \in (0,1]$ to estimate it. This way yields the following result.

Theorem 1. *For any $v \in V_0$, $y^* \in Y^*_{\rm div}(Q_T)$, and $\beta \in (0,1]$ it holds*

$$(1-\beta)\|\nabla e\|^2_{Q_T} + \|e^*\|^2_{Q_T} + \big[\!\big[\, \|e\|^2_\Omega \,\big]\!\big]_0^T \leqslant \mathbb{M}_1(v, y^*, \beta), \tag{4.1}$$

where

$$\mathbb{M}_1(v, y^*, \beta) := \|\nabla v - y^*\|^2_{Q_T} + \frac{1}{\beta} C^2(\Omega)\|\mathrm{R}(v, y^*)\|^2_{Q_T}.$$

It is easy to see that $\mathbb{M}_1(v, y^*, \beta) = 0$ if and only if $v = u$ and $y^* = p^*$.

Remark 1. By (3.3) and (2.5) we can also deduce a simple minorant of the error.

$$(1+\beta)\|\nabla e\|^2_{Q_T} + \|e^*\|^2_{Q_T} + \big[\!\big[\, \|e\|^2_\Omega \,\big]\!\big]_0^T \geqslant \|\nabla v - y^*\|^2_{Q_T} - \frac{1}{\beta} C^2(\Omega)\|\mathrm{R}(v, y^*)\|^2_{Q_T}. \tag{4.2}$$

Also, (4.1) implies the estimate

$$\max_{t\in[0,T]} \big[\!\big[\, \|e(\cdot,t)\|^2_\Omega \,\big]\!\big]_0^T \leqslant \|\nabla v - y^*\|^2_{Q_T} + C^2(\Omega)\|\mathrm{R}(v, y^*)\|^2_{Q_T}.$$

In order to guarantee that $\mathbb{M}_1(v_k, y^*_k, \beta)$ tends to zero we need (3.19) in addition to the standard condition (3.4). A way to avoid this extra requirement is to make a suitable correction of the flux y^*. For this purpose we introduce a "correction function" τ^* and define a modified majorant

$$\mathbb{M}_1(v, y^*, \tau^*, \beta) := \|\nabla v - y^*\|^2_{Q_T} + \frac{1}{\beta} C^2(\Omega)\|\mathrm{R}(v, y^*) + {\rm div}\tau^*\|^2_{Q_T} + \frac{1}{\gamma}\|\tau^*\|^2_{Q_T}.$$

Theorem 2. *For any positive β and γ satisfying $\beta+\gamma \leqslant 1$, $v \in V_0$, $y^* \in Y^*_{\rm div}(Q_T)$, and $\tau^* \in Y^*_{\rm div}(Q_T)$ it holds*

$$(1-\beta-\gamma)\|\nabla e\|^2_{Q_T} + \|e^*\|^2_{Q_T} + [\![\, \|e\|^2_\Omega \,]\!]_0^T \leqslant \mathbb{M}_1(v, y^*, \tau^*, \beta). \tag{4.3}$$

Moreover,

$$\inf_{\tau^* \in Y^*_{\rm div}} \mathbb{M}_1(v, y^*, \tau^*, \beta) \leqslant \mathrm{E}\Big(|\!|\!|(e, e^*)|\!|\!|^2_{Q_T} + \|e_t\|^2_{Q_T} \Big), \tag{4.4}$$

where E is defined in (4.8).

Proof. We modify the last term in (3.3) with the help of τ^*

$$\int\limits_{Q_T} \mathrm{R}(y^*, v)(v-u)dxdt$$

$$= \int\limits_{Q_T} (\mathrm{div}(y^*+\tau^*) + f - v_t)(v-u)dxdt - \int\limits_{Q_T} \tau^* \cdot \nabla(v-u)dxdt,$$

use the estimate

$$2\left| \int\limits_{Q_T} \tau^* \cdot \nabla(v-u)dxdt \right| \leqslant \gamma \|\nabla e\|^2_{Q_T} + \frac{1}{\gamma}\|\tau^*\|^2_{Q_T},$$

and get (4.3).

To prove (4.4) we set $\tau^* = \nabla w_\tau$, where $w_\tau(x,t)$ solves the problem

$$\Delta w_\tau = v_t - f - \mathrm{div} y^*, \qquad w_\tau = 0 \text{ on } S_T. \tag{4.5}$$

for $t \in (0, T]$. Then the second term of $\mathbb{M}_1(v, y^*, \tau^*, \beta)$ vanishes. To estimate the last term, we notice that

$$\int\limits_{Q_T} \nabla w_\tau \cdot \nabla w_\tau dxdt = \int\limits_{Q_T} (\mathrm{div} y^* + f - v_t) w_\tau dxdt = \int\limits_{Q_T} (\mathrm{div} e^* - e_t) w_\tau dxdt.$$

Hence

$$\|\tau^*\|^2_{Q_T} = \|\nabla w_\tau\|^2_{Q_T} \leqslant (C(\Omega)\|e_t\|_{Q_T} + \|e^*\|_{Q_T})\|\nabla w_\tau\|_{Q_T}$$

and we find that

$$\frac{1}{\gamma}\|\tau^*\|^2_{Q_T} \leqslant \frac{2}{\gamma}\left(C^2(\Omega)\|e_t\|^2_{Q_T} + \|e^*\|^2_{Q_T} \right). \tag{4.6}$$

For the sequences satisfying (3.4) this term tends to zero. Since

$$\|\nabla v - y^*\|^2_{Q_T} \leqslant 2(\|\nabla e\|^2_{Q_T} + \|e^*\|^2_{Q_T}), \tag{4.7}$$

we arrive at (4.4) with

$$\mathrm{E} = \frac{2}{\gamma} \max\{1, C^2(\Omega)\}. \tag{4.8}$$

The estimate (4.4) shows that there always exist corrections such that the right hand side of (4.3) tends to zero for any sequence of approximations satisfying (3.4). □

4.2. Estimates generated by (3.15) and (3.18). The identity (3.15) also yields computable bounds of the errors e and e^*.

For the functions $(v, y^*) \in \overset{\circ}{W}{}_2^{1,1+}(Q_T) \times Y_{\mathrm{div}}^{*+}(Q_T)$ we define the functionals

$$\begin{aligned}\mathbb{M}_2(v, y^*, \alpha, \beta, \gamma) := \|\nabla v - y^*\|^2_{Q_T} + 2\mathrm{H}(\nabla v - y^*)\\ + \left(\frac{1}{\alpha} + \frac{C^2(\Omega)}{\beta}\right) \|\mathrm{R}(v, y^*)\|^2_{Q_T} + \frac{1}{\gamma}\|(y^* - \nabla v)_t\|^2_{Q_T}\end{aligned}$$

and

$$\begin{aligned}\mathbb{M}_3(v, y^*, \mu) := \|\nabla v - y^*\|^2_{Q_T} + 2\mathrm{H}(\nabla v - y^*) + \|\mathrm{R}(v, y^*)\|^2_{Q_T}\\ + \frac{1}{\mu}\left(\|(y^* - \nabla v)_t\|^2_{Q_T} + \left(C_P^2(Q_T^m)\|\mathrm{R}(v, y^*)\|^2_{Q_T^m}\right)^{1/2}\right).\end{aligned}$$

Here Q_T^m, $m = 1, 2, ..., M$ denote nonintersecting Lipschitz subdomains of Q_T such that $\overline{Q}_T = \cup_{k=1}^N \overline{Q}_T^m$ and $C_P(Q_T^m)$ are constants in (4.17).

Theorem 3. *For any $v \in \overset{\circ}{W}{}_2^{1,1+}(Q_T)$ and $y^* \in Y_{\mathrm{div}}^{*+}(Q_T)$ and positive α, β, γ, and μ such that $0 < \beta + \gamma \leqslant 1$, $\alpha \leqslant 2$, and $\mu \leqslant 1$ the following estimates hold*

$$\begin{aligned}[\![e]\!]^2_{(1-\beta-\gamma),(2-\alpha),Q_T} + \|e^*\|^2_{Q_T} + \left[\!\left[\|e\|^2_\Omega \right]\!\right]_0^T + \left[\!\left[\|e^*\|^2_\Omega \right]\!\right]_0^T\\ \leqslant \mathbb{M}_2(v, y^*, \alpha, \beta, \gamma),\end{aligned} \tag{4.9}$$

$$\begin{aligned}[\![e]\!]^2_{(1-\beta-\gamma),1,Q_T} + \|e^*\|^2_{\mathrm{div},Q_T} + \left[\!\left[\|e\|^2_\Omega \right]\!\right]_0^T + \left[\!\left[\|e^*\|^2_\Omega \right]\!\right]_0^T\\ \leqslant \mathbb{M}_2(v, y^*, 1, \beta, \gamma).\end{aligned} \tag{4.10}$$

If in addition Q_T is divided into a collection of nonintersecting space-time subdomains Q_T^m, and v and y^* satisfy (4.16), then

$$\begin{aligned}(1 - \mu) [\![e]\!]^2_{(1,1,Q_T} + \|e^*\|^2_{Q_T} + \left[\!\left[\|e\|^2_\Omega \right]\!\right]_0^T + \left[\!\left[\|e^*\|^2_\Omega \right]\!\right]_0^T\\ \leqslant \mathbb{M}_3(v, y^*, \mu).\end{aligned} \tag{4.11}$$

Proof. The last two integrals in (3.15) can be estimated by the Cauchy and Young's inequalities

$$\int\limits_{Q_T} \mathrm{R}(v, y^*) e_t dxdt \leqslant \frac{1}{2\alpha} \|\mathrm{R}(v, y^*)\|^2_{Q_T} + \frac{\alpha}{2} \|e_t\|^2_{Q_T}, \tag{4.12}$$

$$\int\limits_{Q_T} \mathrm{R}(v, y^*) e dxdt \leqslant \frac{C^2(\Omega)}{2\beta} \|\mathrm{R}(v, y^*)\|^2_{Q_T} + \frac{\beta}{2} \|\nabla e\|^2_{Q_T}, \tag{4.13}$$

$$\int\limits_{Q_T} (y^* - \nabla v)_t \cdot \nabla e \, dxdt \leqslant \frac{1}{2\gamma} \|(y^* - \nabla v)_t\|^2_{Q_T} + \frac{\gamma}{2} \|\nabla e\|^2_{Q_T}, \tag{4.14}$$

and we arrive at (4.9).

Next, we use (4.13), (4.14), and (3.18) and obtain the estimate

$$\begin{aligned} [\![e]\!]^2_{(1-\beta-\gamma),1,Q_T} + \|e^*\|^2_{\mathrm{div},Q_T} + \big[\!\big[\|e\|^2_\Omega \big]\!\big]_0^T + \big[\!\big[\|e^*\|^2_\Omega \big]\!\big]_0^T \\ \leqslant \|\nabla v - y^*\|^2_{Q_T} + 2\mathrm{H}(\nabla v - y^*) \\ + \left(1 + \frac{C^2(\Omega)}{\beta}\right) \|\mathrm{R}(v, y^*)\|^2_{Q_T} + \frac{1}{\gamma} \|(y^* - \nabla v)_t\|^2_{Q_T}, \end{aligned}$$

which is (4.10).

It remains to prove (4.11). Using (4.12) with $\alpha = 1$, we can exclude e_t from the right hand side and rewrite (3.15) in the form

$$\begin{aligned} [\![e]\!]^2_{1,1,Q_T} + \big[\!\big[\|e\|^2_\Omega \big]\!\big]_0^T + \|e^*\|^2_{Q_T} + \big[\!\big[\|e^*\|^2_\Omega \big]\!\big]_0^T \\ \leqslant \|\nabla v - y^*\|^2_{Q_T} + 2\mathrm{H}(\nabla v - y^*) + \|\mathrm{R}(v, y^*)\|^2_{Q_T} \\ + 2 \int\limits_{Q_T} (y^* - \nabla v)_t \cdot \nabla e \, dxdt - 2 \int\limits_{Q_T} \mathrm{R}(v, y^*) e \, dxdt, \end{aligned} \tag{4.15}$$

This form is convenient for deriving an error majorant useful for approximations constructed by the domain decomposition method (e.g., see [12, 21]) using space-time decomposition of Q_T. Assume that the approximations v and y^* are integrally balanced and satisfy the conditions

$$\int\limits_{Q_T^m} \mathrm{R}(v, y^*) \, dxdt = 0 \qquad \forall\, m = 1, 2, \ldots, M. \tag{4.16}$$

Then, we can estimate the last two terms in (3.15) by the inequality

$$\|w\|_{Q_T^m} \leqslant C_P(Q_T^m) \, [\![w]\!]_{1,1,Q_T^m}, \tag{4.17}$$

which holds for any $w \in \overset{\circ}{W}{}^{1,1}_2(Q^m_T)$ provided that $\{\!| w |\!\}_{Q^m_T} = 0$. Here $C_P(Q^m_T)$ is the Poincare constant associated with Q^m_T and the norm at the right is reduction of $[\![e]\!]_{1,1,Q_T}$ to this subdomain. We have

$$\int\limits_{Q_T} \mathrm{R}(v,y^*)e\, dxdt \leqslant \sum_{k=1}^{N} C_P(Q^m_T)\|\mathrm{R}(v,y^*)\|_{Q^m_T} [\![e]\!]_{1,1,Q^m_T}$$

$$\leqslant \left(C^2_P(Q^m_T)\|\mathrm{R}(v,y^*)\|^2_{Q^m_T}\right)^{1/2} [\![e]\!]_{1,1,Q_T}.$$

Hence the last two integrals in (4.15) are bounded by the quantity

$$2\left(\|(y^*-\nabla v)_t\|_{Q_T} + \left(C^2_P(Q^m_T)\|\mathrm{R}(v,y^*)\|^2_{Q^m_T}\right)^{1/2}\right) [\![e]\!]_{1,1,Q_T}.$$

By Youngs's inequality with $\gamma \in (0,1]$ we arrive at (4.11). □

4.3. Comments. It is worth adding several comments related to practical applications of the estimates in Theorems 2 and 3.

1. A suitable correction function in (4.3) can be found either by direct minimization of $\mathbb{M}_1(v,y^*,\tau^*,\beta)$ with respect a finite dimensional subspace of Y^*_{div}, or by solving (4.5) approximately and setting τ^* as a suitable reconstruction of the numerical flux (for elliptic problems this way is discussed and tested in [19]).

2. Estimates (4.9) and (4.10) are based on extra differentiability of exact solutions and approximations with respect to t. The corresponding majorants $\mathbb{M}_2$ and $\mathbb{M}_3$ contain an additional term $\|(y^*-\nabla v)_t\|^2_{Q_T}$. Therefore, it is natural to impose an additional condition on sequences of approximations v_k and y^*_k in (3.4), namely

$$\|(y^*_k - \nabla v_k)_t\|_{Q_T} \to 0 \qquad \text{as } k \to +\infty. \tag{4.18}$$

We outline that this condition is related to approximations y^*_k and v_k (which are known), so that in practice (4.18) is fully controllable. It is especially simple to do in the commonly used case of time–incremental approximation, where $[0,T] = \cup^n_{j=1}\overline{I}_j$, $I_j = (t_j, t_{j+1})$, $t_{j+1} > t_j$, $\tau_j = t_{j+1} - t_j$ the approximations are presented in the form

$$v = v_j + \frac{v_{j+1}-v_j}{\tau_j}(t-t_j); \qquad y^* = y^*_j + \frac{y^*_{j+1}-y^*_j}{\tau_j}(t-t_j) \qquad t \in I_j.$$

Hence the contribution to the norm $\|(y^*_k - \nabla v_k)_t\|^2_{Q_T}$ associated with the interval I_j is $\frac{1}{\tau_j}\|(y^*_{j+1} - \nabla v_{j+1} - y^*_j + \nabla v_j\|^2_{\Omega}$. The condition (4.18) will be satisfied provided that the time steps and errors in the relations $y_j^* = \nabla v_j$ are coordinated such that $\|y_j^* - \nabla v_j\|_\Omega$ decreases faster than $\tau_j^{1/2}$.

3. If $Q_T^m = \Omega_m \times (t_{j+1}, t_j)$ and Ω_m is a convex subdomain of Ω, then the constant in (4.17) is easy to estimate

$$C_P(Q_T^m) \leqslant \frac{1}{\pi} \max\{\mathrm{diam}\Omega,\ \tau_j\}.$$

This estimate follows from the well known estimate established in [13].

§5. Error identity for a class of nonlinear problems

5.1. Evolutionary problem generated by convex potentials. Now we consider a class of initial boundary value problems

$$u_t - \mathrm{div} p^* = f \text{ in } Q_T, \tag{5.1}$$

$$u(x,t) = 0 \qquad \text{on } S_T, \tag{5.2}$$

$$u(x,0) = \phi, \tag{5.3}$$

where the vector valued function p^* is joined with ∇u via the relation

$$\mathcal{D}_g(\nabla u, p^*) = 0 \qquad \text{in } Q_T. \tag{5.4}$$

Here $\mathcal{D}_g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a nonnegative functional defined by the relation

$$\mathcal{D}_g(q, q^*) := g(q) + g^*(q^*) - q \cdot q^*,$$

$g : \mathbb{R}^d \to \mathbb{R}$ is a convex function (potential) and g^* is the corresponding dual function (in the sense of Young–Fenchel). The relation (5.4) forms the constitutive relation that connects spatial gradient with the flux. If g is differentiable, then (5.4) amounts $p^* = g'(\nabla u)$. In particular, if $g(q) = \frac{1}{2}|q|^2$, then (5.1)–(5.4) coincides with (1.5)–(1.7).

Let V and Y^* be suitable Banach spaces, which concrete forms depend on the structure and properties of g and g^*. In particular, for $v \in V$ and $y^* \in Y^*$ the functionals

$$\int\limits_{Q_T} g(\nabla v)dxdt \quad \text{and} \quad \int\limits_{Q_T} g^*(y^*)dxdt$$

must be well defined and finite and the product $\nabla v \cdot y^*$ must be summable in Q_T. As before, the space V_0 denotes a subspace of V containing the functions vanishing on S_T.

Assume that the problem is well posed in the sense that the external data and the space V are selected such that these exists a unique generalised solution $u \in V \cap C([0,T], L^2(\Omega))$ that satisfies (5.2), (5.3), (5.4) and the relation

$$\int\limits_{Q_T} (u_t w + p^* \cdot \nabla w) dxdt = \int\limits_{Q_T} f w dxdt \qquad \forall w \in V_0(Q_T) \tag{5.5}$$

5.2. The error identity. Let $v \in V_0$ and $y^* \in Y^*$. We have

$$\int\limits_{Q_T} \mathcal{D}_g(\nabla v, y^*) dt = \int\limits_{Q_T} (g(\nabla v) + g^*(y^*) - \nabla v \cdot y^*) dxdt$$

$$= \int\limits_{Q_T} (g(\nabla u) + g^*(y^*) - \nabla u \cdot y^*) dxdt + \int\limits_{Q_T} (g(\nabla v) + g^*(p^*) - \nabla v \cdot p^*) dxdt$$

$$+ \int\limits_{Q_T} (\nabla u \cdot y^* + \nabla v \cdot p^*) dxdt - \int\limits_{Q_T} (\nabla u \cdot p^* + \nabla v \cdot y^*) dxdt$$

$$= \int\limits_{Q_T} (\mathcal{D}_g(\nabla u, y^*) + \mathcal{D}_g(\nabla v, p^*) dt + \int\limits_{Q_T} \nabla(v-u) \cdot (p^* - y^*) dxdt.$$

Using (5.5) we find that

$$\int\limits_{Q_T} \nabla(v-u) \cdot (p^* - y^*) dxdt = \int\limits_{Q_T} \mathrm{div}(y^* - p^*)(v-u) dxdt$$

$$= \int\limits_{Q_T} (\mathrm{div} y^* + f - u_t)(v-u) dxdt = \int\limits_{Q_T} \mathrm{R}(y^*, v)(v-u) dxdt + \frac{1}{2} [\![\|e\|_\Omega]\!]_0^T$$

and obtain the identity

$$\boxed{\boldsymbol{\mu}(e, e^*) + \frac{1}{2} [\![\|e\|_\Omega]\!]_0^T = \int\limits_{Q_T} \Big(\mathcal{D}_g(\nabla v, y^*) - \mathrm{R}(y^*, v) e \Big) dxdt,} \tag{5.6}$$

where

$$\boldsymbol{\mu}(e, e^*) := \int\limits_{Q_T} \Big(\mathcal{D}_g(\nabla u, y^*) + \mathcal{D}_g(\nabla v, p^*) \Big) dt.$$

It is easy to see that (5.6) is a generalization of (3.3) (in the linear case, the error measure $\boldsymbol{\mu}(e,e^*)$ coincides with $|||(e,e^*)|||^2_{Q_T}$). Also, the identity (5.6) generalizes the identity derived for nonlinear elliptic boundary value problems ([14, 15], see also [18]). The right hand side has two terms: the first term is directly computable and the second one can be estimated by the same method as in Sect. 4. We show this way below with the paradigm of a particular class of problems.

5.3. The hypercircle error identity. Let

$$(v,y^*) \in \mathbb{Q}_f := \Big\{(v,y^*) \in V_0 \times Y^* \Big| \int\limits_{Q_T} (y^* \cdot \nabla w - fw + v_t w)dxdt = 0 \;\forall\, w \in V_0\Big\}.$$

Then (5.6) implies the following identity:

$$\int\limits_{Q_T} (\mathcal{D}_g(\nabla u, y^*) + \mathcal{D}_g(\nabla v, p^*))dt + \frac{1}{2}[\![\, \|e\|_\Omega]\!]_0^T = \int\limits_{Q_T} \mathcal{D}_g(\nabla v, y^*)dxdt. \quad (5.7)$$

It is easy to see that this identity generalizes (3.5).

5.4. Example: α–Laplacian. An interesting particular case of the problem (5.1)–(5.4) corresponds to the power growth functionals

$$g(\nabla u) = \frac{1}{\alpha}|\nabla u|^\alpha \text{ and } g^*(p^*) = \frac{1}{\alpha^*}|p^*|^{\alpha^*} \quad \alpha > 1,\ \alpha^* = \frac{\alpha}{\alpha - 1}. \quad (5.8)$$

In this case, the system (5.1) reads

$$u_t - \operatorname{div}|\nabla u|^{\alpha-2}\nabla u = f \text{ in } Q_T. \quad (5.9)$$

Now the functions u and p^* are joined by the relation

$$p^* = |\nabla u|^{\alpha-2}\nabla u \quad \text{and} \quad \nabla u = |p^*|^{\frac{2-\alpha}{\alpha-1}}p^*. \quad (5.10)$$

Hence the identity (5.6) holds with

$$\begin{aligned}
\mathcal{D}_g(\nabla v, y^*) &= \frac{1}{\alpha}|\nabla v|^\alpha + \frac{1}{\alpha^*}|y^*|^{\alpha^*} - \nabla v \cdot y^*,\\
\mathcal{D}_g(\nabla v, p^*) &= \frac{1}{\alpha}|\nabla v|^\alpha + \frac{1}{\alpha^*}|\nabla u|^{\alpha} - \nabla v \cdot \nabla u|\nabla u|^{\alpha-2},\\
\mathcal{D}_g(\nabla u, y^*) &= \frac{1}{\alpha^*}|y^*|^{\alpha^*} + \frac{1}{\alpha}|p^*|^{\alpha^*} - p^* \cdot y^*|p^*|^{\alpha^*-2}.
\end{aligned}$$

We set

$$V_0 = \overset{\circ}{W}{}^{1,1}_{\alpha}(Q_T) := \left\{ w \in W^1_\alpha(Q_T),\ w = 0 \text{ on } S_T \right\},\ Y^*(Q_T) = L^{\alpha^*}(Q_T, \mathbb{R}^d)$$

and define

$$Y^*_{\alpha,\mathrm{div}}(Q_T) = \{ y^* \in Y^*(Q_T) \ | \ \mathrm{div} y^* \in L^{\alpha^*}(Q_T) \}.$$

Remark 2. We need specific algebraic inequalities related to power growth functionals. The first inequality is

$$\frac{1}{\alpha}|a|^\alpha + \frac{1}{\alpha^*}|b|^\alpha - a \cdot b|b|^{\alpha-2} \geqslant \Sigma_\alpha |a-b|^\alpha \qquad \forall a, b \in \mathbb{R}^d, \tag{5.11}$$

where Σ_α is a positive constant depending on $\alpha \geqslant 2$ only. Without a loss of generality one may rewrite the inequality in terms of the parameters $\lambda \geqslant 0$ and $\theta \in [-1, 1]$ assuming that $|a| = \lambda|b|$, and $a \cdot b = \theta |a||b|$. Then finding Σ_α is reduced to minimization of the quotient

$$\frac{\lambda^\alpha + \alpha - 1 - \alpha\lambda\theta}{\alpha|1 - 2\lambda\theta + \lambda^2|^{\alpha/2}}$$

with respect to all possible λ and θ (except the case $\lambda = \theta = 1$ associated with the case $a = b$). For $\alpha = 2$ the constant is equal to 0.5. Computations show that the value of Σ_α decreases if α is growing ($\Sigma_\alpha \approx 0.195$ for $\alpha = 3$, $\Sigma_\alpha \approx 0.083$ for $\alpha = 4$, $\Sigma_\alpha \approx 0.017$ for $\alpha = 6$, and $\Sigma_\alpha \approx 0.0037$ for $\alpha = 8$.)

For $\alpha \geqslant 2$, we also have the algebraic inequality (e.g., see [4])

$$|a|^\alpha + |b|^\alpha - a \cdot b(|a|^{\alpha-2} + |b|^{\alpha-2}) > \Gamma_\alpha \, |a-b|^\alpha \tag{5.12}$$

which holds with a positive constant $\Gamma_\alpha \geqslant 2\Sigma_\alpha$. It is easy to see that $\Gamma_2 = 1$. For other α this constant can be found numerically. Computations show that for $d = 2$ the constant meets the relation $\Gamma_\alpha \approx 2^{2-\alpha}$.

In view of (5.11), for $\alpha \geqslant 2$ we have the estimate

$$\mu(e, e^*) \geqslant \int\limits_{Q_T} \Big(\frac{1}{\alpha^*}|y^*|^{\alpha^*} + \frac{1}{\alpha}|p^*|^{\alpha^*} - p^* \cdot y^* |p^*|^{\alpha^*-2} \Big) dxdt + \Sigma_\alpha \|\nabla e\|^\alpha_{\alpha, Q_T}. \tag{5.13}$$

From (5.6) and (5.13) we deduce a simple error majorant by arguments close to those used in Section 4. We have

$$\begin{aligned}\int\limits_{Q_T} \mathrm{R}(v,y^*)e\,dxdt &\leqslant \|\mathrm{R}(v,y^*)\|_{\alpha^*,Q_T}\|e\|_{\alpha,Q_T}\\ &\leqslant C_\alpha(\Omega)\|\mathrm{R}(v,y^*)\|_{\alpha^*,Q_T}\|\nabla e\|_{\alpha,Q_T}\\ &\leqslant \frac{C_\alpha^{\alpha^*}(\Omega)}{\alpha^*\gamma}\|\mathrm{R}(v,y^*)\|^{\alpha^*}_{\alpha^*,Q_T}+\frac{\gamma}{\alpha}\|\nabla e\|^{\alpha}_{\alpha,Q_T}. \quad (5.14)\end{aligned}$$

In (5.14), we have used the algebraic inequality $\zeta\zeta^* \leqslant \frac{\gamma}{\alpha}|\zeta|^\alpha + \frac{1}{\alpha^*\gamma}|\zeta^*|^{\alpha^*}$ with positive γ and a generalised version of the Friedrichs inequality

$$\|w\|_{\alpha,\Omega} \leqslant C_\alpha(\Omega)\|\nabla w\|_{\alpha,\Omega} \qquad \forall w \in \overset{\circ}{W}{}^1_\alpha(\Omega),$$

where the constant $C_\alpha(\Omega) > 0$ depends on α, d, and Ω only.

Let $\beta \in (0, \alpha\Sigma_\alpha)$. Then from (5.6), (5.13), and (5.14) it follows that

$$\begin{aligned}&\left(\Sigma_\alpha - \frac{\beta}{\alpha}\right)\|\nabla e\|^\alpha_{Q_T} + \int\limits_{Q_T}\mathcal{D}_g(\nabla u, y^*)dxdt + \frac{1}{2}\|e(\cdot,T)\|^2_\Omega\\ &\leqslant \frac{1}{2}\|v-\phi\|^2_\Omega + \int\limits_{Q_T}\mathcal{D}_g(\nabla v, y^*)dxdt + \frac{C_\alpha^{\alpha^*}(\Omega)}{\alpha^*\beta}\|\mathrm{R}(v,y^*)\|^{\alpha^*}_{\alpha^*,Q_T}. \quad (5.15)\end{aligned}$$

Again, the right hand side contains only known functions and can be directly computed. Notice that the error majorant (4.1) derived for the linear equation is a particular form of (5.15). Indeed, if $\alpha = 2$ then $\alpha^* = 2$, $\Sigma_\alpha = \frac{1}{2}$, $C_\alpha = C$, and $\mathcal{D}_g(\nabla u, y^*) = \frac{1}{2}|p^* - y^*|^2$. Hence multiplying (5.15) by 2 we obtain (4.1). Certainly this simple estimate is not optimal and has the same drawbacks as the estimate (4.1). They could be overcame by the same method as in Section 4. However, a consequent consideration of this question is beyond the framework of the present paper.

5.5. Errors generated by initial data. Finally, we discuss one special application of the error identity (5.6). Assume that the functions $\widetilde{u}$ and $\widetilde{p^*}$ solve the problem (5.1), (5.2), and (5.4) with the initial condition

$$\widetilde{u}(x,0) = \widetilde{\phi}(x) \neq \phi(x).$$

Then,

$$\mathcal{D}_g(\nabla\widetilde{u}, \widetilde{p^*}) = 0 \quad \text{and} \quad \mathrm{div}\widetilde{p^*} + f - \widetilde{u}_t = 0$$

and from (5.6) we obtain

$$\int_0^T (\mathcal{D}_g(\nabla u, \widetilde{p^*}) + \mathcal{D}_g(\nabla \widetilde{u}, p^*))dt + \frac{1}{2}\|(\widetilde{u} - u)(\cdot, T)\|^2_\Omega = \frac{1}{2}\|\widetilde{\phi} - \phi\|^2_\Omega. \quad (5.16)$$

Notice that for any interval (t_1, t_2), $t_2 > t_1$, the quantity

$$\boldsymbol{\mu}_{t_1,t_2}(e, e^*) := \int_{t_1}^{t_2} \int_\Omega (\mathcal{D}_g(\nabla u, \widetilde{p^*}) + \mathcal{D}_g(\nabla u, \widetilde{p^*}))dxdt$$

is nonnegative. Hence the first integral in (5.16) is non–decreasing with respect to T. Since the right hand side does not depend on T, the norm $\|(\widetilde{u} - u)(\cdot, T)\|_\Omega$ is a non-increasing function of T. Moreover, for any positive h the quantity $\boldsymbol{\mu}_{t+h,t}(e, e^*)$ tends to zero as $t \to +\infty$. Using properties of g and g^* associated with a concrete class of problems one can deeper investigate convergence of $\widetilde{u}$ to u by estimating the first term in (5.16).

Consider the functions g and g^* defined by (5.8). Since $\widetilde{u}$ and $\widetilde{p^*}$ satisfy the condition $\widetilde{p^*} = |\nabla \widetilde{u}|^{\alpha-2}\nabla \widetilde{u}$, we have

$$\mathcal{D}_g(\nabla u, \widetilde{p^*}) = \frac{1}{\alpha}|\nabla u|^\alpha + \frac{1}{\alpha^*}|\nabla \widetilde{u}|^\alpha - \nabla \widetilde{u} \cdot \nabla u |\nabla \widetilde{u}|^{\alpha-2},$$
$$\mathcal{D}_g(\nabla \widetilde{u}, p^*) = \frac{1}{\alpha}|\nabla \widetilde{u}|^\alpha + \frac{1}{\alpha^*}|\nabla u|^\alpha - \nabla \widetilde{u} \cdot \nabla u |\nabla u|^{\alpha-2}.$$

Therefore,

$$\mathcal{D}_g(\nabla v, p^*) + \mathcal{D}_g(\nabla u, y^*) = |\nabla v|^\alpha + |\nabla u|^\alpha - \nabla v \cdot \nabla u(|\nabla u|^{\alpha-2} + |\nabla v|^{\alpha-2}).$$

Consider the case $\alpha > 2$. In view of (5.12), we have a simple bound for the error measure

$$\boldsymbol{\mu}(e, e^*) = \int_{Q_T} (\mathcal{D}_g(\nabla v, p^*) + \mathcal{D}_g(\nabla u, y^*))dxdt \;\geqslant\; \Gamma_\alpha \|\nabla(v - u)\|^\alpha_{Q_T}. \quad (5.17)$$

By (5.17) and (5.16) we obtain the estimate

$$\Gamma_\alpha \|\nabla e\|^\alpha_{\alpha, Q_T} + \frac{1}{2}\|e(\cdot, T)\|^2_\Omega \leqslant \frac{1}{2}\|v - \phi\|^2_\Omega, \quad (5.18)$$

whose right hand side does not depend on T. Hence the norm $\|\nabla e\|_{\alpha, Q_T}$ is uniformly bounded with respect to T and, therefore, ∇e (and e) must decrease to zero as $t \to +\infty$.

Assume that $1 < \alpha < 2$. In this case, $\alpha^* > 2$ and it is convenient to rewrite the error identity in terms of fluxes. Using (5.10) and the relation $\nabla\widetilde{u} = |\widetilde{p^*}|^{\frac{2-\alpha}{\alpha-1}}\widetilde{p^*}$, we find that

$$\mathcal{D}_g(\nabla\widetilde{u}, p^*) = \int\limits_{Q_T} (\frac{1}{\alpha}|\widetilde{p^*}|^{\alpha^*} + \frac{1}{\alpha^*}|p^*|^{\alpha^*} - p^* \cdot \widetilde{p^*}|\widetilde{p^*}|^{\alpha^*-2})dxdt$$

$$\mathcal{D}_g(\nabla u, \widetilde{p^*}) = \int\limits_{Q_T} (\frac{1}{\alpha^*}|\widetilde{p^*}|^{\alpha^*} + \frac{1}{\alpha}|p^*|^{\alpha^*} - p^* \cdot \widetilde{p^*}|p^*|^{\alpha^*-2})dxdt$$

and

$$\boldsymbol{\mu}(e, e^*) = \int\limits_{Q_T} \Big(|y^*|^{\alpha^*} + |p^*|^{\alpha^*} - p^* \cdot y^*(|y^*|^{\alpha^*-2} + |p^*|^{\alpha^*-2})\Big)dxdt.$$

In view of (5.12),

$$\boldsymbol{\mu}(e, e^*) \geqslant \Gamma_{\alpha^*}\|e^*\|^{\alpha^*}_{\alpha^*, Q_T} \tag{5.19}$$

From (5.16) and (5.19), we deduce an estimate analogous to (5.18)

$$\Gamma_{\alpha^*}\|e^*\|^{\alpha}_{\alpha, Q_T} + \frac{1}{2}\|e(\cdot, T)\|^2_{\Omega} \leqslant \frac{1}{2}\|v - \phi\|^2_{\Omega}, \tag{5.20}$$

which shows that e^* tends to zero as $t \to +\infty$.

References

1. F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Series in Computational Mathematics, 15, New York, 1991.
2. P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, 1987.
3. K. Kumar, S. Kyas, J. Nordbotten, S. Repin, *Guaranteed and computable error bounds for approximations constructed by an iterative decoupling of the Biot problem.* — Comput. Math. Appl. **91** (2021), 122–149.
4. Ning Ju, *Numerical analysis of parabolic p-laplacian: approximation of trajectories.* — SIAM J. Numer. Anal. **37**, No. 6 (2000), 1861–1884.
5. O. A. Ladyzhenskaya, V. A. Solonnikov, N. N. Uraltseva, *Linear and Quasilinear Equations of Parabolic Type*, Nauka, Moscow, 1967.
6. O. A. Ladyzhenskaya, *The boundary value problems of mathematical physics*, Springer, New York, 1985.
7. U. Langer, S. Matculevich, S. Repin, *Guaranteed error bounds and local indicators for adaptive solvers using stabilised space-time IgA approximations to parabolic problems.* — Comput. Math. Appl. **78**, No. 8 (2019), 2641–2671.
8. S. Matculevich, S. Repin, *Computable estimates of the distance to the exact solution of the evolutionary reaction-diffusion equation.* — Appl. Math. Comput. **247** (2014), 329–347.

9. S. V. Matculevich, S. I. Repin, *Estimates for the difference between exact and approximate solutions of parabolic equations on the basis of Poincare inequalities for traces of functions on the boundary.* — Differ. Equ. **52**, No. 10 (2016), 1355–1365.
10. S. G. Mikhlin, *Variational Methods in Mathematical Physics,* Pergamon Press, Oxford, 1964.
11. W. Prager, J. L. Synge, *Approximations in elasticity based on the concept of functions space.* — Quart. Appl. Math. **5** (1947), 241–269.
12. A. Quarteroni, A. Vali, *Domain decomposition methods for partial differential equations.* Numerical Mathematics and Scientific Computation. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1999.
13. L. E. Payne, H. F. Weinberger, *An optimal Poincaré inequality for convex domains.* — Arch. Rat. Mech. Anal. **5** (1960), 286–292.
14. S. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals.* — Math. Comput., **69(230)** (2000), 481–500.
15. S. Repin, *Two-sided estimates of deviations from exact solutions of uniformly elliptic equations.* — Trudi St.Petersburg Mathematickal Society **9** (2001), 148–179.
16. S. Repin,]it Estimates of deviations from exact solutions initial-boundary value problem for the heat equation. — Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **13**, No. 2 (2002), 121—133.
17. S. Repin, *A posteriori estimates for partial differential equations*, Vol. 4 of Radon Series on Computational and Applied Mathematics. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
18. S. I. Repin, S. A. Sauter, *Accuracy of Mathematical Models. Dimension Reduction, Homogenization, and Simplification*, Vol. 33 of EMS Tracts Math. Berlin: European Mathematical Society (EMS), 2020.
19. С. И. Репин, *Тождество для отклонений от точного решения задачи* $\Lambda^* A \Lambda u + \ell = 0$ *и его следствия.* — Журнал Вычислит. Математики и Мат. Физики, том 61, No. 12 (2021), 22-–45.
20. V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics, 25, Springer, Berlin,2006.
21. A. Toselli, O. Widlund, *Domain decomposition methods–algorithms and theory*, Springer Series in Computational Mathematics, 34, Springer, Berlin, 2005.
22. R. Verfürth, *A review of a posteriori error estimation and adaptive mesh–refinement techniques*, Wiley, Teubner, New-York, 1996.

St.Petesburg Department of Mathematics Institute
Fontanka 27, St.Petersburg 191011, Russia
E-mail: repin@pdmi.ras.ru

Поступило 25 октября 2021 г.