

# ADJOINT CORRECTION AND BOUNDING OF ERROR USING LARGANGE FORM OF TRUNCATION TERM

A.K. Alekseev<sup>a</sup> and I. M. Navon<sup>b</sup>

<sup>a</sup> *Department of Aerodynamics and Heat Transfer, RSC, ENERGIA, Korolev, Moscow Region 141070, Russian Federation, e-mail:Aleksey.Alekseev@relcom.ru*

<sup>b</sup> *Department of Mathematics and C.S.I.T., Florida State University, Tallahassee, FL 32306-4120, USA , e-mail:navon@csit.fsu.edu*

Abstract

The a-posteriori error evaluation based on differential approximation of a finite-difference scheme and adjoint equations is addressed. The differential approximation is composed of primal equations and a local truncation error determined by a Taylor series in Largange form. This approach provides the feasibility of both refining the solution and using the Holder inequality for asymptotic bounding of the remaining error.

*Keywords:* differential approximation, Lagrange truncation term, adjoint problem, a posteriori error estimation, error bound.

## 1. Introduction

Starting with the results of paper [1] the adjoint (dual) equations are widely used for estimation of the a posteriori error of the numerical solution both for finite-element and finite difference discretization methods [2-29]. In a large number of ensuing publications this approach is applied for estimation of the numerical error of some quantities of interest (goal functionals, point-wise parameters etc.) using some form of the residual (truncation error). This approach may be also extended to estimation of model error [4] caused by a difference between fine and coarse models. A broad spectrum of physical models is covered. In Ref. [9] this approach is used for wave equations while in Ref. [12] it is used for transport equation. In Refs. [6-8] a posteriori error estimate is obtained for Navier-Stokes and Euler equations. In these works the Galerkin method is used for the local error estimation while the adjoint equations are used for calculating their weights in the target functional error. A similar approach is used in [19-29] for the refinement of practically useful functionals both by finite-element and finite-difference discretization methods. The local truncation error (residual) was estimated by the action of a differential operator on the interpolated solution, while its contribution to the functional was calculated using an adjoint problem.

The approach considered herein uses another approach for the calculation of the residual as compared with [19-29]. We use a local truncation error determined by the Taylor series with the remainder in Lagrange form. This enables us to both correct the error in a usual way as well as to obtain an asymptotic error bound (based on the Holder inequality) for the refined solution. This approach was used for heat transfer equation in [30] and for the parabolized Navier-Stokes equations in [31].

## 2. The error correction and bounding for finite difference approximation of heat conduction

Let  $L$  be the differential operator determining the problem  $L\tilde{T} = 0$  and let  $L_h$  be the finite difference operator  $L_h T_h = 0$  ( $T_h$  is the grid function). The truncation error produces a field of sources  $\delta T$  disturbing the exact solution  $\tilde{T}$ . The adjoint approach permits to account for the impact of all these sources on the goal functional by summation over the entire computational domain. The truncation error may be estimated via the value of residual  $\delta T = -L T'_h$  engendered by action of differential operator on some extrapolation  $T'_h$  of the numerical solution [19]. As an alternative option we may use a differential approximation of the

finite-difference scheme [32]. Then the truncation source term assumes the form  $\delta T = L_h T_h - LT_h$  and is composed of Taylor series terms with coefficients containing some powers of the grid size. If we use the Lagrange form of Taylor series we may obtain a closed form of the truncation term. This form provides an opportunity to subdivide the truncation error into a computable part containing known values and an incomputable part containing the field of Lagrange coefficients (unknown parameters belonging to the unit interval). This approach enables us both to correct the error and to obtain some asymptotic bounds of the remaining error.

For illustrating this idea we first apply it to the unsteady one dimensional heat conduction equation and its finite-difference approximation. We assume that the solution is smooth enough to have all the required derivatives bounded. Let us consider the estimation of the temperature error at a checkpoint.

$$C \rho \frac{\partial \tilde{T}}{\partial t} - \frac{\partial}{\partial x} \left( \lambda \frac{\partial \tilde{T}}{\partial x} \right) = 0; \text{ in } Q = \Omega \times (0, t_f), \Omega \in R^1. \quad (1)$$

$$\text{Initial conditions: } \tilde{T}(0, x) = T_0(x); T_0(x) \in L_2(\Omega); \quad (2)$$

$$\text{Boundary conditions } \frac{\partial \tilde{T}}{\partial x} \Big|_{x=0} = 0; \frac{\partial \tilde{T}}{\partial x} \Big|_{x=X} = 0. \quad (3)$$

Here  $C = \text{Const}$  is thermal capacity,  $\lambda$  is thermal conductivity,  $\rho = \text{Const}$  is density,  $\tilde{T}$  is temperature (considered here as exact, error-free),  $x$ -coordinate,  $X$ -thickness,  $t$  - time,  $t_f$  - duration of process,  $\Omega$  -domain of calculation  $(0, X)$ . We consider two cases:  $\lambda = \text{Const}$ ,  $\tilde{T}(t, x) \in C^\infty(Q)$  and  $\lambda \in L_2(\Omega)$ ,  $\tilde{T}(t, x) \in H^1(Q)$ . In these spaces the problem is well-posed [33].

Consider a finite-difference approximation of equation (1) having the first order in time and second order in space (for the constant  $\lambda$ ):

$$C \rho \frac{T_k^n - T_k^{n-1}}{\tau} - \lambda \frac{T_{k+1}^n - 2T_k^n + T_{k-1}^n}{h_k^2} = 0; \quad (4)$$

Here  $T$  is the solution of finite difference equation,  $\tau$  is temporal step and  $h_k$  - the spatial step size. The simplicity of the scheme and the low order of approximation are deliberately chosen to illustrate the features of this approach with the simplest mathematical treatment and to obtain an observable (comparing with other sources) truncation error. Herein, we address the impact of this error on the temperature at a certain checkpoint  $T_{est} = T(t_{est}, x_{est})$ . Let us denote the estimated temperature  $T_{est}$  by  $\varepsilon$  and express it as the functional.

$$T_{est} = \varepsilon = \iint_Q T(t, x) \delta(t - t_{est}) \delta(x - x_{est}) dt dx \quad (5)$$

Here  $\delta$  is Dirac's delta function.

The error of the temperature is determined as the sum of contributions of local truncation error with weights depending on the transfer of disturbances and determined by the adjoint parameter. In order to determine the truncation error let us expand the mesh function  $T_k^n$

in a Taylor series and substitute into (4). Herein we imply that there exists a smooth enough function  $T(t, x)$  that coincides with  $T_k^n$  at all grid points. Then equation (4) transforms into equation (6).

$$C \rho \frac{\partial T}{\partial t} - \lambda \frac{\partial^2 T}{\partial x^2} + \delta T = 0; \quad (6)$$

Here  $\delta T = \delta T_t + \delta T_x$  is a local truncation error engendered by the Taylor series remainders. We use here the Lagrange form of remainder, which contains unknown parameters  $\alpha_k^n, \beta_k^n, \gamma_k^n \in (0,1)$ .

$$\delta T_t = -\frac{C\rho}{2} \tau \frac{\partial^2 T(t_n - \alpha_k^n \tau, x_k)}{\partial t^2} \quad (7)$$

$$\delta T_x = -\frac{\lambda}{24} h_k^2 \left( \frac{\partial^4 T(t_n, x_k + \beta_k^n h)}{\partial x^4} + \frac{\partial^4 T(t_n, x_k - \gamma_k^n h)}{\partial x^4} \right) \quad (8)$$

The mathematical properties of the differential approximations are discussed in [32, 34]. According to [32]  $T(t, x) \in C^\infty(Q)$ . Thus, we may consider a finite-difference equation to be equivalent to an approximated equation with an additional perturbation term. By introducing a solution error  $\Delta T$  ( $T = \tilde{T} + \Delta T$ ) we can reformulate (6) as

$$C \rho \frac{\partial (\tilde{T} + \Delta T)}{\partial t} - \lambda \frac{\partial^2 (\tilde{T} + \Delta T)}{\partial x^2} + \delta T = 0; \quad (9)$$

Let us find the error of the functional (5) as a function of the truncation error. For this purpose let us introduce the Lagrangian of the following form

$$\begin{aligned} L = & \iint_Q T(t, x) \delta(t - t_{est}) \delta(x - x_{est}) dt dx + \\ & + \iint_Q C \rho \frac{\partial T}{\partial t} \Psi(t, x) dt dx - \iint_Q \frac{\partial}{\partial x} \left( \lambda \frac{\partial T}{\partial x} \right) \Psi(t, x) dt dx + \iint_Q \delta T \Psi dt dx \end{aligned} \quad (10)$$

Here  $\Psi$  is the adjoint temperature defined by the solution of following adjoint (dual) problem.

$$C \rho \frac{\partial \Psi}{\partial t} + \frac{\partial}{\partial x} \left( \lambda \frac{\partial \Psi}{\partial x} \right) - \delta(t - t_{est}) \delta(x - x_{est}) = 0, (t, x) \in \Omega. \quad (11)$$

$$\text{Boundary conditions: } \frac{\partial \Psi}{\partial x} \Big|_{x=X} = 0, \frac{\partial \Psi}{\partial x} \Big|_{x=0} = 0, \quad (12)$$

$$\text{Initial condition } \Psi(t_f, x) = 0; \quad (13)$$

According to [35] problem (13) is well-posed for  $\Psi(t, x) \in H^{-\alpha}(\Omega)$ ,  $\frac{\alpha}{n} > \frac{1}{2}$ ,  $\Omega \in R^n$ . In the case considered here the problem is well-posed if  $\Psi(t, x) \in H^{-1}(\Omega)$ , however if we smooth the source term according to [3,33], we may obtain a solution  $\Psi_s(t, x) \in H^\beta(\Omega)$ ,  $\beta > 1$  (although containing an error proportional to smoothing parameter  $s, (s > 0)$ , which may be as small as necessary). Finite difference methods for the solution of such equations are presented in [36,37].

It may be shown from this Lagrangian variation [38] that for solutions of primal (1-3) and adjoint (11-13) problems, the variation of the functional caused by the truncation error equals

$$\Delta \mathcal{E} = \Delta T_{est} = T_{est} - T_{exact} = \iint_Q \delta T \Psi(t, x) dt dx \quad (14)$$

### 2.1 Discrete form

Taking into account (14) and the temporal part of truncation error described by (7) we obtain the corresponding part of error  $\Delta T_{est}$  as

$$\Delta \mathcal{E}(\delta T_t) = -\frac{C\rho}{2} \int_Q \left( \tau \frac{\partial^2 T(t_n - \alpha_k^n \tau, x_k)}{\partial t^2} \right) \Psi dx dt \quad (15)$$

Further discussion is significantly devoted to the calculation of the magnitude and bounds of expression (15) and its analogues. Let us present (15) in a discrete form, for example:

$$\Delta \mathcal{E}(\delta T_t) = -\frac{C\rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \left( \tau \frac{\partial^2 T(t_n - \alpha_k^n \tau, x_k)}{\partial t^2} \right) \Psi_k^n h_k \tau \quad (16)$$

Herein  $N_t$  is the number of time steps while  $N_x$  is the number of spatial nodes.

Equation (16) may be expanded in series over  $\alpha_k^n \tau$ ,

$$\Delta \mathcal{E}(\delta T_t) = -\frac{C\rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \left( \tau \frac{\partial^2 T(t_n, x_k)}{\partial t^2} - \tau \alpha_k^n \tau \frac{\partial^3 T(t_n, x_k)}{\partial t^3} + \frac{(\alpha_k^n \tau)^2}{2} \tau \frac{\partial^4 T(t_n, x_k)}{\partial t^4} - \dots \right) \Psi_k^n h_k \tau \quad (17)$$

The first part of sum (17) may be used for correcting of functional (5)

$$\Delta T_t^{corr} = -\frac{C\rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \frac{\partial^2 T(t_n, x_k)}{\partial t^2} \Psi_k^n h_k \tau^2 \quad (18)$$

The second part of (17) contains unknown parameters  $\alpha_k^n$  belonging to the unit interval  $\alpha_k^n \in (0,1)$ , so we may obtain a bound of this expression. If only the first order term over  $\alpha_k^n \tau$  is retained in (17) an upper bound may be obtained of the form

$$\frac{C\rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \alpha_k^n \tau^3 \frac{\partial^3 T(t_n, x_k)}{\partial t^3} \Psi_k^n h_k \leq \frac{C\rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \left| h_k \tau^3 \frac{\partial^3 T(t_n, x_k)}{\partial t^3} \Psi_k^n \right| = \Delta T_{t,1}^{sup} \quad (19)$$

Using this value we can determine the upper bound of the functional error (after refining):

$$\left| T_{est} - \Delta T_t^{corr} - T_{exact} \right| < \Delta T_{t,1}^{sup} \quad (20)$$

Expression (19) is the Holder inequality applied to the scalar product  $(\alpha_k^n, \Theta_k^n)$ ,  $\sum_{k=1, n=2}^{Nx, Nt} \alpha_k^n \Theta_k^n = (\alpha_k^n, \Theta_k^n) \leq \|\alpha_k^n\|_p \|\Theta_k^n\|_q$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ . We consider herein  $p = \infty$  and  $q = 1$ .

Then  $\|\Theta\|_q = \left( |\Theta_1|^q + |\Theta_2|^q + \dots + |\Theta_N|^q \right)^{1/q} = \sum |\Theta_i|$  and  $\|\alpha\|_\infty = \max(|\alpha_i|)$ . Since  $\alpha_k^n \in (0,1)$  this norm of unknown parameters may be easily estimated as equal to unit.

Expression (20) is correct for exact values of adjoint parameter. In reality, the adjoint problem is solved by some finite-difference method, so it contains an approximation error  $\Psi(t, x) = \Psi_{exact}(t, x) + \Delta\Psi(t, x)$ . Hence, the estimation of the functional variation has a component determined by the adjoint problem error.

$$\Delta\varepsilon = \Delta T_{est} = \iint_{\Omega} \delta T \Psi_{exact}(t, x) dt dx + \iint_{\Omega} \delta T \Delta\Psi(t, x) dt dx \quad (21)$$

The second term of (21) corresponds to the remaining error according to [19] and is associated to the errors of approximation of both the adjoint and primal equations. It may be expedient to construct a mesh for the minimization of this term as in [25-27]. As an alternative, we may use the second order adjoint equations [39,40] for calculating this term. If the primal and adjoint problems are solved by methods of order  $O(h^p)$  and  $O(h^a)$  correspondingly, this term is of order  $O(h^{p+a})$ . For schemes of high enough order ( $p \geq 2$  or  $a \geq 2$ ) this term is asymptotically small when compared with the error bounds determined by (19).

The calculation of error caused by spatial approximation is performed similarly. The error correction is as follows

$$\Delta T_x^{corr} = -\frac{\lambda}{12} \sum_{k=1, n=1}^{Nx, Nt} h_k^3 \frac{\partial^4 T(t_n, x_k)}{\partial x^4} \Psi_k^n \tau \quad (22)$$

The incomputable error can be bounded (assuming  $\beta_k^n - \gamma_k^n = 1$ ) in a form:

$$\frac{\lambda}{24} \sum_{k=1, n=2}^{Nx, Nt} h_k^3 \left( \frac{\partial^5 T(t_n, x_k)}{\partial x^5} \beta_k^n - \frac{\partial^5 T(t_n, x_k)}{\partial x^5} \gamma_k^n \right) \Psi_k^n h_k \tau < \Delta T_{x,1}^{sup} = \frac{\lambda}{24} \sum_{k=1, n=2}^{Nx, Nt} h_k^4 \left| \frac{\partial^5 T(t_n, x_k)}{\partial x^5} \Psi_k^n \right| \tau \quad (23)$$

## 2.2 Numerical tests

Let us estimate the approximation error using as a test problem the temperature field evolution engendered by a pointwise heat source ( $t_0, \xi$  - is the initial time and the coordinate of the point source).

$$T_{an}(t, x) = \frac{Q}{2\sqrt{\pi\lambda/(C\rho)}(t-t_0)} \exp\left(-\frac{(x-\xi)^2}{4\lambda/(C\rho)(t-t_0)}\right) \quad (24)$$

We use the data  $f_k = T_0(x_k)$  calculated by (24) as the initial data when solving (4). The length  $X$  of the spatial interval is chosen so as to provide a negligible effect of the boundary condition compared with the effect of approximation. The round-off errors were

estimated by comparing calculations with single and double precision, and the difference was found to be negligible. We should also ascertain that the error  $\iint_Q \delta T \Delta \Psi(t, x) dt dx$  engendered

by adjoint equation approximation is sufficiently small. For calculation of  $\Delta \Psi(t, x)$  the following equation was used

$$C \rho \frac{\partial \Delta \Psi}{\partial t} + \frac{\partial}{\partial x} \left( \lambda \frac{\partial \Delta \Psi}{\partial x} \right) + \delta \Psi_t(t, x) + \delta \Psi_x(t, x) = 0; \quad (25)$$

(second order adjoint equation [39,40]). For  $\lambda = Const$  and  $\delta T(t, x) \in C^\infty(Q)$ , the problem (25) is well-posed for  $\Delta \Psi(t, x) \in C^\infty(Q)$  [33].

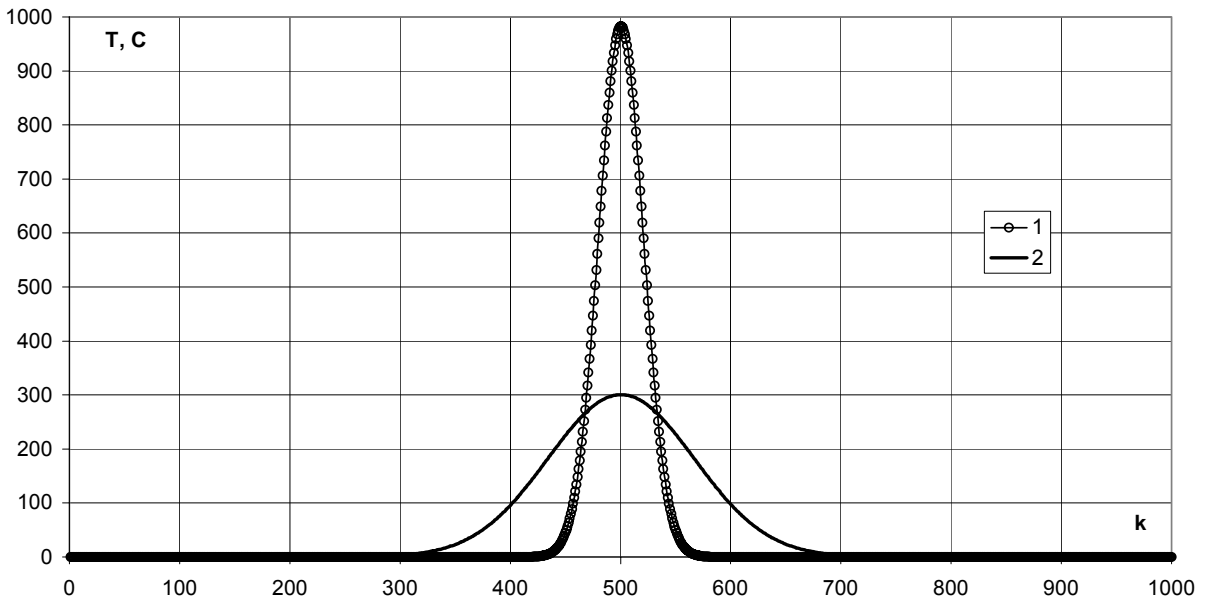


Fig 1. Initial and final temperature distribution. 1 - Initial temperature, 2- Final temperature

The corresponding error of functional has the form :

$$\Delta \varepsilon(\delta T) = -\frac{C \rho}{2} \sum_{k=1, n=2}^{N_x, N_t} \left( \tau \frac{\partial^2 T(t_n, x_k)}{\partial t^2} \right) \Delta \Psi_k^n h_k \tau. \text{ As expected, the computations}$$

demonstrated that the part of error related to the adjoint temperature error (25) is significantly smaller than the main value (related to the adjoint temperature itself).

An implicit method (implemented using the Thomas algorithm) was used for solution of both the heat transfer equation and the adjoint equations of first and second orders. The spatial grid consisted of 100–1000 nodes, while the temporal integration consisted of 100–10000 steps. Thermal conductivity was  $\lambda=10^{-4}$  kW/(m·K) and the volume heat capacity was equal to  $C\rho=500$  kJ/(m<sup>3</sup>·K). The initial and final temperature distributions are presented in Figure 1. The temperature errors were estimated via adjoint equations and compared with the deviation of the numerical solution from the analytical one (24).

### 2.3 The error caused by the truncation error of time approximation

Estimates of temperature calculation error as a function of the time step are presented in Fig. 2 (central point at the final moment). The spatial step is chosen to be enough small ( $h=0.0001$  m) so as to provide a small impact of the spatial discretization error in comparison with the temporal one. The error caused by adjoint temperature approximation was calculated using equation (25) and was significantly smaller than the temporal one.

The correction term is of first order of accuracy and successfully eliminated most of the approximation error. The bounding term is of second order and is significantly greater than the remaining error (discrepancy of refined solution and analytical value). The observable orders of both correction and bounding terms are in a good agreement with expressions (18,19).

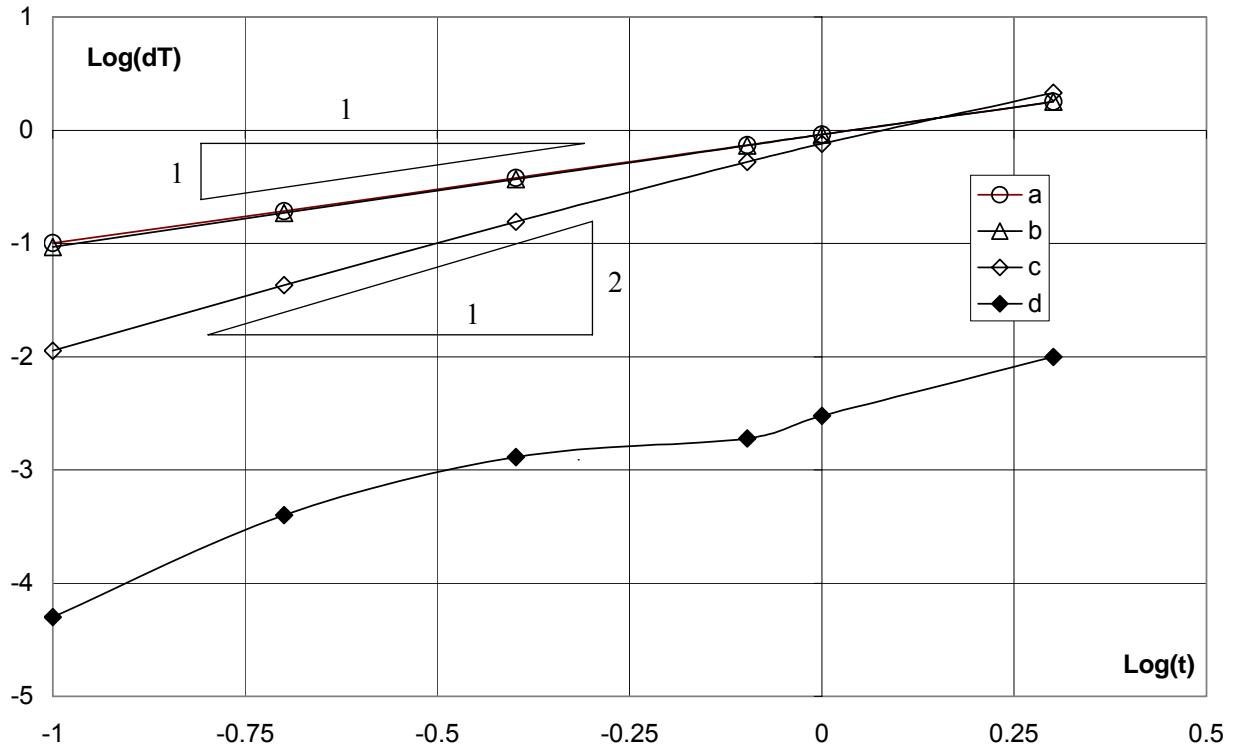


Fig. 2. Variation of errors as a function of temporal step (in decimal logarithm scale). a – deviation of calculated temperature from analytical value ( $T - T_{an}$ ), b- correction of temperature  $\Delta T_x^{corr}$  (18), c- refined solution error bound  $\Delta T_{x,1}^{sup}$  (19), d-discrepancy between refined and analytic solutions ( $T_x^{corr} - T_{an}$ )

Figure 3 illustrates the comparison between initial finite-difference and corrected finite-difference solutions and the error bounds (all normalized by analytic value), ( $h=0.0001$  m,  $\tau=1.0$  sec) related with results of Fig 1.

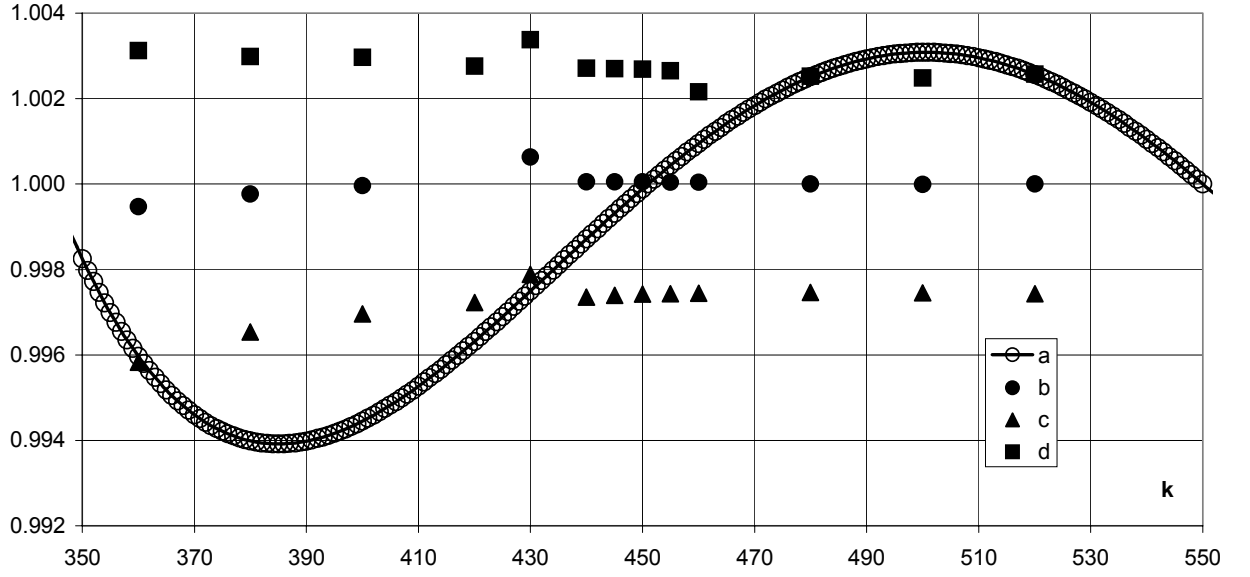


Fig. 3. The comparison of numerical and refined solutions (all divided by analytical value).  
a- numerical, b-refined solution, c- lower bound, d- upper bound

#### 2.4 The error of temperature calculation engendered by the spatial discretization.

Let us consider the error caused by the truncation error of the spatial approximation. In order to observe this error, we should provide a small contribution of truncation error of the temporal approximation. With this purpose the following second order time approximation scheme was used.

$$\begin{aligned}
 C\rho \frac{T_k^{n-1/2} - T_k^{n-1}}{\tau} - \frac{1}{2} \lambda \frac{T_{k+1}^{n-1} - 2T_k^{n-1} + T_{k-1}^{n-1}}{h_k^2} &= 0; \\
 C\rho \frac{T_k^n - T_k^{n-1/2}}{\tau} - \frac{1}{2} \lambda \frac{T_{k+1}^n - 2T_k^n + T_{k-1}^n}{h_k^2} &= 0;
 \end{aligned} \tag{26}$$

It may be demonstrated in a manner similar to previous treatments that the error caused by the temporal approximation is of second order in  $\tau$ .

$$\Delta \varepsilon(\delta T_t) = -\frac{C\rho}{12} \sum_{k=1, n=2}^{N_x, N_t} \frac{\partial^3 T(t_n, x_k)}{\partial t^3} \Psi_k^n h_k \tau^3 \tag{27}$$

A bound on the incomputable error caused by temporal step is:

$$\Delta T_t^{\text{sup}} = \Delta \varepsilon(\delta T) = \frac{C\rho}{4} \sum_{k=1, n=2}^{N_x, N_t} \left| \frac{\partial^4 T(t_n, x_k)}{\partial t^4} \Psi_k^n \right| h_k \tau^4 \tag{28}$$

The error caused by the spatial approximation retains its previous form (22,23). Numerical tests demonstrated that the error caused by the time step (27) was not greater



than  $2 \cdot 10^{-5}$  and was significantly smaller than the error caused by the spatial approximation. The error caused by the adjoint equation approximation  $\iint_Q \delta T \Delta \Psi(x,t) dt dx$  was even smaller

by several orders of magnitude. The temperature error estimations as a function of the spatial step size are presented in Fig. 4 (for central point at the final time).

The comparison of deviations of the solution from the analytical one and the correcting term demonstrates that the refinement by  $\Delta T_x^{corr}$  (22) enables us to eliminate a significant part of the error. Comparison of the remaining error  $T_x^{corr} - T_{an}$  and  $\Delta T_{x,1}^{sup}$  demonstrates a reliable bounding by expression (23). The remaining error  $T_x^{corr} - T_{an}$  contains all uncontrolled errors including those caused by boundary terms, errors of upper orders etc, so it exhibits a slightly irregular behavior.

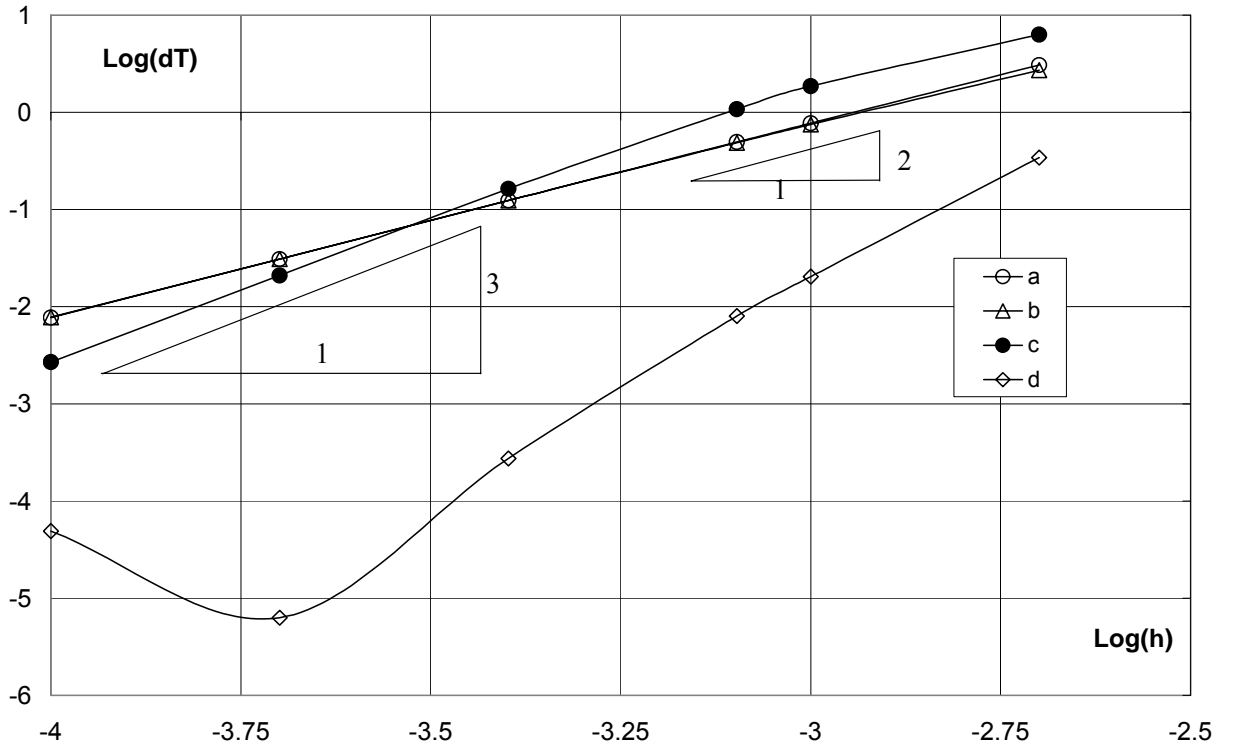


Fig. 4. Variation of errors as a function of spatial step (in logarithm scale).

a – Deviation of calculated temperature from analytical value ( $T - T_{an}$ ), b- Correction of temperature  $\Delta T_x^{corr}$ , c- Refined solution error bound  $\Delta T_{x,1}^{sup}$ , d- Deviation of refined solution from analytical value ( $T_x^{corr} - T_{an}$ )

The quadratic character of  $\Delta T_x^{corr}$  and the third order of  $\Delta T_x^{sup}$  should be noted as coinciding with the formal order of (22,23). The convergence rate of  $\Delta T_x^{corr}$  and  $\Delta T_x^{sup}$  demonstrates that the discontinuities of high order derivatives for equation (1) under initial conditions (24) and boundary conditions (3) did not engender any visible effect (they are located in zones of small  $\Psi$ ).

## 2.5 The effect of discontinuities of the derivatives

The above considered solutions possessed an infinite number of bounded derivatives which was the reason for good agreement between observed and nominal convergence rates. If the physical field is specified by small number of bounded derivatives, the order of convergence may differ from the nominal one.

Let us consider this problem at a heuristic level for some function  $\rho(t, x)$  having  $m$  bounded spatial derivatives (the derivative of  $m - th$  order has a finite number of jump discontinuities).

We consider an approximation of the derivative of the order  $p$  by finite differences of a formal order of accuracy  $j$ . We denote the finite difference scheme as  $\frac{D^p \rho(t, x)}{Dx^p}$ . The value

$\lim_{h \rightarrow 0} \sum_{\Omega} \left( h^j \frac{D^{p+j} \rho(t, x)}{Dx^{p+j}} \right) h \Psi$  corresponds to the correction term. Consider its asymptotic

form. The derivative of order  $m+1$  has an asymptotic  $(\rho_+^{(m)} - \rho_-^{(m)})/h \sim \Delta/h$  for the jump discontinuity, while the derivative of order  $m+2$  has the asymptotic  $(\Delta/h - 0/h)/h \sim \Delta/h^2$ , correspondingly the derivative of the order  $p+j$  has the asymptotic  $\frac{\Delta}{h^{p+j-m}}$ . Thus

$\lim_{h \rightarrow 0} \left( h^j \frac{D^{p+j} \rho(t, x)}{Dx^{p+j}} \right) \sim \lim_{h \rightarrow 0} \left( h^j \frac{\Delta}{h^{p+j-m}} \right)$ . Only a limited number of nodes participate in the

summation in the vicinity of discontinuity, so the multiplier  $h$  (appearing during summation)

should be taken into account, yielding  $\sum_{k=1, n=n_r, -n_s}^{k=Nx, n=n_r+n_s} \left( h^j \frac{D^{p+j} \rho(t, x)}{Dx^{p+j}} \right) h \Psi \sim h^{m-p+1}$ .

Thus, the terms of the  $j$ -th formal order of accuracy contain a component of  $j$ -th order (appearing due to integration over the smooth part of the solution) and a component having the order  $i = m - p + 1$  (engendered by the jump discontinuity of the  $m - th$  order derivative). The picture is complicated by the dipole nature of the error caused by the discontinuity that may be compensated by summation. If we have a stepwise discontinuity of the first derivative at point  $k$  ( $\rho_{k-1} = 0, \rho_k = \Delta, \rho_{k+1} = \Delta$ ) then we obtain mutually compensated singularities:

$$\text{At point } k: h \frac{\partial^2 \rho}{\partial x^2} \sim h \frac{\rho_{k+1} - 2\rho_k + \rho_{k-1}}{h^2} = -\frac{\Delta}{h},$$

$$\text{At point } k-1: h \frac{\partial^2 \rho}{\partial x^2} \sim h \frac{\rho_k - 2\rho_{k-1} + \rho_{k-2}}{h^2} = +\frac{\Delta}{h}.$$

The behavior of observable convergence rate of these terms may be more complicated since they are calculated as terms of the numerical solution that only asymptotically approximate the exact values. Under these conditions we can not have an expectation of obtaining similar results to Figs. 2 and 4 for situations where discontinuous derivatives are present.

For example let us carry out numerical tests to study the asymptotic dependence of the error on the space step size for a temperature gradient discontinuity. In order to deal with the discontinuity we used a divergent integro-interpolation method [41] well suited for the calculation of temperature gradient discontinuities.

Fig. 5 presents temperature error estimates (for central point at the final moment) depending on the spatial step for the thermal conductivity coefficient having a 10% jump at the center of the grid and initial temperature of Fig. 1 (unfortunately, an analytical solution is not

available).

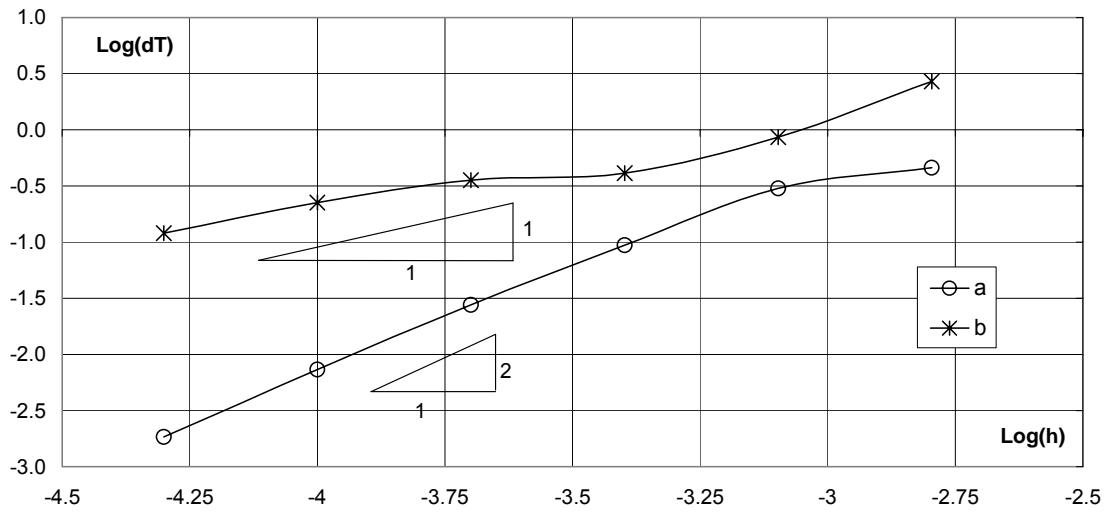


Fig. 5. Variation of errors as a function of the spatial step (in logarithm scale) for break in first derivative caused by discontinuous conductivity.

a - correction of temperature, b- error bound of refined solution

As another test we consider the evolution of the initial temperature distribution of a step shape. The initial, the final distribution of temperature and the location of estimated points are presented in Fig 6. The break of thermal conductivity is located at the center point ( $x_s = X/2$ ) and coincides with a stepwise discontinuity in the initial temperature.

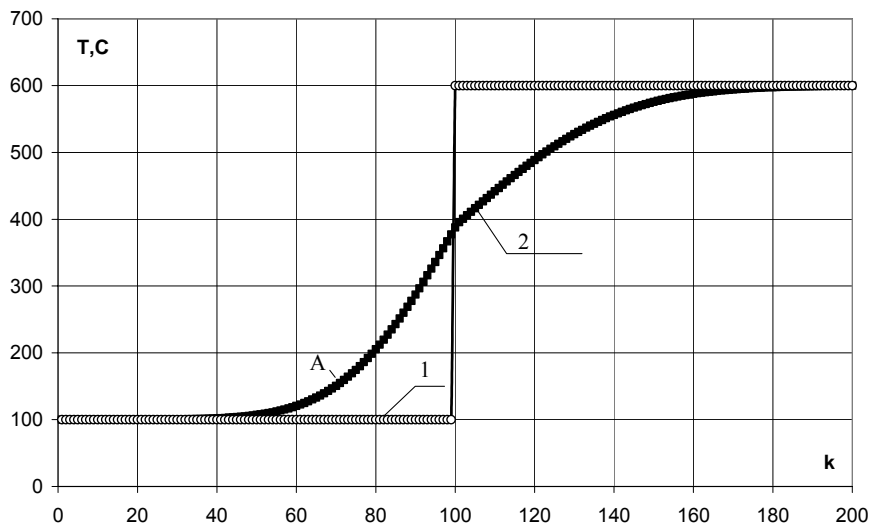


Fig. 6. Initial and final temperature distribution.

1- Initial temperature, 2- Final temperature, A- point of estimation

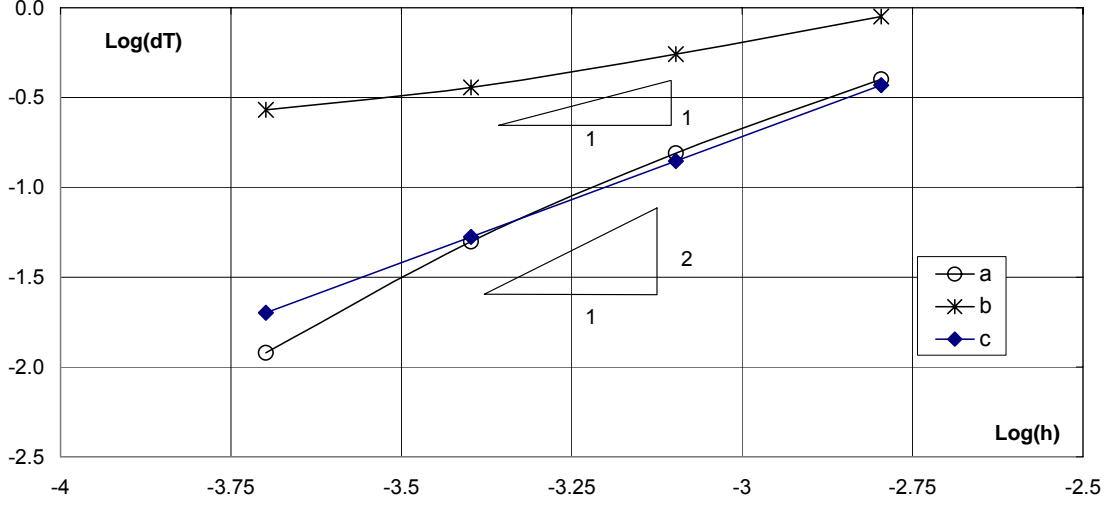


Fig. 7. Variation of errors as a function of the spatial step (in logarithm scale) for break in first derivative caused by discontinuous conductivity.

a - Correction of temperature, b- error bound of refined solution, c-deviation of numerical solution from analytical value

Fig. 7 presents the temperature error estimates depending on the spatial step. The rate of convergence of  $T_{est} - T_{an}$  and  $\Delta T_x^{corr}$  is close to second order despite the influence of discontinuity. This is caused by a mutual compensation of errors (dipole nature of error) in the vicinity of the discontinuity as confirmed by an analysis of local distribution of error density  $-\frac{\lambda}{12} h_k^3 \frac{\partial^4 T(t_n, x_k)}{\partial x^4} \Psi_k^n \tau$  (engendering  $\Delta T_x^{corr}$  in accordance with (22)). The order of  $\Delta T_{x,1}^{sup}$  is close to one (slightly below), which corresponds to the expected influence of the temperature gradient discontinuity and is in contrast with the formal third order expected from (23)). Thus, the calculation of approximation errors by the method considered is strongly affected by the number of bounded derivatives of the solution.

### 3. Error correction and bounding for viscous flow

The heat transfer equation is providing a favorable example for our approach due to the great smoothness of the solutions. Let us consider the method discussed above for the pointwise error in a two dimensional supersonic viscous flow. The nondivergent form of the parabolized Navier-Stokes equations (PNS) is used. The flow is calculated by marching along the  $X$  axis.

$$\frac{\partial(\rho U)}{\partial X} + \frac{\partial(\rho V)}{\partial Y} = 0 \quad (29)$$

$$U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} + \frac{1}{\rho} \frac{\partial P}{\partial X} - \frac{1}{\text{Re}} \frac{\partial^2 U}{\partial Y^2} = 0 \quad (30)$$

$$U \frac{\partial V}{\partial X} + V \frac{\partial V}{\partial Y} + \frac{1}{\rho} \frac{\partial P}{\partial Y} - \frac{4}{3 \rho \text{Re}} \frac{\partial^2 V}{\partial Y^2} = 0 \quad (31)$$

$$U \frac{\partial e}{\partial X} + V \frac{\partial e}{\partial Y} + (\gamma - 1)e \left( \frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} \right) - \frac{1}{\rho} \frac{\gamma}{\text{Re Pr}} \frac{\partial^2 e}{\partial Y^2} - \frac{1}{\rho} \frac{4}{3 \text{Re}} \left( \frac{\partial U}{\partial Y} \right)^2 = 0 \quad (32)$$

$$P = \rho RT; \quad e = C_v T = \frac{RT}{\gamma - 1}; \quad (X, Y) \in \Omega = (0 < X < X_{max}; 0 < Y < Y_{max});$$

On  $\Gamma_{in}$  we have  $f^i = f_{in}^i(X, Y)$ , on  $\Gamma_{out}$  we have  $\partial f^i / \partial Y = 0$ , ( $f^i = (\rho, U, V, e)$ ).

The boundary  $\Gamma = \Gamma_{in} \cup \Gamma_{out}$ ,  $\Gamma_{in}$  is the inflow boundary,  $\Gamma_{out}$  is the outflow boundary.

The density at some point is considered as an estimated parameter. Let us write the estimated value  $\rho(X^{est}, Y^{est})$  in the form of a functional.

$$\rho^{est} = \varepsilon = \int_{\Omega} \rho(X, Y) \delta(Y - Y^{est}) \delta(X - X^{est}) dXdY \quad (33)$$

We calculate the variation of the functional with respect to local disturbances (truncation error)  $\delta f^i$  using the adjoint equations in the form described in [31,42].

### 3.1 Adjoint problem

$$\begin{aligned} & U \frac{\partial \Psi_{\rho}}{\partial X} + V \frac{\partial \Psi_{\rho}}{\partial Y} + (\gamma - 1) \frac{\partial (\Psi_{\rho} e / \rho)}{\partial Y} + (\gamma - 1) \frac{\partial (\Psi_{\rho} U / \rho)}{\partial X} - \\ & - \frac{\gamma - 1}{\rho} \left( \frac{\partial e}{\partial Y} \Psi_{\rho} + \frac{\partial e}{\partial X} \Psi_{\rho} \right) + \left( \frac{1}{\rho^2} \frac{\partial P}{\partial X} - \frac{1}{\rho^2 \text{Re}} \frac{\partial^2 U}{\partial Y^2} \right) \Psi_{\rho} + \frac{1}{\rho^2} \left( \frac{\partial P}{\partial Y} - \frac{4}{3 \text{Re}} \frac{\partial^2 V}{\partial Y^2} \right) \Psi_{\rho} - \\ & - \frac{1}{\rho^2} \left( \frac{\gamma}{\text{Re Pr}} \frac{\partial^2 e}{\partial Y^2} + \frac{4}{3 \text{Re}} \left( \frac{\partial U}{\partial Y} \right)^2 \right) \Psi_{\rho} - \delta(X - X^{est}) \delta(Y - Y^{est}) = 0 \end{aligned} \quad (34)$$

The source in (34) corresponds to the location of the estimated parameter.

$$\begin{aligned} & U \frac{\partial \Psi_U}{\partial X} + \frac{\partial (\Psi_U V)}{\partial Y} + \rho \frac{\partial \Psi_{\rho}}{\partial X} - \left( \frac{\partial V}{\partial X} \Psi_{\rho} + \frac{\partial e}{\partial X} \Psi_{\rho} \right) + \frac{\partial}{\partial X} \left( \frac{P}{\rho} \Psi_{\rho} \right) + \\ & + \frac{\partial^2}{\partial Y^2} \left( \frac{1}{\rho \text{Re}} \Psi_{\rho} \right) - \frac{\partial}{\partial Y} \left( \frac{8}{3 \text{Re}} \frac{\partial U}{\partial Y} \Psi_{\rho} \right) = 0 \end{aligned} \quad (35)$$

$$\begin{aligned} & \frac{\partial (U \Psi_{\rho})}{\partial X} + V \frac{\partial \Psi_{\rho}}{\partial Y} - \left( \frac{\partial U}{\partial Y} \Psi_{\rho} + \frac{\partial e}{\partial Y} \Psi_{\rho} \right) + \\ & + \rho \frac{\partial \Psi_{\rho}}{\partial Y} + \frac{\partial}{\partial Y} \left( \frac{P}{\rho} \Psi_{\rho} \right) + \frac{4}{3 \text{Re}} \frac{\partial^2}{\partial Y^2} \left( \frac{\Psi_{\rho}}{\rho} \right) = 0 \end{aligned} \quad (36)$$

$$\frac{\partial (U \Psi_e)}{\partial X} + \frac{\partial (V \Psi_e)}{\partial Y} - \frac{\gamma - 1}{\rho} \left( \frac{\partial \rho}{\partial Y} \Psi_{\rho} + \frac{\partial \rho}{\partial X} \Psi_{\rho} \right) - (\gamma - 1) \left( \frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} \right) \Psi_e +$$

$$+ (\gamma - 1) \frac{\partial \Psi_V}{\partial Y} + (\gamma - 1) \frac{\partial \Psi_U}{\partial X} + \frac{\gamma}{\text{Re Pr}} \frac{\partial^2}{\partial Y^2} \left( \frac{\Psi_e}{\rho} \right) \quad (37)$$

The parameters  $(\Psi_\rho, \Psi_U, \Psi_V, \Psi_e)$  are the adjoint analogs of density, velocity components, and energy, respectively.

$$\text{On boundary } \Gamma_{out} : \Psi_{\rho, U, V, e} \Big|_{\Gamma_{out}} = 0, \text{ on } \Gamma_{in} : \frac{\partial \Psi_f}{\partial Y} = 0;$$

The adjoint problem is calculated in the reverse direction along  $X$ .

Using the solution of above adjoint problem we may express the variation of the target functional as a function of the truncation error in the following form:

$$\delta \varepsilon = \iint_{\Omega} (\delta \rho \Psi_\rho + \delta U \Psi_U + \delta V \Psi_V + \delta e \Psi_e) dX dY \quad (38)$$

Here  $\delta \rho$  etc. are the truncation errors.

### 3.2 Taking into account the viscosity impact

In the tests presented below we should compare the results of finite-difference calculations of Parabolized Navier Stokes equations with the analytical solutions available for inviscid gas flows. These numerical results contain the influence of viscosity in addition to the impact of truncation error. On the other hand, some considered solutions contain shock waves, so using a viscous statement may be necessary from a computational point of view [24].

In this context, the influence of viscous terms in equations (29-32) on an estimated parameter is of interest. We will consider the solution of equations without viscosity as a non-perturbed one. Let the viscous terms disturb this solution. For example, for the longitudinal velocity undisturbed values are governed by equation  $U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} = 0$ , while the disturbed ones are governed by  $U \frac{\partial \tilde{U}}{\partial X} + V \frac{\partial \tilde{U}}{\partial Y} - \frac{1}{\text{Re } \rho} \frac{\partial^2 \tilde{U}}{\partial Y^2} = 0$ . Then the variation of the target functional due to viscous terms assumes the form

$$\delta \varepsilon = - \int_{\Omega} \left( \frac{1}{\text{Re } \rho} \frac{\partial^2 \tilde{U}}{\partial Y^2} \Psi_U + \dots \right) d\Omega \quad (39)$$

In contrast to (34-37), the corresponding adjoint equations have no viscous terms. This approach may be viewed as some variant of the estimation of model error [4] caused by the difference between two models. Certainly, this approach is valid only when the influence of viscous terms is small enough, i.e. when they do not cause a radical change of the flow structure.

Another reason for this technique development arises from discontinuities that are typical of supersonic flows described by Euler equations, for example. The approach based on differential approximation is not formally applicable for supersonic Euler equations due to unbounded derivatives. Nevertheless, we may use the parabolized Navier-Stokes for basic flow calculation, consider viscous terms as a perturbation, and calculate the effect of this perturbation on the solution. This may enable us to expand the applicability of the differential approximation approach to discontinuous flows as described by the Euler equations.

### 3.3. Finite difference scheme

Herein we use a first order finite-difference scheme based on upwind differences [43]. It contains two steps, predictor and corrector. Both steps are calculated implicitly, using the three point Thomas algorithm. The tilde marks parameters computed at the first step. This scheme is rather simple, has a large enough truncation error and is monotonic. The last feature is very important for calculation of derivatives that approximate the truncation terms. The scheme (for  $V_k^n > 0$  option) is presented below

#### Predictor

$$U_k^n \frac{\tilde{\rho}_k^{n+1} - \rho_k^n}{h_x} + \rho_k^n \frac{U_k^n - U_{k-1}^{n-1}}{h_x} + V_k^n \frac{\tilde{\rho}_k^{n+1} - \tilde{\rho}_{k-1}^{n+1}}{h_y} + \rho_k^n \frac{V_{k+1}^n - V_{k-1}^n}{2h_y} = 0 \quad (40)$$

$$U_k^n \frac{\tilde{U}_k^{n+1} - U_k^n}{h_x} + V_k^n \frac{\tilde{U}_k^{n+1} - \tilde{U}_{k-1}^{n+1}}{h_y} + \frac{P_k^n - P_{k-1}^{n-1}}{h_x \rho_k^n} - \frac{1}{\text{Re} \rho_k^n} \frac{\tilde{U}_{k+1}^{n+1} - 2\tilde{U}_k^{n+1} - \tilde{U}_{k-1}^{n+1}}{h_y^2} = 0. \quad (41)$$

$$U_k^n \frac{\tilde{V}_k^{n+1} - V_k^n}{h_x} + V_k^n \frac{\tilde{V}_k^{n+1} - \tilde{V}_{k-1}^{n+1}}{h_y} + \frac{P_{k+1}^n - P_{k-1}^n}{2h_y \rho_k^n} - \frac{4}{3\text{Re} \rho_k^n} \frac{\tilde{V}_{k+1}^{n+1} - 2\tilde{V}_k^{n+1} - \tilde{V}_{k-1}^{n+1}}{h_y^2} = 0. \quad (42)$$

$$U_k^n \frac{e_k^{n+1} - e_k^n}{h_x} + V_k^n \frac{e_k^{n+1} - e_{k-1}^{n+1}}{h_y} + (\gamma - 1)e_k^n \frac{U_k^n - U_{k-1}^{n-1}}{h_x} + (\gamma - 1)e_k^n \frac{V_{k+1}^n - V_{k-1}^n}{2h_y} - \frac{\gamma}{\text{Re} \rho_k^n} \frac{e_{k+1}^{n+1} - 2e_k^{n+1} - e_{k-1}^{n+1}}{h_y^2} = 0 \quad (43)$$

#### Corrector

$$\tilde{U}_k^{n+1} \frac{\rho_k^{n+1} - \rho_k^n}{h_x} + \tilde{\rho}_k^{n+1} \frac{\tilde{U}_k^{n+1} - U_k^n}{h_x} + \tilde{V}_k^{n+1} \frac{\rho_k^{n+1} - \rho_{k-1}^{n+1}}{h_y} + \tilde{\rho}_k^{n+1} \frac{\tilde{V}_{k+1}^{n+1} - \tilde{V}_{k-1}^{n+1}}{2h_y} = 0. \quad (44)$$

$$\tilde{U}_k^{n+1} \frac{U_k^{n+1} - U_k^n}{h_x} + \tilde{V}_k^{n+1} \frac{U_k^{n+1} - U_{k-1}^{n+1}}{h_y} + \frac{\tilde{P}_k^{n+1} - P_k^n}{h_x \tilde{\rho}_k^{n+1}} - \frac{1}{\text{Re} \tilde{\rho}_k^{n+1}} \frac{U_{k+1}^{n+1} - 2U_k^{n+1} - U_{k-1}^{n+1}}{h_y^2} = 0. \quad (45)$$

$$\tilde{U}_k^{n+1} \frac{V_k^{n+1} - V_k^n}{h_x} + \tilde{V}_k^{n+1} \frac{V_k^{n+1} - V_{k-1}^{n+1}}{h_y} + \frac{\tilde{P}_{k+1}^{n+1} - \tilde{P}_{k-1}^{n+1}}{2h_y \tilde{\rho}_k^{n+1}} - \frac{4}{3\text{Re} \tilde{\rho}_k^{n+1}} \frac{V_{k+1}^{n+1} - 2V_k^{n+1} - V_{k-1}^{n+1}}{h_y^2} = 0. \quad (46)$$

$$\tilde{U}_k^{n+1} \frac{e_k^{n+1} - e_k^n}{h_x} + \tilde{V}_k^{n+1} \frac{e_k^{n+1} - e_{k-1}^{n+1}}{h_y} + (\gamma - 1)\tilde{e}_k^{n+1} \frac{\tilde{U}_k^{n+1} - U_k^n}{h_x} + (\gamma - 1)\tilde{e}_k^{n+1} \frac{\tilde{V}_{k+1}^{n+1} - \tilde{V}_{k-1}^{n+1}}{2h_y} - \frac{\gamma}{\text{Re} \tilde{\rho}_k^{n+1}} \frac{e_{k+1}^{n+1} - 2e_k^{n+1} - e_{k-1}^{n+1}}{h_y^2} = 0 \quad (47)$$

For the adjoint system a similar finite difference scheme was used. The main numerical feature of this system is engendered by the presence of a singular source term  $\delta(X - X^{est})\delta(Y - Y^{est})$ , which is related to the location of the estimated point. A mollification (smooth approximation of  $\delta$ -function) was used for the approximation of this term in part of the calculations in the form of  $\delta(x) \approx \exp(-X^2 / \sigma^2 - Y^2 / \sigma^2)$ .

### 3.4. Refining and bounding the error

Total expression for refinement of the functional determined by all first order terms of finite-difference scheme is derived using above discussed method and follows:

$$\begin{aligned}
\Delta\rho^{corr} = & -\frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 \rho(X_n, Y_k)}{\partial X^2} U_k^n \Psi_{\rho, k}^n h_{y, k} h_{x, n}^2 - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 \rho(X_n, Y_k)}{\partial Y^2} |V_k^n| \Psi_{\rho, k}^n h_{x, k} h_{y, n}^2 \\
& - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 U(X_n, Y_k)}{\partial X^2} \rho_k^n \Psi_{\rho, k}^n h_{y, k} h_{x, n}^2 \\
& - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 U(X_n, Y_k)}{\partial X^2} U_k^n \Psi_{U, k}^n h_{y, k} h_{x, n}^2 - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 U(X_n, Y_k)}{\partial Y^2} |V_k^n| \Psi_{U, k}^n h_{x, k} h_{y, n}^2 \\
& - \frac{\gamma-1}{2\rho_k^n} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 \rho(X_n, Y_k)}{\partial X^2} e_k^n \Psi_{U, k}^n h_{y, k} h_{x, n}^2 - \frac{\gamma-1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 e(X_n, Y_k)}{\partial X^2} \Psi_{U, k}^n h_{y, k} h_{x, n}^2 \\
& - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 V(X_n, Y_k)}{\partial X^2} U_k^n \Psi_{V, k}^n h_{y, k} h_{x, n}^2 - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 V(X_n, Y_k)}{\partial Y^2} |V_k^n| \Psi_{V, k}^n h_{x, k} h_{y, n}^2 \\
& - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 e(X_n, Y_k)}{\partial X^2} U_k^n \Psi_{e, k}^n h_{y, k} h_{x, n}^2 - \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 e(X_n, Y_k)}{\partial Y^2} |V_k^n| \Psi_{e, k}^n h_{x, k} h_{y, n}^2 \\
& - \frac{\gamma-1}{2} \sum_{k=1, n=2}^{N, Nx} \frac{\partial^2 U(X_n, Y_k)}{\partial X^2} e_k^n \Psi_{e, k}^n h_{y, k} h_{x, n}^2
\end{aligned} \tag{48}$$

Total expression for error bound has the form:

$$\begin{aligned}
\Delta\rho^{sup} = & \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 \rho(X_n, Y_k)}{\partial X^3} U_k^n \Psi_{\rho, k}^n \right| h_{y, k} h_{x, n}^3 + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 \rho(X_n, Y_k)}{\partial Y^3} V_k^n \Psi_{\rho, k}^n \right| h_{x, k} h_{y, n}^3 \\
& + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 U(X_n, Y_k)}{\partial X^3} \rho_k^n \Psi_{\rho, k}^n \right| h_{y, k} h_{x, n}^3 \\
& + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 U(X_n, Y_k)}{\partial X^3} U_k^n \Psi_{U, k}^n \right| h_{y, k} h_{x, n}^3 + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 U(X_n, Y_k)}{\partial Y^3} V_k^n \Psi_{U, k}^n \right| h_{x, k} h_{y, n}^3 \\
& + \frac{\gamma-1}{2\rho_k^n} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 \rho(X_n, Y_k)}{\partial X^3} e_k^n \Psi_{U, k}^n \right| h_{y, k} h_{x, n}^3 + \frac{\gamma-1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 e(X_n, Y_k)}{\partial X^3} \Psi_{U, k}^n \right| h_{y, k} h_{x, n}^3 \\
& + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 V(X_n, Y_k)}{\partial X^3} U_k^n \Psi_{V, k}^n \right| h_{y, k} h_{x, n}^3 + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 V(X_n, Y_k)}{\partial Y^3} V_k^n \Psi_{V, k}^n \right| h_{x, k} h_{y, n}^3 \\
& + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 e(X_n, Y_k)}{\partial X^3} U_k^n \Psi_{e, k}^n \right| h_{y, k} h_{x, n}^3 + \frac{1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 e(X_n, Y_k)}{\partial Y^3} V_k^n \Psi_{e, k}^n \right| h_{x, k} h_{y, n}^3 \\
& + \frac{\gamma-1}{2} \sum_{k=1, n=2}^{N, Nx} \left| \frac{\partial^3 U(X_n, Y_k)}{\partial X^3} e_k^n \Psi_{e, k}^n \right| h_{y, k} h_{x, n}^3
\end{aligned} \tag{49}$$

A bound on the refined functional error may be determined by these expressions as:

$$\left| \rho - \Delta\rho^{corr} - \rho_{exact} \right| < \Delta\rho^{sup} \tag{50}$$

This bound does not account for errors of adjoint problem solution, errors caused by boundary condition approximation etc. It also uses derivatives whose boundedness can not be proven at present. So, it should be investigated by means of numerical tests.



### 3.5. Numerical tests

First, we consider a smooth flow. The error of flow density past the expansion fan (Prandtl-Mayer flow) is addressed (freestream Mach number  $M=4$ , angle of flow deflection  $\alpha = 10^\circ$ ). Let us consider the related results for inviscid flow.

Fig. 8 presents the deviation of the finite-difference solution (density) from the analytic one and the correction of error in accordance with (48) (all divided by analytical value of the density). The refinement of the solution using adjoint parameters according to (48) enables the elimination of a major part of the discretization error. The first order of computable error may be detected if we analyze Fig. 8. Calculations demonstrated a good agreement of the refined solution with the analytical one and reliability of the error bound estimate. However, the order of the bound is slightly smaller than the second order of accuracy provided by (49). This is due to the slow growth of third derivatives of the calculated flow parameters as the step size decreases. It may be caused either by some properties of the finite-difference scheme or by the formation of weak discontinuities in the flowfield.

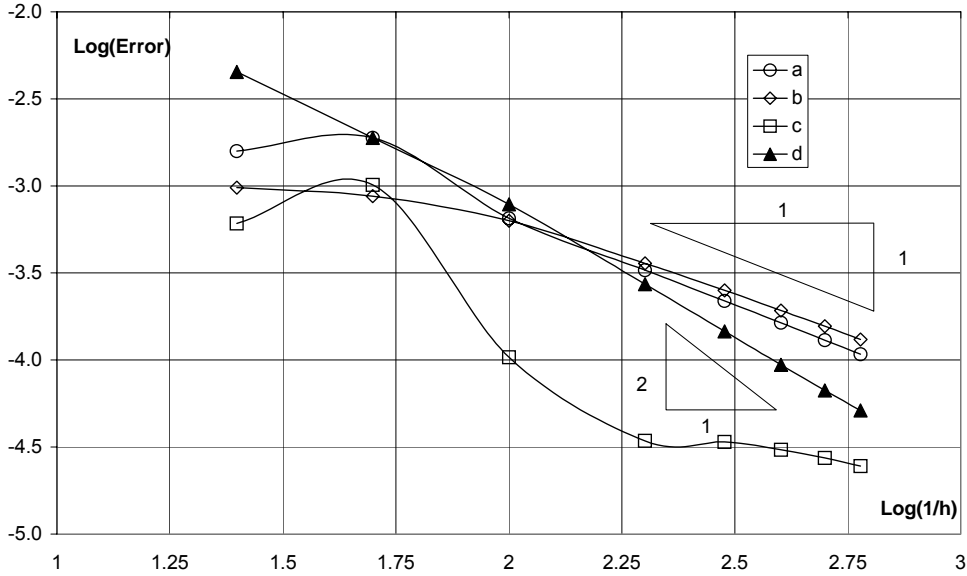


Fig. 8. The errors as functions of the reciprocal of mesh step (Logarithm scale).  
a-deviation of finite-difference solution from analytical one, b- error correction according (48).  
c- error of refined solution, d-bound of refined solution error

For comparison, let us consider the residual based approach closely related to [19] for an estimation of computable error without explicitly using the differential approximation. The truncation source term driving the error estimation has a formal appearance  $\delta\rho = L_h^{(1)}\rho_h - L\rho_h$  if we use the differential approximation. It may be estimated in other fashion as the residual  $\delta\rho = -L\rho_h'$  engendered by action of the differential operator on some extrapolation of the numerical solution [19]. Herein we use a different approach and estimate it as  $\delta\rho' = -L_h^{(2)}\rho_h$ . Here  $L_h^{(1)}$  is the finite difference operator of basic (low) precision,  $L_h^{(2)}$  is the finite difference operator of high precision and  $L$  is the differential operator. The main difference between this approach and the one in [19] is in the residual calculation. We do not use an interpolation of flow parameters from grid points to total domain. Instead, we apply a higher order scheme on

the same numerical solution.

Thus, the lower term of differential approximation may be estimated via the residual obtained from using a high order stencil on the solution calculated via main finite-difference scheme. Fig. 9 presents the deviation of the finite-difference solution from the analytic one, residual based correction of error and the error of refined solution.

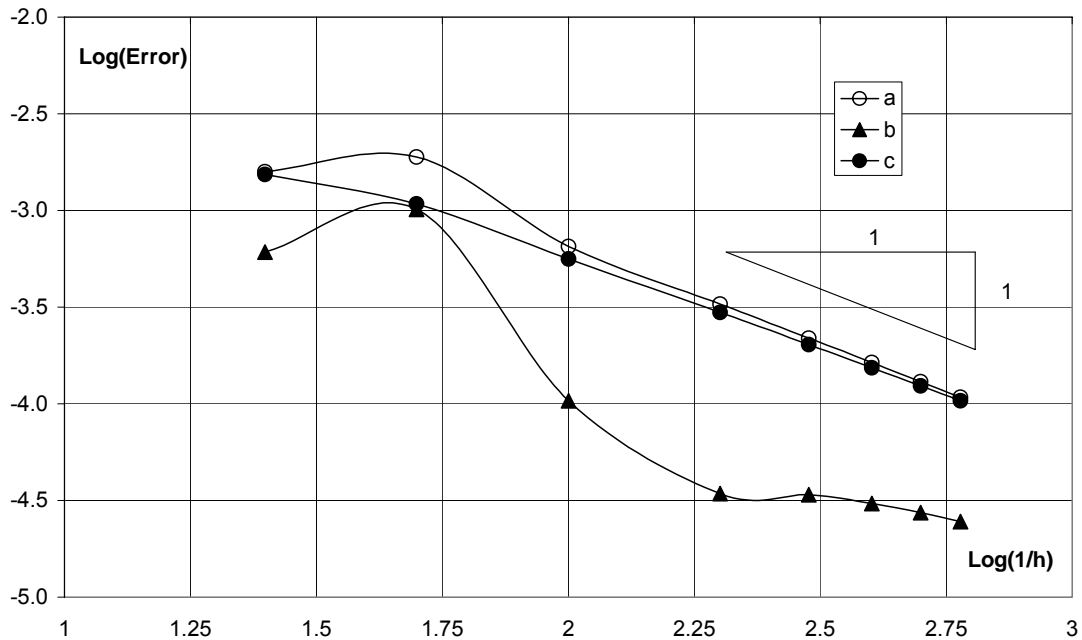


Fig. 9. The errors as functions of the reciprocal of mesh step (inviscid flow). a-deviation of finite-difference solution from analytical one, b- error of refined solution, c- residual based error estimation

The comparison of Figs. 8 and 9 demonstrates these two approaches to be very similar in as far as correction of numerical error is concerned. However, the differential approximation approach additionally yields an upper bound of the refined solution error.

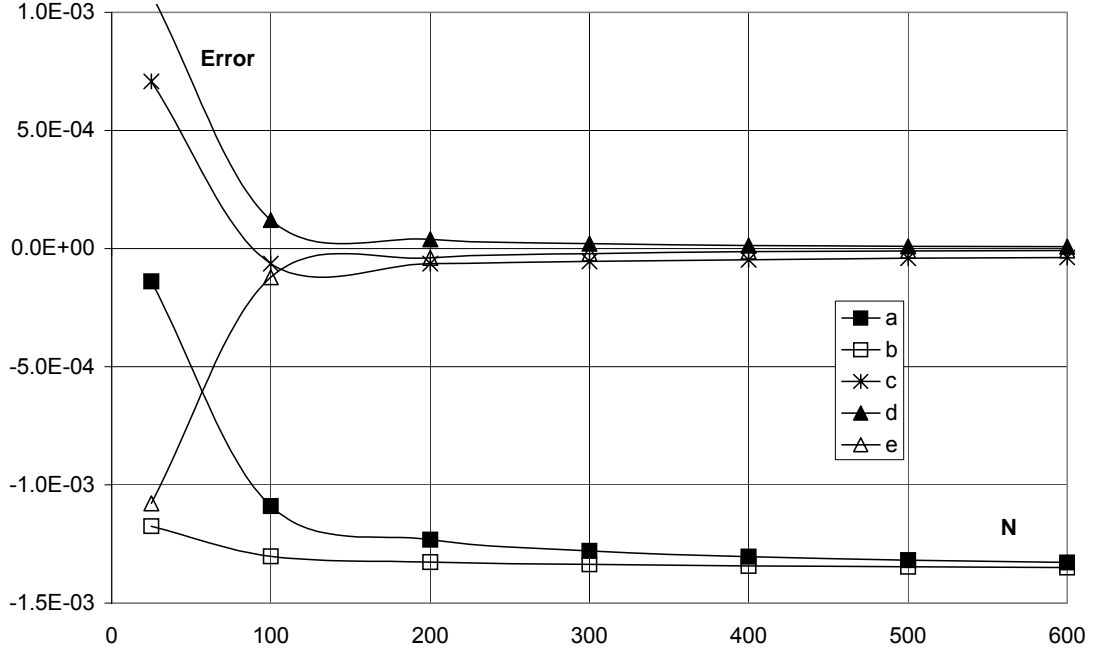


Fig. 10. The errors as functions of the number of grid points (viscous flow,  $Re=1000$ ). a- deviation on numerical from exact value, b-error due to viscous terms, c-deviation of refined solution from analytical one, d- upper bound of refined solution error, e-low bound of refined solution error

Let us consider results corresponding to calculations taking into account the viscosity. Fig. 10 presents the relative error of flow density calculation via PNS for  $Re=1000$  as a function of the number of nodes in  $Y$  direction. The part of error caused by viscous terms (39), relative deviation  $\frac{\rho - \Delta\rho^{corr} - \Delta\rho_{visc} - \rho_{exact}}{\rho}$  of refined solution from the analytical one, and bound

of refined solution error (49) are presented. It can be seen that the main part of error is determined by viscosity and it may be computed and eliminated. Figure 10 demonstrates that the estimation of viscosity impact using adjoint equations enables us to obtain a result close to the inviscid computation as far as accuracy is concerned. Thus, there exists the feasibility for calculation of inviscid flow (Euler equations) and a posteriori error estimation on the basis of PNS equations. This extends the applicability of the considered method which is not directly applicable to the supersonic Euler equations due to the existence of discontinuous solutions. In general, for a smooth flow the errors for both inviscid flow and for viscous flow (refined via adjoint parameters) are close.

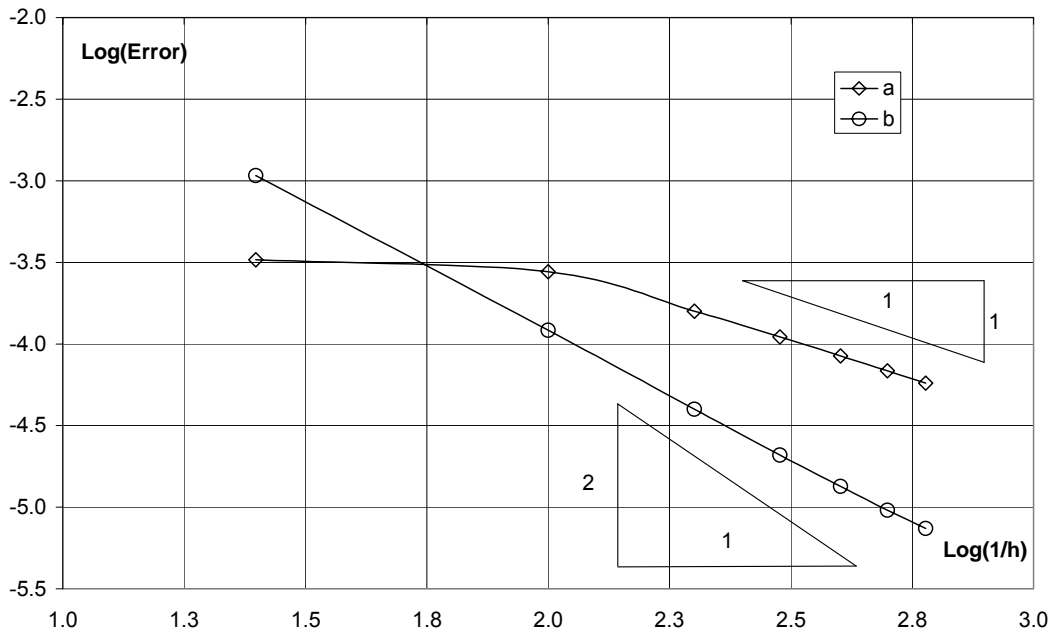


Fig. 11. The errors as functions of the reciprocal of mesh step (viscous flow,  $Re=1000$ ).  
a- error correction, b- bound of error

As can be seen from Figs. 10 and 11 the correction term has a first order of accuracy, the bound order is slightly less than two, however the error remaining after taking into account the viscous term is greater than the bound for fine enough meshes. So, the account of viscosity impact is not accurate enough in that it is limiting the comparison of calculations and analytical data for viscous flowfield.

As another test, the error of the density past crossing shocks ( $\alpha = \pm 22.23^\circ$ ,  $M=4$ ,  $Re=1000$ ) is evaluated. Fig. 12 presents the density isolines within flowfield, Fig. 13 shows the spatial distribution of error bound according (49). This information may be considered as the spatial distribution of the incomputable numerical error and used as guidance for choice of mesh refining.

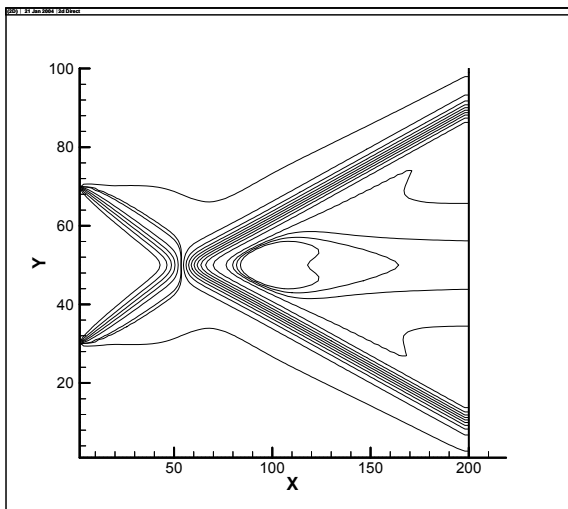


Fig. 12. Isolines of density (crossing shocks)

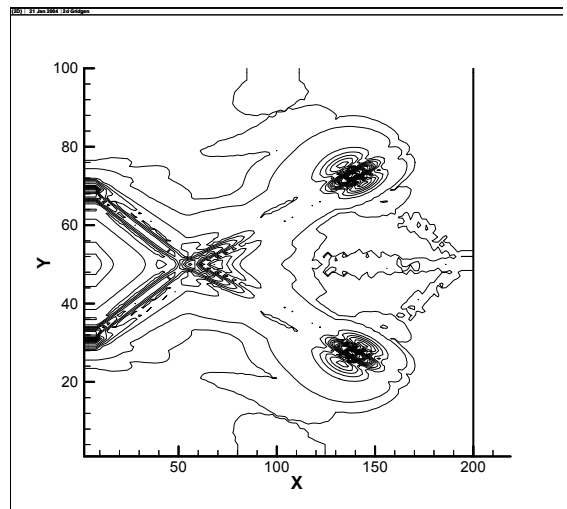


Fig. 13. Isolines of error bound density (36)

This test is more complicated due to presence of unbounded derivatives of gasdynamics parameters for inviscid flow. The presence of viscosity enables us to calculate flows with shocks, while at the same time it introduces an error proportional to  $1/Re$ .

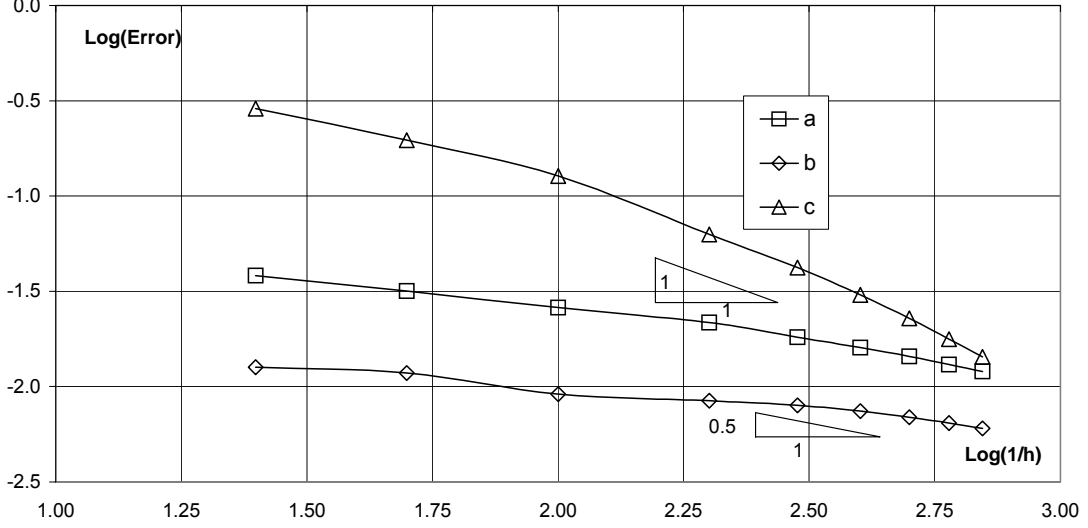


Fig. 14. The errors as functions of the reciprocal of mesh step (viscous shocked flow).  
a- deviation of refined solution from analytical one, b- error correction term (48), c-error bound (49)

Fig. 14 presents results for  $Re=1000$  as a function of the spatial step size. The adjoint correction and the deviation of numerical solution from exact one has an order less 0.5 that provides restrictions for adjoint bounding. So, the calculation of errors for shocked flows poses a significant challenge for further analysis.

### 3.6 Divergent Euler equations

A common way to handle discontinuous flows is the use of conservative form of equations and divergent finite-difference scheme. Unfortunately, the differential approximation based error correction and bounds converges only in the one-dimensional case. Let us now consider a two-dimensional problem.

The following systems of divergent Euler equations (steady, two-dimensional) and related adjoint equations were used in numerical tests.

$$\frac{\partial(\rho U^k)}{\partial X^k} = 0; \quad (38)$$

$$\frac{\partial(\rho U^k U^i + P \delta_{ik})}{\partial X^k} = 0; \quad (39)$$

$$\frac{\partial(\rho U^k h_0)}{\partial X^k} = 0; \quad (40)$$

Here  $U^1 = U, U^2 = V$ ,  $h(\rho, P) = \gamma e$  is the enthalpy,  $h_0 = (U^2 + V^2)/2 + h$  is the total enthalpy.

Adjoint equations:

$$U^k \frac{\partial \Psi_\rho}{\partial X^k} + U^k U^i \frac{\partial \Psi_i}{\partial X^k} + \frac{\gamma-1}{\gamma} \frac{\partial \Psi_k}{\partial X^k} (h_0 - U_n U_n / 2) + U^k h_0 \frac{\partial \Psi_h}{\partial X^k} - \delta(X - X^{est}) \delta(Y - Y^{est}) = 0 \quad (41)$$

$$U^i \frac{\partial \Psi_i}{\partial X^k} + U^i \frac{\partial \Psi_k}{\partial X^i} + \frac{\partial \Psi_\rho}{\partial X^k} - \frac{\gamma-1}{\gamma} \frac{\partial \Psi_n}{\partial X^n} U_k + h_0 \frac{\partial \Psi_h}{\partial X^k} = 0; \quad (42)$$

$$U^k \frac{\partial \Psi_h}{\partial X^k} + \frac{\gamma-1}{\gamma} \frac{\partial \Psi_k}{\partial X^k} = 0; \quad (43)$$

The equations (41-43) do not contain any derivatives of the field parameters in contrast to system (34-37) and thus should provide for a better performance for discontinuous solutions.

Two dimensional first order finite-difference schemes were used namely ("donor cells" [43] and a scheme of Courant-Isaacson-Rees [44]) with practically identical results. The expressions for truncation error are obtained in a way similar to (48,49) and are omitted herein due to their very bulky form. As expected, the deviation of the finite-difference solution from the analytic one for divergent scheme is significantly smaller compared with the nondivergent one and the solution is monotonic enough. Nevertheless, the error estimates do not converge. This is caused by the fact that the error estimates use derivatives that are also unbounded in the divergent case (excluding one-dimensional flow).

If we introduce viscosity terms into the systems (38-40) and (41-43), we can obtain convergent estimates of the error for the divergent scheme too (Fig. 15). The comparison of Figures 14 and 15 demonstrates the improved behavior of the divergent system when compared with (29-32) and (34-37).

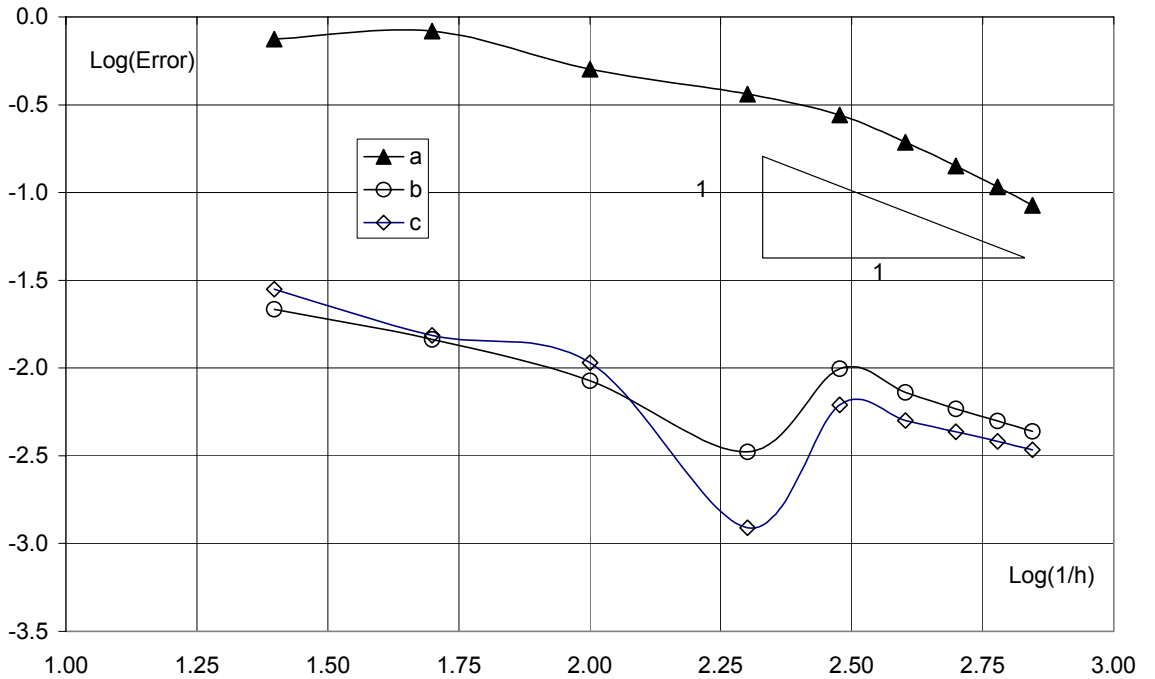


Fig. 15. The error of calculation as a function of the reciprocal of mesh step (viscous flow, divergent scheme).

a -error bound, b- deviation of refined solution from analytical one, c-adjoint error correction

#### 4. Discussion

The calculation of discretization errors using differential approximation and adjoint equations requires the existence of bounded derivatives of a relatively high order. They do not always exist, so, for supersonic Euler equations, these estimates may be calculated only for smooth solutions.

The second order convergence predicted by formal analysis was found in numerical tests only for inviscid continuous flows. This may be related to the lack of solution smoothness for both the PNS and Euler equations. For solutions with an infinite number of bounded derivatives (heat conduction) similar error estimates exhibited the predicted order of convergence.

For discontinuous flow the use of viscosity enables us to carry out these error estimates, although numerical tests revealed a very small order of convergence. The viscosity engenders its own component of error, which may also be eliminated using adjoint equations.

In general, the calculation of error for discontinuous flows poses a significant challenge and requires further research and analysis.

For justification of error estimates we should verify that the unaccounted error component induced by approximation error of adjoint equations is small enough. This condition is satisfied asymptotically if the order of approximation of both primal and adjoint problem is high enough. On other hand, we can solve second order adjoint equations [39] for calculation of this component in a manner similar to [30].

The computed fields used for error estimations may have numerical oscillations providing the growth of norm of high order derivatives. Thus, for non-monotonic finite-difference schemes the error bounds may be too large.

#### 5. Conclusion

The presentation of the truncation error in Lagrange form provides an opportunity for subdivision of approximation error into computable and incomputable parts. The computable part enables refinement of the solution using adjoint equations. The asymptotic bound of the refined solution error may be determined simultaneously using Holder inequality.

The method is directly applicable for continuous solutions and monotonic finite-difference schemes.

Numerical tests demonstrated the efficiency of this method for pointwise error estimation on examples of heat conduction equation and parabolized Navier-Stokes.

## References

1. I. Babushka and A. D. Miller, The post-processing approach in the finite element method, iii: A posteriori error estimation and adaptive mesh selection, *Int. J. Numer. Meth. Eng.*, **20**, pp. 2311–2324 (1984).
2. M Ainsworth. and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*. Wiley – Interscience, NY. (2000).
3. J.T. Oden and S. Prudhomme, Goal-Oriented Error Estimation and Adaptivity for the Finite Element Method, *Computers&Mathematics with Applic.*, **41**, 735 (2001).
4. J.T. Oden and S. Prudhomme, Estimation of modeling error in computational mechanics. *Journal of Computational Physics*, **182**, 496 (2002).
5. S. Prudhomme and J. T. Oden, On goal-oriented error estimation for elliptic problems: Application to the control of pointwise errors, *Computer Methods in Applied Mechanics and Engineering*, **176**, 313-331 (1999).
6. C. Johnson. On Computability and Error Control in CFD. *International J. for Numerical Methods in Fluids*, **20**, 777 (1995).
7. J. Hoffman and C. Johnson, Computability and adaptivity in CFD, in *Encyclopedia of Computational Mechanics Vol 3: Fluids*, Erwin Stein, Rene de Borst and Thomas J.R. Hughes (eds), Wiley & Sons: (2004 ).
8. R. Hartmann and P. Houston. Goal-Oriented A Posteriori Error Estimation for Compressible Fluid Flows. In *Numerical Mathematics and Advanced Applications*, Brezzi F., Buffa A., Corsaro S., and Murli A. (eds), Springer-Verlag, 775 (2003).
9. W. Bangerth, R. Rannacher, Finite Element Approximation of the Acoustic Wave Equation: Error Control and Mesh Refinement, *East-West J. of Numer. Math.*; **7/4**: 263-282 (1996).
10. R. Becker, R. Rannacher An optimal control approach to a posteriori error estimation in finite element methods. In A. Iserles, editor, *Acta Numerica*, Cambridge Univ. Press, 1-102 (2001).
11. V. Heuveline, R. Rannacher, Duality-based adaptivity in the hp-finite element method, *J. Numer. Math.*, **11**, No. 2, 1–18 (2003).
12. P. Houston, R. Rannacher, and E. Suli. A posteriori error analysis for stabilized finite element approximations of transport problems. *Comput. Methods Appl. Mech. Engrg.*, **190** (11-12), 1483–1508 (2000).
13. P. Monk and E. Suli. The adaptive computation of far field patterns by a posteriori error estimates of linear functionals. *SIAM J. Numer. Anal.*, **36**(1), 251–274. (1998).
14. E. Suli, P. Houston, Adjoint error correction for integral outputs, in: *Adaptive Finite Element Approximation of Hyperbolic Problems*, , Springer, (2002).
15. E. Suli and P. Houston. Finite element methods for hyperbolic problems: a posteriori error analysis and adaptivity. In: I. Duff and G.A. Watson, eds., *The State of the Art in Numerical Analysis*. Oxford University Press, 441-471 (1997).
16. E. Suli. A posteriori error analysis and adaptivity for finite element approximations of hyperbolic problems. In: D. Krüner, M. Ohlberger and C Rohde (Eds.) *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws. Lecture Notes in Computational Science and Engineering*. Volume 5,. Springer-Verlag, 123-194 (1998).
17. L. Ferm and Per Lötstedt, Adaptive error control for steady state solutions of inviscid flow, *SIAM J. Sci. Comput.*, **23** 1777-1798 (2002).
18. D. J. Estep, M. G. Larson, and R. D. Williams, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, *Memoirs A.M.S.* **146**, 1-109 (2000).
19. M. B. Giles On adjoint equations for error analysis and optimal grid adaptation in CFD. In *Computing the Future II: Advances and Prospects in Computational Aerodynamics*. Hafez M. and Caughey D. A. (eds) Wiley, New York, p. 155 (1998).
20. M. Giles and N.A Pierce, Improved Lift and Drag Estimates Using Adjoint Euler Equations.



- Technical Report 99-3293, AIAA, Reno, NV, 1 (1999).*
21. N. A. Pierce. and M. B. Giles. Adjoint recovery of superconvergent functionals from PDE approximations. *SIAM Rev.*, **42**, 247 (2000).
  22. M. B. Giles and E. Suli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, 145 –236 (2002).
  23. M.B. Giles, N.A. Pierce and E. Suli. Progress in adjoint error correction for integral functionals, *Computing and Visualisation in Science*, **6**, 2 (2004).
  24. N.A. Pierce, M.B. Giles, Adjoint and defect error bounding and correction for functional estimates. *Journal of Computational Physics*; **200**, 769-794 (2004).
  25. D. Venditti and D. Darmofal. Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. *J. Comput. Phys.* **164**, 204 (2000).
  26. D. Venditti and D. Darmofal. Grid Adaptation for Functional Outputs: Application to Two-Dimensional Inviscid Flow. *J. Comput. Phys.* **176**, 40 (2002).
  27. D. Darmofal and D. Venditti. Anisotropic grid adaptation for functional outputs: application to two-dimensional viscous flows. *J. Comput. Phys.*, **187**, 22 (2003).
  28. M. A. Park, Three-Dimensional Turbulent RANS Adjoint-Based Error Correction, *AIAA Paper 2003-3849*, 1 (2003).
  29. M. A. Park, Adjoint-Based, Three-Dimensional Error Prediction and Grid Adaptation. *AIAA Paper 2002-3286*, 1 (2002).
  30. A.K. Alekseev, I.M. Navon, On a-posteriori pointwise error estimation using adjoint temperature and Lagrange remainder. *Computer Methods in Applied Mechanics and Engineering*, **194/18-20** p. 2211-2228, (2005).
  31. A.K. Alekseev, I.M. Navon A posteriori pointwise error estimation for compressible fluid flows using adjoint parameters and Lagrange remainder, *International Journal for Numerical Methods in Fluids*, **47**, (1) p. 45–74, (2005).
  32. Yu.I. Shokin, *Method of differential approximation*. Springer-Verlag (1983).
  33. O. A. Ladyzenskaja, V.A. Solonnikov, and N.N.Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Trans. Math. Monograph 23, American Mathematical Society, Providence, RI, (1968).
  34. G.I. Marchuk and V.V. Shaidurov, *Difference methods and their extrapolations*. Springer: N.Y. (1983).
  35. J.L. Lions, Pointwise Control of Distributed Systems, in H.T. Banks, ed., *Control and estimation in distributed Parameter systems. Frontiers in applied mathematics*, SIAM, **11**, 1-39 (1992).
  36. A.K. Tornberg and B. Engquist. Regularization techniques for numerical approximation of PDEs with singularities. *J. of Sci. Comput.*, **19**, 527-552 (2003).
  37. J. Walden. On the approximation of singular source terms in differential equations. *Numer. Meth. Part. D*, **15** 503-520 (1999).
  38. G. I. Marchuk. *Adjoint Equations and Analysis of Complex Systems*. Kluwer Academic Publishers, Dordrecht, Hardbound (1995).
  39. Z. Wang, I.M. Navon, F.X. Le Dimet, and X. Zou. The Second Order Adjoint Analysis: Theory and Applications. *Meteorol. Atmos. Phys.*, **50**, 3 (1992).
  40. A. Alekseev. and I.M. Navon. On Estimation of Temperature Uncertainty Using the Second Order Adjoint Problem. *Int. Journal of Comput. Fluid Dyn.*, **16 (2)**, 113 (2002).
  41. A. A. Samarskii, *Theory of difference schemes*, Marcel Dekker, NY, (2001).
  42. A.K. Alekseev. 2D Inverse Convection Dominated problem for Estimation of Inflow Parameters from Outflow Measurements. *Inverse Problems in Eng.* **8**, 413 (2000)
  43. P. Roache. *Computational Fluid dynamics*. Hermosa publisher (1976).
  44. A.G. Kulikovskii, N.V. Pogorelov, and A.Yu. Semenov. *Mathematical Aspects of Numerical Solution of Hyperbolic Systems*. Monographs and Surveys in Pure and Applied Mathematics, **188**, Chapman&Hall/CRC, Boca Raton, Fl (2001).

### Figure Captions

- Fig. 1. Initial and final temperature distribution. 1 - Initial temperature, 2- Final temperature
- Fig. 2. Variation of errors as a function of temporal step (in decimal logarithm scale). a – deviation of calculated temperature from analytical value ( $T - T_{an}$ ), b- correction of temperature  $\Delta T_x^{corr}$  (18), c- refined solution error bound  $\Delta T_{x,1}^{sup}$  (19), d-discrepancy between refined and analytic solutions ( $T_x^{corr} - T_{an}$ ).
- Fig. 3. The comparison of numerical and refined solutions (all divided by analytical value). a- numerical, b-refined solution, c- lower bound, d- upper bound
- Fig. 4. Variation of errors as a function of spatial step (in logarithm scale). a – Deviation of calculated temperature from analytical value ( $T - T_{an}$ ), b- Correction of temperature  $\Delta T_x^{corr}$ , c- Refined solution error bound  $\Delta T_{x,1}^{sup}$ , d- Deviation of refined solution from analytical value ( $T_x^{corr} - T_{an}$ )
- Fig. 5. Variation of errors as a function of spatial step (in logarithm scale) for break in first derivative caused by discontinuous conductivity a - Correction of temperature, b- error bound of refined solution
- Fig. 6. Initial and final temperature distribution. 1- Initial temperature, 2- Final temperature, A- point of estimation
- Fig. 7. Variation of errors as a function of spatial step (in logarithm scale) for break in first derivative caused by discontinuous conductivity. a - Correction of temperature, b- error bound of refined solution, c-deviation of numerical solution from analytical value
- Fig. 8. The errors as functions of the reciprocal of mesh step (Logarithm scale). a-deviation of finite-difference solution from analytical one, b- error correction according (48). c- error of refined solution, d-bound of refined solution error
- Fig. 9. The errors as functions of the reciprocal of mesh step (inviscid flow). a-deviation of finite-difference solution from analytical one, b- error of refined solution, c- residual based error estimation
- Fig. 10. The errors as functions of the number of grid points (viscous flow,  $Re=1000$ ). a- deviation on numerical from exact value, b-error due to viscous terms, c-deviation of refined solution from analytical one, d- upper bound of refined solution error, e-low bound of refined solution error
- Fig. 11. The errors as functions of the reciprocal of mesh step (viscous flow,  $Re=1000$ ). a- error correction, b- bound of error
- Fig. 12. Isolines of density (crossing shocks)
- Fig. 13. Isolines of error bound density (36)
- Fig. 14. The errors as functions of the reciprocal of mesh step (viscous shocked flow). a- deviation of refined solution from analytical one, b- error correction term (48), c-error bound (49)
- Fig. 15. The error of calculation as a function of the reciprocal of mesh step (viscous flow, divergent scheme). a -error bound, b- deviation of refined solution from analytical one, c-adjoint error correction

