

# Discretization of the stationary convection-diffusion-reaction equation

**Citation for published version (APA):**

Hof, van 't, B., Thijs Boonkkamp, ten, J. H. M., & Mattheij, R. M. M. (1996). *Discretization of the stationary convection-diffusion-reaction equation*. (RANA : reports on applied and numerical analysis; Vol. 9602). Eindhoven: Technische Universiteit Eindhoven.

**Document status and date:**

Published: 01/01/1996

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

EINDHOVEN UNIVERSITY OF TECHNOLOGY  
Department of Mathematics and Computing Science

RANA 96-02  
February 1996

Discretization of the Stationary  
Convection-Diffusion-Reaction Equation

by

B. van 't Hof  
J.H.M. ten Thije Boonkamp  
R.M.M. Mattheij



Reports on Applied and Numerical Analysis  
Department of Mathematics and Computing Science  
Eindhoven University of Technology  
P.O. Box 513  
5600 MB Eindhoven  
The Netherlands  
ISSN: 0926-4507

# Discretization of the Stationary Convection-Diffusion-Reaction Equation

B. van 't Hof, J.H.M. ten Thije Boonkamp, R.M.M. Mattheij

*Eindhoven University of Technology  
Department of Mathematics and Computing Science  
P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

## Abstract

In this paper we present a new finite volume discretization method based on an exponential flux approximation scheme. It can be applied to problems in which convective and diffusive terms, as well as a production term play a role; equations of this type are found among others in combustion theory. We show that the errors only depend on the variation of the coefficients within grid size, and not on the magnitude of local Peclet numbers. The scheme is shown to be second order accurate uniformly in Peclet numbers; which we believe to be unique for this method.

**Key words.** convection-diffusion-reaction equation, finite volume method, flux computation.

**AMS subject classifications.** Primary: 65N06, Secondary: 76M20, 65N12

## 1 Introduction

In many combustion problems, including laminar flames, the flow of a reacting gas mixture has to be modelled along with the chemical reactions between the constituent species. The conservation laws for reacting gas flow and the theory of chemical kinetics form the basis of combustion theory. These equations describe conservation of mass, momentum and energy of the mixture and change of mass of each individual species. Most of these equations contain a convection term, a diffusion or conduction term and a source term representing the production of a species or the production of heat due to chemical reactions. Equations of this type are referred to as convection-diffusion-reaction equations.

The combustion equations are so complicated that they can only be solved numerically. The computation of a numerical solution of the combustion equations requires the discretization of these equations and subsequently the iterative solution of the resulting set of algebraic equations. In this paper we focus on the discretization of the convection-diffusion-reaction equation.

The discretization of the convection-diffusion-reaction equation in this paper is based on a scheme presented by Thiart [14, 15]. This scheme has already been used successfully in laminar flame computations (see [6] or [10]). Thiart's method is essentially a finite volume method, combined with an exponential scheme for the flux computation. In fact, the fluxes are computed locally from the conservation equations.

The scheme we propose has the following three properties. First, it is second order accurate both for diffusion dominated and for convection dominated flows. Secondly, it does not produce oscillations in the vicinity of steep gradients when convection is dominant. Thirdly, it uses a 3-point,  $3 \times 3$ -point or  $3 \times 3 \times 3$ -point stencil for 1, 2 or 3 dimensional problems respectively. No other discretization method we know of unifies these three qualities. Comparable exponential schemes [14, 15, 11, 8] acquire second order accuracy only when the mesh size and local Peclet numbers are sufficiently small. Central schemes produce oscillations when convection is dominant,

and upwind schemes lose accuracy when convection is dominant. Higher order upwind schemes, like the QUICK-scheme by Leonard [7] use a larger molecule than the proposed scheme. Methods based on the use of flux-limiters [12, 13], besides using a larger discretization molecule, produce non-linear equations and are therefore not comparable to the current method.

We have organized this paper as follows. In Section 2, the finite volume method is summarized for rectangular grids. The computation of the fluxes is presented in Section 3. Furthermore, it is demonstrated that the numerical flux is second order accurate. Combination of the finite volume discretization and the numerical fluxes gives the Modified Thiart Scheme, which is presented in Section 4. Limiting cases of the scheme, when convection or diffusion are dominant, are given in Section 5. Next the global discretization error is derived in Section 6 and finally, in Section 7, a few numerical examples are given.

## 2 Finite Volume Discretization

Consider the stationary conservation law for a quantity  $\phi$ , subject to convection, diffusion and a (chemical) source term. In order to formulate a conservation law, let  $\Omega \subset \mathbf{R}^n$ ,  $n = 1, 2, 3$  denote the domain in which the process takes place. Let  $V \subset \Omega$  be an arbitrary subdomain,  $\delta V$  the boundary of  $V$  and  $\mathbf{n}$  the outward unit vector normal to  $\delta V$ . Let  $\rho v$  denote the mass flux,  $\Gamma$  a general diffusion coefficient  $\Gamma \geq \Gamma_{min} > 0$ , and  $s$  the (chemical) source term. The integral formulation of this conservation law then reads:

$$\oint_{\delta V} (\rho v \phi - \Gamma \nabla \phi) \cdot \mathbf{n} \, d\gamma = \int_V s \, d\Omega. \quad (2.1)$$

If all variables are sufficiently smooth, this equation is equivalent to the differential equation:

$$\nabla \cdot (\rho v \phi - \Gamma \nabla \phi) = s. \quad (2.2)$$

In combustion problems, the source term  $s$  can be interpreted as a production/consumption term for chemical species or a production term for heat. In this case  $s$ , the diffusion coefficient  $\Gamma$  and the mass flux  $\rho v$  are nonlinear functions of  $\phi$ . However, since this functional dependence has no influence on the discretization itself, the dependence of  $s$ ,  $\Gamma$  and  $\rho v$  on  $\phi$  is ignored and they are assumed to be given functions of  $\mathbf{x}$ . In dealing with (2.1) and (2.2), it is convenient to introduce the flux vector  $\mathbf{f}$ :

$$\mathbf{f} := \rho v \phi - \Gamma \nabla \phi, \quad (2.3)$$

after which (2.1) simplifies to

$$\oint_{\delta V} \mathbf{f} \cdot \mathbf{n} \, d\gamma = \int_V s \, d\Omega. \quad (2.4)$$

In this paper we will discuss finite volume methods for (2.4), thereby restricting ourselves to the one- and two-dimensional cases.

In finite volume methods, the computational domain  $\Omega$  is covered with a finite set of  $N$  control volumes  $\mathcal{V} = \{V_1, \dots, V_N\}$ , and (2.4) is imposed for each of these control volumes. The integrals occurring in (2.4) are subsequently approximated by quadrature rules. Note that a meaningful approximation of (2.4) should have the equivalent of the property

$$\oint_{\delta V_i} \mathbf{f} \cdot \mathbf{n} \, d\gamma + \oint_{\delta V_j} \mathbf{f} \cdot \mathbf{n} \, d\gamma = \oint_{\delta \tilde{V}} \mathbf{f} \cdot \mathbf{n} \, d\gamma, \quad (2.5)$$

with  $\tilde{V} = V_i \cup V_j$  the union of two arbitrary adjacent control volumes. Equation (2.5) implies conservation for  $\phi$  in  $\tilde{V}$  if  $\phi$  is conserved in  $V_i$  and  $V_j$ . This means that a discrete conservation law holds for any subdomain of  $\Omega$  that is simply covered by elements of  $\mathcal{V}$ , like it is in the original physics of the problem. Therefore also global conservation for  $\phi$  on  $\Omega$  holds, which is why finite volume methods are called conservative.

Let us now restrict our attention to two-dimensional problems. The flux vector  $\mathbf{f}$  consists of two components, so we write  $\mathbf{f} =: (F, G)^T$ . When  $\Omega$  is covered by rectangular control volumes and when the integrals in (2.4) are approximated by the midpoint rule, we obtain the following discrete conservation law:

$$\Delta y(F_e - F_w) + \Delta x(G_n - G_s) = s_P \Delta x \Delta y, \quad (2.6)$$

where  $s_P$  denotes  $s(\mathbf{x}_P)$ ,  $F_e$  denotes the numerical approximation of  $F(\mathbf{x}_e)$  etc.; see Figure 1. The finite volume discretization has to be completed with the numerical calculation of  $F_e, G_n$  etc.

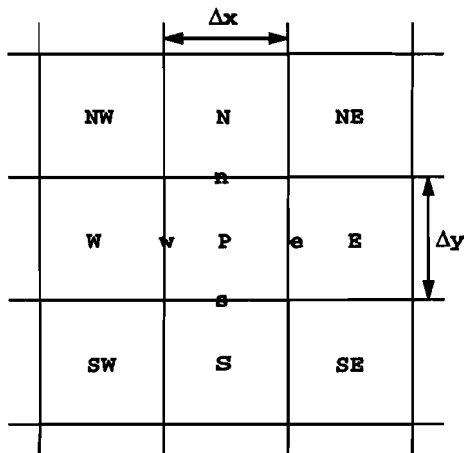


Figure 1: The position names in the grid.

This will be carried out in Section 3.

In finite volume methods, the choice of the control volume set  $\mathcal{V}$  is essential, and even for the case of uniform, structured grids there are several possibilities. The options are known as the cell-centered and cell-vertex approaches. In the cell-centered approach, the domain is divided into control-volumes, after which a grid point is assigned to every control volume. In the cell-vertex approach, the reverse is done: a mesh is generated, after which control volumes are constructed, sometimes one around every grid point [1], sometimes using the grid points as cell vertices [4] and sometimes both is done at the same time [16], [9]. In Figures 2 to 5 it is seen that these approaches differ indeed. The cell-vertex approach has advantages when discretizing the Euler equations because a four-point stencil is sufficient for a second-order discretization [4]. Also, stable discretizations have been obtained for the Navier-Stokes equations on a non-uniform or a curvilinear grid, where the cell-centered approach had failed [16].

In combustion simulation, local grid refinement is absolutely necessary. This can be done in several ways (see for two examples [4] and [2]). One of these is the (smooth) transformation of the uniform grid, so that grid lines will become closer where more detail is needed. Another method is to keep the uniform grid in tact as much as possible and add detail where necessary: this approach produces locally uniform grids (see Figures 6 and 7).

### 3 Flux Computation

Once we have chosen grid points and control volumes, the integrals in (2.4) have to be approximated. Often, these integrals are approximated by the (second order) mid-point rule, so that the discretization for a rectangular control volume is given by (2.6). However, it is still not clear how  $F_e, F_w, G_n$  and  $G_s$  have to be calculated. In this section we derive formulas for these fluxes, which are second order accurate both for strongly convective and strongly diffusive flows. More accurate

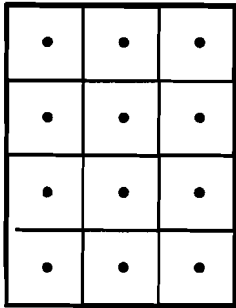


Figure 2: Cell-centered grid.

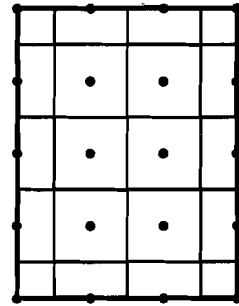


Figure 3: Cell-vertex grid

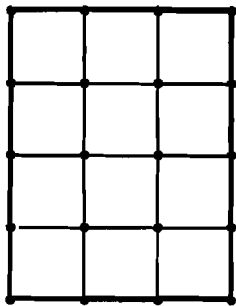


Figure 4: Cell-vertex grid for Euler equations, according to Hall [4].

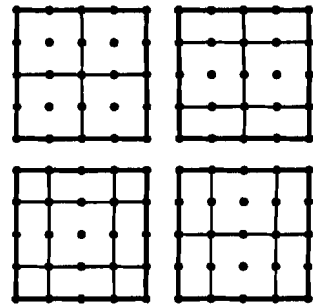


Figure 5: Cell-vertex grid with four sets of control volumes, according to Wagener [16].

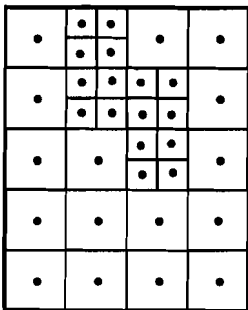


Figure 6: Locally uniform grid refinement in a cell-centered grid.

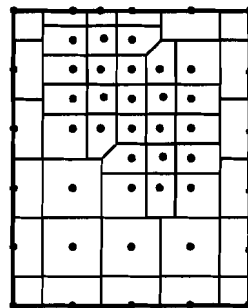


Figure 7: Locally uniform grid refinement in a cell-vertex grid.

approximations would be meaningless because of the second order accuracy of the approximations of the integrals in (2.4).

We first consider the one-dimensional conservation law. Generalizing the result which we will obtain to two-dimensional problems will be discussed later. Our goal is to find an accurate approximation of  $F(x_e)$ , given the values of  $\phi$  at the points  $x_P$  and  $x_E$ . For this it makes sense to consider a (1-D) two-point boundary value problem first

$$\begin{aligned} \frac{d}{dx} \left( \rho u \phi - \Gamma \frac{d\phi}{dx} \right) &= s \quad , \quad x_P < x < x_E, \\ \phi(x_P) &= \phi_P \quad , \quad \phi(x_E) = \phi_E. \end{aligned} \quad (3.1)$$

The flux  $F$  corresponding to (3.1) reads

$$F = \rho u \phi - \Gamma \frac{d\phi}{dx}. \quad (3.2)$$

Define  $\lambda$ ,  $P$  and  $S$  by

$$\lambda := \frac{\rho u}{\Gamma}, \quad P := \lambda \Delta x, \quad S(x) := \int_{x_e}^x s(\xi) d\xi. \quad (3.3)$$

Recall that  $\Delta x = x_E - x_P$  and that  $x_e = \frac{1}{2}(x_P + x_E)$ . We can derive the following result about the flux  $F$ .

**Lemma 3.1** . *The flux  $F(x_e)$  is given exactly by the following integral formulation:*

$$F(x_e) = \frac{\exp\left(\int_{x_P}^{x_E} \lambda dx\right) \phi_P - \phi_E}{\int_{x_P}^{x_E} \frac{1}{\Gamma} \exp\left(\int_x^{x_E} \lambda d\xi\right) dx} - \frac{\int_{x_P}^{x_E} \frac{1}{\Gamma} S(x) \exp\left(\int_x^{x_E} \lambda d\xi\right) dx}{\int_{x_P}^{x_E} \frac{1}{\Gamma} \exp\left(\int_x^{x_E} \lambda d\xi\right) dx}. \quad (3.4)$$

**Proof.** The differential equation (3.1) is linear with variable coefficients. For this class of boundary value problems, there is a standard solution approach, so that we can find an integral representation for  $\phi$ . However, since our object is not to calculate  $\phi$ , but to calculate  $F(x_e)$ , we do the manipulations a little differently from the standard approach. We integrate (3.1) and substitute (3.2) to obtain

$$\frac{dF}{dx} = s \quad \Leftrightarrow \quad F(x) - F(x_e) = S(x) \quad \Leftrightarrow \quad \Gamma \left( \lambda \phi - \frac{d\phi}{dx} \right) = F(x_e) + S(x).$$

Here, the flux  $F(x_e)$ , which is quantity to be calculated, enters the formula. We find a first-order ordinary differential equation for  $\phi$ . When we apply the boundary condition  $\phi(x_P) = \phi_P$ , we obtain the following expression for  $\phi(x)$ :

$$\begin{aligned} \frac{d\phi}{dx}(x) &= \lambda(x)\phi(x) - \frac{1}{\Gamma(x)}F(x_e) - \frac{S(x)}{\Gamma(x)} \quad \Rightarrow \\ \phi(x) &= \exp\left(\int_{x_P}^x \lambda(\xi)d\xi\right) \phi_P - F(x_e) \int_{x_P}^x \frac{1}{\Gamma(\xi)} \exp\left(\int_{\xi}^x \lambda(\zeta)d\zeta\right) d\xi - \int_{x_P}^x \frac{S(\xi)}{\Gamma(\xi)} \exp\left(\int_{\xi}^x \lambda(\zeta)d\zeta\right) d\xi. \end{aligned}$$

Applying the other boundary condition  $\phi(x_E) = \phi_E$ , we obtain an equation from which we can solve  $F(x_e)$  and obtain (3.4). This concludes the proof of Lemma 3.1.  $\square$

We use this lemma to construct a flux approximation scheme. To this end define the functions

$$\begin{aligned} B(z) &:= \frac{z}{e^z - 1}, & W(z) &:= \frac{e^z - 1 - z}{z(e^z - 1)}, \\ \mathcal{F}^h(m, P, a, b) &:= m \frac{B(-P)a - B(P)b}{P}, & \mathcal{F}^i(P, s) &:= \left(\frac{1}{2} - W(P)\right)s. \end{aligned} \quad (3.5)$$



Furthermore, for a generic function  $f$  we introduce the interpolation values  $f_e$  and  $f_{\lambda_e}$ :

$$f_e := f(x_P) + \frac{1}{2}(f(x_E) - f(x_P)) \quad , \quad f_{\lambda_e} := f(x_P) + W(P_e)(f(x_E) - f(x_P)). \quad (3.6)$$

Note, however, that this notation is not used for the flux function  $F$ , so  $F_e$  still denotes the numerical approximation of  $F(x_e)$ . Let  $s_U$  denote the upwinded value of  $s$ :

$$s_U := \begin{cases} s_P & \text{if } u_e \geq 0 \\ s_E & \text{if } u_e < 0 \end{cases}. \quad (3.7)$$

Then the 1-D Modified Thiert Flux Approximation Scheme can be written as

$$F_e := F_e^h + F_e^i, \quad (3.8)$$

where

$$F_e^h := \mathcal{F}^h(\Gamma_{\lambda_e} \lambda_{\lambda_e}, P_e, \phi_P, \phi_E) \quad , \quad F_e^i := \mathcal{F}^i(P_e, s_U \Delta x). \quad (3.9)$$

For this approximation we have the following result:

**Theorem 3.2** . *Let  $\lambda, \Gamma \in C^2$ . Let  $s, \phi \in C^1$ . The numerical flux  $F_e$  is a second order approximation of  $F(x_e)$ . We can write*

$$F(x_e) = F_e + \eta_F(x_e) \Delta x^2,$$

where  $\eta_F$  is a bounded function.

**Proof.** Define the function  $w_\lambda$  by

$$w_\lambda(x) := \exp\left(\int_x^{x_E} \lambda(\xi) d\xi\right). \quad (3.10)$$

Recalling the standard scalar product for functions:

$$\langle f, g \rangle := \int_{x_P}^{x_E} f(\xi) g(\xi) d\xi,$$

it is easy to see that  $w'_\lambda(x) = -\lambda(x)w_\lambda(x)$ ; so

$$\langle \lambda, w_\lambda \rangle = w_\lambda(x_P) - w_\lambda(x_E) = \exp\left(\int_{x_P}^{x_E} \lambda(x) dx\right) - 1 = \exp(\langle \lambda, 1 \rangle) - 1.$$

Now, after some manipulations using the formulas above, the flux  $F(x_e)$  is written as:

$$F(x_e) = F^h + F^i(x_e), \quad (3.11)$$

where  $F^h$  is the homogeneous part, linear in  $\phi_P$  and  $\phi_E$ , and  $F^i$  is the inhomogeneous part, linear in  $s$ . They are given by

$$F^h := \mathcal{F}^h\left(\frac{\langle \lambda, w_\lambda \rangle}{\langle \Gamma^{-1}, w_\lambda \rangle}, \langle \lambda, 1 \rangle, \phi_P, \phi_E\right) \quad , \quad F^i(x_e) := -\frac{\langle \Gamma^{-1} S, w_\lambda \rangle}{\langle \Gamma^{-1}, w_\lambda \rangle}. \quad (3.12)$$

Now we must find an appropriate way to approximate (3.12). One could do this, for instance, by assuming that  $\Gamma, \rho u$  and  $s$  are constant in the interval  $(x_P, x_E)$ . By using the function  $W$ , the integrals in the terms  $F^h$  and  $F^i(x_e)$  can be evaluated analytically, yielding:

$$F_{const}^h = \mathcal{F}^h(\Gamma_e \lambda_e, P_e, \phi_P, \phi_E) \quad , \quad F_{const}^i(x_e) = \mathcal{F}^i(P_e, s_U \Delta x), \quad (3.13)$$

which is obviously very similar to (3.9). In [14] and [15], Thiart uses a formula very similar to (3.13) to approximate fluxes. We will see, however, that (3.13) loses its second order accuracy for larger Peclet numbers, a problem which can be mended by the small correction which is given by (3.9).

It is a logical step to approximate the integrals in (3.12) by means of the -second order accurate-trapezoidal rule. This rule integrates functions accurately if their second derivative is moderate. So when  $\lambda$  is large, this causes  $w_\lambda$  to have a large second derivative, which rules out this method. The trapezoidal rule is therefore inappropriate for the integrals of the form  $\langle f, w_\lambda \rangle$  (with  $f = \Gamma^{-1}, \lambda$  or  $\Gamma^{-1}S$ ), but not for the integral  $\langle \lambda, 1 \rangle$ . The trapezoidal rule for the latter integral can be formulated as:

$$\frac{1}{\Delta x} \langle \lambda, 1 \rangle = \lambda_e - \frac{1}{12} \Delta x^2 \lambda''(\xi_1) \text{ for some } \xi_1 \in (x_P, x_E).$$

We use the local Peclet number  $P_e$  to approximate  $\langle \lambda, 1 \rangle$ :  $P_e = \langle \lambda, 1 \rangle + \frac{1}{12} \Delta x^3 \lambda''(\xi_1)$ .

A similar integration formula can be constructed for the other integrals. We now look for a good approximation of quotients of the following general form:  $\frac{\langle f, w_\lambda \rangle}{\langle 1, w_\lambda \rangle}$ . First of all, we see that, since  $w_\lambda$  is strictly positive,

$$\frac{\langle f, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = f(\xi_2) \quad \text{for some } \xi_2 \in (x_P, x_E). \quad (3.14)$$

We will approximate (3.14) by replacing  $f$  by its linear interpolant  $\hat{f}$ :

$$\hat{f}(x) = f(x_P) + \frac{f(x_E) - f(x_P)}{\Delta x} (x - x_P). \quad (3.15)$$

The difference between  $f$  and  $\hat{f}$  is given by

$$f(x) - \hat{f}(x) = -\frac{1}{2} (x - x_P)(x_E - x) f''(\xi_3) \quad \text{for some } \xi_3 \in (x_P, x_E). \quad (3.16)$$

Using (3.15), we obtain

$$\frac{\langle \hat{f}, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = f(x_P) + \frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} \frac{f(x_E) - f(x_P)}{\Delta x}. \quad (3.17)$$

By virtue of the fact that  $(x - x_P)(x_E - x)$  is positive for  $x \in (x_P, x_E)$ , we obtain

$$\frac{\langle f - \hat{f}, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = -\frac{1}{2} \frac{\langle (x - x_P)(x_E - x), w_\lambda \rangle}{\langle 1, w_\lambda \rangle} f''(\xi_4) \quad \text{for some } \xi_4 \in (x_P, x_E).$$

First, the factor  $\frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle}$  must be approximated. If  $\lambda$  is constant, it is given by:

$$\frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = W(P) \Delta x.$$

Let  $\kappa$  be a function such that  $\kappa(x) \geq \lambda(x)$  for all  $x \in [x_P, x_E]$ . Then, we can derive the following inequality:

$$\begin{aligned} \langle x - x_P, w_\kappa \rangle \langle 1, w_\lambda \rangle - \langle x - x_P, w_\lambda \rangle \langle 1, w_\kappa \rangle &= \int_{x_P}^{x_E} \int_{x_P}^{x_E} (x - y) w_\kappa(x) w_\lambda(y) dx dy = \\ &= \int_{x_P}^{x_E} \int_y^{x_E} (x - y) (w_\kappa(x) w_\lambda(y) - w_\kappa(y) w_\lambda(x)) dx dy = \\ &= \int_{x_P}^{x_E} \int_y^{x_E} (x - y) w_\kappa(x) w_\lambda(y) \left( 1 - e^{\int_y^x (\kappa(\eta) - \lambda(\eta)) d\eta} \right) dx dy \leq 0, \end{aligned}$$

or, equivalently,

$$\frac{\langle x - x_P, w_\kappa \rangle}{\langle 1, w_\kappa \rangle} \leq \frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle}.$$

Let  $\lambda_- := \min_{x \in [x_P, x_E]} \lambda(x)$  and  $\lambda_+ := \max_{x \in [x_P, x_E]} \lambda(x)$ , then

$$W(\lambda_+ \Delta x) \Delta x = \frac{\langle x - x_P, w_{\lambda_+} \rangle}{\langle 1, w_{\lambda_+} \rangle} \leq \frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} \leq \frac{\langle x - x_P, w_{\lambda_-} \rangle}{\langle 1, w_{\lambda_-} \rangle} = W(\lambda_- \Delta x) \Delta x.$$

By continuity of  $W$  and  $\lambda$ , we find that

$$\frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = W(P(\xi_5)) \Delta x \quad \text{for some } \xi_5 \in (x_P, x_E). \quad (3.18)$$

In the same way we can derive that

$$\frac{\langle (x - x_P)(x_E - x), w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = \frac{\langle x - x_P, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} \Delta x - \frac{\langle (x - x_P)^2, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} =$$

$$(W(P(\xi_5)) - W_2(P(\xi_6))) \Delta x^2, \quad \text{for some } \xi_6 \in (x_P, x_E),$$

where  $W_2$  can be found by taking  $\lambda$  constant:

$$W_2(z) := 2 \frac{e^z - (1 + z + \frac{1}{2}z^2)}{(e^z - 1)z^2}.$$

Using (3.17) and (3.18), we can approximate  $\frac{\langle f, w_\lambda \rangle}{\langle 1, w_\lambda \rangle}$  by (cf. (3.14))

$$\frac{\langle f, w_\lambda \rangle}{\langle 1, w_\lambda \rangle} = f_{\lambda_e} + \eta_1(x_e) \Delta x^2. \quad (3.19)$$

Let  $W_3 = W - W_2$ . The error term  $\eta_1(x_e)$  can then be written as

$$\eta_1(x_e) = -\frac{1}{2} W_3(P_e) f''(x_e) + \mathcal{O}(\Delta x).$$

Recall that  $f_{\lambda_e} = f(x_P) + W(P_e)(f(x_E) - f(x_P))$ .

Similarly to the definition of  $f_e$ ,  $f_{\lambda_e}$  is a second-order approximation of  $f(x_{\lambda_e})$ . The point  $x_{\lambda_e}$  will be slightly upwinded from  $x_e$ , i.e.  $x_P < x_{\lambda_e} \leq x_e$  if  $u_e \geq 0$ , and  $x_e \leq x_{\lambda_e} < x_E$  if  $u_e \leq 0$ . It is noteworthy that the trapezoidal rule can be seen as a special case of the integration rule (3.19), because

$$\frac{\langle \lambda, 1 \rangle}{\Delta x} = \frac{\langle \lambda, w_0 \rangle}{\langle 1, w_0 \rangle} = \lambda_e + \eta_1(x_e) \Delta x^2,$$

with the error  $\eta_1$  given by

$$\eta_1(x_e) = -\frac{1}{12} \lambda''(x_e) + \mathcal{O}(\Delta x) = -\frac{1}{2} W_3(0) \lambda''(x_e) + \mathcal{O}(\Delta x).$$

Now we have all the tools we need: an accurate integration rule and an analytical formula for  $F(x_e)$  which we want to approximate. The function  $\mathcal{F}^h$  can be rewritten as

$$\mathcal{F}^h(m, P, a, b) = m \frac{a+b}{2} - \frac{m \Delta x}{P} \frac{P/2}{\tanh(P/2)} \frac{b-a}{\Delta x}.$$

We exactly know  $\phi(x_P)$  and  $\phi(x_E)$ , which play the roles of  $a$  and  $b$  in the formula above. Since  $\langle \lambda, 1 \rangle = P_e + \mathcal{O}(\Delta x^3)$ , we can compute the term  $\frac{\langle \lambda, 1 \rangle / 2}{\tanh(\langle \lambda, 1 \rangle / 2)}$  to third order accuracy too. This leaves us to find the errors in the approximation of  $\frac{\langle \lambda, w_\lambda \rangle \Delta x}{\langle \Gamma^{-1}, w_\lambda \rangle \langle \lambda, 1 \rangle}$  and  $\frac{\langle \lambda, w_\lambda \rangle}{\langle \Gamma^{-1}, w_\lambda \rangle}$ , which play

the roles of  $\frac{m\Delta x}{P}$  and  $m$  respectively; see (3.12). Using the new integration formula, we find the following approximation:

$$\frac{\langle \lambda, w_\lambda \rangle_{\Delta x}}{\langle \Gamma^{-1}, w_\lambda \rangle} = \frac{\lambda_{\lambda_e}}{\lambda_e(\Gamma^{-1})_{\lambda_e}}(1 + \eta_2(x_e)\Delta x^2) = \Gamma_{\lambda_e} \frac{\lambda_{\lambda_e}}{\lambda_e}(1 + \eta_2(x_e)\Delta x^2 - \eta_3(x_e)\Delta x^2),$$

with error terms  $\eta_2(x_e)$  and  $\eta_3(x_e)$  given by

$$\eta_2(x_e) = \frac{W_3(P_e)}{2} \frac{(\Gamma^{-1})''}{\Gamma^{-1}}(x_e) + \mathcal{O}(\Delta x) \quad , \quad \eta_3(x_e) = \left( \frac{\Gamma'(x_e)}{\Gamma(x_e)} \right)^2 W(P_e)(1 - W(P_e)) + \mathcal{O}(\Delta x).$$

Let us investigate when these errors are small. The error  $\eta_3\Delta x^2$  is an estimate for the relative difference between  $(\Gamma_{\lambda_e})^{-1}$  and  $(\Gamma^{-1})_{\lambda_e}$ . Since  $\Gamma \geq \Gamma_{min}$ , this error must be small if  $\Gamma$  is represented accurately by the grid. The error  $\eta_2\Delta x^2$  is an estimate of the relative difference between  $\Gamma^{-1}$  and the linear interpolant  $\hat{\Gamma}^{-1}$  in the point  $x^* = x_P + W_3(P_e)\Delta x$ . Our conclusion is that both errors are small if  $\Gamma$  is well represented by the grid.

We use this integration rule again to find the following approximation:

$$\frac{\langle \lambda, w_\lambda \rangle}{\langle \Gamma^{-1}, w_\lambda \rangle} = \frac{\lambda_{\lambda_e}}{(\Gamma^{-1})_{\lambda_e}} (1 + (\eta_2(x_e) - \eta_4(x_e))\Delta x^2) = \lambda_{\lambda_e}\Gamma_{\lambda_e} (1 + (\eta_2(x_e) - \eta_4(x_e) - \eta_3(x_e))\Delta x^2),$$

with error  $\eta_4(x_e)$  given by:

$$\eta_4(x_e) = \frac{1}{2}W_3(P_e) \frac{\lambda''(x_e)}{\lambda(x_e)} + \mathcal{O}(\Delta x).$$

This error  $\eta_4\Delta x^2$  is an estimate for the difference between  $\lambda$  and the linear interpolant  $\hat{\lambda}$  in the point  $x^*$ . Therefore, this error will be small if  $\lambda$  is well represented by the grid, except in regions where  $\lambda$  has large relative changes because it becomes very small relative to the usual level of  $\lambda$ . The error is then not small relative to itself, but to this overall level.

We now have the following approximation for the homogeneous part of the flux  $F^h$ :

$$F^h = \mathcal{F}^h(\lambda_{\lambda_e}\Gamma_{\lambda_e}, P_e, \phi_P, \phi_E) + \eta_5(x_e)\Delta x^2, \quad (3.20)$$

with error term  $\eta_5$  given by

$$\eta_5(x_e) = F^h(x_e)(\eta_2(x_e) - \eta_3(x_e)) - (\rho\nu)(x_e)\eta_4(x_e)\phi(x_e) + \mathcal{O}(\Delta x).$$

We use the integration formula once more to obtain

$$F^i(x_e) = -\frac{\langle \Gamma^{-1}S, w_\lambda \rangle}{\langle \Gamma^{-1}, w_\lambda \rangle} = -S_{\lambda_e} - \eta_6(x_e)\Delta x^2, \quad (3.21)$$

with error term  $\eta_6$  given by

$$\eta_6(x_e) = \frac{\Gamma'(x_e)}{\Gamma(x_e)}s(x_e)(W(P_e)(1 - W(P_e)) - W_3(P_e)) - \frac{1}{2}s'(x_e)W_3(P_e) + \mathcal{O}(\Delta x).$$

The term  $S_{\lambda_e}$  can be written as:

$$S_{\lambda_e} = -\int_{x_P}^{x_e} s(x)dx + W(P_e) \int_{x_P}^{x_E} s(x)dx.$$

We approximate the first integral by the trapezoidal rule, and the second integral by the midpoint rule. Then we find:

$$S_{\lambda_e} = \Delta x \left( -\frac{1}{4}(s(x_P) + s(x_e)) + W(P_e)s(x_e) \right) + \mathcal{O}(\Delta x^3) = -\Delta x \left( \frac{1}{2} - W(P_e) \right) s_U + \eta_7(x_e)\Delta x^2, \quad (3.22)$$

with error term  $\eta_7$  given by

$$\eta_7(x_e) = \frac{1}{8} (4W(|P_e|) - 1) s'(x_e) + \mathcal{O}(\Delta x).$$

Using (3.21) and (3.22), we find:

$$F^i(x_e) = \mathcal{F}^i(P_e, s_U \Delta x) - (\eta_6(x_e) + \eta_7(x_e)) \Delta x^2. \quad (3.23)$$

Combining (3.20) and (3.23), we find

$$F(x_e) = F_e + (\eta_5(x_e) - \eta_6(x_e) - \eta_7(x_e)) \Delta x^2. \quad (3.24)$$

So,  $\eta_F(x_e) = \eta_5(x_e) - \eta_6(x_e) - \eta_7(x_e)$ .  $\square$

For the generalization to two-dimensional problems, we write the problem in its quasi one-dimensional formulation

$$\frac{\partial}{\partial x} \left( \rho u \phi - \Gamma \frac{\partial \phi}{\partial x} \right) = \hat{s} \quad , \quad \frac{\partial}{\partial y} \left( \rho v \phi - \Gamma \frac{\partial \phi}{\partial y} \right) = \tilde{s},$$

with  $\hat{s}$  and  $\tilde{s}$  defined as

$$\hat{s} := s - \frac{\partial G}{\partial y} \quad , \quad \tilde{s} := s - \frac{\partial F}{\partial x}.$$

Recall that  $G$  denotes the vertical flux component:

$$G = \rho v \phi - \Gamma \frac{\partial \phi}{\partial y}.$$

Let us introduce the following notation, necessary for the formulation of two-dimensional flux approximations:

$$\kappa := \frac{\rho v}{\Gamma} \quad , \quad Q := \kappa \Delta y.$$

Recall that  $\Delta y = y_N - y_P$  (see Figure 1),  $\lambda = \frac{e^u}{\Gamma}$  and  $P = \lambda \Delta x$ . For a generic function  $f$ , introduce the interpolated values  $f_n$  and  $f_{\kappa n}$ :

$$f_n := f(x_P) + \frac{1}{2}(f(x_N) - f(x_P)) \quad , \quad f_{\kappa n} := f(x_P) + W(Q_n)(f(x_N) - f(x_P)).$$

Recall that  $f_e = f(x_P) + \frac{1}{2}(f(x_E) - f(x_P))$ ,  $f_{\lambda e} := f(x_P) + W(P_e)(f(x_E) - f(x_P))$ . With these notations we can define a 2-D flux approximation scheme. First, we calculate the homogeneous fluxes  $F_e^h$  and  $G_n^h$ :

$$F_e^h := \mathcal{F}^h(\Gamma_{\lambda e} \lambda_{\lambda e}, P_e, \phi_P, \phi_E) \quad , \quad G_n^h := \mathcal{F}^h(\Gamma_{\kappa n} \kappa_{\kappa n}, Q_n, \phi_P, \phi_N), \quad (3.25)$$

and  $F_w^h, G_s^h$  similarly. With these, we can calculate the quasi 1-D source terms. For in order to calculate  $F(x_e)$  with second order accuracy, it is necessary to know  $\hat{s}(x_U)$  to first order accuracy. Fortunately, we have

$$G(x_n) = G_n^h + \mathcal{O}(\Delta y),$$

so we may ignore the inhomogeneous term  $G^i$  for the approximation of the so-called cross-flux differential  $\frac{\partial G}{\partial y}$ . We shall approximate the cross-flux differential  $\frac{\partial G}{\partial y}$  by the difference of the homogeneous flux term  $G^h$ :

$$\hat{s}_P := s_P - \frac{G_n^h - G_s^h}{\Delta y} \quad , \quad \tilde{s}_P := s_P - \frac{F_e^h - F_w^h}{\Delta x}, \quad (3.26)$$

so the upwinded quasi 1-D source terms  $\tilde{s}_U$  and  $\hat{s}_U$  are given by:

$$\hat{s}_U := \begin{cases} \hat{s}_P & \text{if } u_e \geq 0 \\ \hat{s}_E & \text{if } u_e < 0 \end{cases}, \quad \tilde{s}_U := \begin{cases} \tilde{s}_P & \text{if } v_n \geq 0 \\ \tilde{s}_N & \text{if } v_n < 0 \end{cases}, \quad (3.27)$$

and then the inhomogeneous flux terms:

$$F_e^i := \mathcal{F}^i(P_e, \hat{s}_U \Delta x), \quad G_n^i := \mathcal{F}^i(Q_n, \tilde{s}_U \Delta y). \quad (3.28)$$

The 2-D Modified Thiant Flux Approximation Scheme is given by

$$F_e := F_e^h + F_e^i, \quad G_n := G_n^h + G_n^i. \quad (3.29)$$

As for the one-dimensional case, we can define error functions  $\eta_F(\mathbf{x}_e)$  and  $\eta_G(\mathbf{x}_n)$ . These are very similar to (3.24), though a term has to be added because the quasi 1-D source terms  $\hat{s}$  and  $\tilde{s}$  are not known exactly and have to be approximated.

## 4 The Modified Thiant Scheme

In this section we will show how, using the Modified Thiant Flux Approximation, one can construct the so-called *Modified Thiant Scheme*, a discretization for the convection-diffusion-reaction equation (2.1). We give the discretization by defining the discretization molecule for the differential operator in the interior domain. In order to do this, we define coefficients  $a$  and  $b$  by

$$\begin{aligned} a_E &:= \lambda_{\lambda e} \Gamma_{\lambda e} \frac{B(P_e)}{P_e}, & a_P &:= \lambda_{\lambda e} \Gamma_{\lambda e} \frac{B(-P_e)}{P_e} + \lambda_{\lambda w} \Gamma_{\lambda w} \frac{B(P_w)}{P_w}, & a_W &:= \lambda_{\lambda w} \Gamma_{\lambda w} \frac{B(-P_w)}{P_w}, \\ b_N &:= \kappa_{\kappa n} \Gamma_{\kappa n} \frac{B(Q_n)}{Q_n}, & b_P &:= \kappa_{\kappa n} \Gamma_{\kappa n} \frac{B(-Q_n)}{Q_n} + \kappa_{\kappa s} \Gamma_{\kappa s} \frac{B(Q_s)}{Q_s}, & b_S &:= \kappa_{\kappa s} \Gamma_{\kappa s} \frac{B(-Q_s)}{Q_s}. \end{aligned} \quad (4.1)$$

Hence the homogeneous flux differences,  $F_e^h - F_w^h$  and  $G_n^h - G_s^h$  can be written in the following way

$$F_e^h - F_w^h = a_P \phi_P - a_E \phi_E - a_W \phi_W, \quad G_n^h - G_s^h = b_P \phi_P - b_N \phi_N - b_S \phi_S. \quad (4.2)$$

We will also need the homogeneous flux terms in the points  $\mathbf{x}_{Ne}$ ,  $\mathbf{x}_{nE}$  etcetera, which are located on the cell faces of the neighboring cells. For these calculations, we define  $a$  and  $b$  at the other points of the 9-point stencil implicitly by

$$\begin{aligned} F_{Ne}^h - F_{Nw}^h &= a_N \phi_N - a_{NE} \phi_{NE} - a_{NW} \phi_{NW}, & G_{nE}^h - G_{sE}^h &= b_E \phi_E - b_{NE} \phi_{NE} - b_{SE} \phi_{SE}, \\ F_{Se}^h - F_{Sw}^h &= a_S \phi_S - a_{SE} \phi_{SE} - a_{SW} \phi_{SW}, & G_{nW}^h - G_{sW}^h &= b_W \phi_W - b_{NW} \phi_{NW} - b_{SW} \phi_{SW}. \end{aligned}$$

In the further analysis, it we will assume the coefficients  $a_*$  and  $b_*$  two be non-negative. We first show that this is a reasonable assumption. The coefficients  $a_*$  contain quotients of the form  $\lambda_{\lambda e}/\lambda_e$ , which depend on  $P(\mathbf{x}_P)$  and  $P(\mathbf{x}_E)$  only. When all such quotients are positive, it follows that all coefficients  $a_*$  are positive too. Figure 8 shows that  $a_* > 0$ , unless  $P(\mathbf{x}_P)$  and  $P(\mathbf{x}_E)$  differ very much from each other; not only in a relative sense, but also in an absolute sense. It is safe to say, therefore, that the coefficients  $a_*$  and  $b_*$  are positive if the grid describes the mass flow with any degree of accuracy.

Because  $F^i$  and  $G^i$  are calculated upwind, define

$$\begin{aligned} c_W &:= \max(0, x_w - x_{\lambda w}), & c_P &:= \max(0, x_e - x_{\lambda e}) + \max(0, x_{\lambda w} - x_w), & c_E &:= \max(0, x_{\lambda e} - x_e), \\ d_S &:= \max(0, y_s - y_{\kappa s}), & d_P &:= \max(0, y_n - y_{\kappa n}) + \max(0, y_{\kappa s} - y_s), & d_N &:= \max(0, y_{\kappa n} - y_n), \\ c_{NW} &:= c_W =: c_{SW}, & c_N &:= c_P =: c_S, & c_{NE} &:= c_E =: c_{SE}, \\ d_{NW} &:= d_N =: d_{NE}, & d_W &:= d_P =: d_E, & d_{SW} &:= d_S =: d_{SE}. \end{aligned}$$

It follows that  $c_*$  and  $d_*$  are non-negative. The inhomogeneous parts of the flux differences  $F_e^i - F_w^i$  and  $G_n^i - G_s^i$  are given by:

$$F_e^i - F_w^i = c_P \hat{s}_P - c_E \hat{s}_E - c_W \hat{s}_W, \quad G_n^i - G_s^i = d_P \tilde{s}_P - d_N \tilde{s}_N - d_S \tilde{s}_S.$$

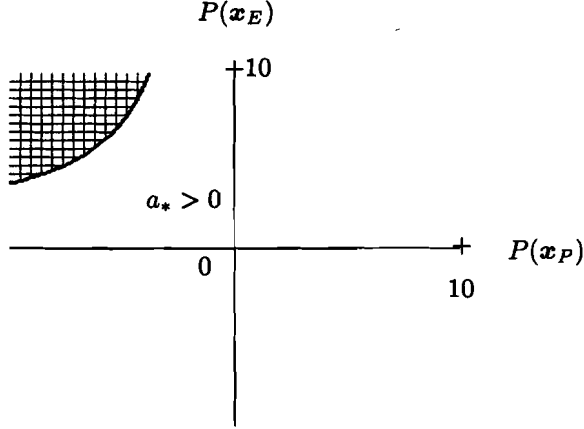


Figure 8: The coefficients  $a_*$ , depending on  $P(\mathbf{x}_P)$  and  $P(\mathbf{x}_E)$ .

The approximated conservation law (2.6) is now given by

$$\frac{(F_e^h - F_w^h) + (F_e^i - F_w^i)}{\Delta x} + \frac{(G_n^h - G_s^h) + (G_n^i - G_s^i)}{\Delta y} = S_P. \quad (4.3)$$

Using the coefficients defined above, the conservation law in the interior domain is approximated by

$$D_P \phi_P = \sum_{I \in \mathcal{N}_8} D_I \phi_I + B_P s_P + \sum_{I \in \mathcal{N}_4} B_I s_I, \quad (4.4)$$

where  $\mathcal{N}_8$  denotes the set of eight neighboring grid points  $\mathcal{N}_8 := \{N, E, S, W, NE, NW, SE, SW\}$ ,  $\mathcal{N}_4$  denotes the set of four directly neighboring grid points  $\mathcal{N}_4 := \{N, E, S, W\}$ , and where we have defined the following coefficients

$$D_I := \frac{\Delta x - c_I}{\Delta x} \frac{b_I}{\Delta y} - \frac{d_I}{\Delta y} \frac{a_I}{\Delta x} \quad I = N, S, \quad D_I := \frac{\Delta y - d_I}{\Delta y} \frac{a_I}{\Delta x} - \frac{c_I}{\Delta x} \frac{b_I}{\Delta y} \quad I = E, W,$$

$$D_P := \frac{\Delta y - d_P}{\Delta y} \frac{a_P}{\Delta x} + \frac{\Delta x - c_P}{\Delta x} \frac{b_P}{\Delta y}, \quad D_I := \frac{d_I}{\Delta y} \frac{a_I}{\Delta x} + \frac{c_I}{\Delta x} \frac{b_I}{\Delta y} \quad \forall I \in \mathcal{N}_8 \setminus \mathcal{N}_4,$$

$$B_P := 1 - \frac{c_P}{\Delta x} - \frac{d_P}{\Delta y}, \quad B_I := \frac{c_I}{\Delta x} \quad I = E, W, \quad B_I := \frac{d_I}{\Delta y} \quad I = N, S.$$

The discretization (4.4) is, however, not yet our desired scheme. For that we need one more modification. In a stationary flow the constant function  $\phi(\mathbf{x}) \equiv 1$  is a solution to the continuous problem (2.1) with  $s \equiv 0$ , because  $\nabla \cdot (\rho \mathbf{v}) = 0$ . In the discretized equation (4.4) this is not the case! In order to preserve the constant solution, we modify the discretization molecule to obtain the Modified Thiant Scheme:

$$\sum_{I \in \mathcal{N}_8} D_I \phi_P = \sum_{I \in \mathcal{N}_8} D_I \phi_I + B_P s_P + \sum_{I \in \mathcal{N}_4} B_I s_I. \quad (4.5)$$

We have not yet paid any attention to the discretization of the boundary conditions. The boundary conditions are no essential part of the Modified Thiant Scheme, and any consistent treatment of the boundary conditions will do to complete the system. For the sake of completeness,

we give some examples of how this can be done. A Dirichlet boundary condition of the form  $\phi(\mathbf{x}) = \phi^0(\mathbf{x})$  can be discretized by

$$\phi_P = \phi^0(\mathbf{x}_P).$$

Neumann boundary conditions of the form  $\frac{\partial \phi}{\partial n} = \phi^1(\mathbf{x})$  can be discretized by introducing a mirror point  $\mathbf{x}_M$ :

$$\frac{\phi_P - \phi_M}{h} = \pm \phi^1(\mathbf{x}_m), \quad h = \Delta x \text{ or } h = \Delta y, \quad \mathbf{x}_m = \frac{1}{2}(\mathbf{x}_M + \mathbf{x}_P).$$

To conclude this section, it will be shown that the proposed discretization (4.5) is second order consistent. First, let us look at the accuracy of (4.4). The local discretization error  $\eta_8$  is found when the exact solution  $\phi$  is substituted in (2.6). It is then given by

$$\eta_8 = \frac{F_e - F_w}{\Delta x} + \frac{G_n - G_s}{\Delta y} - s_P.$$

We use the discretization error for  $F$  which was derived in Section 3 to evaluate  $\eta_8$ :

$$\begin{aligned} \eta_8 &= \frac{(F(\mathbf{x}_w) - F_w) - (F(\mathbf{x}_e) - F_e)}{\Delta x} + \frac{(G(\mathbf{x}_s) - G_s) - (G(\mathbf{x}_n) - G_n)}{\Delta y} + \\ &\quad \frac{F(\mathbf{x}_e) - F(\mathbf{x}_w)}{\Delta x} + \frac{G(\mathbf{x}_n) - G(\mathbf{x}_s)}{\Delta y} - s_P. \end{aligned}$$

It can easily be seen that this is equal to

$$\eta_8 = -\Delta x^2 \frac{\partial \eta_F}{\partial x} - \Delta y^2 \frac{\partial \eta_G}{\partial y} + \frac{\Delta x^2}{24} \frac{\partial^3 F}{\partial x^3} + \frac{\Delta y^2}{24} \frac{\partial^3 G}{\partial y^3}.$$

Obviously, the scheme is second order consistent. Consistency of (4.5) follows because it differs from (4.4) by second order terms only. The difference between (4.5) and (4.4) is in the left hand side of the equations. This difference is given by:

$$\left( D_P - \sum_{I \in \mathcal{N}_8} D_I \right) \phi_P = \mathcal{O}(\Delta x^2 + \Delta y^2).$$

This means that (4.5) and (4.4) have only second order differences and that (4.5) is second order consistent.

## 5 Limiting Cases

It is interesting to see how the Modified Thiant Scheme compares to other, conventional schemes. We shall do this by considering two limiting cases: the case where  $P$  is small ( $P \rightarrow 0$ ), and the case where  $P$  is very large ( $P \rightarrow +\infty$ ).

When  $P$  is very small, we can approximate  $W$  and  $B$  by

$$W(x) \doteq \frac{1}{2}, \quad B(x) \doteq 1 - \frac{x}{2}.$$

Doing so, we find that (4.5) becomes:

$$\frac{1}{\Delta x} \left( \lambda_e \Gamma_e \frac{\phi_P + \phi_E}{2} - \lambda_w \Gamma_w \frac{\phi_P + \phi_W}{2} \right) - \frac{1}{\Delta x} \left( \Gamma_e \frac{\phi_E - \phi_P}{\Delta x} - \Gamma_w \frac{\phi_P - \phi_W}{\Delta x} \right) = s_P.$$

This can be seen as a rather standard central difference scheme.

Next, we consider very large  $P$ . Now we can approximate  $W$  and  $B$  by

$$W(x) \doteq 0, \quad B(x) \doteq 0.$$



Again we apply this to (4.5) and find:

$$\frac{1}{\Delta x} ((\rho u)_P \phi_P - (\rho u)_W \phi_W) = \frac{1}{2} (s_P + s_W).$$

Again, this can be understood as a central difference scheme. For this, we may view it from the perspective of the point  $x_w$ . Then it appears to be a finite volume discretization derived from applying the trapezoidal rule on the interval  $(x_W, x_P)$ . Similarly, one can analyze the limiting cases where  $P \rightarrow -\infty$ . The analysis can even be extended to the analogous two-dimensional limiting cases, but formulas become larger then.

Though most exponential schemes make a transition from central difference schemes for diffusion-dominated flow to upwind schemes for convection-dominated flow, these small analyses show that the Modified Thiant scheme makes a transition from one central difference scheme to another.

## 6 Global Discretization Error

The Modified Thiant Scheme (4.5) can be written in matrix-vector form, and will then look like

$$D\phi = Bs + \phi^0 + \phi^1.$$

Here the vectors  $\phi^0$  and  $\phi^1$  contain the boundary values for  $\phi$  and  $\frac{\partial \phi}{\partial n}$  which occur in the Dirichlet and Neumann boundary conditions respectively. In this section, we will show that  $D$  is often monotone which means that its inverse  $D^{-1}$  exists and has only non-negative entries. The monotonicity of  $D$  is now formulated in the following way:

**Theorem 6.1** . Define the function  $f$  by

$$f(x, y) := \frac{(1/2 - W(y))(B(-x) + B(x))}{(1/2 + W(x))B(-y)}.$$

For problems with constant coefficients  $\Gamma$  and  $\rho v$ , the discretization matrix  $D$  is monotone if the local Peclet numbers  $P$  and  $Q$  fulfil

$$f(|P|, |Q|) < \frac{\Delta x^2}{\Delta y^2} \quad \text{and} \quad f(|Q|, |P|) < \frac{\Delta y^2}{\Delta x^2}. \quad (6.1)$$

**Proof.** For the proof of Theorem 6.1, we use [3], Section 4.3, where it is proven that a matrix  $A$  is monotone if the following conditions are satisfied:

- $A$  is irreducible,
- $a_{ii} > 0$ ,
- $a_{ij} \leq 0 \quad \forall j \neq i$ ,
- $\sum_j a_{ij} \geq 0$  for all  $i \in \{1, \dots, N\}$ ,
- $\sum_j a_{ij} > 0$  for all  $i$  in a nonempty subset  $L \subset \{1, \dots, N\}$ .

The last of these conditions is verified automatically by the discretization of the Dirichlet boundary condition. The other ones are verified if  $D_I, I \in \mathcal{N}_4$  are positive and that  $D_I, I \in \mathcal{N}_8 \setminus \mathcal{N}_4$  are all nonnegative. The coefficients  $c_*$  and  $d_*$  have been constructed so that they would not be negative. The coefficients  $a_*$  and  $b_*$  are positive in the constant coefficient case, so  $D_I \geq 0 \quad \forall I \in \mathcal{N}_8 \setminus \mathcal{N}_4$ . It can easily be seen that  $d_S = 0$  when  $Q < 0$ , resulting in  $D_S > 0$ . If  $Q \geq 0$  we see that

$$D_S = (1/2 + W(|P|)) \frac{\Gamma}{\Delta y^2} B(-Q) - (1/2 - W(Q)) \frac{\Gamma}{\Delta x^2} (B(P) + B(-P)).$$

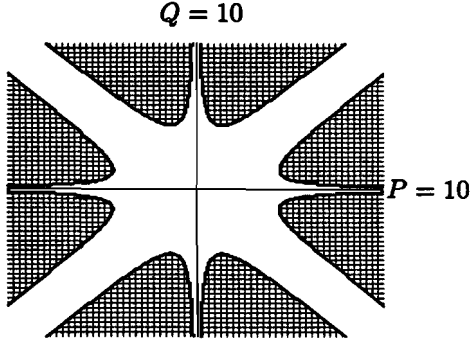


Figure 9: Peclet numbers  $P, Q$  for which  $D$  is monotone if  $\Delta x = \Delta y$ .

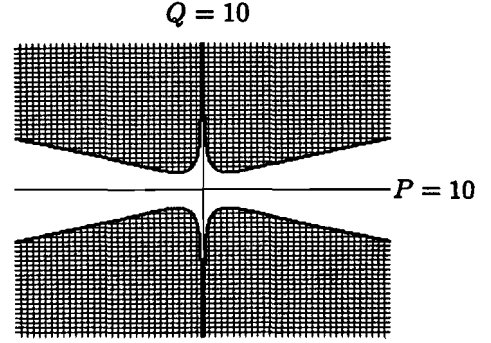


Figure 10: Peclet numbers  $P, Q$  for which  $D$  is monotone if  $\Delta x = 2\Delta y$ .

Obviously,  $D_S > 0$  if

$$f(|P|, |Q|) < \frac{\Delta x^2}{\Delta y^2}.$$

When the same conditions are written out for  $D_N, D_W$  and  $D_E$ , we find exactly the condition (6.1).  $\square$

A graphical representation of (6.1) can be made by shading the area in the  $(P, Q)$ -plane where the coefficients are all positive. Figures 9 and 10 show these areas. We see that these areas are not only very large compared to the area  $|P|, |Q| < 2$ , which is the area where central difference schemes have monotone discretization matrices, but they can also be adjusted to the specific problem we are trying to solve by choosing the appropriate  $\Delta x : \Delta y$  ratio. A simple monotonicity condition like (6.1) can only be given for the constant coefficient case. In general, we will find monotone discretization matrices in a wide variety of cases.

From the monotonicity of  $D$ , we can define an upper bound for the global discretization error.

**Theorem 6.2** . Consider the 2-D boundary value problem:

$$\begin{aligned} \mathcal{A}\phi(x, y) &= s(x, y) & \text{for all } (x, y) \in \Omega, \\ \phi(x, y) &= \phi^0(x, y) & \text{for all } (x, y) \in L^0 \subset \delta\Omega, \\ \frac{\partial \phi}{\partial n}(x, y) &= \phi^1(x, y) & \text{for all } (x, y) \in L^1 = \delta\Omega \setminus L^0. \end{aligned} \tag{6.2}$$

Let  $\phi^*$  and  $s$  be the restrictions to the grid of  $\phi$  and  $s$  respectively. Let  $\phi^0$  and  $\phi^1$  be the restrictions to the discrete Dirichlet and Neumann boundaries  $L^0$  and  $L^1$  of  $\phi^0$  and  $\phi^1$ . Furthermore, let  $\phi$  be the numerical approximation of  $\phi^*$ :

$$D\phi = Bs + \phi^0 + \phi^1,$$

and let  $\eta_\phi$  be the local discretization error:

$$\eta_\phi := Bs + \phi^0 + \phi^1 - D\phi^*.$$

Let  $m$  be defined by

$$m := \min \left( 1, \min_i \sum_j b_{ij} \right).$$

Let  $\psi$  denote the solution of (6.2) with  $s(x, y) \equiv 1$ ,  $\phi^0(x, y) \equiv 1$ ,  $\phi^1(x, y) \equiv 1$ . Let  $\psi^*$  denote the restriction to the grid hereof, and  $\eta_\psi$  the local discretization error of  $\psi$ . If  $\|\eta_\psi\|_\infty < m$ , we have the following upper bound for the global discretization error:

$$\|\phi - \phi^*\|_\infty \leq \frac{\|\psi^*\|_\infty \|\eta_\phi\|_\infty}{m - \|\eta_\psi\|_\infty}.$$

**Proof.** We can derive:

$$\min_i (D\psi^*)_i \geq \min_i (m - (\eta_\psi)_i) \geq m - \|\eta_\psi\|_\infty.$$

Let  $e$  denote the vector with only one-entries. Using the inequality above and the monotonicity of  $D$ , we find:

$$0 < (D^{-1}e)_i < \frac{1}{m - \|\eta_\psi\|_\infty} (\psi^*)_i,$$

and

$$\|\phi - \phi^*\|_\infty = \|D^{-1}\eta_\phi\|_\infty \leq \|D^{-1}e\|_\infty \|\eta_\phi\|_\infty \leq \frac{\|\psi^*\|_\infty \|\eta_\phi\|_\infty}{m - \|\eta_\psi\|_\infty}.$$

□

This theorem is applied to the discretized convection-diffusion-reaction equation as follows. First, let us look at the scaling of the Modified Thuart Scheme (4.5). In other words, let us focus on  $m$ . The sum of the  $B$ -factors is given by:

$$B_P + B_E + B_W + B_S + B_N = 1 - \frac{c_P - c_E - c_W}{\Delta x} - \frac{d_P - d_S - d_N}{\Delta y},$$

which will be larger than  $\frac{1}{2}$ , unless Peclet number variations are very large within grid size (Figure 11; cf. Figure 8). Therefore, the scheme is well-scaled. We shall assume that  $m > \frac{1}{2}$  (see Figure 11). Furthermore, it is not necessary to know the function  $\psi$ . It is enough to know that it exists, is

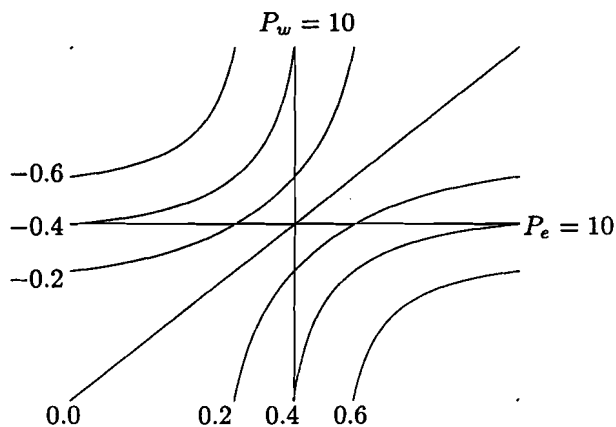


Figure 11: The value of  $\frac{c_P - c_E - c_W}{\Delta x}$ , depending on  $P_e$  and  $P_w$ .

bounded ( $\|\psi\|_\infty = M < +\infty$ ), and that its local discretization error will be under  $\frac{1}{4}$  for sufficiently smooth meshes, unless the problem is very ill posed. Then the theorem tells us that the numerical solution will not differ more from the exact solution than  $4M$  times the local discretization error, which was shown to be  $\mathcal{O}(\Delta x^2 + \Delta y^2)$ .

## 7 Numerical Examples

In order to test the 1-D Modified Thiart Scheme, we construct the following boundary value problem:

$$\frac{d}{dx} \left( m\phi - (1+x-x^2) \frac{d\phi}{dx} \right) = 4\operatorname{sech}^2(4x-2)(m-1+2x+8(1+x-x^2)\tanh(4x-2)),$$

$$\phi(0) = \tanh(-2) \quad , \quad \phi(1) = \tanh(2),$$

with

$$\phi(x) = \tanh(4x-2).$$

The Peclet number can be varied by varying  $m$ . We calculate numerical solutions on grids of  $N$  grid points, using finite volume methods based on the following four flux approximation schemes:

1. The upwind flux:

$$F_e^1 = m\phi_P - \Gamma_e \frac{\phi_E - \phi_P}{\Delta x},$$

2. The central difference flux:

$$F_e^2 = m\phi_e - \Gamma_e \frac{\phi_E - \phi_P}{\Delta x},$$

3. The homogeneous flux (similar to [14]):

$$F_e^3 = \mathcal{F}^h(m, P_e, \phi_P, \phi_E),$$

4. Modified Thiart Flux Approximation Scheme, which for 1-D problems reduces to

$$F_e^4 = \mathcal{F}^h(m, P_e, \phi_P, \phi_E) + \mathcal{F}^i(P_e, s_U \Delta x)$$

(which is also very similar to [15]).

Tables 1 and 2 show  $\epsilon_N$ , the 2-norm of the global discretization error, for a diffusion dominated problem ( $m = 1$ , Table 1) and for a convection dominated problem ( $m = 10^5$ , Table 2). In the diffusion dominated problem all methods, except upwind, are second order and almost equally accurate. In the convection dominated problem, however, there is much more to be seen. First of all, the homogeneous flux approximation seems to be only first order convergent. Enough refinement, however, will reduce Peclet numbers and increase convergence, a process which has already been started, because convergence is slightly over 2 for the finest meshes. The central difference scheme produces an oscillation on the coarsest meshes, causing an error of about the same magnitude as the solution itself. These oscillations disappear quickly. Only Modified Thiart Scheme seems to work properly here. There is second order convergence from the start, and for course meshes as well as for fine meshes this method gives the most accurate results. It is illustrative to see that Modified Thiart Scheme on a 10-point mesh obtains a comparable accuracy to the central difference scheme on a 40-point mesh.

We conduct a similar test in the 2-D setting. For this, we use the following analytical solution of the convection-diffusion-reaction equation:

$$\begin{aligned} \rho u &= \frac{27(1-x)x(1-y)}{6x+2} \quad , \quad \rho v = \left( \frac{y-1}{1/3+x} \right)^2 + \frac{9}{4}y(2-y), \\ \Gamma &= \Gamma_0(1+10x(1-x)y(1-y)) \quad , \quad \phi(x,y) = 1 + \tanh \left( \frac{25}{4}(2y-1) - 15x^2 \right), \end{aligned}$$

from which the source term can be computed by the evaluation of the right-hand side in (2.2). We applied inhomogeneous Dirichlet boundary conditions to the boundaries  $y = 0$  and  $y = 1$ ,

N	Upwind		Central		Homogeneous Flux		Modified Thiert	
	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$
10	4.1e-3	1.45	7.0e-3	4.00	6.9e-3	4.00	6.4e-3	3.91
20	2.8e-3	1.60	1.8e-3	3.97	1.7e-3	3.97	1.6e-3	3.93
40	1.8e-3	1.80	4.4e-4	3.98	4.3e-4	3.98	4.1e-4	3.96
80	9.8e-3	1.90	1.1e-4	3.99	1.1e-4	3.99	1.0e-4	3.98
160	5.2e-4	1.95	2.8e-5	3.99	2.7e-5	3.99	2.6e-5	3.99
320	2.7e-4	1.98	7.0e-6	4.00	6.8e-6	4.00	6.6e-6	3.99
640	1.3e-4	1.99	1.7e-6	4.00	1.7e-6	4.00	1.7e-6	4.00
1280	6.8e-5	1.99	4.4e-7	4.00	4.3e-7	4.00	4.1e-7	4.00
2560	3.4e-5	2.00	1.1e-7	4.00	1.1e-7	4.00	1.0e-7	4.00
5120	1.7e-5		2.7e-8		2.7e-8		2.6e-8	

Table 1: Errors for diffusion dominated flow ( $m = 1$ ).

N	Upwind		Central		Homogeneous Flux		Modified Thiert	
	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$
10	9.8e-2	1.94	2.1e00	15.9	9.8e-2	1.94	6.8e-3	3.91
20	5.0e-2	1.97	1.3e-1	15.5	5.0e-2	1.97	1.7e-3	3.95
40	2.6e-2	1.98	8.5e-3	14.7	2.6e-2	1.99	4.4e-4	3.97
80	1.3e-2	1.99	5.8e-4	8.93	1.3e-2	2.00	1.1e-4	3.98
160	6.5e-3	2.00	6.5e-5	4.60	6.4e-3	2.00	2.8e-5	3.99
320	3.2e-3	2.00	1.4e-5	4.05	3.2e-3	2.01	6.9e-6	3.99
640	1.6e-3	2.00	3.5e-6	4.01	1.6e-3	2.03	1.7e-6	3.99
1280	8.1e-4	2.00	8.7e-7	4.00	7.8e-4	2.07	4.3e-7	4.01
2560	4.1e-4	2.00	2.2e-7	4.00	3.8e-4	2.15	1.1e-7	4.09
5120	2.0e-4		5.4e-8		1.8e-4		2.6e-8	

Table 2: Errors for convection dominated flow ( $m = 10^5$ ).

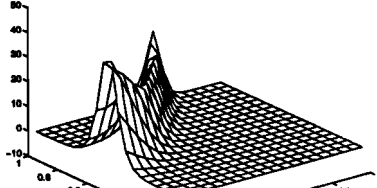


Figure 12: The source term  $s$  for diffusion dominated flow ( $\Gamma_0 = 0.1$ ).

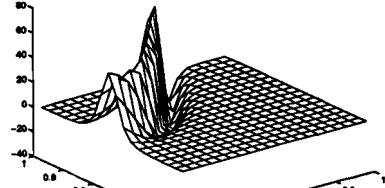


Figure 13: The source term  $s$  for convection dominated flow ( $\Gamma_0 = 0.005$ ).

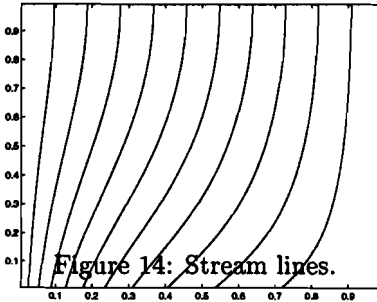


Figure 14: Stream lines.

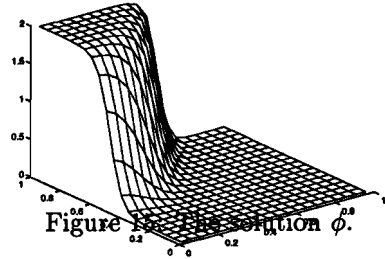


Figure 15: The solution  $\phi$ .

and inhomogeneous Neumann boundary conditions to the other boundaries. The nature of the problem, convection dominated flow or diffusion dominated flow, can be chosen by varying  $\Gamma_0$ . Figures 12, 13, 14 and 15 give graphical representations of the source term, the stream function and the solution.

Again, four finite volume methods are used to discretize the equation: the Modified Thiant Scheme, the central difference scheme, and two second order exponential schemes which are obtained as minor simplifications from Modified Thiant Scheme:

1. A flux approximation given by

$$F_e^3 = \mathcal{F}^h((\rho u)_e, P_e, \phi_P, \phi_E) + \mathcal{F}^i(P_e, \hat{\sigma}_U \Delta x),$$

in which the term  $\lambda_{\lambda e} \Gamma_{\lambda e}$ , which occurs in Modified Thiant Scheme, is replaced by the simpler form  $(\rho u)_e$ . We shall call this the 'constant mass-flux' approximation.

2. The homogeneous flux approximation, given by

$$F_e^4 = \mathcal{F}^h(\lambda_{\lambda e} \Gamma_{\lambda e}, P_e, \phi_P, \phi_E).$$

We apply these schemes on grids of size  $N \times N$ , and compute  $\epsilon_N$ , the 2-norm of the global discretization error, which can be found in Tables 3 and 4. Again we solve a problem with dominant convection ( $\Gamma_0 = 0.005$ , Table 4) and one with dominant diffusion ( $\Gamma_0 = 0.1$ , Table 3). It can be seen one more that for diffusion dominated flow it does not matter very much which method is used because all three methods are comparably accurate and converge quadratically.

N	Modified Thiaert		Central		Constant Mass-Flux		Homogeneous Flux	
	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$
10	4.0e-2	6.0	6.0e-2	5.2	3.6e-2	5.6	4.9e-2	4.1
20	6.7e-3	3.5	1.2e-2	4.0	6.4e-3	3.5	1.2e-2	3.7
40	1.9e-3	3.7	2.9e-3	4.0	1.8e-3	3.7	3.2e-3	3.8
80	5.1e-4	3.8	7.2e-4	4.0	5.0e-4	3.8	8.3e-4	3.9
160	1.3e-4		1.8e-4		1.3e-4		2.1e-4	

Table 3: Errors for diffusion dominated flow ( $\Gamma_0 = 0.1$ ).

N	Modified Thiaert		Central		Constant Mass-Flux		Homogeneous Flux	
	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$	$\epsilon_N$	$\epsilon_N/\epsilon_{2N}$
10	6.8e-2	4.5	2.3e-1	13.0	6.6e-2	3.7	1.6e-1	2.0
20	1.5e-2	4.9	1.8e-2	4.4	1.8e-2	3.6	8.3e-2	2.3
40	3.1e-3	6.0	4.0e-3	4.1	4.9e-3	2.9	3.7e-2	2.7
80	5.4e-4	6.5	9.9e-4	4.0	1.7e-3	2.9	1.4e-2	2.7
160	9.3e-5		2.4e-4		5.7e-4		4.4e-3	

Table 4: Errors for convection dominated flow ( $\Gamma_0 = 0.005$ ).

When we try to solve the problem with dominant convection, we run into a new difficulty. When  $\Gamma_0$  is very small, the discretization matrix for the central difference scheme becomes so ill-conditioned that our solver, GMRES(n) with a tridiagonal preconditioner, cannot find the answer. Since 2-D problems increase in size much more rapidly than 1-D problems, we cannot refine the meshes so often that we can really see the asymptotic behavior. The results in Table 4 are therefore not quite as pronounced as in Table 2. Still it can be observed that the central difference scheme suffers from large errors due to spurious oscillations, that the Modified Thiaert Scheme is superior to the other schemes, and that the methods 2 and 3 do not quite converge quadratically. The Modified Thiaert Scheme shows a convergence rate which is faster than expected, which we assume is non-asymptotic behavior.

## References

- [1] Z. CAI AND S. MCCORMICK, *On the Accuracy of the Finite Volume Element Method for Diffusion Equations on Composite Grids*, SIAM J. Numer. Anal. 27 (1990), pp. 636-655.
- [2] P.J.J. FERKET, *The Finite Difference Based Fast Adaptive Composite Grid Method*, RANA 95-19, Department of Mathematics and Computing Science, Eindhoven University of Technology, December 1995 (to appear in Computing).
- [3] W. HACKBUSH, *Elliptic Differential Equations, Theory and Numerical Treatment*, Springer, Berlin, 1992.
- [4] M.G. HALL, *Cell-Vertex Multigrid Schemes for Solution of the Euler Equation*. Controller HMSO London, 1985.
- [5] G. HELLWIG, *Partial Differential Equations, an Introduction*. Blaisdell Publishing Company, 1960.
- [6] H.C. DE LANGE, *Modelling of Premixed Laminar Flames*, Ph.D. Thesis, Eindhoven University of Technology, 1992.

- [7] B.P. LEONARD, *A Stable and Accurate Convective Modelling Procedure Based on Quadratic Upstream Interpolation*, Comp. Meths. Appl. Mech. Eng. 19 (1979), pp. 59-98.
- [8] S.V. PATANKAR, *Numerical Heat Transfer and Fluid Flow*, McGraw-Hill, New York 1980.
- [9] R.C. RADESPIEL, C. ROSSOW AND R.C. SWANSON, *Efficient Cell-Vertex Multigrid Scheme for the Three-Dimensional Navier-Stokes Equations*, AIAA J. 28 (1990), pp. 1464-1472.
- [10] L.M.T. SOMERS, *The Simulation of Flat Flames with Detailed and Reduced Chemical Models*, Ph. D. Thesis, Eindhoven University of Technology, 1994.
- [11] D.B. SPALDING, *A Novel Finite Difference Formulation For Differential Expressions Involving Both First and Second Derivatives*, Int. J. Numer. Meths. Eng. 4 (1972), pp. 551-559.
- [12] P.K. SWEBY, *Flux Limiters*, Proc. INRIA Workshop, Rocquencourt/France 1983 (1985), pp. 48-65.
- [13] P.K. SWEBY, *High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws*, SIAM J. Numer. Anal. 21 (1984), pp. 995-1011.
- [14] G.D. THIART, *Finite Difference Scheme for the Numerical Solution of Fluid Flow and Heat Transfer Problems on Nonstaggered Grids*, Numer. Heat Transfer, Part B 17 (1990), pp 41-62.
- [15] G.D. THIART, *Improved Finite-Difference Scheme for the Solution of Convection-Diffusion Problems with the SIMPLEN algorithm*, Numer. Heat Transfer, Part B 18 (1990), pp 81-95.
- [16] F.O.O WAGENER, *Cell-Vertex Discretization of the Driven Cavity Flow Without Artificial Diffusion*, ARD 02 RUG (9317) (1993).



PREVIOUS PUBLICATIONS IN THIS SERIES:

Number	Author(s)	Title	Month
95-15	P.W.C. Vosbeek R.M.M. Mattheij	Contour Dynamics with Symplectic Time Integration	October '95
95-16	P.M.E.J. Wijckmans R.M.M. Mattheij	Discontinuities in Mechanical Systems	October '95
95-17	L.P.H. de Goey J.H.M. ten Thijsse Boonkkamp	A Generalized Definition of Flame Stretch for 3D Instationary Pre- mixed Flames	November '95
95-18	M.M.A. de Rijcke	Translatable strict LF-spaces of $\mathcal{D}_-(\mathbb{R})$ -type and Realization theory	November '95
95-19	P.J.J. Ferket	The Finite Difference Based Fast Adaptive Composite Grid Method	December '95
95-20	S.J.L. van Eijndhoven M.M.A. de Rijcke	Convolution algebras translation in- variant operators	December '95
95-21	P.J.P.M. Simons R.M.M. Mattheij	The cooling of molten glass in a mould	December '95
96-01	M. Günther G. Prokert	Existence results for the quasista- tionary motion of a free capillary liq- uid drop	January '96
96-02	B. van 't Hof J.H.M. ten Thijsse Boonkkamp R.M.M. Mattheij	Discretization of the Stationary Convection-Diffusion-Reaction Equation	February '96

