1992

R.R.P. van Nooyen

A Petrov-Galerkin mixed finite element method
with exponential fitting

Department of Numerical Mathematics Report NM-R9204 March

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

A Petrov-Galerkin Mixed Finite Element Method with Exponential Fitting

R. R. P. van Nooyen

CWI,
P. O. box 4079, 1009 AB Amsterdam,
The Netherlands

We discuss a Petrov-Galerkin mixed finite element formulation of the semiconductor continuity equations on a rectangular domain. We give error estimates for equations that are in principle degenerate in the singularly perturbed case. We give arguments that indicate that the method is also effective in the singularly perturbed case. We develop a discretisation that gives a higher order accurate solution for use in an a posteriori error estimator. The methods under discussion is a modified version of the method discussed in CWI report NM-R9005.

1990 CR category: G. 1.8 Finite Element Methods
1980 AMS Subject Classification: 65N30
Key Words and Phrases: Mixed Finite Elements, Semiconductor Continuity Equation, Exponential Fitting, A posteriori error estimates, Convection Diffusion Equation, Current Continuity Equation.
Note: This report will be submitted for publication elsewhere.

1 Introduction.

The use of a form of exponential fitting for the semiconductor continuity equation is suggested by the success of the Scharfetter-Gummel discretisation[1] in one dimension and variations on that discretisation in two dimensions. Numerous derivations of Scharfetter-Gummel type discretisations are given in the literature, for instance by Selberherr[2], Markowich[3], Bank et al.[4], Brezzi et al.[5], and others. This paper extends a one dimensional exponential fitting technique, discussed by Hemker[6], to the two dimensional problem.

In section 2 we introduce a model equation for the semiconductor continuity equations. We introduce several bilinear forms, related to the coefficients in this equation. In section 3 and 4 we treat the discretisation. In section 5 we collect some technical results and in section 6 we derive two error estimates. These error estimates are based on the techniques used by Douglas and Roberts[7]. The proofs in section 6 take all characteristics of our special discrete system into account, in particular the quadrature rule for the approximation of certain integrals in the discrete system. Note that the error estimates in section 6 are degenerate if the problem is singularly perturbed, i.e. if the convection dominates in the problem. On the other hand, an indication for good behaviour of the method for singular problems is that - for constant coefficients - it reproduces reproduces the solution $C\exp(-\beta_1 x_1 - \beta_2 x_2)$ exactly. In section 9, we develop an a posteriori error estimator. In the last section we discuss our findings.

2 The equation.

We consider the following problem, find $u \in \mathrm{H}^2(\Omega)$ such that:

$$-\operatorname{div}(\frac{1}{\alpha}(\mathbf{grad}\, u + u\boldsymbol{\beta})) + \gamma u = f \quad \text{on} \quad \Omega \ \text{and} \tag{1}$$

$$u = -g \ \text{ on } \partial\Omega\,,$$

Report NM-R9204
ISSN 0169-0388
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

coefficients:

$$\alpha \in W_1^\infty(\Omega) \text{ and } \exists\ A \in \mathbb{R} : \alpha \geqslant A > 0 \text{ on } \Omega\,, \tag{2}$$

$$\frac{1}{\alpha} \in W_1^\infty(\Omega) \text{ on } \Omega\,, \tag{3}$$

$$\boldsymbol{\beta} = (\beta_1,\beta_2)^T \text{ with } \beta_1,\beta_2 \in W_1^\infty(\Omega)\,, \tag{4}$$

$$\gamma \in W_1^\infty(\Omega) \text{ and } \gamma \geqslant 0 \text{ on } \Omega\,, \tag{5}$$

where $W_1^\infty(\Omega)$, $H^2(\Omega)$ are the usual Sobolev spaces[8], and

$$H(\text{div},\Omega) := \{\ \boldsymbol{\tau} \in L^2(\Omega)^2 \mid \text{div}\,\boldsymbol{\tau} \in L^2(\Omega)\ \}\,,$$

with scalar product

$$(\boldsymbol{\sigma},\boldsymbol{\tau})_{H(\text{div},\Omega)} = \int_\Omega \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,d\mu + \int_\Omega \text{div}\,\boldsymbol{\sigma}\ \text{div}\,\boldsymbol{\tau}\,d\mu\,,$$

is a Hilbert space (see also Girault and Raviart, [9] formula 2.15 in section 2.2). We assume, that the equation has a solution and that $f \in L^2(\Omega)$, $g \in H^{3/2}(\partial\Omega)$.

The stationary semiconductor continuity equations take the form (1). Here $\boldsymbol{\beta}$ corresponds to the electric field, the term γu corresponds to a linear approximation to the recombination term and $1/\alpha$ corresponds to the electron or hole mobility. The exact correspondence depends on the choice of scaling[10].

To formulate the weak mixed form of this equation, we use the following bilinear forms

$$(s,t) = \int_\Omega s\,t\,d\mu \quad \forall\ s,t \in L^2(\Omega)\,,$$

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) = \int_\Omega \alpha\boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,d\mu \quad \forall\ \boldsymbol{\sigma},\boldsymbol{\tau} \in H(\text{div},\Omega)\,,$$

$$b(\boldsymbol{\sigma},t) = \int_\Omega \boldsymbol{\beta}\cdot\boldsymbol{\sigma}\,t\,d\mu \quad \forall\ \boldsymbol{\sigma} \in H(\text{div},\Omega)\,,\ t \in L^2(\Omega)\,,$$

$$c(s,t) = \int_\Omega \gamma\,s\,t\,d\mu \quad \forall\ s,t \in L^2(\Omega)\,,$$

$$< g,h > \ = \int_{\partial\Omega} g\,h\,d\lambda \quad \forall\ g,h \in L^2(\partial\Omega)\,.$$

Given these definitions, we see immediately, that any solution $u \in H^2(\Omega)$ of (1) generates a solution $(\boldsymbol{\sigma},u) \in H(\text{div},\Omega) \times L^2(\Omega)$ of

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}) - (\text{div}\,\boldsymbol{\tau},u) + b(\boldsymbol{\tau},u) = \ < g,\boldsymbol{\tau}\cdot\mathbf{n}_{\partial\Omega} > \qquad \forall\ \boldsymbol{\tau} \in H(\text{div},\Omega)\,, \tag{6a}$$

$$(\text{div}\,\boldsymbol{\sigma},t) + c(u,t) = (f,t) \quad \forall\ t \in L^2(\Omega)\,. \tag{6b}$$

Where $\boldsymbol{\sigma} = -\frac{1}{\alpha}(\mathbf{grad}\,u + u\boldsymbol{\beta})$.

To simplify the notation, we denote the Cartesian product of a normed linear space E with itself by E in bold faced type, $\mathbf{E} := E \times E$. We define

$$\|(\mu_1,\mu_2)^T\|_{\mathbf{E}} := (\sum_{i=1}^{2} \|\mu_i\|_E^2)^{1/2} \quad \forall\ (\mu_1,\mu_2)^T. \in \mathbf{E}\,.$$

3 Preparations.

We introduce a partition of the domain and we define the adjoint problem of (1), which we use in the derivation of one of our error estimates. Next, we introduce several special projections, that are needed in the definition of our approximation spaces and in the derivation of the error estimates. Finally we give an error estimate for the projections.

3.1. The partitioning of the domain.

We assume, that our domain Ω is rectangular. On Ω, we use Cartesian coordinates, with the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ parallel to the edges of Ω. We define $\tau_i := \boldsymbol{\tau}\cdot\mathbf{e}_i$ for $\boldsymbol{\tau} \in \mathbf{L}^2(\Omega)$ and $x_i := \mathbf{x}\cdot\mathbf{e}_i$ for $\mathbf{x} \in \mathbb{R}^2$. Before we treat our discretisation, we define our approximation space. We assume that our partition is the cartesian product of partitions

$$P = \{ 0 = p_0 < p_1 < \cdots < p_{N_1} = L_1 \} , \tag{7}$$

and

$$Q = \{ 0 = q_0 < q_1 < \cdots < q_{N_2} = L_2 \} \tag{8}$$

of the sides of our domain. We define the index set K,

$$K = \{ (i+½,j+½) \mid i=0,1,\ldots,N_1-1 \, , j=0,1,\ldots,N_2-1 \} ,$$

with the obvious index pair for a given cell,

$$\Omega_{i+½,j+½} = \{ \mathbf{x} \mid p_i < x_1 < p_{i+1} \, , q_j < x_2 < p_{j+1} \} .$$

We define $\mathbf{x}_k$ to be the centre of Ω_k and $\mathbf{h}_k$ to be the diagonal of Ω_k. We use the notation χ_k for the characteristic function of Ω_k. (The characteristic function of a set is the function that is equal to one in all points of the set and zero elsewhere). The edges of Ω_k are the sets:

$$\Gamma_{k,i,j} = \{ \mathbf{x} \in \overline{\Omega}_k \mid \mathbf{x}\cdot\mathbf{e}_i = (\mathbf{x}_k+(j-½)\mathbf{h}_k)\cdot\mathbf{e}_i \} \quad \text{for} \quad i=1,2 \, , j = 0,1 . \tag{9}$$

$\chi_{k,i,j}$ is the characteristic function of edge $\Gamma_{k,i,j}$. So $(i,j)=(1,0),(1,1),(2,0),(2,1)$ denote the left, right, bottom and top edges.

3.2. The adjoint problem.

We use the following definition for the adjoint problem of (1) (cf. Douglas and Roberts [7]),

$$w \in \mathrm{H}^2(\Omega) \, , \tag{10}$$

$$-\operatorname{div}(\frac{1}{\alpha}\,\mathbf{grad}\, w) + \frac{\boldsymbol{\beta}}{\alpha}\cdot\mathbf{grad}\, w + \gamma w = f \text{ on } \Omega ,$$

$$w = 0 \text{ on } \partial\Omega .$$

The adjoint problem is called regular, if there is a unique solution w for every $f \in \mathrm{L}^2(\Omega)$ and this solution satisfies $\|w\|_{\mathrm{H}^2(\Omega)} \leqslant C\|f\|_{\mathrm{L}^2(\Omega)}$ for every $f \in \mathrm{L}^2(\Omega)$.

Both in the above equation and in the rest of this report, the upper case C, without a subscript, denotes a generic constant. It may have a different value at each appearance.

The weak mixed form of the adjoint problem is:

$$(\boldsymbol{\tau},w) \in \mathrm{H}(\mathrm{div},\Omega)\times \mathrm{L}^2(\Omega) \, , \tag{11}$$

$$a(\boldsymbol{\tau},\boldsymbol{\sigma})-(\operatorname{div}\boldsymbol{\sigma},w) = 0 \quad \forall \;\; \boldsymbol{\sigma} \in \mathrm{H}(\mathrm{div},\Omega) \text{ and} \tag{11a}$$

$$(\operatorname{div}\boldsymbol{\tau},t)-b(\boldsymbol{\tau},t)+c(w,t) = (f,t) \quad \forall \;\; t \in \mathrm{L}^2(\Omega) . \tag{11b}$$

Any solution $w \in \mathrm{H}^2(\Omega)$ of (10) generates a solution $(-\frac{1}{\alpha}\,\mathbf{grad}\, w,w)$ of this problem. If (9) is regular, then this solution satisfies $\|w\|_{\mathrm{H}^2(\Omega)} + \|\boldsymbol{\tau}\|_{\mathrm{H}^1(\Omega)} \leqslant C\|f\|_{\mathrm{L}^2(\Omega)}$.

3.3. Some projections.

We introduce several local projections, we use these to define four global mappings, P_h, $\mathbf{P}_h$, Π_h and $\tilde{\Pi}_h$ that map function spaces to finite dimensional function spaces. First, we define $P[\Omega_k]$ to be the orthogonal projection from $\mathrm{L}^2(\Omega_k)$ to the space of constant functions on Ω_k, and we define $P[\Gamma_{k,i,j}]$ to be the orthogonal projection from $\mathrm{L}^2(\Gamma_{k,i,j})$ to the space of constant functions on $\Gamma_{k,i,j}$.

We use $P[\Omega_k]$ to create two global mappings, P_h: $\mathrm{L}^2(\Omega)\rightarrow \mathrm{L}^2(\Omega)$,

$$P_h f = \sum_{k \in K} \chi_k P[\Omega_k](f) \quad \forall \;\; f \in \mathrm{L}^2(\Omega) \, , \tag{12a}$$

and $\mathbf{P}_h:\mathbf{L}^2(\Omega)\to\mathbf{L}^2(\Omega)$,

$$\mathbf{P_h}\boldsymbol{\beta} = \sum_{k\in K} \chi_k \left[P[\Omega_k](\boldsymbol{\beta}\cdot\mathbf{e}_1)\mathbf{e}_1 + P[\Omega_k](\boldsymbol{\beta}\cdot\mathbf{e}_2)\mathbf{e}_2\right] \quad \forall \ \boldsymbol{\beta}\in\mathbf{L}^2(\Omega) . \tag{12b}$$

Next, we introduce two mappings, based on $P[\Gamma_{k,i,j}]$. These mappings have as their domain the space Σ,

$$\Sigma := \{ \ \boldsymbol{\tau}\in \mathrm{H}(\mathrm{div},\Omega) \ | \ \boldsymbol{\tau}|_{\partial\Omega_k}\cdot\mathbf{n}_{\partial\Omega_k} \in \mathrm{L}^2(\partial\Omega_k) \quad \forall \ k\in K \ \} .$$

This space is similar to the one introduced by Roberts and Thomas in formula (1.10) of their report[11].

To simplify the definition of these mappings, we introduce local coordinates on each cell Ω_k,

$$\vec{\xi}_k := \begin{bmatrix} \dfrac{x_1 - x_{k,1}}{h_{k,1}} + \dfrac{1}{2} \\ \dfrac{x_2 - x_{k,2}}{h_{k,2}} + \dfrac{1}{2} \end{bmatrix} . \tag{13}$$

The mappings are defined as follows:

$$\Pi_h\boldsymbol{\tau} = \sum_{k\in K} \chi_k \sum_{i=1}^{2} \left[(1-\xi_{k,i})P[\Gamma_{k,i,0}](\tau_i) + \xi_{k,i}P[\Gamma_{k,i,1}](\tau_i)\right]\mathbf{e}_i , \tag{14}$$

$$\tilde{\Pi}_h\boldsymbol{\tau} = \sum_{k\in K} \chi_k \sum_{i=1}^{2} \left[(1-\zeta_{k,i}) P[\Gamma_{k,i,0}](\tau_i) + \zeta_{k,i}P[\Gamma_{k,i,1}](\tau_i)\right]\mathbf{e}_i , \tag{15}$$

where

$$\zeta_{k,i} = \begin{cases} \dfrac{\exp(\xi_{k,i}h_{k,i}P[\Omega_k](\beta_i)) - 1}{\exp(h_{k,i}\ P[\Omega_k](\beta_i)) - 1} & \text{if } P[\Omega_k](\beta_i) \neq 0 , \\ \xi_{k,i} & \text{if } P[\Omega_k](\beta_i) = 0 . \end{cases}$$

So, for $\Pi_h\boldsymbol{\tau}$ we get the i^{th} component on Ω_k by linear interpolation between the projections of this component on the two sides orthogonal to $\mathbf{e}_i$. For $\tilde{\Pi}_h\boldsymbol{\tau}$ however, we obtain the same component by using an exponential function to interpolate between the projections of this component on the two sides orthogonal to $\mathbf{e}_i$.

Now we introduce the following finite dimensional function spaces as the ranges of the above projections,

$$V_h = \Pi_h(\Sigma) , \ W_h = P_h(\mathrm{L}^2(\Omega)) \ \text{ and } \ X_h = \tilde{\Pi}_h(\Sigma) .$$

$V_h \times W_h$ is the lowest order Raviart-Thomas-Nedelec space for rectangles. This space and the above projections were described by Douglas and Roberts, [7] Raviart and Thomas[12] and, for $\Omega \subset \mathbb{R}^3$, by Nedelec[13]. In this paper we use the usual space, $V_h \times W_h$, as the trial function space and $X_h \times W_h$ as the test function space. In effect, we use exponential test functions instead of the usual linear test functions. Thus, we obtain a Petrov-Galerkin mixed finite element discretisation.

3.4. Error estimates for projections.

We prove a lemma on the accuracy of our projections. Considering the number and diversity of articles on error estimates, e.g. the classical projection estimates from Ciarlet and Raviart[14], this may seem superfluous, but we shall see that the relative simplicity of the case under consideration makes it possible to derive sharp error estimates under minimal assumptions.

Lemma 1.
If f is a square integrable function with square integrable derivatives on a rectangle

$\Omega=[0,h_1]\times[0,h_2]$

with sides $\Gamma_{1,1}=\{\ h_1\ \}\times[0,h_2]$, $\Gamma_{2,1}=[0,h_1]\times\{\ h_2\ \}$, $\Gamma_{1,0}=\{\ 0\ \}\times[0,h_2]$ and $\Gamma_{2,0}=[0,h_1]\times\{\ 0\ \}$, then the following inequalities hold,

$$||f-P[A]f||_{L^2(\Omega)} \leq (2h_1^2+2h_2^2)^{1/2}\, ||\,\mathbf{grad} f\,||_{L^2(\Omega)} . \tag{16a}$$

If s is a continuous function with domain $[0,h_1]$ and range $[0,1]$, then we have,

$$||f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f||_{L^2(\Omega)} \leq (2h_1^2+2h_2^2)^{1/2}\, ||\,\mathbf{grad} f\,||_{L^2(\Omega)} . \tag{16b}$$

If $f\in L^\infty(\Omega)$, $\mathbf{grad} f\in \mathbf{L}^\infty(\Omega)$, then

$$||f-P[A]f||_{L^\infty(\Omega)} \leq (h_1+h_2)\, ||\,\mathbf{grad} f\,||_{L^\infty(\Omega)} . \tag{16c}$$

Proof.

We start by proving the above inequalities for $f\in C^1(A)$. We can then extend them by the usual density argument to $H^1(\Omega)$. To prove the first inequality, we write,

$$||f-P[A]f||^2_{L^2(\Omega)} = \int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{1}{h_1h_2}\int_{w=0}^{h_1}\int_{z=0}^{h_2} f(x,y)-f(w,z)dwdz\right]^2 dxdy ,$$

by definition,

$$f(x,y)-f(w,z) = \int_{a=w}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)da .$$

If we substitute this into the above expression, then we find

$$||f-P[\Omega]f||^2_{L^2(\Omega)} = \int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{1}{h_1h_2}\int_{w=0}^{h_1}\int_{z=0}^{h_2}\left[\int_{a=w}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)db\right]dwdz\right]^2 dxdy .$$

We apply the Hölder inequality to the inner integrals and extend the integrations over a and b where appropriate,

$$||f-P[\Omega]f||^2_{L^2(\Omega)} \leq$$

$$\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{h_1^{1/2}}{h_2^{1/2}}||\partial f/\partial x_1||_{L^2(\Omega)} + h_2^{1/2}\left[\int_{b=0}^{h_2}\left[\frac{\partial f}{\partial b}(x,b)\right]^2 db\right]^{1/2}\right]^2 dxdy .$$

We use $(|A|+|B|)^2 \leq 2(A^2+B^2)$ to write this as,

$$||f-P[\Omega]f||^2_{L^2(\Omega)} \leq$$

$$2\int_{x=0}^{h_1}\int_{y=0}^{h_2}\frac{h_1}{h_2}||\partial f/\partial x_1||^2_{L^2(\Omega)}dxdy + 2\int_{y=0}^{h_2} h_2||\partial f/\partial x_2||^2_{L^2(\Omega)}dy .$$

This reduces to,

$$||f-P[\Omega]f||^2_{L^2(\Omega)} \leq$$

$$2h_1^2\,||\partial f/\partial x_1||^2_{L^2(\Omega)} + 2h_2^2\,||\partial f/\partial x_2||^2_{L^2(\Omega)} .$$

Now, we consider the second inequality, (16b), we write,

$$||f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f||^2_{L^2(\Omega)} =$$

$$\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{1}{h_2}\int_{z=0}^{h_2}\Big[(1-s(x))(f(x,y)-f(0,z)) + s(x)(f(x,y)-f(h_1,z))\Big]dz\right]^2 dxdy .$$

We use partial derivatives to rewrite the expression,

$$\|f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f\|^2_{L^2(\Omega)} =$$

$$\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{1}{h_2}\int_{z=0}^{h_2}\left[(1-s(x))\left\{\int_{a=0}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)db\right\} + s(x)\left\{\int_{a=h_1}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)db\right\}\right]dz\right]^2 dxdy\ .$$

We use Hölder and extend the integrals where appropriate,

$$\|f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f\|^2_{L^2(\Omega)} \leq$$

$$\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{h_1^{1/2}}{h_2^{1/2}}\|\partial f/\partial x_1\|_{L^2(\Omega)} + h_2^{1/2}\left[\int_{b=0}^{h_2}\left[\frac{\partial f}{\partial b}(x,b)\right]^2 db\right]^{1/2}\right]^2 dxdy\ .$$

We use $(|A|+|B|)^2 \leq 2(A^2+B^2)$ to write this as,

$$\|f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f\|^2_{L^2(\Omega)} \leq$$

$$2\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\frac{h_1}{h_2}\|\partial f/\partial x_1\|^2_{L^2(\Omega)} + h_2\int_{b=0}^{h_2}\left[\frac{\partial f}{\partial b}(x,b)\right]^2 db\right]dxdy\ .$$

This reduces to,

$$\|f-(1-s)\Pi[\Gamma_{1,0}]f-s\Pi[\Gamma_{1,1}]f\|^2_{L^2(\Omega)} \leq$$
$$2h_1^2\|\partial f/\partial x_1\|^2_{L^2(\Omega)} + 2h_2^2\|\partial f/\partial x_2\|^2_{L^2(\Omega)}\ .$$

Lastly we verify (16c),

$$f(x,y)-f(w,z) = \int_{a=w}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)da\ .$$

So,

$$\frac{1}{h_1h_2}\int_{x=0}^{h_1}\int_{y=0}^{h_2} f(x,y)-f(w,z)dxdy =$$

$$\frac{1}{h_1h_2}\int_{x=0}^{h_1}\int_{y=0}^{h_2}\left[\int_{a=w}^{x}\frac{\partial f}{\partial a}(a,z)da + \int_{b=z}^{y}\frac{\partial f}{\partial b}(x,b)da\right]dxdy \leq (h_1+h_2)\|\mathbf{grad} f\|_{L^\infty(\Omega)}\ .$$

□

Note that the above inequalities imply,

$$\|\boldsymbol{\sigma}-\Pi_h\boldsymbol{\sigma}\|_{L^2(\Omega)} \leq \max_{k\in K}(2h_{k,1}^2+2h_{k,2}^2)^{1/2}\|\boldsymbol{\sigma}\|_{H^1(\Omega)}\ , \qquad (17a)$$

$$\|\boldsymbol{\sigma}-\tilde{\Pi}_h\boldsymbol{\sigma}\|_{L^2(\Omega)} \leq \max_{k\in K}(2h_{k,1}^2+2h_{k,2}^2)^{1/2}\|\boldsymbol{\sigma}\|_{H^1(\Omega)}\ , \qquad (17b)$$

$$\|u-P_hu\|_{L^2(\Omega)} \leq \max_{k\in K}(2h_{k,1}^2+2h_{k,2}^2)^{1/2}\|u\|_{H^1(\Omega)}\ , \qquad (17c)$$

for suitable u and $\boldsymbol{\sigma}$.

4 The discretisation.

We describe our discretisation. The basic idea of mixed finite elements with a lowest order Raviart-Thomas trial space and an exponentially fitted test subspace for the vector valued functions is complicated by the use of a quadrature rule, needed to keep the M-matrix property for the system without Lagrange multipliers for non-zero γ. This quadrature rule is discussed in section 4.1. Another complication is the approximation of the coefficients by piecewise constant functions, as described below. In section 4.2 we give the resulting discretisation.

We replace the coefficients α, $\boldsymbol{\beta}$ and γ by two dimensional step functions. To write our modified problem in weak form, we need to define three new bilinear forms,

$$\overline{a}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \int_{\Omega_k} \boldsymbol{\sigma}\cdot\boldsymbol{\tau} P_h\alpha \, d\mu \quad \forall \ \boldsymbol{\sigma},\boldsymbol{\tau} \in \Sigma ,$$

$$\overline{b}(\boldsymbol{\sigma},t) := \int_{\Omega} t\boldsymbol{\sigma}\cdot \mathbf{P}_h\boldsymbol{\beta} \, d\mu \quad \forall \ \boldsymbol{\sigma} \in \Sigma, t \in L^2(\Omega) ,$$

$$\overline{c}(s,t) := \int_{\Omega} st P_h\gamma \, d\mu \quad \forall \ s,t \in L^2(\Omega) .$$

The bar on the bilinear forms denotes that the coefficients are replaced by their cell-wise averages. We then replace $\overline{a}$ by $\overline{a}_q$, the subscript q indicates that a - not yet specified - quadrature rule will be used in the evaluation of this bilinear form.

4.1. The quadrature rule.

We construct a quadrature rule $\overline{a}_{h,1}$ by imposing the condition that, if α, $\boldsymbol{\beta}$ are constant, $\gamma \equiv 0$, and the solution satisfies $u = C\exp(-\beta_1 x_1 - \beta_2 x_2) + K$, with $C,K \in \mathbb{R}$, then the discrete solution should satisfy $\boldsymbol{\sigma}_h = \Pi_h\boldsymbol{\sigma}$ and $u_h = P_h u$. We see that for the u given above $\boldsymbol{\sigma} = -K\boldsymbol{\beta}/\alpha$, so $\boldsymbol{\sigma}$ is constant. We define α_h separately for each basis function $\boldsymbol{\eta}_{i,j+1/2}$ where

$$\boldsymbol{\eta}_{i,j+1/2} = \begin{cases} \zeta_{(i-1/2,j+1/2),1}\mathbf{e}_1 & \text{on } \Omega_{i-1/2,j+1/2} , \\ (1-\zeta_{(i+1/2,j+1/2),1})\mathbf{e}_1 & \text{on } \Omega_{i+1/2,j+1/2} , \\ 0 & \text{elsewhere} , \end{cases}$$

and $\boldsymbol{\eta}_{i+1/2,j}$ where

$$\boldsymbol{\eta}_{i+1/2,j} = \begin{cases} \zeta_{(i+1/2,j-1/2),2}\mathbf{e}_2 & \text{on } \Omega_{i+1/2,j-1/2} , \\ (1-\zeta_{(i+1/2,j-1/2),2})\mathbf{e}_2 & \text{on } \Omega_{i+1/2,j+1/2} . \\ 0 & \text{elsewhere} . \end{cases}$$

We denote the set of all possible indices for the basis functions $\boldsymbol{\eta}$ by

$$E = \{ e=(i,j-1/2) \mid i=0,1,2,\ldots,N_1 , j=1,2,\ldots,N_2 \} \cup$$
$$\{ e=(i-1/2,j) \mid i=1,2,\ldots,N_1 , j=0,1,2,\ldots,N_2 \} .$$

Our quadrature rule should satisfy the following condition,

$$\overline{a}_{h,1}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) = \overline{a}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) , \tag{A}$$

for all constant $\boldsymbol{\sigma}$ and all $r \in E$. Due to our assumption that the coefficients are constant, we have $a=\overline{a}$ and $b=\overline{b}$. The above condition guarantees that, for constant coefficients and constant $\boldsymbol{\sigma}$,

$$a(\boldsymbol{\sigma},\boldsymbol{\tau}_h) - (u, \operatorname{div}\boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h,u) = \overline{a}_{h,1}(\Pi_h\boldsymbol{\sigma},\boldsymbol{\tau}_h) - (P_h u, \operatorname{div}\boldsymbol{\tau}_h) + b(\boldsymbol{\tau}_h,P_h u) \quad \forall \ \boldsymbol{\tau}_h \in X_h$$

and we also have,

$$(\operatorname{div}\boldsymbol{\sigma},t) = (\operatorname{div}\Pi_h\boldsymbol{\sigma},t) = 0 \quad \forall \ t \in L^2(\Omega) .$$

So our condition (A) on $\overline{a}_{h,1}$ is sufficient for our purposes. We now select the quadrature rule by taking the following definition for $\overline{a}_{h,1}$,

$$\overline{a}_{h,1}(\sigma,\tau) = \tag{18a}$$
$$\sum_{k\in K}\sum_{i=1}^{2} \mu(\Omega_k)P[\Omega_k](\alpha)\left[P[\Omega_k](\zeta_{k,i})P[\Gamma_{k,i,1}](\sigma_i\tau_i) + P[\Omega_k](1-\zeta_{k,i})P[\Gamma_{k,i,0}](\sigma_i\tau_i)\right] .$$

We introduce a new problem dependent norm on X_h

$$\|\tau_h\|_h = \sum_{k\in K}\sum_{i=1}^{2} \mu(\Omega_k)\left[P[\Omega_k](\zeta_{k,i})P[\Gamma_{k,i,1}](\tau_{h,i}^2) + P[\Omega_k](1-\zeta_{k,i})P[\Gamma_{k,i,0}](\tau_{h,i}^2)\right]^{1/2} . \tag{18b}$$

From this point onwards, we take $\overline{a}_q = \overline{a}_{h,1}$.

4.2. The discrete system.

We approximate the solution (σ,u) of (6) by $(\sigma_h,u_h)\in V_h\times W_h$, where

$$\overline{a}_q(\sigma_h,\tau) - (u_h,\operatorname{div}\tau) + \overline{b}(\tau,u_h) = \;<\tau\cdot\mathbf{n}_{\partial\Omega},g>\quad \forall\ \tau\in X_h\ , \tag{19a}$$

$$(\operatorname{div}\sigma_h,t) + \overline{c}(u_h,t) = (f,t)\quad\forall\ t\in W_h\ . \tag{19b}$$

If we use $\overline{a}$ in stead of $\overline{a}_q$, then that means that our discrete problem does not always yield an M-matrix for u_h. Consider, for instance, the corresponding discretisation on a uniform mesh with mesh width h in one dimension with $\alpha = 1$, $\beta = \vec{0}$ and γ constant. If $\alpha\gamma h^2/6 > 1$, then the off-diagonal elements of the discretisation matrix for u_h after elimination of σ_h through static condensation have the same sign as the elements on the diagonal.

The idea of using linear trial functions and exponential test functions was used by Hemker for singularly perturbed two point boundary problems[6]. For the one dimensional case, the introduction of exponential test functions follows from the requirement that it must be possible to approximate the Green's function of the problem by the test functions. For finite elements in one dimension the the singularly perturbed case was studied by O'Riordan and Stynes[15-20] and Reinhardt[21]. For finite element in two dimensions O'Riordan and Stynes derive a uniformly convergent estimate[22] but only for problems with a strictly positive zero-order term.

In the following sections, we prove, that the solution of our discretisation (19) is an $\mathcal{O}(h)$ approximation to the solution of our original problem.

5 Several technical results.

This section contains some technical results, collected for later reference.

Lemma 2.

$$\tilde{\Pi}_h\circ\Pi_h = \tilde{\Pi}_h\ , \tag{20a}$$

$$\Pi_h\circ\tilde{\Pi}_h = \Pi_h\ , \tag{20b}$$

$$(\operatorname{div}\sigma,P_h t) = (\operatorname{div}\Pi_h\sigma,t)\quad\forall\ \sigma\in\Sigma\ ,\ t\in L^2(\Omega)\ , \tag{20c}$$

$$\Pi_h\tau\cdot\vec{n}_{\partial\Omega} = \tilde{\Pi}_h\tau\cdot\vec{n}_{\partial\Omega}\quad\forall\ \tau\in\Sigma\ . \tag{20d}$$

Proof.
Both mappings are based on the same projections $P[\Gamma_{k,i,j}]$, so (20a) and (20b) are trivial.

To prove (20c) we use a special case of Green's theorem:

$$\int_{\Omega_k}\operatorname{div}\sigma\,d\mu = \sum_{i=1}^{2}\frac{\mu(\Omega_k)}{h_{k,i}}\left[P[\Gamma_{k,i,1}](\sigma_i) - P[\Gamma_{k,i,0}](\sigma_i)\right] .$$

If we combine this with the definition of Π_h, the proof of (20c) is complete. Equation (20d) follows immediately from the definitions. □

Lemma 3.
If $\sigma\in\Sigma$ and we define $a_{k,i} = P[\Gamma_{k,i,0}](\sigma_i)$ and $b_{k,i} = P[\Gamma_{k,i,1}](\sigma_i)$, then the following inequalities

hold for $\|\Pi_h \boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega_k)}$ and $\|\tilde{\Pi}_h \boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega_k)}$,

$$\frac{\mu(\Omega_k)}{6}\sum_{i=1}^{2}(a_{k,i}^2 + b_{k,i}^2) \leq \|\Pi_h \boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega_k)}^2 \leq \frac{\mu(\Omega_k)}{2}\sum_{i=1}^{2}(a_{k,i}^2 + b_{k,i}^2) . \tag{21a}$$

$$\|\tilde{\Pi}_h \boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega_k)}^2 \leq 2\|\tilde{\Pi}_h \boldsymbol{\sigma}\|_h^2 \leq 12\|\Pi_h \boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega)}^2 . \tag{21b}$$

Proof.
Formula (21a) follows immediately from

$$(\Pi_h \boldsymbol{\sigma}, \Pi_h \boldsymbol{\sigma}) = \sum_{i=1}^{2}\sum_{k\in K}\int_{\Omega_k}\left[(1-\xi_{k,i})a_{k,i} + \xi_{k,i}b_{k,i}\right]^2 d\mu .$$

Next, we derive (21b) from,

$$(\tilde{\Pi}_h \boldsymbol{\sigma}, \tilde{\Pi}_h \boldsymbol{\sigma}) = \sum_{i=1}^{2}\sum_{k\in K}\int_{\Omega_k}\left[(1-\zeta_{k,i})a_{k,i} + \zeta_{k,i}b_{k,i}\right]^2 d\mu .$$

We see immediately that

$$\int_{\Omega_k}\left[(1-\zeta_{k,i})a_{k,i} + \zeta_{k,i}b_{k,i}\right]^2 d\mu \leq \int_{\Omega_k} 2(1-\zeta_{k,i})^2 a_{k,i}^2 + 2\zeta_{k,i}^2 b_{k,i}^2 d\mu \leq$$

$$2\int_{\Omega_k}(1-\zeta_{k,i})a_{k,i}^2 + \zeta_{k,i}b_{k,i}^2 d\mu = 2\mu(\Omega_k)\left[P[\Omega_k](1-\zeta_{k,i})a_{k,i}^2 + P[\Omega_k](\zeta_{k,i})b_{k,i}^2\right] .$$

This implies (21b). □

Lemma 4 shows, that $\bar{a}$ is $\mathbf{L}^2(\Omega)$-bounded and $\mathbf{L}^2(\Omega)$-elliptic.

Lemma 4.
Let $\alpha \in W_1^\infty(\Omega)$, $\alpha \geq A > 0$ on Ω and $\bar{a}(\boldsymbol{\sigma},\boldsymbol{\tau}) := \int_\Omega P_h(\alpha)\, \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\, d\mu \quad \forall\ \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathbf{L}^2(\Omega)$, then

$$\bar{a}(\boldsymbol{\sigma},\boldsymbol{\tau}) \leq \|\alpha\|_{L^\infty(\Omega)}\|\boldsymbol{\sigma}\|_{\mathbf{L}^2(\Omega)}\|\boldsymbol{\tau}\|_{\mathbf{L}^2(\Omega)} \quad \forall\ \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathbf{L}^2(\Omega) , \tag{22a}$$

and

$$\bar{a}(\boldsymbol{\tau},\boldsymbol{\tau}) \geq A\|\boldsymbol{\tau}\|_{\mathbf{L}^2(\Omega)}^2 \quad \forall\ \boldsymbol{\tau} \in \mathbf{L}^2(\Omega) . \tag{22b}$$

Proof.
From (2) it follows that,

$$A \leq \frac{\int_{\Omega_k}\alpha\, d\mu}{\mu(\Omega_k)} \leq \|\alpha\|_{L^\infty(\Omega_k)} ,$$

together with the definitions of P and $\bar{a}$ this implies (22a) and (22b). □.

We introduce the minimum mesh width $h_{\min}$ and the maximum mesh width $h_{\max}$,

$$h_{\min} = \min_{k\in K}\ \min_{i=1,2}\ |h_{k,i}| , \tag{23a}$$

$$h_{\max} = \max_{k\in K}\ \max_{i=1,2}\ |h_{k,i}| . \tag{23b}$$

5.1. The properties of $\bar{a}_q$.

We discuss the properties of the quadrature rule $\bar{a}_q$. We assume that $\bar{a}_q = \bar{a}_{h,1}$, where $\bar{a}_{h,1}$ is given by (18a). We discuss the interaction between Π, $\tilde{\Pi}$ and $\bar{a}_q$. We show, that $\bar{a}_q$ is $\mathbf{L}^2(\Omega)$-bounded on V_h, and we also show, that $\bar{a}_q$ is $\mathbf{L}^2(\Omega)$-elliptic on V_h and X_h.

Lemma 5.
If $\boldsymbol{\sigma},\boldsymbol{\tau} \in \Sigma$, then

$$\bar{a}_q(\Pi_h\boldsymbol{\sigma},\Pi_h\boldsymbol{\tau}) = \bar{a}_q(\Pi_h\boldsymbol{\tau},\Pi_h\boldsymbol{\sigma}) = \bar{a}_q(\boldsymbol{\sigma},\Pi_h\boldsymbol{\tau}) = \bar{a}_q(\Pi_h\boldsymbol{\sigma},\boldsymbol{\tau}) = \tag{24a}$$
$$\bar{a}_q(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) = \bar{a}_q(\tilde{\Pi}_h\boldsymbol{\sigma},\boldsymbol{\tau}) = \bar{a}_q(\tilde{\Pi}_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) ,$$

$$\|\alpha\|_{L^\infty(\Omega)} \|\tilde{\Pi}_h\boldsymbol{\sigma}\|_h^2 \geqslant \bar{a}_q(\tilde{\Pi}_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\sigma}) \geqslant ½\bar{a}(\tilde{\Pi}_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\sigma}) \geqslant \frac{A}{2}\|\tilde{\Pi}_h\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 , \tag{24b}$$

$$\bar{a}_q(\Pi_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) \leqslant 6\|\alpha\|_{L^\infty(\Omega)} \|\Pi_h\boldsymbol{\sigma}\|_{L^2(\Omega)} \|\tilde{\Pi}_h\boldsymbol{\tau}\|_h , \tag{24c}$$

$$A\|\tilde{\Pi}_h\boldsymbol{\tau}\|_h^2 \leqslant \bar{a}_q(\boldsymbol{\tau},\tilde{\Pi}_h\boldsymbol{\tau}) \leqslant \|\alpha\|_{L^\infty(\Omega)} \|\tilde{\Pi}_h\boldsymbol{\tau}\|_h^2 . \tag{24d}$$

Proof.
The definitions of Π_h, $\tilde{\Pi}_h$ and $\bar{a}_q$ imply (24a). Inequality (24b) follows immediately from (18a), (18b) and (21b). To prove (24c), we need some auxiliary variables, $a_{k,i} = P[\Gamma_{k,i,0}](\boldsymbol{\sigma})$, $b_{k,i} = P[\Gamma_{k,i,1}](\boldsymbol{\sigma})$, $c_{k,i} = P[\Gamma_{k,i,0}](\boldsymbol{\tau})$ and $d_{k,i} = P[\Gamma_{k,i,1}](\boldsymbol{\tau})$. We use Cauchy-Schwartz twice to obtain

$$\bar{a}_q(\Pi\boldsymbol{\sigma},\tilde{\Pi}\boldsymbol{\tau}) = \sum_{k\in K} P[\Omega_k](\alpha)\mu(\Omega_k)\sum_{i=1}^{2} (P[\Omega_k](1-\zeta_{k,i})a_{k,i}c_{k,i} + P[\Omega_k](\zeta_{k,i})b_{k,i}d_{k,i}) \leqslant$$
$$\sum_{k\in K} P[\Omega_k](\alpha)\mu(\Omega_k)\left[\sum_{i=1}^{2}(a_{k,i}^2 + b_{k,i}^2)\right]^{½}\left[\sum_{i=1}^{2}[P[\Omega_k](1-\zeta_{k,i})^2c_{k,i}^2 + P[\Omega_k](\zeta_{k,i})^2d_{k,i}^2]\right]^{½} .$$

We use

$$P[\Omega_k](f)^2 \leqslant P[\Omega_k](f^2)$$

to rewrite the term in c and d and we use (21a) to replace the term in a and b by $\|\Pi_h\boldsymbol{\sigma}\|_{L^2(\Omega_k)}$,

$$\bar{a}_q(\Pi_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) \leqslant$$
$$\sum_{k\in K} P[\Omega_k](\alpha)\mu(\Omega_k)6\frac{\|\Pi_h\boldsymbol{\sigma}\|_{L^2(\Omega_k)}}{\mu(\Omega_k)^{½}}\left[\sum_{i=1}^{2}(P[\Omega_k]((1-\zeta_{k,i})^2)c_{k,i}^2 + P[\Omega_k]((\zeta_{k,i})^2)d_{k,i}^2)\right]^{½} .$$

We see immediately that this implies,

$$\bar{a}_q(\Pi_h\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) \leqslant$$
$$6\|\alpha\|_{L^\infty(\Omega)} \|\Pi_h\boldsymbol{\sigma}\|_{L^2(\Omega)} \|\tilde{\Pi}_h\boldsymbol{\tau}\|_h .$$

This proves (24c). Inequality (24d) follows immediately from (18).

□

5.2. The difference between $\bar{a}$ and $\bar{a}_q$.

For our error estimates, we need an upper bound for the difference between the value of $a(\boldsymbol{\sigma}_h,\boldsymbol{\tau})$ and that of $\bar{a}_q(\boldsymbol{\sigma}_h,\boldsymbol{\tau})$ for $\boldsymbol{\sigma}_h \in V_h$, $\boldsymbol{\tau} \in H^1(\Omega)$. As we already know from (16c) (see also Lemmas 8 and 9) that,

$$|a(\boldsymbol{\sigma},\boldsymbol{\tau}_h) - \bar{a}(\boldsymbol{\sigma},\boldsymbol{\tau}_h)| \leqslant 2h_{\max}\|\alpha\|_{W_1^\infty(\Omega)}\|\boldsymbol{\sigma}\|_{L^2(\Omega)}\|\boldsymbol{\tau}_h\|_{L^2(\Omega)} ,$$

an estimate for $|\bar{a}(\boldsymbol{\sigma},\boldsymbol{\tau}_h) - \bar{a}_q(\boldsymbol{\sigma},\boldsymbol{\tau}_h)|$ suffices. Such an estimate is derived in lemma 6.

Lemma 6.
Let $\boldsymbol{\tau}_h \in X_h$ and $\boldsymbol{\sigma} \in H^1(\Omega)$, then

$$|\bar{a}(\boldsymbol{\sigma},\boldsymbol{\tau}_h) - \bar{a}_q(\boldsymbol{\sigma},\boldsymbol{\tau}_h)| \leqslant 2\|\alpha\|_{L^\infty(\Omega)}\, h_{\max}\, \|\boldsymbol{\tau}_h\|_h \|\boldsymbol{\sigma}\|_{H^1(\Omega)} . \tag{25}$$

Proof.
To simplify our notation, we introduce $a_{k,i} = P[\Gamma_{k,i,0}](\tau_h)$, $b_{k,i} = P[\Gamma_{k,i,1}](\tau_h)$, $\sigma_{k,i,0} = P[\Gamma_{k,i,0}](\sigma_i)$ and $\sigma_{k,i,1} = P[\Gamma_{k,i,1}](\sigma_i)$. We prove the lemma for $\boldsymbol{\sigma}$ with $\sigma_1,\sigma_2 \in C^1(\bar{\Omega})$, and extend by density.

We consider the difference between the two forms on one subdomain Ω_k with $P[\Omega_k](\alpha) = 1$.

$$\left|\int_{\Omega_k} \sigma\cdot\tau_h \, d\mu - \mu(\Omega_k) \sum_{i=1}^{2} \left[P[\Omega_k](1-\zeta_{k,i})P[\Gamma_{k,i,0}](\sigma_i\tau_{h,i}) + P[\Omega_k](\zeta_{k,i})P[\Gamma_{k,i,1}](\sigma_i\tau_{h,i})\right]\right| =$$

$$\left|\int_{\Omega_k} \sum_{i=1}^{2} \left[(1-\zeta_{k,i})a_{k,i} + \zeta_{k,i}b_{k,i}\right]\sigma_i \, d\mu - \mu(\Omega_k) \sum_{i=1}^{2} \left[P[\Omega_k](1-\zeta_{k,i})P[\Gamma_{k,i,0}](a_{k,i}\sigma_i) + P[\Omega_k](\zeta_{k,i})P[\Gamma_{k,i,1}](b_{k,i}\sigma_i)\right]\right| =$$

$$\left|\int_{\Omega_k} \sum_{i=1}^{2} \left[(1-\zeta_{k,i})a_{k,i}\sigma_i + \zeta_{k,i}b_{k,i}\sigma_i - P[\Omega_k](1-\zeta_{k,i})a_{k,i}\sigma_{k,i,0} - P[\Omega_k](\zeta_{k,i})b_{k,i}\sigma_{k,i,1}\right] d\mu\right| =$$

$$\left|\int_{\Omega_k} \sum_{i=1}^{2} \left[(1-\zeta_{k,i})a_{k,i}(\sigma_i-\sigma_{k,i,0}) + \zeta_{k,i}b_{k,i}(\sigma_i-\sigma_{k,i,1})\right] d\mu\right| ,$$

The application of the Cauchy-Schwartz inequality to this last term and insertion of α yields the following result,

$$|\bar{a}(\sigma,\tau_h) - \bar{a}_q(\sigma,\tau_h)| \leq h_{\max} \|\alpha\|_{L^\infty(\Omega)} \|\tau_h\|_h \left[\sum_{k\in K}\sum_{i=1}^{2}\sum_{j=0}^{1} \|\sigma_i-\sigma_{i,k,j}\|^2_{L^2(\Omega_k)}\right]^{1/2} .$$

If we take $s \equiv j$ in (16b) then this implies,

$$|\bar{a}(\sigma,\tau_h) - \bar{a}_q(\sigma,\tau_h)| \leq \|\alpha\|_{L^\infty(\Omega)} \|\tau_h\|_h \left[\sum_{i=1}^{2} 4h^2_{\max} \|\mathbf{grad}\,\sigma_i\|^2_{L^2(\Omega)}\right]^{1/2} \leq 2h_{\max}\|\alpha\|_{L^\infty(\Omega)}\|\tau_h\|_h\|\sigma\|_{H^1(\Omega)} .$$

Because $C^1(\bar{\Omega})$ is dense in $H^1(\Omega)$, the formula also holds for $\sigma_1,\sigma_2 \in H^1(\Omega)$.

□

6 The error estimates.

We use the standard estimates for $\|\sigma-\Pi_h\sigma\|_{L^2(\Omega)}$ and $\|u-P_hu\|_{L^2(\Omega)}$, as described in section 3.4, to reduce the problem to deriving bounds for $\|P_hu-u_h\|_{L^2(\Omega)}$ and $\|\Pi_h\sigma-\sigma_h\|_{L^2(\Omega)}$. We discuss two possible derivations of an $\mathcal{O}(h)$ error bound. The first derivation needs the assumption, that $h_{\max}$ is "small enough", the second derivation places a condition on an approximation of the discrete version of the adjoint problem.

6.1. Errors due to approximation of the bilinear forms.

As preparation for the derivation of a priori error estimates, we derive some upper bounds on the errors caused by the piecewise constant approximation of the coefficients α,β and γ. We use the following well-known notation. If V and W are normed linear spaces, then $\mathcal{L}(V,W;\mathbb{R})$ is the space of bounded bilinear forms on V and W. the standard norm of an element $b \in \mathcal{L}(V,W;\mathbb{R})$ is given by

$$\|b\|_{\mathcal{L}(V,W;\mathbb{R})} = \sup_{v\in V}\sup_{w\in W} \frac{|b(v,w)|}{\|v\|_V\|w\|_W} .$$

Lemma 7.
If $\alpha \in W^\infty_1(\Omega)$ then

$$\|a-\bar{a}_q\|_{\mathcal{L}(H^1(\Omega),(X_h,\|.\|_h);\mathbb{R})} \leq 6h_{\max}\|\alpha\|_{W^\infty_1(\Omega)} ,$$

where $(X_h, \|.\|_h)$ is a normed linear space with as elements the elements of X_h but with $\|.\|_h$ as norm.
Proof.
From equation (16c) and (21b) it follows that,

$$|a(\sigma,\tau_h) - \bar{a}(\sigma,\tau_h)| \leq 4h_{\max}\|\alpha\|_{W^\infty_1(\Omega)}\|\sigma\|_{L^2(\Omega)}\|\tau_h\|_h .$$

When combined with lemma 6, this implies

$$\| a - \bar{a}_q \|_{\mathscr{L}(\mathrm{H}^1(\Omega),(X_h, \|\cdot\|_h);\mathbf{R})} \leqslant 6h_{\max} \| \alpha \|_{\mathrm{W}_1^\infty(\Omega)} .$$

□

Lemma 8.
If $\boldsymbol{\beta} \in \mathrm{W}_1^\infty(\Omega)$ then

$$\| b - \bar{b} \|_{\mathscr{L}(\mathrm{L}^2(\Omega),\, \mathrm{L}^2(\Omega);\mathbf{R})} \leqslant 4h_{\max} \| \boldsymbol{\beta} \|_{\mathrm{W}_1^\infty(\Omega)} .$$

Proof.
This follows immediately from (16c).

□

Lemma 9.
If $\gamma \in \mathrm{W}_1^\infty(\Omega)$ then

$$\| c - \bar{c} \|_{\mathscr{L}(\mathrm{L}^2(\Omega),\, \mathrm{L}^2(\Omega);\mathbf{R})} \leqslant 2h_{\max} \| \gamma \|_{\mathrm{W}_1^\infty(\Omega)} .$$

Proof.
This follows immediately from (16c).

□

6.2. An a priori error estimate.

The following two lemmas show nice properties of our discretisation. We need these properties to derive the error bound.

Lemma 10.
Let $\boldsymbol{\tau} \in \Sigma$, $t \in \mathrm{L}^2(\Omega)$, then

$$\bar{b}(\tilde{\Pi}_h \boldsymbol{\tau}, t - P_h t) - (\operatorname{div} \tilde{\Pi}_h \boldsymbol{\tau}, t - P_h t) = 0 . \tag{26}$$

Proof.
A straightforward calculation shows that $\mathbf{P}_h(\boldsymbol{\beta}) \cdot \tilde{\Pi}_h \boldsymbol{\tau} - \operatorname{div} \tilde{\Pi}_h \boldsymbol{\tau}$ is constant on Ω_k. From this (26) easily follows. □

Lemma 11.
If $(\boldsymbol{\sigma}, u)$ is a solution of (6) and $(\boldsymbol{\sigma}_h, u_h)$ is a solution of (19), then

$$(\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), P_h t) + c(u - u_h, P_h t) = 0 \quad \forall \; t \in \mathrm{L}^2(\Omega) . \tag{27}$$

Proof.
We take (19b),

$$(\operatorname{div} \boldsymbol{\sigma}_h, P_h t) + \bar{c}(u_h, P_h t) = (f, P_h t) ,$$

$\bar{c}$ is derived by orthogonal $\mathrm{L}^2(\Omega_k)$ projection, so this implies

$$(\operatorname{div} \boldsymbol{\sigma}_h, P_h t) + c(u_h, P_h t) = (f, P_h t) .$$

If we subtract this from (6b), $(\operatorname{div} \boldsymbol{\sigma}, P_h t) + c(u, P_h t) = (f, P_h t)$, then we find (27). □

We are now ready to give an estimate for $\| \Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_h$.

Theorem 1.
If $(\boldsymbol{\sigma},u)$ is the solution of (6), $(\boldsymbol{\sigma}_h,u_h)$ is the solution of (19) and $(\boldsymbol{\sigma},u) \in \mathrm{H}^1(\Omega)\times \mathrm{H}^2(\Omega)$, then there exist positive real numbers C and D such that

$$C < \frac{12}{A}\max(1, \|\alpha\|_{W_1^\infty(\Omega)}, \|\boldsymbol{\beta}\|_{W_1^\infty(\Omega)}, \|\gamma\|_{W_1^\infty(\Omega)}) \max(1, \|\boldsymbol{\sigma}\|_{H^1(\Omega)}, \|u\|_{L^2(\Omega)}) , \tag{28}$$

$$D < 2\frac{\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}}{A} \ , \ :$$

$$\|\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_h^2 \leqslant Ch_{\max}(\|\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_h + \|P_h u - u_h\|_{L^2(\Omega)}) + D\|\Pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_h \|P_h u - u_h\|_{L^2(\Omega)} .$$

Proof.
According to (24d), $A\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_h^2 \leqslant \bar{a}_q(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h))$. This is the starting point for the derivation of our error bound. Equations (6a) and (19a) imply, that

$$\bar{a}_q(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) = (\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + a(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) - \bar{a}_q(\boldsymbol{\sigma}_h,\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) =$$
$$(\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + (\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) - b(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) +$$
$$< g,\mathbf{n}_{\partial\Omega}\cdot\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) > + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u_h) - (\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u_h) - < g,\mathbf{n}_{\partial\Omega}\cdot\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) > =$$
$$(\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + (\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u)_{L^2(\Omega)} - (b-\bar{b})(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) - \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) +$$
$$\bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u_h) - (\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u_h) .$$

Where we give $b-\bar{b}$, $\bar{a}_q-a$ etc. their obvious meaning. If we use lemma 10, we find:

$$A\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_h^2 \leqslant (\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) - (b-\bar{b})(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) +$$
$$(\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h u-u_h)_{L^2(\Omega)} - \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h u-u_h) .$$

If we use (20b) and (20c) to prepare the way, then the application of lemma 11 to this expression results in:

$$A\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_{L^2(\Omega)}^2 \leqslant (\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) - (b-\bar{b})(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) -$$
$$c(u-u_h,P_h u-u_h) - \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h u-u_h) .$$

As γ is non-negative according to (5), we may add $c(P_h u-u_h,P_h u-u_h)$ on both sides of the inequality, we find,

$$A\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_{L^2(\Omega)}^2 + c(P_h u-u_h,P_h u-u_h) \leqslant$$
$$(\bar{a}_q-a)(\boldsymbol{\sigma},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) - (b-\bar{b})(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),u) -$$
$$(c-\bar{c})(u-P_h u,P_h u-u_h) - \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),Phu-u_h) .$$

We use lemmas 7, 8 and 9 to reduce this to,

$$A\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_h^2 \leqslant$$
$$h_{\max}\left[6\|\alpha\|_{W_1^\infty(\Omega)}\|\boldsymbol{\sigma}\|_{H^1(\Omega)} + 4\|\boldsymbol{\beta}\|_{W_1^\infty(\Omega)}\|u\|_{L^2(\Omega)}\right]\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_h +$$
$$2h_{\max}\|\gamma\|_{W_1^\infty(\Omega)}\|u-P_h u\|_{L^2(\Omega)}\|P_h u-u_h\|_{L^2(\Omega)} +$$
$$2\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_{L^2(\Omega)}\|Pu-u_h\|_{L^2(\Omega)} .$$

Note that for all $u \in \mathrm{L}^2(\Omega)$, $\|u-P_h u\|_{L^2(\Omega)} \leqslant \|u\|_{L^2(\Omega)}$ and $\|\Pi_h\boldsymbol{\sigma}\|_h = \|\tilde{\Pi}_h\boldsymbol{\sigma}\|_h$.

□

Next, we prepare for the second part of our error estimate.

Lemma 12.

If $(\boldsymbol{\sigma},u)$ is the solution of (6), $(\boldsymbol{\sigma}_h,u_h)$ is a solution of (19) and $(\boldsymbol{\tau},q)$ is the solution of the adjoint problem for an arbitrary right hand side $p \in L^2(\Omega)$, then

$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) =$$
$$a(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) - \bar{a}_q(\boldsymbol{\sigma}_h,\tilde{\Pi}_h\boldsymbol{\tau}) + (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h)\,.$$

Proof.
We start by replacing b by $\bar{b}$,

$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) =$$
$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - \bar{b}(\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h)\,.$$

We use (20a) and (20c) to get,

$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) =$$
$$(\operatorname{div}\tilde{\Pi}_h\boldsymbol{\tau},P_h u-u_h) - \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau},P_h u-u_h) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h)\,.$$

We use lemma 10 to find,

$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) =$$
$$(\operatorname{div}\tilde{\Pi}_h\boldsymbol{\tau},u-u_h) - \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau},u-u_h) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h)\,.$$

We use equation (6a) and equation (19a),

$$(\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) =$$
$$a(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) - \langle g,\tilde{\Pi}_h\boldsymbol{\tau}\cdot\mathbf{n}_{\partial\Omega}\rangle - \bar{a}_q(\boldsymbol{\sigma}_h,\tilde{\Pi}_h\boldsymbol{\tau}) +$$
$$\langle g,\tilde{\Pi}_h\boldsymbol{\tau}\cdot\mathbf{n}_{\partial\Omega}\rangle + (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) +$$
$$\bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h)\,.$$

□

Lemma 13.
If $(\boldsymbol{\sigma},u)$ is the solution of (6), $(\boldsymbol{\sigma}_h,u_h)$ is a solution of (19) and $(\boldsymbol{\tau},w)$ is the solution of the adjoint problem for an arbitrary right hand side $p \in L^2(\Omega)$, then

$$c(P_h w,u-u_h) = -a(\boldsymbol{\tau},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w)\,.$$

Proof.
According to lemma 11,

$$c(P_h w,u-u_h) = -(\operatorname{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h w)\,,$$

according to (20b) and (20c) we can rewrite the right hand side,

$$c(P_h w,u-u_h) = -(\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h w)\,.$$

We wish to use equation (26) from lemma 10 to remove P_h. To do this we must add and subtract a term $\bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),P_h w)$ on the right hand side of our equation. We apply lemma 10 and gather terms in $\bar{b}$ together,

$$c(P_h w,u-u_h) = -(\operatorname{div}\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w)\,.$$

Finally, we use (11a),

$$c(P_h w,u-u_h) = -a(\boldsymbol{\tau},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w)\,.$$

□

Theorem 2.
Assume the adjoint problem (11) has a unique solution for all square integrable right hand sides and assume that there is a constant C_r such that, if $(\boldsymbol{\tau},w)$ is the solution of (11) for a given right hand side f, then

$$\|\boldsymbol{\tau}\|_{H^1(\Omega)} + \|w\|_{H^1(\Omega)} \leqslant C_r \|f\|_{L^2(\Omega)} .$$

Now, if $(\sigma,u) \in H^1(\Omega)\times H^2(\Omega)$ is the solution of (6), and $(\boldsymbol{\sigma}_h,u_h)$ is a solution of (19) then there are constants

$$0 < C,D,E < 4C_r(1+2h_{\max})\max(\|\alpha\|_{W_1^\infty(\Omega)}, \|\boldsymbol{\beta}\|_{W_1^\infty(\Omega)}, \|\gamma\|_{W_1^\infty(\Omega)}) ,$$

such that

$$\|P_h u-u_h\|_{L^2(\Omega)} \leqslant Ch_{\max}(\|u\|_{L^2(\Omega)} + \|\boldsymbol{\sigma}_{H^1(\Omega)}) + Dh_{\max}\|\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_h + Eh_{\max}\|P_h u-u_h\|_{L^2(\Omega)} .$$

Proof.
If we have an estimate for $(P_h u-u_h,p)$ for all $p \in L^2(\Omega)$, then we can use

$$\|t\|_{L^2(\Omega)} = \sup_{p \in L^2(\Omega), p\neq 0} \frac{(p,t)}{\|p\|_{L^2(\Omega)}} ,$$

to find $\|P_h u - u_h\|_{L^2(\Omega)}$. We use the regularity of the adjoint problem (11) to find a solution $(\boldsymbol{\tau},w) \in H^1(\Omega)\times L^2(\Omega)$ of (11) for a given right hand side $p \in L^2(\Omega)$. We may write,

$$(p,P_h u-u_h) = (\operatorname{div}\boldsymbol{\tau},P_h u-u_h) - b(\boldsymbol{\tau},P_h u-u_h) + c(w,P_h u-u_h) .$$

If we apply lemma 12, we find,

$$\begin{gathered}(p,P_h u-u_h) = \\ a(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) - \bar{a}_q(\boldsymbol{\sigma}_h,\tilde{\Pi}_h\boldsymbol{\tau}) + \\ (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h) + c(w-P_h w,P_h u-u_h) + c(P_h w,P_h u-u_h) .\end{gathered}$$

We use lemma 13,

$$\begin{gathered}(p,P_h u-u_h) = \\ a(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) - \bar{a}_q(\boldsymbol{\sigma}_h,\tilde{\Pi}_h\boldsymbol{\tau}) + \\ (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h) + \\ c(w-P_h w,P_h u-u_h) - a(\boldsymbol{\tau},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w) .\end{gathered}$$

We can write this as follows,

$$\begin{gathered}(p,P_h u-u_h) = \\ (a-\bar{a}_q)(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) + \bar{a}_q(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\tilde{\Pi}_h\boldsymbol{\tau}) + \\ (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h) + \\ c(w-P_h w,P_h u-u_h) - a(\boldsymbol{\tau},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w) .\end{gathered}$$

We use (24a) to write this as,

$$\begin{gathered}(p,P_h u-u_h) = \\ (a-\bar{a}_q)(\boldsymbol{\sigma},\tilde{\Pi}_h\boldsymbol{\tau}) - (a-\bar{a}_q)(\boldsymbol{\tau},\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)) + \\ (b-\bar{b})(\tilde{\Pi}_h\boldsymbol{\tau},u) + \bar{b}(\tilde{\Pi}_h\boldsymbol{\tau}-\boldsymbol{\tau},P_h u-u_h) + (\bar{b}-b)(\boldsymbol{\tau},P_h u-u_h) + \\ c(w-P_h w,P_h u-u_h) + \bar{b}(\tilde{\Pi}_h(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),w-P_h w) .\end{gathered}$$

We can use the regularity of the adjoint problem (11), lemma 7, 8 and 9 and the projection error estimates (16a,b,c), to obtain

$$\|P_h u-u_h\|_{L^2(\Omega)} \leqslant$$

$$C_r(1+2h_{\max})2h_{\max}\,\|\alpha\|_{W_1^\infty(\Omega)}\left[\|\sigma\|_{H^1(\Omega)} + \|\tilde{\Pi}_h(\sigma-\sigma_h)\|_h\right] +$$
$$4C_rh_{\max}\,\|\beta\|_{W_1^\infty(\Omega)}(1+2h_{\max})\|u\|_{L^2(\Omega)} + 2C_rh_{\max}\,\|\beta\|_{L^\infty(\Omega)}\|P_hu-u_h\|_{L^2(\Omega)} +$$
$$2C_rh_{\max}\left[h_{\max}\,\|\gamma\|_{W_1^\infty(\Omega)}\|P_hu-u_h\|_{L^2(\Omega)} + \|\beta\|_{L^\infty(\Omega)}\|\tilde{\Pi}_h(\sigma-\sigma_h)\|_{L^2(\Omega)}\right].$$

This can be written as,

$$\|P_hu-u_h\|_{L^2(\Omega)} \leqslant$$
$$\tilde{C}h_{\max}(1+h_{\max}) + \tilde{D}nh_{\max}(1+h_{\max})\|\tilde{\Pi}_h(\sigma-\sigma_h)\|_h + \tilde{E}h_{\max}(1+h_{\max})\|P_hu-u_h\|_{L^2(\Omega)}.$$

□

If $h_{\max}$ is small enough, theorem 1 and theorem 2 together give an $\mathcal{O}(h_{\max})$ error estimate. An important limit on $h_{\max}$ is implied by the form of the estimates in theorem 1 and 2. The main problem is, that large values of $\|\alpha\|_{W_1^\infty(\Omega)}$, $\|\beta\|_{W_1^\infty(\Omega)}$ and $\|\gamma\|_{W_1^\infty(\Omega)}$ decrease the range of $h_{\max}$ for which the estimate is valid. This problem can be avoided if we make an extra assumption. We discuss this in the next section.

6.3. A different approach.

To improve our estimate of $\|P_hu - u_h\|_{L^2(\Omega)}$, we consider the adjoint of the discrete problem. This means, that we look for $(\tau_h,v_h) \in X_h \times W_h$, such that

$$\bar{a}_q(\tau_h,\sigma_h) - (\operatorname{div}\sigma_h,v_h) = 0 \quad \forall\ \sigma_h \in V_h\,, \tag{29a}$$

$$(\operatorname{div}\tau_h,t_h) - \bar{b}(\tau_h,t_h) + \bar{c}(v_h,t_h) = (f,t_h) \quad \forall\ t_h \in W_h\,. \tag{29b}$$

We call this system regular, if there is at least one solution for each $f \in P_h(L^2(\Omega))$, and that all solutions for a particular f satisfy

$$\|\tau_h\|_h + \|v_h\|_{L^2(\Omega)} \leqslant C\,\|P_hf\|_{L^2(\Omega)}\,, \tag{29c}$$

with C independent of the mesh size. This is a somewhat less stringent regularity condition than that given for the continuous adjoint problem (10). Note, that $\tau_h \in X_h$, so $\tau_{h,i}$ is a piecewise exponential function on Ω_k for $i=1,2$.

An example of a general condition under which this system is regular is the following:

$$\alpha \geqslant A > 0\,,\ \gamma \geqslant C_0 > 0 \ \text{and}\ AC_0 - \|\beta\|^2_{L^\infty(\Omega)} \geqslant C_1 > 0\,. \tag{30}$$

To show this, we need the following relations,

$$\int_\Omega \frac{P_h(\alpha)}{4}\tau_h\cdot\tau_h - \mathbf{P}_h(\beta)\cdot\tau_hv_h + P_h(\gamma)v_hv_h\,d\mu = \tag{31}$$

$$\int_\Omega \frac{P_h(\alpha)}{4}\left[\tau_h - \frac{2\mathbf{P}_h(\beta)}{P_h(\alpha)}v_h\right]^2 + \left[P_h(\gamma) - \frac{\mathbf{P}_h(\beta)^2}{P_h(\alpha)}\right]v_hv_h\,d\mu \geqslant \tag{31a}$$

$$\int_\Omega P_h(\gamma)\left[v_h - \frac{\mathbf{P}_h(\beta)\cdot\tau_h}{2P_h(\gamma)}\right]^2 + \left[P_h(\alpha) - \frac{\mathbf{P}_h(\beta)^2}{P_h(\gamma)}\right]\frac{\tau_h\cdot\tau_h}{4}\,d\mu\,. \tag{31b}$$

We know, that $(\operatorname{div}\tilde{\Pi}_h\sigma,P_ht) = (\operatorname{div}\Pi_h\sigma,P_ht)$, so, if we take the sum of (29a) and (29b) with $\sigma = \Pi_h\tau_h$ and $t = v_h$, we find

$$\bar{a}_q(\tau_h,\Pi_h\tau_h) - \bar{b}(\tau_h,v_h) + \bar{c}(v_h,v_h) = (f,v_h)\,. \tag{32}$$

According to (24a), $\bar{a}_q(\tau_h,\Pi_h\tau_h) = \bar{a}_q(\tilde{\Pi}_h\tau_h,\tilde{\Pi}_h\tau_h)$, and by (24b) we have

$$\frac{1}{4}\bar{a}(\tilde{\Pi}_h\sigma,\tilde{\Pi}_h\sigma) \leqslant \bar{a}_q(\tilde{\Pi}_h\sigma,\tilde{\Pi}_h\sigma)\,.$$

Hence

$$\int_\Omega \frac{P_h(\alpha)}{4}\tau_h\cdot\tau_h - \mathbf{P}_h(\boldsymbol\beta)\cdot\tau_h v_h + P_h(\gamma)v_h v_h \, d\mu \leqslant \int_\Omega P_h(f) v_h \, d\mu \,. \tag{33}$$

This expression is identical to (31), so (31a) is smaller than (f,v_h), combined with (30) this implies

$$\frac{C_1}{A}\|v_h\|_{L^2(\Omega)} \leqslant \|f\|_{L^2(\Omega)} \,. \tag{34a}$$

In the same way, we find, that (31b) is smaller than (f,v_h), together with (30) and (34a), this implies

$$\frac{C_1}{(AC_0)^{1/2}}\|\tau_h\|_{L^2(\Omega)} \leqslant \|f\|_{L^2(\Omega)} \,. \tag{34b}$$

From (32) we see that this implies,

$$A\|\tau_h\|_h^2 \leqslant a_q(\tau_h,\tau_h) \leqslant$$
$$\|f\|_{L^2(\Omega)}\|v_h\|_{L^2(\Omega)} + \|\boldsymbol\beta\|_{L^\infty(\Omega)}\|\tau_h\|_{L^2(\Omega)}\|v_h\|_{L^2(\Omega)} + \|\gamma\|_{L^\infty(\Omega)}\|v_h\|^2_{L^2(\Omega)} \,,$$

this implies that there is a C such that

$$\|\tau_h\|_h \leqslant C\|f\|_{L^2(\Omega)} \,.$$

Theorem 3.
If we assume, that (29c) holds, then

$$\|P_h u - u_h\|_{L^2(\Omega)} \leqslant \tag{35}$$
$$h_{\max}\left[6\|\alpha\|_{W_1^\infty(\Omega)}\|\boldsymbol\sigma\|_{H^1(\Omega)} + + 2(\|\boldsymbol\beta\|_{W_1^\infty(\Omega)} + \|\gamma\|_{W_1^\infty(\Omega)})\|u\|_{L^2(\Omega)}\right] .$$

Proof.
We use (29b),

$$(P_h u - u_h, P_h f) = (\operatorname{div}\tau_h, P_h u - u_h) - \bar b(\tau_h, P_h u - u_h) + \bar c(P_h u - u_h, v_h) \,.$$

Hence, according to lemma 10 and the definition of $\bar c$,

$$(P_h u - u_h, P_h f) = (\operatorname{div}\tau_h, u - u_h) - \bar b(\tau_h, u - u_h) + \bar c(u - u_h, v_h) \,.$$

We use (6a) and (19a) to find

$$(P_h u - u_h, P_h f) = (\operatorname{div}\tau_h, u - u_h) - (\bar b - b)(\tau_h, u) - b(\tau_h, u) + \bar b(\tau_h, u_h) + \bar c(u - u_h, v_h) =$$
$$(\bar b - b)(\tau_h, u) + a(\boldsymbol\sigma, \tau_h) - \bar a_q(\boldsymbol\sigma_h, \tau_h) + (\bar c - c)(u, v_h) + c(u - u_h, v_h) \,.$$

According to (24a) and lemma 11, this implies

$$(P_h u - u_h, P_h f) =$$
$$(\bar b - b)(\tau_h, u) + (a - \bar a_q)(\boldsymbol\sigma, \tau_h) + \bar a_q(\Pi_h\boldsymbol\sigma - \boldsymbol\sigma_h, \tau_h) + (\bar c - c)(u, v_h) - (\operatorname{div}(\Pi_h\boldsymbol\sigma - \boldsymbol\sigma_h), v_h) \,.$$

Now, (29a) implies,

$$(P_h u - u_h, P_h f) = (\bar b - b)(\tau_h, u) + (a - \bar a_q)(\boldsymbol\sigma, \tau_h) + (\bar c - c)(u, v_h) \,.$$

Finally, we use lemma 7, 8, and 9 and (29c) to obtain our error estimate (35). □

7 A verification of the local maximum principle.

We use the discrete adjoint problem to show that, for this quadrature rule, the matrix after elimination of $\boldsymbol\sigma$ by static condensation is an M-matrix. The discrete adjoint problem is defined in (29).

We assume a regular uniform mesh. We denote the matrix corresponding to (29), after elimination of $\boldsymbol\sigma_h$, by A. We see, that the matrix A has non-positive off-diagonal elements. We shall show, that A is an M-matrix. To do this, we use theorem 5.12, chapter 5, page 124 of [23]. This theorem states, that, for irreducible matrices with non-positive off-diagonal elements, the M-matrix property is equivalent to the existence of a positive vector with a non-negative image, that is not identically zero. In our case, the vector $(1, 1, \ldots, 1)^T$ has a such an image, because all row sums are non-negative, and any row corresponding to an edge or corner has a positive row-sum.

The fact, that the matrix A is irreducible follows from theorem 3.6, [23] which states that, for a square matrix, irreducibility is equivalent to its di-graph being strongly connected. Inspection shows, that the di-graph of the matrix under consideration is indeed strongly connected.

According to theorem 5.6[23], A^T is an M-matrix too. This implies, that the discrete equations for the original u_h satisfy a local maximum principle.

The M-matrix property implies that the system for u_h has a unique solution. From the form of the equations for $\boldsymbol{\sigma}_h$, we see that a given u_h induces a unique $\boldsymbol{\sigma}_h$. this implies that our system is always uniquely solvable. A quick calculation of the coefficients of u_h in (19a) shows that, for constant coefficients and large $\boldsymbol{\beta}$, i.e. with large convection diffusion ratios, we get a relation between σ_h and u_h where the "upwind" point is weighed more heavily. If $\boldsymbol{\beta}/\alpha$ remains bounded and we go to the limit $|\beta_1| + |\beta_2| \rightarrow \infty$ then we get a first order upwind scheme. This suggests that the scheme, in which the coefficients are continuously dependent on this ratio, remains useful close to such a limit.

8 An a-posteriori estimator.

We use a special quadrature rule and obtain a higher order discretisation. We seek an $\overline{a}_{h,3}(.,.)$. that minimises $\overline{a} - \overline{a}_{h,3}$. To do this, we choose a special quadrature rule for each $\overline{a}(.,\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ is one of the basis functions introduced earlier. Due to the nature of our test functions, the quadrature rule is essentially a one-dimensional rule.

8.1. The derivation of the quadrature rule.

For $\boldsymbol{\eta}_{i,j+\frac{1}{2}}$ we proceed as follows. We replace the two dimensional integral by a repeated integral, we integrate exactly in the $\mathbf{e}_2$ direction and then use a three point rule to approximate the remaining integral. As nodes for the last integration we take either the centres of $\Gamma_{i-\frac{1}{2},j+\frac{1}{2},0}$, $\Gamma_{i+\frac{1}{2},j+\frac{1}{2},0}$ and $\Gamma_{i+\frac{1}{2},j+\frac{1}{2},1}$. Or, if we are at a boundary, the edge centre on the boundary and the two next closest edge centres. We choose the weights as follows,

$$\overline{a}_{h,3}(\Pi_h\boldsymbol{\sigma},\boldsymbol{\eta}_{i,j+\frac{1}{2}}) = \overline{a}(\boldsymbol{\sigma},\boldsymbol{\eta}_{i,j+\frac{1}{2}}) \, ,$$

for all $\boldsymbol{\sigma}$ with x_1-components that are second order polynomials in x_1. I.e. for all $p,q,r \in \mathbb{R}$, and all $\boldsymbol{\eta}_{i,j+\frac{1}{2}}$, we have

$$\overline{a}_{h,3}(\Pi_h((px_1^2+qx_1+r)\mathbf{e}_1),\boldsymbol{\eta}_{i,j+\frac{1}{2}}) = \overline{a}((px_1^2+qx_1+r)\mathbf{e}_1,\boldsymbol{\eta}_{i,j+\frac{1}{2}}) \, ,$$

In a similar manner, we define the rule for $\boldsymbol{\eta}_{i+\frac{1}{2},j}$.

8.2. An estimator for the local discretisation error and a lower bound for the global error.

If we assume that $c=\overline{c}$, $b=\overline{b}$ and $a=\overline{a}$, then we can use this rule to obtain an a-posteriori estimator for the local discretisation error and a lower bound for the global error as follows. It is immediately obvious, that

$$\overline{a}_{h,3}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) - \overline{a}_{h,1}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) \geqslant \mathcal{O}(h_{\max}^2) \, ,$$

where r is a possible index-tuple. Moreover,

$$\overline{a}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) - \overline{a}_{h,3}(\boldsymbol{\sigma},\boldsymbol{\eta}_r) = \mathcal{O}(h_{\max}^3) \, ,$$

if $\boldsymbol{\sigma}$ is smooth enough. If

$$|\overline{a}_{h,3}(\rho_h,\boldsymbol{\eta}_r) - (\operatorname{div}\boldsymbol{\eta}_r - \boldsymbol{\beta}\cdot\boldsymbol{\eta}_r, w_h)| \geqslant K \, ,$$

then we have either

$$\| w_h \|_{L^\infty(\Omega)} \geqslant C_1 K \, ,$$

or

$$\| \boldsymbol{\rho}_h \|_{L^\infty(\Omega)} \geqslant C_2 K ,$$

We see immediately that, if $(\boldsymbol{\sigma}_h, u_h)$ is the solution of (19) with $\bar{a}_q = \bar{a}_{h,1}$ then

$$\bar{a}_{h,1}(\Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) - ((\operatorname{div} - \boldsymbol{\beta})\boldsymbol{\eta}_r, P_h u - u_h) = \mathcal{O}(h^k) ,$$

with $k = 1$ or 2 depending on the coefficients in (1) and

$$\bar{a}_{h,3}(\Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) - ((\operatorname{div} - \boldsymbol{\beta})\boldsymbol{\eta}_r, P_h u - u_h) = \mathcal{O}(h_{\max}^{k+2}) + \bar{a}_{h,1}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) - \bar{a}_{h,3}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) .$$

So, $(a_{h,1} - a_{h,3})(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r)$ is an estimate for the local discretisation error. Moreover this implies, that there is a constant C such that

$$\| \Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{L^\infty(\Omega)} + \| P_h u - u_h \|_{L^\infty(\Omega)} \geqslant C \, | \bar{a}_{h,1}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) - \bar{a}_{h,3}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) | + \mathcal{O}(h_{\max}^{k+2}) .$$

If we assume that

$$\| \Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{L^\infty(\Omega)} + \| P_h u - u_h \|_{L^\infty(\Omega)} = \mathcal{O}(h_{\max}^{k}) ,$$

we see that, for $h_{\max}$ small enough,

$$\| \Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{L^\infty(\Omega)} + \| P_h u - u_h \|_{L^\infty(\Omega)} \geqslant \tfrac{1}{2} C \, | \bar{a}_{h,1}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) - \bar{a}_{h,3}(\boldsymbol{\sigma}_h, \boldsymbol{\eta}_r) | .$$

This provides a lower bound on the global discretisation error. We expect the solution for $\bar{a}_{h,3}$ to be two orders of magnitude more accurate than the solution for $\bar{a}_{h,1}$.

9 Numerical results.

We consider problem (1) with

$$u = \tanh(8(x_1 - x_2)) ,$$
$$\alpha = 100 , \ \beta_1 = \beta_2 = 100 ,$$
$$\Gamma_1 = \partial\Omega , \ g = u|_{\partial\Omega} ,$$
$$f = -\frac{\operatorname{div}(\operatorname{grad} u + u\boldsymbol{\beta})}{\alpha} .$$

We find the following results for the two discretisations. The two components of the error vectors for the fluxes were identical up to the accuracy given. We use the 2-norm as norm for the error vectors,

$$\| v \| = \left\{ \frac{1}{|I|} \sum_{i \in I} v_i^2 \right\} ,$$

where $|I|$ is the number of elements in the index set.

the $\log_2$ of the errors for $\bar{a}_q = \bar{a}_{h,1}$,		
meshwidth	$\log_2 \| P_h u - u_h \|$	$\log_2 \| (\Pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{e}_1 \|$
1 / 4	-1.5	-1.3
1 / 8	-1.9	-1.9
1 / 16	-2.6	-2.6
1 / 32	-3.8	-3.7
1 / 64	-5.4	-5.4

the $\log_2$ of the errors for $\bar{a}_q = \bar{a}_{h,3}$,		
meshwidth	$\log_2 \| P_h u - u_h \|$	$\log_2 \| (\Pi_h \sigma - \sigma_h) \cdot e_1 \|$
1 / 4	-3.0	-3.1
1 / 8	-3.9	-4.5
1 / 16	-6.0	-6.7
1 / 32	-8.5	-10.0
1 / 64	-10.9	-13.7

We see that the order of convergence is indeed higher for the second method. we also see that the difference in order for the fluxes approaches 2. Deviations from the expected order may be caused by the steepness of the solution relative to the mesh.

10 Conclusions.

The Petrov Galerkin mixed finite element method with exponentially fitted test functions for the fluxes has several nice properties. For instance, just as for a finite volume method, if the true solution σ is divergence-free, then the same holds for σ_h. Furthermore we have a formal a-priori error estimate, and after elimination of σ_h by static condensation the two dimensional discretisation results in an M-matrix for u_h. We can extend the method to three dimensions without additional difficulties. Section 9 suggest that the scheme with the three point quadrature rule $\bar{a}_{h,3}$ can serve as a source for a-posteriori error estimates. To judge the effectiveness of the method for singularly perturbed problems is very difficult. However the fact that it incorporates exponential fitting, copes well with the exponential solution of the constant coefficient case and approaches a two-dimensional upwind scheme if the convection goes to infinity suggests the method based on $\bar{a}_{h,1}$ can be applied to such problems.

References

1. D. L. Scharfetter and H. K. Gummel, "Large-Signal Analysis of a Silicon Read Diode Oscillator," *IEEE Transactions on Electron Devices*, vol. ED-16, no. 1, pp. 64-77, 1969.
2. Siegfried Selberherr, *Analysis and simulation of semiconductor devices,* Springer-verlag, Wien New York, 1984.
3. Peter A. Markowich, *The Stationary Semiconductor Device Equations,* Springer-Verlag, Wien New York, 1986.
4. Wolfgang Fichtner, Donald J. Rose, and Randolph E. Bank, "Numerical methods for semiconductor device simulation," *SIAM journal of scientific and statistical computing*, vol. 4, no. 3, pp. 416-435, 1983.
5. F. Brezzi, L. D. Marini, and P. Pietra, "Mixed exponential fitting schemes for current continuity equations," in *Proceedings of the sixth international NASECODE conference*, ed. J. J. H. Miller, Boole Press Ltd, 1989.
6. P. W. Hemker, *A Numerical Study of Stiff Two-Point Boundary Problems,* Mathematical Centre Tracts, 80, Mathematical Centre, Amsterdam, 1977.
7. J. Douglas, Jr. and J. E. Roberts, "Global estimates for mixed methods for second order elliptic equations," *Mathematics of computation*, vol. 44, no. 169, pp. 39-52, 1985.
8. Robert A. Adams, *Sobolev Spaces,* Academic Press, 1975.
9. V. Girault and P. Raviart, *Finite Element Methods for Navier-Stokes Equations,* Springer series in computational mathematics, 5, Springer-Verlag, 1986.
10. S. J. Polak, C. den Heijer, H. A. Schilders, and P. Markowich, "Semiconductor device modelling from the numerical point of view," *International Journal for Numerical Methods in Engineering*, vol. 24, pp. 763-838, 1987.

11. Jean E. Roberts and Jean-Marie Thomas, "Mixed and Hybrid Finite Element Methods," RR 737, INRIA, Rocquencourt, October 1987.

12. P. A. Raviart and J. M. Thomas, "A mixed finite element method for 2-nd order elliptic problems," in *Mathematical aspects of the finite element method*, Lecture Notes in Mathematics, vol. 606, pp. 292-315, Springer, 1977.

13. J. C. Nedelec, "Mixed Finite Elements in $\mathbf{R}^3$," *Numer. Math.*, vol. 35, pp. 315-341, 1980.

14. P. G. Ciarlet and P. A. Raviart, "General Lagrange and Hermite Interpolation in $\mathbf{R}^n$ with Applications to Finite Element Methods.," *Arch. Rational Mech. Anal.*, vol. 46, pp. 177-199, 1972.

15. Eugene O'Riordan, "Singularly Perturbed Finite Element methods," *Numerische Mathematik*, vol. 44, pp. 425-434, Springer-Verlag, 1984.

16. Eugene O'Riordan and Martin Stynes, "A finite element method for a singularly perturbed boundary value problem in conservative form," in *BAIL III*, ed. J. J. H. Miller, pp. 271-275, Boole Press, Dublin, 1984.

17. Martin Stynes and Eugene O'Riordan, "A Finite Element Method for a Singular Perturbed Boundary Value Problem," *Numerische Mathematik*, vol. 50, pp. 1-15, Springer-Verlag, 1986.

18. Eugene O'Riordan and Martin Stynes, "A Uniformly Accurate Finite-Element Method for a Singularly Perturbed One-Dimensional Reaction-Diffusion Problem," *Mathematics of Computation*, vol. 47, no. 176, pp. 555-570, 1986.

19. Eugene O'Riordan and Martin Stynes, "An analysis of a Superconvergence Result for a Singularly Perturbed Boundary Value Problem," *Mathematics of Computation*, vol. 46, no. 173, pp. 81-92, 1986.

20. Eugene O'Riordan and Martin Stynes, "A uniform finite element method for a conservative singularly perturbed problem," *Journal of Computational and Applied Mathematics*, vol. 18, pp. 163-174, 1987.

21. H.-J. Reinhardt, "A-Posteriori Error Analysis and Adaptive Finite Element Methods for Singularly Perturbed Convection-Diffusion Equations," *Math. Meth. in the Appl. Sci.*, vol. 4, pp. 529-548, 1982.

22. Eugene O'Riordan and Martin Stynes, "A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in two dimensions," *Mathematics of Computation*, vol. 57, no. 195, pp. 47-62, 1991.

23. M. Fiedler, *Special matrices and their applications in numerical mathematics*, Martinus Nijhoff, Dordrecht, 1986.