To the Graduate Council:

I am submitting herewith a dissertation written by Michael Plexousakis entitled "An Adaptive Nonconforming Finite Element Method for the Nonlinear Schrödinger Equation". I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

O. Karakashian, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of The Graduate School

An Adaptive Nonconforming Finite Element Method for the Nonlinear Schrödinger Equation

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Michael Plexousakis
December, 1996

Στην Ντινα

και στους γονεις μου

ACKNOWLEDGEMENTS

I am particularly indebted to my thesis advisor, Professor Ohannes Karakashian for his continuous support and encouragement throughout my graduate studies and his valuable guidance and help in completing this work. I would also like to express my special thanks to Professors Steven Serbin, Vasilios Alexiades and Michael Guidry for reading the entire thesis, providing valuable comments, and for serving in my committee. Further, I want to express my gratitude to Professor Samuel Jordan for being a constant source of help and support throughtout my studies in the University of Tennessee, Knoxville, and Professors Vassilios Dougalis and Georgios Akrivis for their guidance and encouragement during my undergraduate studies.

ABSTRACT

A fully discrete approximation to the solution of the two–dimensional Nonlinear Schrödinger Equation is presented and analyzed. The spatial discretization is based on non–conforming finite elements and the temporal discretization is achieved by a variable step Crank–Nicolson scheme. The numerical method presented here is aimed at approximating general (i.e. not necessarily radially symmetric solutions) of the Nonlinear Schrödinger equation which emanate from general initial profiles (vanishing at infinity). In particular, we apply it to study the formation of singularities. The numerical scheme employs a spatial refinement strategy which allows the mesh to be refined locally and without compromising the overall quality of the mesh. The adaptive refinement strategy is effected by a simple and computationaly efficient test, based on inverse estimates on the approximating subspace. Several numerical experiments concerning the formation of singularities are presented.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Nonlinear cubic Schrödinger Equation (henceforth referred to as the NLS equation)

$$u_t = i\Delta u + i|u|^2 u, \qquad x \in \mathbb{R}^d,\ t \geq 0, \tag{1.1a}$$

where u is a complex-valued function of $x \in \mathbb{R}^d$ and the temporal variable t, is a basic evolution model for nonlinear waves in various branches of physics such as hydrodynamics, plasma physics and nonlinear optics. The initial–value problem for the NLS, consisting of (1.1a) and the initial condition

$$u(x,0) = u^0(x), \qquad x \in \mathbb{R}^d, \tag{1.1b}$$

where u^0 is a complex–valued function defined on $\mathbb{R}^d$, describes, for $d = 2$, the propagation of light beams in nonlinear, dispersive media, [14], [42], [26]. In this case, the time variable t is replaced by the axial coordinate along the beam. For $d = 1$, (1.1a)–(1.1b) has been used in the study of pulse propagation along optical fibers, [24], and as an envelope equation in the theory of water waves, [32], [45]. For $d = 3$, the NLS equation is obtained as a limiting case of the Zakharov model of Langmuir waves [47]. The expository paper [34] describes the properties of a class of nonlinear Schrödinger equations and discusses various other physical interpretations.

The initial–value problem (1.1a)–(1.1b) has been the subject of intensive theoretical and numerical studies over the last several years. The one dimensional initial–value problem admits global solutions provided that the initial data are smooth enough and decay sufficiently fast at infinity, [40]. The one dimensional NLS equation is a special case, since it admits *solitons*, i.e. solitary waves that preserve their shapes after mutual interaction, [1]. For $d = 2$, global existence occurs only under suitable smallness conditions on the initial data, [44] and, moreover, singular solutions exist that blow up in finite time, [40], [22], [21]. The case $d = 3$ is 'super-critical'. The initial–value problem (1.1a)–(1.1b) is locally well–posed and there exist solutions that blow up in finite time, [47], [21].

The evolution of spike singularities in evolution equations has long been observed in numerical simulations and in the laboratory. The detailed dynamics of blow–up have both academic and practical interest. For example, in nonlinear optics, understanding of the dynamics is essential for studying the properties of the beam focus, which often has a damaging effect on the experimental apparatus. The nature of the blow–up singularity has been studied both theoretically and numerically, but mainly for radially symmetric solutions (cf., for example, [48] and [29]). It was shown in [29] that in three dimensions, the peak amplitude of singular radial solutions grows like $(t^* - t)^{-\frac{1}{2}}$ as t approaches t^*, the blow–up time. The determination of the blow–up rate for two–dimensional radially symmetric solutions is more difficult since $d = 2$ is the critical exponent for the cubic NLS equation. Early numerical experiments and formal asymptotic reasoning favored a rate of $(t^* - t)^{-\frac{2}{3}}$ in [48] and [41] while [46] and [43] favored a rate of $[\ln(\frac{1}{t^*-t})/(t^* - t)]^{\frac{1}{2}}$. In [27] it was concluded that the blow–up rate is $[\ln\ln(\frac{1}{t^*-t})/(t^* - t)]^{\frac{1}{2}}$ using numerical techniques ("dynamical rescaling", see also Chapter 5) and asymptotic analysis. In [28] the numerical methods introduced in [27] were extended to cover the non radially symmetric case.

We refer the reader to [40] for a review of the known existence results on the Cauchy problem for the NLS equation, as well as to [34] for a collection of references to the vast Russian literature on the subject. We note that in physics, the 3-d evolution of a spike singularity is usually referred to as "wave collapse" and the corresponding 2-d phenomenon as "self-focusing".

In this work, we shall be concerned with the numerical solution of the following initial- and boundary-value problem for the NLS equation: Let Ω be a bounded domain in $\mathbb{R}^2$ with smooth boundary $\partial\Omega$. For $0 < t^* < \infty$ we seek a complex-valued function u satisfying

$$\begin{cases} u_t = i\Delta u + i\lambda|u|^2 u, & (x,t) \in \overline{\Omega} \times [0,t^*], \\ u = 0, & (x,t) \in \partial\Omega \times [0,t^*], \\ u(x,0) = u^0(x), & x \in \overline{\Omega} \end{cases} \tag{1.2}$$

where λ is a nonzero real number and u^0 is a given complex-valued function defined on $\overline{\Omega}$. We shall assume that a solution of (1.2) exists and is sufficiently smooth to guarantee the error estimates established below.

The mathematical literature contains a large number of works devoted to the numerical approximation of (1.2) and similar initial- and boundary-value problems. Numerical experiments in the case $d = 1$ were presented in [17], [23], [37], [38], [4]. For $d \geq 2$, the articles [29] and [30] introduced the method of "dynamic rescaling" as a tool for studying the form of the singularity. This method was extended in [27] to cover anisotropic solutions, and to follow the evolution of initial profiles u^0 that are not radially symmetric. In both cases, it was assumed that the modulus of the solution approached infinity at exactly one point. In [5] the initial- and boundary-value problem (1.2) was discretized in space by a standard Galerkin–finite element method and in time by the Crank–Nicolson method. The resulting fully discrete schemes were applied with adaptive space and time meshes to approximate radial solutions of the NLS that blow up in finite time. The construction of adaptive Galerkin methods in [5] was heavily based on the first two invariants I_1 and I_2 of (1.1a). It is known (see for example, [40]) that the NLS equation has the invariant functionals

$$I_1 = \int_{\mathbb{R}^d} |u(t)|^2 \, dx \tag{1.3}$$

and

$$I_2 = \int_{\mathbb{R}^d} \left\{ \frac{1}{2}|\nabla u(t)|^2 - \frac{\lambda}{4}|u(t)|^4 \right\} dx \tag{1.4}$$

The first is obtained by multiplying (1.1a) by $\overline{u}$ and taking imaginary parts, the second by multiplying (1.1a) by $\overline{u}_t$ and taking real parts. The time step reduction mechanism was motivated by the need to control the change in a weighted version of I_2, while the spatial grid refinement mechanism was motivated by a local L^∞–L^2 "inverse" property. Notice that under the hypothesis of radial symmetry, the initial–value problem (1.1a)–(1.1b) may be cast in the form

$$\begin{cases} u_t = i(u_{rr} + \frac{d-1}{r}u_r) + i|u|^2u, & r > 0,\ t > 0, \\ u(r,0) = u^0(r), & r > 0 \end{cases}$$

where $u = u(r,t)$ and $r = (x_1^2 + \cdots + x_d^2)^{1/2}$, leading to a problem involving only one spatial dimension.

The computational part of the work at hand makes no assumptions on the radial symmetry of the solution or the location of the blow-up. We allow both isotropic and anisotropic solutions, but we assume that the restriction of the solution $u(x,t)$ on the boundary $\partial\Omega$ is negligibly small for $t \geq 0$, so that we may impose the Dirichlet boundary condition $u = 0$ on $\partial\Omega$. Moreover, initial profiles with several peaks are also allowed. We discretize (1.2) in space by a nonconforming finite element method. The elements of the approximating space reduce to piecewise polynomial functions which are, in general, discontinuous across interelement boundaries. The triangulations on which the approximating spaces are defined are required to satisfy only mild assumptions. In particular, highly non uniform meshes are allowed and no global quasi–uniformity assumption is made. The temporal discretization is achieved by means of a Crank–Nicolson type method, or more precisely by the one–step, Gauss–Legendre, Implicit Runge–Kutta method. Time stepping with this method requires solving a nonlinear system of equations at each step. A modified Newton method, the so called Implicit–Explicit scheme, is used to solve the nonlinear systems. The linearization of the fully discrete scheme by the Implicit–Explicit scheme leads to large, complex, sparse linear systems. The solution of these linear systems by means of iterative methods is discussed in detail in Chapter 4.

The analysis of the convergence of the fully discrete scheme is undertaken in Chapter 3. We show the existence of the resulting approximations and prove optimal order error estimates in $L^\infty(L^2)$. The existence and convergence of the fully discrete scheme is carried out in three steps. First, a *modified scheme* is introduced whereby the nonlinear term $|u|^2u$ is replaced by a globally Lipschitz continuous function which coincides with $|u|^2u$ in a sufficiently large ball. Then the existence of solutions to the modified scheme is established, by means of a fixed point-type result, and finally the convergence of the modified scheme is exhibited. The main error estimate in $L^\infty(L^2)$ may be stated as follows: Let $0 = t^0 < t^1 < \cdots < t^N = t^*$ be a partition of $[0, t^*]$, and $I_n = (t^n, t^{n+1}]$. With each of these intervals associate a triangulation $\mathcal{T}_h^n$ of the domain Ω satisfying the mesh conditions stated in Chapter 2 and a finite dimensional subspace V_h^n whose elements reduce to polynomial functions of degree $\leq r-1$ on each element of T_h^n. Let u denote the solution of (1.2) and let U^n be the solution of the modified scheme at t^n. With $E^n = u(t^n) - U^n$ we have,

$$\|E^n\| \leq C_n\Big\{\|h_0^r u^0\|_{r,h} + \Big(\sum_{m=0}^{n-1}\Big[|||h_m^r u_t|||_m^2 + k_m^4|||u^{(3)}|||_m^2 + k_m^4|||\Delta u^{(2)}|||_m^4\Big]\Big)^{1/2}$$
$$+\Big(\sum_{m=0}^{n-1}\Big[k_m^5 + k_m\|h_m^r u(t^n)\|^2\Big]\Big)^{1/2} + \Big(\mathcal{N}_C(n-1)\sum_{m=1}^{n-1}\|J[\omega^m]\|^2\Big)^{1/2}\Big\}.$$

Here, $J[\omega^n] = (P_E^n - P_E^{n-1})u(t^n)$ is the jump of the elliptic projection $\omega(t) = P_E^n u(t^n)$ (see Chapter 2) and $\mathcal{N}_C$ denotes the number of times where $V_h^n \neq V_h^{n-1}$. Since the term $||J[\omega^m]||$ is always bounded by $||h_m^r u||_{r,h}$ this establishes an optimal order error estimate, provided that the space does not change very often. As a final step of the convergence analysis it is shown that the solution of the modified scheme is also a solution of the original scheme.

A second component of the computational part of this thesis was the development of a finite element code so that the fully discrete scheme discussed above, could be implemented with an adaptive space mesh and variable time step. This necessitated the development of appropriate data structures to handle nonconforming triangulations and a dynamically refined mesh. We refined the triangles of the mesh using only regular refinement (the midpoints of the three sides are joined to form four triangles) but allowed each triangle to have at most one nonconforming node per side. This, obviously, limits the number of neighbors of each triangle and simplifies the data structures necessary to represent such a triangulation. The mechanism for the temporal grid refinement was similar to the one used in [5], while the spatial grid refinement mechanism consisted of a computationally efficient and inexpensive test, motivated by a local H^1–L^2 inverse property on each of the triangles of the mesh. The last chapter of this thesis contains the results of experiments performed with this adaptive code. We followed the evolution in time of initial profiles with or without radial symmetry, and we also experimented with initial profiles having several peaks. The code was successful in refining the mesh near the evolving singularities and achieved amplitude magnifications in the range of $O(10^7)$ to $O(10^8)$.

The remaining of this thesis is organized as follows: In Chapter 2 we introduce the notation that will be used in subsequent chapters, the approximating subspaces and the assumptions on the triangulations of the domain Ω. Chapter 3 contains the consistency, stability and convergence of the fully discrete approximations. In particular, optimal order error estimates in $L^\infty(L^2)$ are proved. Chapter 4 describes the details of a finite element code implementing a fully discrete scheme for the solution of (1.2), as well as, an analysis of the performance of the Implicit–Explicit scheme used for solving the systems of nonlinear equations arising in the fully discrete scheme. Moreover, an adaptive version of the code which features automatic grid refinement in both space and time is also presented and analyzed. Finally, Chapter 5 gives the results of numerical experiments performed with the adaptive version of our code, on approximating solutions of the NLS equation that blow up in finite time.

Chapter 2

Notation and Preliminaries

In this chapter we introduce the notation that will be used in the remaining of the thesis and construct the finite dimensional spaces where the solution u of (1.2) will be sought. The elements of these spaces are functions defined on suitable partitions of the domain Ω. For example, in the framework of finite elements, a partition of the domain Ω will correspond to a triangulation of Ω satisfying certain mild conditions.

2.1 Function spaces

Let Ω be a domain in $\mathbb{R}^d$. For a subdomain $D \subset \Omega$ and $1 \leq p \leq \infty$, $L^p(\Omega)$ denotes the usual space of complex-valued measurable functions v on D such that

$$||v||_{L^p(D)} = \left\{ \int_D |v|^p \, dx \right\}^{1/p} < \infty, \qquad 1 \leq p < \infty$$
$$||v||_{L^\infty(D)} = \operatorname*{ess\,sup}_{x \in D} |v(x)| < \infty, \qquad p = \infty$$

For $p = 2$, $L^2(D)$ is a Hilbert space with inner product *defined* by

$$(v, w)_D = \int_D v\overline{w} \, dx, \qquad v, w \in L^2(D),$$

where dx denotes the Lebesgue measure in $\mathbb{R}^d$. If $D = \Omega$, we shall write $(\cdot, \cdot)$ in place of $(\cdot, \cdot)_\Omega$ and $|| \cdot ||$ instead of $|| \cdot ||_{L^2(\Omega)}$. For $m \geq 0$ integer, and $1 \leq p \leq \infty$, we denote by $W^{m,p}(D)$ the Sobolev space of complex-valued functions on D with norm

$$||v||_{m,p,D} = \left\{ \sum_{|\alpha|\leq m} \int_D |\partial^\alpha v|^p \, dx \right\}^{1/p}, \qquad 1 \leq p < \infty$$

$$||v||_{m,\infty,D} = \max_{|\alpha|\leq m} ||\partial^\alpha v||_{L^\infty(D)}, \qquad p = \infty$$

For $p = 2$, $H^m(D) = W^{m,2}(D)$, is a Hilbert space with inner product

$$(v, w)_{m,D} = \sum_{|\alpha|\leq m} \int_D \partial^\alpha v \, \partial^\alpha \overline{w} \, dx, \qquad v, w \in H^m(D)$$

and norm $|| \cdot ||_{m,D}$. We shall also make use of the seminorms

$$|v|_{m,D} = \left\{ \sum_{|\alpha|=m} \int_D |\partial^\alpha v|^2 \, dx \right\}^{1/2}$$

Assuming that the subdomain D has Lipschitz-continuous boundary ∂D, we let, $H_0^1(D) = \left\{ v \in H^1(D) : v_{|\partial D} = 0 \right\}$, where the values on ∂D are understood in the sense of the trace, cf. [2].

2.2 Partitions of the domain Ω and the finite dimensional spaces

Let $0 < h \leq 1$, be a spatial discretization parameter and denote by $\mathcal{T}_h = \{\Omega_i^h, i = 1, \ldots, d_h\}$ a partition of Ω consisting of triangles. A peripheral triangle Ω_i^{*h}, i.e. one intersecting the boundary $\partial\Omega$, may have a curved side. We shall assume that the elements of this partition satisfy certain requirements. Let $\mu_1, \ldots, \mu_4$ be positive constants independent of the family of triangulations $\{\mathcal{T}_h\}$. Denote the boundary of Ω_i^h by $\partial\Omega_i^h$, and set $\partial\Omega_{ij}^h = \partial\Omega_i^h \cap \partial\Omega_j^h$, $i, j = 1, \ldots, d_h$, $\partial\Omega_i^{*h} = \partial\Omega_i^h \cap \partial\Omega$, and let $\mathcal{N}_i = \{j : \partial\Omega_{ij}^h \text{ is a line segment}\}$, i.e. $|\mathcal{N}_i|$ is the number of neighbors of Ω_i^h. Also, we let h_i denote the diameter of the smallest disk containing Ω_i^h, and $\underline{h}_i$ denote the diameter of the largest disk contained in Ω_i^h.

(i) Each Ω_i^h is a nonempty open, connected subset of Ω.

(ii) $\bigcup_{i=1}^{d_h} \overline{\Omega}_i^h = \overline{\Omega}$ and $\Omega_i^h \cap \Omega_j^h = \emptyset$, $\forall i \neq j$.

(iii) Each Ω_i^h has Lipschitz-continuous boundary $\partial\Omega_i^h$ and contains a point z_i^h such that such that $(\mathbf{x} - \mathbf{z}_i^h) \cdot \mathbf{n}_i^h \geq \mu_1 h_i$ for almost all x on $\partial\Omega_i^h$, where n_i^h is the unit outward normal vector to $\partial\Omega_i^h$ and $\mathbf{x}, \mathbf{z}_i^h$ are the position vectors of x and z_i^h.

(iv) There exists a constant $\mu_2 > 0$, independent of h and i such that

$$h_i \leq \mu_2 \underline{h}_i, \quad i = 1, \ldots, d_h.$$

(v) If Ω_i^h and Ω_j^h are neighboring triangles then

$$\mu_3^{-1} \leq \frac{h_i}{h_j} \leq \mu_3, \quad \mu_3 \geq 1.$$

(vi) Each Ω_i^h contains a subset S_i^h with respect to which Ω_i^h is starlike and $m(\Omega_i^h) \leq \mu_4\, m(S_i^h)$, $i = 1, \ldots, d_h$. Here $m(D)$ denotes the measure of D.

It is clear that under general conditions on Ω, e.g. Ω convex, or $\partial\Omega$ smooth, or h sufficiently small, we can construct triangulations $\mathcal{T}_h$ that will satisfy the above assumptions.

We shall also need to introduce some notation for inner products and norms on subsets of the boundaries $\partial\Omega_i^h$. For $i = 1, \ldots, d_h$, let $\sigma^{(i)}$ denote the Lebesgue measure on $\partial\Omega_i^h$. For a subset Γ_i^h of $\partial\Omega_i^h$ of positive measure, we denote by $L^2(\Gamma_i^h)$ the usual Hilbert space of square integrable functions on Γ_i^h equipped with the inner product $\langle v, w\rangle_{\Gamma_i^h} = \int_{\Gamma_i^h} v\overline{w}\, d\sigma^{(i)}$ and the norm $|v|_{\Gamma_i^h} = \left[\langle v, v\rangle_{\Gamma_i^h}\right]^{1/2}$. In addition, we shall denote by $\frac{\partial}{\partial n^{(i)}}$ the normal derivative operator on Ω_i^h, the normal derivative being understood in the distributional sense. Moreover, for $v \in H^m(\Omega)$ we let $v^{(i)}$ denote the restriction of v on Ω_i^h and for $m \geq 1$, $v\big|_{\partial\Omega_i^h}$ the trace of v on $\partial\Omega_i^h$.

For D a subdomain of Ω and $r \geq 1$ integer, we let $\mathbb{P}_r(D)$ denote the space of polynomials in d variables of degree less than on equal to $r - 1$ on D. We also let

$\mathbb{P}_0 = \{0\}$ and

$$V_h^r = \mathbb{P}_r(\Omega_1^h) \times \ldots \times \mathbb{P}_r(\Omega_{d_h}^h)$$

The space V_h^r possesses optimal approximation properties. We quote the following result form [7]:

PROPOSITION 2.1 *Let $m \geq 0$. There exists a constant c such that if $v \in H^m(\Omega_i^h)$, there exists a $\chi \in \mathbb{P}_m(\Omega_i^h)$ such that*

$$\|v - \chi\|_{j,\Omega_i^h} \leq c h_i^{m-j} |v|_{m,\Omega_i^h}, \qquad 0 \leq j \leq m. \tag{2.2}$$

From (2.2) we can immediately get global approximation on V_h^r. Indeed, if $v \in H^m(\Omega)$ and $1 \leq m \leq r$, there exists $\chi \in V_h^r$ such that

$$\sum_{i=1}^{d_h} \|v - \chi\|_{j,\Omega_i^h}^2 \leq c \sum_{i=1}^{d_h} h_i^{2(m-j)} |v|_{m,\Omega_i^h}^2, \qquad 0 \leq j \leq m \tag{2.3}$$

Our assumptions on the partition $\mathcal{T}_h$ yield the following local inverse property: There is a constant c such that for all $\chi \in \mathbb{P}_r(\Omega_i^h)$ we have,

$$|\chi|_{m,\Omega_i^h} \leq c h_i^{-(m-s)} |\chi|_{s,\Omega_i^h}, \qquad i = 1, \ldots, d_h, \quad 0 \leq s \leq m \leq r. \tag{2.4}$$

Finally, to estimate boundary and interface integrals we shall use the following result from [7]:

$$|v|_{0,\partial\Omega_i^h}^2 \leq c\{h_i^{-1} \|v\|_{0,\Omega_i^h}^2 + \|v\|_{0,\Omega_i^h} |v|_{1,\Omega_i^h}\}, \tag{2.5}$$

with c depending on the dimension d and the constant μ_1.

2.3 The bilinear form $a_h^\gamma(\cdot, \cdot)$

We denote by E_h the energy space $E_h = H^2(\Omega_1^h) \times \ldots \times H^2(\Omega_{d_h}^h)$ and equip it with the norm

$$\|v\|_{1,h}^2 = \sum_{i=1}^{d_h} \left(\|\nabla v^{(i)}\|_{0,\Omega_i^h}^2 + h_i \left| \frac{\partial v^{(i)}}{\partial n} \right|_{\partial\Omega_{ij}^*}^2 + h_i^{-1} |v^{(i)}|_{\partial\Omega_{ij}^*}^2 \right. \\ \left. + \sum_{j \in \mathcal{N}_i} \tau_{ij} \left[h_i \left| \frac{\partial v^{(i)}}{\partial n} \right|_{\partial\Omega_{ij}^h}^2 + h_i^{-1} |v^{(i)} - v^{(j)}|_{\partial\Omega_{ij}^h}^2 \right] \right) \tag{2.6}$$

Define the bilinear form $a_h^\gamma(\cdot,\cdot) : E_h \times E_h \to \mathbb{C}$ by

$$
\begin{aligned}
(2.7)\quad a_h^\gamma(u,v) = \sum_{i=1}^{d_h} \Bigg\{ \left(\nabla u^{(i)}, \nabla v^{(i)}\right)_{\Omega_i^h} + \sum_{j\in\mathcal{N}_i} \tau_{ij} \Big[-\langle \frac{\partial u^{(i)}}{\partial n}, v^{(i)} - v^{(j)} \rangle_{\partial\Omega_{ij}^h} \\
-\langle \frac{\partial v^{(i)}}{\partial n}, u^{(i)} - u^{(j)} \rangle_{\partial\Omega_{ij}^h} + \gamma h_i^{-1} \langle u^{(i)} - u^{(j)}, v^{(i)} - v^{(j)} \rangle_{\partial\Omega_{ij}^h} \Big] \\
-\langle \frac{\partial u^{(i)}}{\partial n}, v^{(i)} \rangle_{\partial\Omega_{ij}^{*h}} - \langle \frac{\partial v^{(i)}}{\partial n}, u^{(i)} \rangle_{\partial\Omega_{ij}^{*h}} + \gamma h_i^{-1} \langle u^{(i)}, v^{(i)} \rangle_{\partial\Omega_{ij}^{*h}} \Bigg\},
\end{aligned}
$$

where γ is a constant to be chosen appropriately and (τ_{ij}) denotes the matrix with elements

$$
\tau_{ij} = \begin{cases} 1, & i > j \\ 0, & i \le j \end{cases} \qquad i,j = 1,\ldots,d_h.
$$

The bilinear form $a_h^\gamma(\cdot,\cdot)$ generalizes the standard Dirichlet form $a(u,v) = (\nabla u, \nabla v)_{0,\Omega}$ in the sense that

$$
(2.8)\qquad a_h^\gamma(u,v) = -(\Delta u, v), \quad \forall\, u \in H^2(\Omega) \cap H_0^1(\Omega),\ v \in E_h
$$

Also, the third and sixth terms in the definition of $a_h^\gamma(\cdot,\cdot)$ are added to produce symmetry. Indeed, it is clear that a_h^γ is Hermitian and as a consequence,

$$
(2.9)\qquad a_h^\gamma(v,v) \text{ is real for } v \in V_h^r + (H^2(\Omega) \cap H_0^1(\Omega)).
$$

Moreover, the fourth and seventh terms are penalty terms which will induce coercivity of the form a_h^γ with respect to the norm $||\cdot||_{1,h}$. The bilinear form a_h^γ is *continuous*, i.e.,

$$
(2.10)\qquad |a_h^\gamma(u,v)| \le (1+\gamma)||u||_{1,h}||v||_{1,h}, \qquad \forall\, u,v \in E_h.
$$

Moreover, it is *coercive* in that there exist constants c_a, γ_0 such that for all $\gamma \ge \gamma_0$

$$
(2.11)\qquad a_h^\gamma(v,v) \ge c_a ||v||_{1,h}^2, \qquad \forall\, v \in V_h^r,
$$

where $c_a = \frac{1}{2}\min\{1, \frac{1}{2c}\}$, $\gamma_0 = 2\max\{c_a, \frac{1}{c_a}\}$ with c some constant depending on the constants of (2.4) and (2.5). The proofs of these two crucial properties are straightforward generalizations of the proofs concerning the case of real E_h that can be found in [7].

With a_h^γ we associate an *elliptic projection* operator $P_E : E_h \longrightarrow V_h^r$ by

$$(2.12) \qquad a_h^\gamma(P_E v, \chi) = a_h^\gamma(v, \chi), \qquad \forall\, \chi \in V_h^r.$$

That P_E is well-defined follows easily from (2.10) and (2.11). Also, using standard techniques of finite element theory, we can obtain the following approximation result

$$(2.13) \qquad ||v - P_E v||_{1,h} \le c||h^{s-1}v||_{r,h}, \quad v \in H^r(\Omega) \cap H_0^1(\Omega), \quad 2 \le s \le r,$$

where $||\cdot||_{r,h}$ denotes the "mesh dependent" norm

$$||h^{s-1}v||_{r,h} = \left(\sum_{K \in \mathcal{T}_h} h^s ||v||_{r,K}^2 \right)^{1/2}.$$

Furthermore, using Nitsche's duality argument, we can obtain the optimal order L^2-estimate $||v - P_E v|| = O(h^r)$. However, in analogy with (2.13) we assume that there holds

$$(2.14) \qquad ||v - P_E v|| \le c||h^s v||_{r,h}, \quad v \in H^r(\Omega) \cap H_0^1(\Omega), \quad 2 \le s \le r.$$

For standard (conforming) finite elements (2.14) is known to hold in one dimension, cf. [12, Chapter 0], [6]. In higher dimensions, (2.14) holds under local quasiuniformity conditions on the mesh, cf. [18],[19].

Remark. In the remaining of this thesis and especially in Chapter 3, where the time stepping scheme is introduced, the triangulation $\mathcal{T}_h$ of the domain Ω and, consequently, the corresponding finite dimensional spaces, will be allowed to change from one time step to the next. For this reason, the triangulation $\mathcal{T}_h$, the corresponding finite dimensional subspace V_h^r and all quantities defined on V_h^r, such as the bilinear form a_h^γ will acquire a superscript or subscript of n, denoting the current time level.

Chapter 3

Error Estimates

3.1 The fully discrete scheme

Given the initial value problem

$$
\begin{cases} y' = f(y,t), & 0 < t \le t^*, \\ y(0) = y^0 \end{cases} \tag{3.1}
$$

the *midpoint* or, more precisely, the one-stage Gauss–Legendre method generates approximations y^n to $y(t^n)$, for $0 \le n \le J$, where $k = t^*/J$ is the (constant) stepsize and $t^n = nk$, by

$$
y^{n,1} = y^n + \frac{k}{2} f(y^{n,1}, t^{n,1}) \tag{3.2}
$$

and

$$
y^{n+1} = 2y^{n,1} - y^n \tag{3.3}
$$

Here we have set $t^{n,1} = t^n + \frac{k}{2}$. This is the first member of the infinite class of Gauss–Legendre methods, cf. [16]. As such, it has a *classical order* ν of 2 and a *stage order* p of 1, and is algebraicaly stable. We may now describe our fully discrete scheme for the solution of (1.2). Let $0 = t^0 < t^1 < \cdots < t^N = t^*$ be a partition of $[0, t^*]$, and

$$
I_n = (t^n, t^{n+1}], \quad k_n = t^{n+1} - t^n. \tag{3.4}
$$

We associate a partition $\mathcal{T}_{h_n}$ of Ω and a finite element space V_h^n to each interval I_n as follows:

$$(3.5) \qquad V_h^n = \{\chi \in L^2(\Omega) : \chi|_K \in \mathbb{P}_{r-1}(K), K \in \mathcal{T}_{h_n}\}.$$

We also associate a space V_h^{-1} with $\{t^0\}$ but for simplicity we take $V_h^{-1} = V_h^0$. Changing slightly the notation introduced in Chapter 2, we shall denote by K the elements of the partition $\mathcal{T}_{h_n}$. Also h_K stands for the diameter of the element K, and $h_n = \max_{K\in\mathcal{T}_{h_n}} h_K$, $\underline{h}_n = \min_{K\in\mathcal{T}_{h_n}} h_K$.

The one-stage Gauss–Legendre method for (1.2) that we consider may then be defined as follows: Let $U^0 = u^0$ and assume that for each $0 \le n \le N-1$ approximations $U^j \in V_h^{j-1}$ to $u(t^j)$, $j = 0, \dots, n$ are given. With Π^n denoting an appropriate projection operator into V_h^n, we let $U^{n+} = \Pi^n U^n$ and we define the quantity $U^{n,1} \in V_h^n$ by, cf. (3.2),

$$(3.6) \qquad (U^{n,1}, \psi) + \mathrm{i}\,\frac{k_n}{2} a_{h_n}^{\gamma}(U^{n,1}, \psi) - \mathrm{i}\,\lambda\,\frac{k_n}{2}(|U^{n,1}|^2 U^{n,1}, \psi) = (U^{n+}, \psi),$$

$\forall\, \psi \in V_h^n$. Then, $U^{n+1} \in V_h^n$ is defined by

$$(3.7) \qquad U^{n+1} = 2\,U^{n,1} - U^{n+}.$$

Two particular choices of Π^n, are the L^2–projection and the elliptic projection into V_h^n. Thus if $V_h^{n-1} \subseteq V_h^n$ then $U^{n+} = U^n$.

We shall make frequent use of the following notation: We denote by $|\!|\!|\cdot|\!|\!|_n$ the norm of $L^2(I_n, L^2(\Omega))$ i.e.,

$$(3.8) \qquad |\!|\!|v|\!|\!|_n := \Big(\int_{I_n} \|v(t)\|^2 dt\Big)^{1/2}.$$

Also for $s, m = 0, 1, \dots,$ and $v \in H^m(\Omega)$ we let

$$(3.9) \qquad \|h_n^s v\|_{m,h} := \Big(\sum_{K\in\mathcal{T}_{h_n}} h_K^{2s}\,\|v\|_{m,K}^2\Big)^{1/2}.$$

We introduce the *modified scheme* whereby (3.6) is replaced by

$$(3.10) \qquad (U^{n,1}, \psi) + \mathrm{i}\,\frac{k_n}{2} a_{h_n}^{\gamma}(U^{n,1}, \psi) - \mathrm{i}\,\lambda\,\frac{k_n}{2}(f(U^{n,1}), \psi) = (U^{n+}, \psi),$$

where $f = f(z)$, $f : \mathbb{C} \to \mathbb{C}$, is an appropriate globally Lipschitz continuous function, such that $f(z) = |z|^2 z$ in a ball with a radius that depends on the solution u of (1.2). In particular, u is in that ball. We quote the following result form [25].

LEMMA 3.11 *Let* $M := \sup_{[0,t^*]} ||u||_\infty + 1$. *Then there exists a function* $f : \mathbb{C} \to \mathbb{C}$ *such that*

$$f(z) = |z|^2 z, \quad \text{if } |z| \le M, \tag{3.12a}$$

$$|f(z)| \le c_1 |z|, \quad c_1 > 0, \quad \forall\, z \in \mathbb{C} \tag{3.12b}$$

$$|f(z) - f(w)| \le c_2 |z - w|, \quad c_2 > 0, \quad \forall\, z, w \in \mathbb{C}. \tag{3.12c}$$

Indeed, as done in [25] we may consider a function of the form $f(z) = \gamma(|z|^2)z$, where the function $\gamma : [0, \infty) \to [0, 4M^2]$ is smooth with bounded derivatives of order up to two. In this case, it can be shown that $c_1 = c_1' M^2$ and $c_2 = c_2' M^2$. Furthermore,

$$(f(z), z) \text{ is real } \forall\, z \in \mathbb{C}. \tag{3.12d}$$

As a first step towards proving the convergence of (3.6) to (1.2) we analyze the stability and convergence of (3.10). In Section 3.2 we establish the existence of solutions to the modified scheme. In Section 3.3 we present the basic error analysis in $L^\infty(L^2(\Omega))$ for (3.10) with $\Pi^n = P^n$, for all n, where P^n is the L^2-projection into V_h^n. We shall also make use of the notation

$$k = \max_{0 \le n \le N-1} k_n, \; h = \max_{\substack{K \in \mathcal{T}_{h_n} \\ 0 \le n \le N-1}} h_K, \; \underline{h} = \min_{\substack{K \in \mathcal{T}_{h_n} \\ 0 \le n \le N-1}} h_K, \; C_N = c|\ln(\underline{h})|.$$

3.2 Existence of solutions to the modified scheme

To prove existence of solutions we use the following well known Brouwer-type fixed-point result, cf. [13].

LEMMA 3.13 *Let $(H,(\cdot,\cdot))$ be a finite dimensional inner product space and $\|\cdot\|$ the associated norm. Let $g: H \to H$ be continuous and assume that there exists $\beta > 0$ such that for every $z \in H$ with $\|z\| = \beta$ there holds $Re(g(z), z) \geq 0$. Then there exists a $z^* \in H$ such that $g(z^*) = 0$ and $\|z^*\| \leq \beta$.*

THEOREM 3.14 *Let U^n be given in V_h^{n-1}. Then, there exists $U^{n,1} \in V_h^n$ satisfying* (3.10). *Furthermore $U^{n,1}$ is unique if k_n is sufficiently small.*

Proof. The vector space V_h^n is a finite dimensional Hilbert space equipped with the L^2 inner product. Consider the map $F: V_h^n \to V_h^n$ given by

$$(F(v),\psi) = (v,\psi) + \mathrm{i}\,\frac{k_n}{2} a_{h_n}^{\gamma}(v,\psi) - \mathrm{i}\,\lambda\frac{k_n}{2}(f(v),\psi) - (U^{n+},\psi), \tag{3.15}$$

$\forall \psi \in V_h^n$. Our task is to show that F has a zero in V_h^n. Since $f: \mathbb{C} \to \mathbb{C}$ is continuous, F is continuous and therefore it suffices to show the existence of a constant $\beta > 0$ such that

$$\text{Re }(F(v),v) \geq 0, \quad \text{for all } v \in V_h^n \text{ s.t. } \|v\| = \beta\ . \tag{3.16}$$

Set $\psi = v$ in (3.15) to obtain, in view of (3.12d),

$$\text{Re }(F(v),v) \geq \|v\|^2 - \|v\|\,\|U^{n+}\|.$$

We see that (3.16) holds for all v satisfying

$$\|v\| = \beta := \|U^{n+}\|.$$

The proof of existence is now complete. Uniqueness follows easily from (3.12c). □

3.3 Error Analysis in $L^\infty(L^2(\Omega))$

In this section we shall operate under the assumption that $\Pi^n = P^n$, the L^2-projection operator onto V_h^n. As usual, we split the error $U - u$ into $(U - W) + (W - u)$, where $W \in V_h^n$ is an appropriately chosen function, and we estimate $E = U - W$ and $u - W$.

We now introduce the functions ω, η

$$\omega(x,t) = P_E^n u(x,t), \quad \eta = u - \omega, \quad (x,t) \in \Omega \times I_n, \quad n = 0, \ldots, N-1.$$

Obviously these functions are continuous with respect to t in each time interval I_n, and have jump discontinuities at the point t^n, only if $V_h^{n-1} \neq V_h^n$.

3.3.1 The basic error equation

We define the functions

$$\begin{aligned}
W^n &= P_E^{n-1} u(t^n), \\
W^{n+} &= P_E^n u(t^n), \\
W^{n+1} &= P_E^n u(t^{n+1}), \\
W^{n,1} &= \frac{1}{2}(W^{n+1} + W^{n+}) = \frac{1}{2} P_E^n (u(t^{n+1}) + u(t^n)), \\
E^{n,1} &= U^{n,1} - W^{n,1}, \\
E^{n+} &= U^{n+} - W^{n+}, \\
E^{n+1} &= U^{n+1} - W^{n+1} = 2\, E^{n,1} - E^{n+}.
\end{aligned}$$

In view of (2.13) and (2.14), we need only consider the term $U - W$. It is easy to see from (3.10), (1.2), (2.8) and the definition of W that the error $E = U - W$ satisfies the equation

$$\begin{aligned}
(E^{n,1} - E^{n+}, \psi) + \mathrm{i}\,\frac{k_n}{2} a_{h_n}^{\gamma}(E^{n,1}, \psi) - \mathrm{i}\,\lambda \frac{k_n}{2}(f(U^{n,1}) - f(W^{n,1}), \psi) \\
= (\Theta^n + A^n + B^n, \psi) + \mathrm{i}\,\lambda \frac{k_n}{2}(f(W^{n,1}) - f(u(t^{n,1}), \psi),
\end{aligned} \tag{3.17}$$

where

$$\Theta^n = \frac{1}{2}(\eta(t^{n+1}) - \eta(t^{n+})), \quad (\eta(t^{n+}) = \lim_{t \to t^{n+}} \eta(t)), \tag{3.18}$$

$$A^n = -\frac{1}{2}(u(t^{n+1}) - u(t^n) - k_n u_t(t^{n,1})), \tag{3.19}$$

$$B^n = \frac{1}{4} k_n \Delta[u(t^{n+1}) + u(t^n) - 2u(t^{n,1})]. \tag{3.20}$$

Our next task is to bound the terms Θ^n, A^n and B^n. In the following lemma, we derive appropriate estimates for these terms, which allow us to obtain the optimal order convergence result of our scheme.

LEMMA 3.21 *For any n, $0 \le n \le N-1$ there holds*

$$\|\Theta^n\| \le c k_n^{1/2} \|h_n^r u_t\|_n, \tag{3.22}$$

$$\|A^n\| \le c k_n^{5/2} \|u^{(3)}\|_n, \tag{3.23}$$

$$\|B^n\| \le c k_n^{5/2} \|\Delta u^{(2)}\|_n\,, \tag{3.24}$$

where Θ^n, A^n and B^n are defined in (3.18)- (3.20).

Proof. The estimate (3.22) follows easily from

$$\Theta^n = \frac{1}{2}\int_{t^n}^{t^{n+1}} \eta_t(s)ds,$$

and (2.14). Using Taylor's theorem we have,

$$A^n = -\frac{1}{4}\int_{t^n}^{t^{n+1}} (t^{n+1}-s)^2 u^{(3)}\,ds + \frac{k_n}{2}\int_{t^n}^{t^{n+1}} (t^{n,1}-s)\; u^{(3)}\,ds,$$

from which (3.23) follows easily. (3.24) is obtained in a similar way. □

3.3.2 Convergence of the modified scheme

We now set $\psi = E^{n,1}$ in (3.17) and take real parts to obtain

$$\begin{aligned} &\|E^{n,1}\|^2 - \mathrm{Re}(E^{n+}, E^{n,1}) + \lambda\frac{k_n}{2}\mathrm{Im}\,(f(U^{n,1}) - f(W^{n,1}), E^{n,1}) \\ &= \mathrm{Re}\,(\Theta^n + A^n + B^n, E^{n,1}) + \lambda\frac{k_n}{2}\mathrm{Im}\,(f(u(t^{n,1})) - f(W^{n,1}), E^{n,1}). \end{aligned} \tag{3.25}$$

Using (3.12c) we obtain

$$\left|(f(U^{n,1}) - f(W^{n,1}), E^{n,1})\right| \le c\|E^{n,1}\|^2, \tag{3.26}$$

and

$$\left|(f(u(t^{n,1})) - f(W^{n,1}), E^{n,1})\right| \le c\|u(t^{n,1}) - W^{n,1}\|\|E^{n,1}\|. \tag{3.27}$$

Now using Taylor's theorem and (2.14), we obtain

$$
\begin{aligned}
\|u(t^{n,1}) - W^{n,1}\| &\leq \|u(t^{n,1}) - \frac{1}{2}[u(t^n) + u(t^{n+1})]\| \\
&\qquad + \frac{1}{2}\|[I - P_E^n][u(t^n) + u(t^{n+1})]\| \\
&\leq ck_n^2 + c\|h_n^r[u(t^n) + u(t^{n+1})]\|.
\end{aligned}
\tag{3.28}
$$

Using (3.22)–(3.24) and (3.26)–(3.28) in (3.25) yields

$$
\begin{aligned}
\|E^{n,1}\|^2 &\leq \|E^{n+}\|\|E^{n,1}\| + ck_n\|E^{n,1}\|^2 \\
&\quad + ck_n^{1/2}\left\{|||h_n^r u_t|||_n + k_n^2|||u^{(3)}|||_n + k_n^2|||\Delta u^{(2)}|||_n\right\}\|E^{n,1}\| \\
&\quad + ck_n\left\{k_n^2 + \|h_n^r[u(t^n) + u(t^{n+1})]\|\right\}\|E^{n,1}\|.
\end{aligned}
\tag{3.29}
$$

Thus,

$$
\|E^{n,1}\|^2 \leq c\|E^{n+}\|^2 + ck_n\mathcal{E}_n,
\tag{3.30}
$$

where $\mathcal{E}_n$ denotes the optimal (consistency) term

$$
\mathcal{E}_n = |||h_n^r u_t|||_n^2 + k_n^4|||u^{(3)}|||_n^2 + k_n^4|||\Delta u^{(2)}|||_n^2 + ck_n\left\{k_n^4 + \|h_n^r[u(t^n) + u(t^{n+1})]\|^2\right\}.
$$

In order to obtain an estimate for E^{n+1}, we note that

$$
\|E^{n,1}\|^2 - \mathrm{Re}(E^{n+}, E^{n,1}) = \frac{1}{4}\|E^{n+1}\|^2 - \frac{1}{4}\|E^{n+}\|^2.
$$

Hence,

$$
\begin{aligned}
\|E^{n+1}\|^2 &\leq \|E^{n+}\|^2 + ck_n\|E^{n,1}\|^2 + c\left\{|||h_n^r u_t|||_n^2 + k_n^4|||u^{(3)}|||_n^2 + k_n^4|||\Delta u^{(2)}|||_n^2\right\} \\
&\quad + ck_n\left\{k_n^4 + \|h_n^r[u(t^n) + u(t^{n+1})]\|^2\right\} \\
&\leq (1 + ck_n)\|E^{n+}\|^2 + c\mathcal{E}_n.
\end{aligned}
\tag{3.31}
$$

We next consider the term $\|E^{n+}\|$. Note that since $\Pi^n = P^n$ here,

$$
E^{n+} = U^{n+} - W^{n+} = P^n E^n - P^n J[\omega^n],
\tag{3.32}
$$

where $J[\omega^n] = (P_E^n - P_E^{n-1})u(t^n)$ is the "jump" in the elliptic projection of u across t^n. Hence it follows from (3.31) and (3.32) that

$$
\|E^{n+1}\|^2 \leq (1 + ck_n + \beta_n)\|E^n\|^2 + c\mathcal{E}_n + M_n\|J[\omega^n]\|^2,
\tag{3.33}
$$

where M_n is a number depending on n and which will be specified in the sequel and

$$\beta_n = \begin{cases} 0 & \text{if } V_h^n = V_h^{n-1} \\ \frac{1}{M_n - 1} & \text{otherwise.} \end{cases} \qquad n = 1, \ldots, N-1. \tag{3.34}$$

We are now ready to prove the main convergence result for the modified scheme.

THEOREM 3.35 *Let u be the solution of* (1.2). *If $\{U^n\}_{n=0}^N$ are given by* (3.10), *then*

$$\begin{aligned} \|E^n\| \le C_n \Big\{ \|h_0^r u^0\|_{r,h} &+ \Big(\sum_{m=0}^{n-1} \Big[|||h_m^r u_t|||_m^2 + k_m^4 |||u^{(3)}|||_m^2 + k_m^4 |||\Delta u^{(2)}|||_m^2 \Big] \Big)^{1/2} \\ &+ \Big(\sum_{m=0}^{n-1} \Big[k_m^5 + k_m \|h_m^r u(t^n)\|^2 \Big] \Big)^{1/2} + \Big(\mathcal{N}_C(n-1) \sum_{m=1}^{n-1} \|J[\omega^m]\|^2 \Big)^{1/2} \Big\}, \end{aligned} \tag{3.36}$$

$n = 0, \ldots, N-1$. *Here $C_n = ce^{ct^n}$, and $\mathcal{N}_C(n)$ denotes the number of times where $V_h^j \neq V_h^{j-1}$, $j = 1, \ldots n$, $(\mathcal{N}_C = \mathcal{N}_C(N-1))$.*

Proof. From (3.33) we have

$$\begin{aligned} \|E^{n+1}\|^2 \le \prod_{j=0}^{n} (1 + \beta_j + ck_j) \|E^0\|^2 &+ c \sum_{m=0}^{n} \Big(\prod_{j=m+1}^{n} (1 + \beta_j + ck_j) \Big) \\ &\times \Big(\mathcal{E}_m + M_m \|J[\eta^m]\|^2 \Big). \end{aligned} \tag{3.37}$$

Now fix n and choose $M_m = M = \mathcal{N}_C(n)$, $m = 1, \ldots, n$, where $\mathcal{N}_C(n)$ is the number of times where $V_h^j \neq V_h^{j-1}$, $j = 1, \ldots n$ (in the case where $\mathcal{N}_C(n) = 0$ or 1 we take $M = 2$.) Then, $\beta_j = \beta = \frac{1}{M-1}$, for $V_h^j \neq V_h^{j-1}$, cf. (3.34),

$$\begin{aligned} \prod_{j=0}^{n} (1 + \beta_j + ck_j) &\le \prod_{\substack{j=0 \\ \beta < ck_j}}^{n} (1 + 2ck_j) \prod_{\substack{j=0 \\ \beta \ge ck_j}}^{n} (1 + 2\beta) \\ &\le \prod_{j=0}^{n} (1 + 2ck_j)(1 + 2\beta)^M \le e^{2ct^{n+1}} 3e^2. \end{aligned} \tag{3.38}$$

Set $C_n := c \left(3e^{2ct^n + 2} \right)^{1/2}$. Then since $J[\omega^0] = 0$ and $E^0 = u^0 - P_E^0 u^0$, from (2.14) we obtain

$$\|E^{n+1}\| \le C_{n+1} \Big\{ \|h_0^r u^0\|_{r,h} + \Big(\sum_{m=0}^{n} \mathcal{E}_m \Big)^{1/2} + \sqrt{\mathcal{N}_C(n)} \Big(\sum_{m=1}^{n} \|J[\omega^m]\|^2 \Big)^{1/2} \Big\}.$$

which is the required result. □

3.3.3 Convergence analysis for the fully discrete scheme

We let $\psi = E^{n,1}$ in (3.17) and take imaginary parts. From the coercivity of the bilinear form $a^{\gamma}_{h_n}(\cdot,\cdot)$ and (3.29), (3.30) it follows that

$$
\begin{aligned}
k_n||E^{n,1}||^2_{1,h} &\le c\{||E^{n,1}||^2 + ||E^{n+}||\,||E^{n,1}|| + \mathcal{E}_n\} \\
&\le c\{||E^{n+}||^2 + \mathcal{E}_n\}.
\end{aligned}
$$

Using (3.33) and (3.36) we have

$$
k_n||E^{n,1}||^2_{1,h} \le C_n \max_{0\le m\le n}\{k_m^4 + h_m^{2r}\}. \tag{3.39}
$$

Using the inverse inequality (3.43) and the above estimate we get

$$
||E^{n,1}||_\infty \le ck_n^{-1/2}|\ln \underline{h}_n| \max_{0\le m\le n}\{k_m^4 + h_m^{2r}\}. \tag{3.40}
$$

We next bound the term $||W^{n,1}||_\infty$. For this, let $u^{n,1} := \frac{1}{2}[u(t^n) + u(t^{n+1})]$ and let $\chi^{n,1}$ be an element of V_h^n to be chosen appropriately. From the triangle inequality we have,

$$
\begin{aligned}
||W^{n,1}||_\infty &\le ||u^{n,1}||_\infty + ||[I - P_E^n]u^{n,1}||_\infty \\
&\le ||u^{n,1}||_\infty + ||u^{n,1} - \chi^{n,1}||_\infty + ||\chi^{n,1} - P_E^n u^{n,1}||_\infty.
\end{aligned}
$$

Using the inverse inequality (3.43) and the triangle inequality we have,

$$
\begin{aligned}
||W^{n,1}||_\infty &\le ||u^{n,1}||_\infty + ||u^{n,1} - \chi^{n,1}||_\infty + c|\ln h_n|\,||\chi^{n,1} - P_E^n u^{n,1}||_{1,h} \\
&\le ||u^{n,1}||_\infty + ||u^{n,1} - \chi^{n,1}||_\infty \\
&\quad + c|\ln \underline{h}_n| \left\{||\chi^{n,1} - u^{n,1}||_{1,h} + ||u^{n,1} - P_E^n u^{n,1}||_{1,h}\right\}
\end{aligned}
$$

We now choose $\chi^{n,1}$ so that

$$
||\chi^{n,1} - u^{n,1}||_{1,h} \le c||h_n^{r-1}u^{n,1}||_{r,h} \tag{3.41}
$$

$$
||\chi^{n,1} - u^{n,1}||_\infty \le ch_n^r||u^{n,1}||_{r,\infty} \tag{3.42}
$$

The fact that we may chose $\chi^{n,1}$ having the above approximation properties follows from the approximation properties of V_h^n and the trace inequality (2.5). The approximation property of $\chi^{n,1}$ in the $||\cdot||_\infty$-norm follows from similar results in [7]. We then conclude that

$$||U^{n,1}||_\infty \le ||W^{n,1}||_\infty + ||E^{n,1}||_\infty \le M+1$$

for h and k sufficiently small. This in turn implies that the solution U of (3.10) is also a solution of (3.6) which also satisfies the error estimate of Theorem 3.35.

3.4 Some technical lemmas

The main purpose of this section is to establish the inverse inequality

$$(3.43) \qquad ||\chi||_{L^\infty(\Omega)} \le |\ln \underline{h}_n|^{1/2}\, ||\chi||_{1,h}, \quad \forall\, \chi \in V_h^n$$

used in Section 3.3.3. The proof proceeds along the lines of a proof of a similar inequality that in known to hold in the case $E_h \subseteq H_0^1(\Omega)$. To start we establish the interpolation-type inequality (3.55) for elements of the energy space E_h. In view of the nature of the norm $||\cdot||_{1,h}$, certain trace inequalities are made use of. Next the L^∞-L^p inverse inequality (3.75) is proved, thereby extending a similar result to elements having one curved side. The desired inequality (3.43) then follows easily by appropriately choosing p.

We start by establishing some useful trace estimates for functions in $W^{1,r}(D)$, where D is a Lipschitz domain. We are particularly interested in assessing the dependence of the constants appearing in these bounds on D and r. Moreover, the proofs that are provided, which are elementary and for the most part well-known, require no more regularity on the part of D than we have already assumed in properties (i)–(vi) in Chapter 2. Also, density arguments will enable us to consider only elements of $C^1(\overline{D})$. Indeed, for $v \in C^1(\overline{D})$, integration by parts gives

$$(3.44) \qquad \begin{aligned} \int_{\partial D} (\mathbf{x}-\mathbf{z})\cdot\mathbf{n}\, v^r\, d\sigma &= \sum_{i=1}^{2} \int_D \frac{\partial}{\partial x_i}\left[(x_i - z_i)v^r\right]\, dx \\ &= \sum_{i=1}^{2} \int_D \left[u^r + (x_i - z_i)\, r v^{r-1}\frac{\partial v}{\partial x_i}\right]\, dx, \end{aligned}$$

for any rational $r = \frac{m}{n} > 1$ with m even and n odd, and $\mathbf{z} \in D$. If $\mathbf{z}$ is as in property (iii), using Hölder's inequality in (3.44) we obtain

$$\mu_1 h |v|^r_{L^r(\partial D)} \leq 2||v||^r_{L^r(D)} + rh2^{1/2}||v||^{r-1}_{L^{(r-1)q}(D)}||\nabla v||_{L^{q'}(D)}, \tag{3.45}$$

with $h = \operatorname{diam}(D)$, $q' = \frac{q}{q-1}$ and $q \geq 1$ satisfies $q(r-1) \geq 1$. Thus,

$$|v|_{L^r(D)} \leq \left(\frac{2}{\mu_1}\right)^{1/r} h^{-\frac{1}{r}}||v||_{L^r(D)} + \left(\frac{r2^{1/2}}{\mu_1}\right)^{1/r} ||v||^{\frac{r-1}{r}}_{L^{2(r-1)}(D)}||\nabla v||^{1/r}_{L^2(D)}. \tag{3.46}$$

Since such rationals are dense in the set of real numbers, (3.46) extends by continuity to all $r \in (1, \infty]$. Letting $r = q = 2$ in (3.46), we obtain

$$|v|^r_{L^r(\partial D)} \leq c\left\{h_i^{-1/2}||v||_{L^2(D)} + ||v||^{1/2}_{L^2(D)}||\nabla v||^{1/2}_{L^2(D)}\right\}. \tag{3.47}$$

On the other hand, with $r = p' = \frac{p}{p-1}$, $q = p$, we obtain

$$\begin{aligned} |v|_{L^{p'}(D)} &\leq ch_i^{-1+1/p}||v||_{L^{p'}(D)} + c||v||^{1/p}_{L^{p'}(D)}||\nabla v||^{1-1/p}_{L^{p'}(D)} \\ &\leq ch_i^{-1+1/p}||v||_{L^{p'}(D)} + ch_i^{1/p}||\nabla v||_{L^{p'}(D)}. \end{aligned} \tag{3.48}$$

Now Hölder's inequality with $m = \frac{2p}{p-2}$ yields

$$||v||_{L^{p'}(D)} \leq ch||v||_{L^m(D)}. \tag{3.49}$$

Using this in (3.48) gives the estimate

$$|v|_{L^{p'}(D)} \leq ch^{1/p}\left\{||v||_{L^m(D)} + ||\nabla v||_{L^{p'}(D)}\right\}. \tag{3.50}$$

As our next result, we shall establish the estimate,

$$\begin{aligned} |v^{(i)} - v^{(j)}|_{L^p(\partial\Omega^h_{ij})} \leq{}& ch^{-1/m}|v^{(i)} - v^{(j)}|_{\partial\Omega^h_{ij}} \\ &+ ch_i^{1/p}\left(||\nabla v^{(i)}||_{\partial\Omega^h_i} + ||\nabla v^{(j)}||_{\partial\Omega^h_j}\right), \end{aligned} \tag{3.51}$$

where m is as above and Ω^h_i and Ω^h_j are two adjacent elements. For simplicity only, we assume that Ω^h_i and Ω^h_j are triangles and that they share a full edge $\partial\Omega^h_{ij}$. Now let $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be an invertible affine function that maps Ω^h_j onto Ω^h_i and such that $F(x) = x, \quad \forall\, x \in \partial\Omega_{ij}^h$. Using F, we extend $v^{(j)}$ into a function $\tilde{v}^{(j)}$ defined on $\overline{\Omega_i^h \cap \Omega_j^h}$. It follows from the mesh property (iv) that for some constant c independent of h_i,

$$|\tilde{v}^{(j)}|_{s,\Omega_i^h} \leq c|v^{(j)}|_{s,\Omega_j^h}, \quad s = 0, 1.$$

Letting $e = v^{(i)} - v^{(j)} = v^{(i)} - \tilde{v}^{(j)}$ and taking $r = p$, $q = 2$ in (3.46) we obtain,

$$(3.52)\ |e|_{L^p(\partial\Omega_{ij}^h)} \leq ch_i^{-1/p}||e||_{L^p(\Omega_i^h)} + c||e||_{L^{2(p-1)}(\Omega_i^h)}^{\frac{p-1}{p}}||\nabla e||_{L^2(\Omega_i^h)}^{1/p}$$
$$\leq ch_i^{-1/p}||e||_{L^p(\Omega_i^h)} + ch^{-\frac{1}{p(p-1)}}||e||_{L^{2(p-1)}(\Omega_i^h)} + ch^{1/p}||\nabla e||_{L^2(\Omega_i^h)}^{1/p}.$$

To estimate the terms $||e||_{L^p(\Omega_i^h)}$ and $||e||_{L^{2(p-1)}(\Omega_i^h)}$, we use the family of imbedding inequalities

$$(3.53) \qquad ||v||_{L^q(D)} \leq m(D)^{-\frac{q-2}{2q}}||v||_{L^2(D)} + cq||v||_{L^2(D)}^{1-\frac{q-2}{q}}||\nabla v||_{L^2(D)}^{\frac{q-2}{q}}$$
$$\leq ch^{-1+\frac{2}{q}}||v||_{L^2(D)} + cqh^{\frac{2}{q}}||\nabla v||_{L^2(D)},$$

valid for $q \in [2, \infty)$ (cf. [2]). The linear dependence on q can be deduced from the proofs presented in the above reference. Now using (3.51) with $q = 2(p-1)$, we obtain

$$||v||_{L^{2(p-1)}(D)} \leq ch^{-\frac{p-2}{p-1}}||v||_{L^2(D)} + cph^{\frac{1}{p-1}}||\nabla v||_{L^2(D)}.$$

Using this together with (3.50) and $q = p$, we get

$$|e|_{L^p(\partial\Omega_{ij}^h)} \leq ch_i^{-1+1/p}||e||_{L^2(\Omega_i^h)} + cph_i^{1/p}||\nabla e||_{L^2(\Omega_i^h)}.$$

To complete the proof, we use the Poincaré type inequality

$$||e||_{L^2(\Omega_i^h)} \leq ch^{1/2}|e|_{L^2(\partial\Omega_{ij}^h)} + ch_i||\nabla e||_{L^2(\Omega_i^h)},$$

whose proof follows easily form (3.44).

We are now ready to prove the following estimate for the elements of E_h:

PROPOSITION 3.54 *Let $2 \leq p < \infty$. Then,*

$$(3.55) \qquad ||v||_{L^p(\Omega)} \leq cp\,||v||_{1,h}, \qquad \forall\, v \in E_h,$$

for some constant c independent of p and h.

Proof. In view of the imbedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$, valid for $d \le 3$, we have that $E_h \subset L^q(\Omega)$, for all $q \ge 1$. Now for $p \in [2,\infty)$ and $v \in E_h$, consider the boundary value problem

$$(3.56) \qquad \begin{cases} -\Delta\Phi = |v|^{p-2}v, & \text{in } \Omega, \\ \Phi = 0, \quad \text{on } \partial\Omega. \end{cases}$$

Since Ω is bounded, $|v|^{p-2}v \in L^q(\Omega)$, $\forall\, q \ge 1$. Thus a unique solution to (3.56) exists which is in $W^{2,q}(\Omega) \cap H_0^1(\Omega)$, $1 < q < \infty$, (cf. [20]). Furthermore, for some constant $c = c(\Omega)$,

$$(3.57) \qquad ||\Phi||_{2,p',\Omega} \le cp\,||\,|v|^{p-2}v\,||_{L^{p'}(\Omega)} = cp\,||v||^{p-1}_{L^p(\Omega)}, \quad p' = \frac{p}{p-1} \in (1,2].$$

The linear dependence of the constant in (3.57) on p is a consequence of estimates on Riesz potentials (cf. e.g. [39]). It follows form (3.56) that

$$(3.58) \qquad ||v||^p_{L^p(\Omega)} = a_h^\gamma(\Phi, v) = \sum_{i=1}^{d_h}\Bigg\{(\nabla\Phi^{(i)}, \nabla v^{(i)})_{\Omega_i^h} - \sum_{j\in\mathcal{N}_i} \tau_{ij}\langle\frac{\partial\Phi^{(i)}}{\partial n}, v^{(i)} - v^{(j)}\rangle_{\partial\Omega_{ij}^h} - \langle\frac{\partial\Phi^{(i)}}{\partial n}, v^{(i)}\rangle_{\partial\Omega_i^{*h}}\Bigg\}.$$

Using Schwarz's inequality and the imbedding $W^{2,p'}(\Omega) \hookrightarrow H^1(\Omega)$, we get

$$(3.59) \qquad \sum_{i=1}^{d_h}(\nabla\Phi^{(i)}, \nabla v^{(i)})_{\Omega_i^h} \le ||\Phi||_1||v||_{1,h} \le c\,||\Phi||_{2,p',\Omega}||v||_{1,h}.$$

Furthermore, using Hölder's inequality, (3.51) and (3.50) with $m = \frac{2p}{p-2} \in (2,\infty]$,

$$(3.60) \qquad \begin{aligned} \langle\frac{\partial\Phi^{(i)}}{\partial n}, v^{(i)} - v^{(j)}\rangle_{\partial\Omega_{ij}^h} &\le \left|\frac{\partial\Phi^{(i)}}{\partial n}\right|_{L^{p'}(\partial\Omega_{ij}^h)} |v^{(i)} - v^{(j)}|_{L^p(\partial\Omega_{ij}^h)} \\ &\le ch_i^{\frac{1}{p}}\left[||\Phi^{(i)}||_{1,m,\Omega_i^h} + ||\Phi^{(i)}||_{2,p',\Omega_i^h}\right] \times \\ &\times \left[h_i^{-\frac{1}{m}}|v^{(i)} - v^{(j)}|_{\partial\Omega_{ij}^h} + h_i^{\frac{1}{p}}\left(||\nabla v^{(i)}||_{\Omega_i^h} + ||\nabla v^{(j)}||_{\Omega_j^h}\right)\right]. \end{aligned}$$

Using Hölder's inequality with $m' = \frac{m}{m-1} = \frac{2}{p+2} \in [1,2]$,

$$(3.61) \qquad \sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{1}{p}-\frac{1}{m}}||\Phi^{(i)}||_{1,m,\Omega_i^h}|v^{(i)} - v^{(j)}|_{\partial\Omega_{ij}^h} \le$$
$$\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}||\Phi^{(i)}||^m_{1,m,\Omega_i^h}\right)^{1/m}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{(\frac{1}{p}-\frac{1}{m})m'}|v^{(i)} - v^{(j)}|^{m'}_{\partial\Omega_{ij}^h}\right)^{1/m'}.$$

Further use of Hölder's inequality yields

$$(3.62)\quad \left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{(\frac{1}{p}-\frac{1}{m})m'}|v^{(i)}-v^{(j)}|^{m'}_{\partial\Omega^h_{ij}}\right)^{1/m'}\leq\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^2\right)^{1/p}\times$$
$$\times\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{-1}|v^{(i)}-v^{(j)}|^2_{\partial\Omega^h_{ij}}\right)^{1/2}.$$

Since the cardinality of $\mathcal{N}_i$ is uniformly bounded,

$$(3.63)\qquad \left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}||\Phi||_{1,m,\Omega}\right)^{1/m}\leq c||\Phi||_{2,p',\Omega},$$

using the imbedding $W^{2,p'}(\Omega)\hookrightarrow W^{1,m}(\Omega)$, valid for $d\leq 3$. It also follows from the mesh properties (iv) and (v) that

$$(3.64)\qquad \sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^2\leq c\,m(\Omega).$$

Using (3.62), (3.63) and (3.64) in (3.61), we get

$$(3.65)\qquad \sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{1}{p}-\frac{1}{m}}||\Phi^{(i)}||_{1,m,\Omega_i^h}|v^{(i)}-v^{(j)}|_{\partial\Omega^h_{ij}}\leq c||\Phi||_{2,p',\Omega}||v||_{1,h}.$$

Similarly, using the inequality $\left(\sum_i|a_i|^q\right)^{1/q}\leq\left(\sum_i a_i^2\right)^{1/2},\quad q\geq 2,$

$$(3.66)\quad \sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{1}{p}-\frac{1}{m}}||\Phi||_{2,p',\Omega}|v^{(i)}-v^{(j)}|_{\partial\Omega^h_{ij}}\leq$$
$$\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}||\Phi||^{p'}_{2,p',\Omega}\right)^{1/p'}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{2-\frac{p}{2}}|v^{(i)}-v^{(j)}|_{\partial\Omega^h_{ij}}\right)^{1/p}$$
$$\leq c||\Phi||_{2,p',\Omega}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{2-\frac{p}{2}}|v^{(i)}-v^{(j)}|^p_{\partial\Omega^h_{ij}}\right)^{1/p}\leq c||\Phi||_{2,p',\Omega}||v||_{1,h}.$$

Furthermore, from (3.63) and using Hölder's inequality,

$$(3.67)\quad \sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}}||\Phi||_{1,m,\Omega_i^h}||\nabla v^{(i)}||_{\Omega_i^h}\leq$$
$$\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}||\Phi||^m_{1,m,\Omega_i^h}\right)^{1/m}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}m'}||\nabla v^{(i)}||^{m'}_{\Omega_i^h}\right)^{1/m'}$$
$$\leq c||\Phi||_{2,p',\Omega}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}m'}||\nabla v^{(i)}||^{m'}_{\Omega_i^h}\right)^{1/m'}$$

Using Hölder's inequality,

(3.68)

$$\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}m'}||\nabla v^{(i)}||_{\Omega_i^h}^{m'}\right)^{1/m'} \le \left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^2\right)^{1/p}\left(\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}|\nabla v^{(i)}||_{\Omega_i^h}^2\right)^{1/2}$$

$$\le c[m(\Omega)]^{1/p}||v||_{1,h}$$

Hence,

$$\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}}||\Phi^{(i)}||_{1,m,\Omega_i^h}||\nabla v^{(i)}||_{\Omega_i^h} \le cp\,||\Phi||_{2,p',\Omega}||v||_{1,h} \tag{3.69}$$

Similar arguments yield

$$\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}}||\Phi^{(i)}||_{2,p',\Omega_i^h}||\nabla v^{(i)}||_{\Omega_i^h} \le cp\,||\Phi||_{2,p',\Omega}||v||_{1,h} \tag{3.70}$$

and

$$\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}h_i^{\frac{2}{p}}\left(||\Phi^{(i)}||_{1,m,\Omega_i^h}+||\Phi^{(i)}||_{2,p',\Omega_i^h}\right)||\nabla v^{(j)}||_{\Omega_j^h} \le c||\Phi||_{2,p',\Omega}||v||_{1,h} \tag{3.71}$$

Using (3.69), (3.70) and (3.71), we get

$$\sum_{i=1}^{d_h}\sum_{j\in\mathcal{N}_i}\tau_{ij}\langle\frac{\partial\Phi^{(i)}}{\partial n}, v^{(i)}-v^{(j)}\rangle_{\partial\Omega_i^h} \le cp\,||v||_{L^p(\Omega)}^{p-1}||v||_{1,h}. \tag{3.72}$$

In an entirely analogous manner it can be shown that

$$\sum_{i=1}^{d_h}\langle\frac{\partial\Phi^{(i)}}{\partial n}, v^{(i)}\rangle_{\partial\Omega_i^h} \le c||\Phi||_{2,p',\Omega}||v||_{1,h}. \tag{3.73}$$

Using (3.72), (3.73) and (3.59) in (3.58), completes the proof. □

As a final result, we shall establish an inverse inequality satisfied by the elements of the finite dimensional subspaces $\mathbb{P}_r(D)$.

THEOREM 3.74 *Let D be an element with one curved side and let T_1 and T_2 be the inscribed, resp., circumscribed triangles. Then there exists a constant c depending on r such that*

$$||\chi||_{L^\infty(D)} \le ch^{-2/p}||\chi||_{L^p(D)}, \qquad \forall\,\chi\in\mathbb{P}_r(D) \tag{3.75}$$

where $h = diam(T_1)$.

Proof. We shall denote by $\widehat{T}_1$ the triangle with vertices $(0,0),(1,0)$ and $(0,1)$ and by $\widehat{T}_2$ the triangle with vertices $(0,0),(2,0)$ and $(0,2)$. Moreover, denote by $\widehat{D}$ the element with vertices $(0,0),(1,0)$ and $(0,1)$. Let $F : \hat{x} \in R^2 \mapsto F(\hat{x}) = B\hat{x} + b$ be an affine function mapping $\widehat{D}$ onto D and $\widehat{T}_i$ onto T_i, $i = 1,2$ (see Figure 3.1). Also, let $\chi(\hat{x}) = \chi(x)$ for $\chi \in \mathbb{P}_r(D)$. It is easy to see that

$$||\chi||_{L^\infty(D)} = ||\widehat{\chi}||_{L^\infty(\widehat{D})} \leq ||\chi||_{L^\infty(\widehat{T}_2)}. \tag{3.76}$$

We shall next establish that for some constant c, independent of χ

$$||\widehat{\chi}||_{L^\infty(\widehat{T}_2)} \leq c\,||\widehat{\chi}||_{L^\infty(\widehat{T}_1)}. \tag{3.77}$$

Assume not. Then there exists a sequence $\{\widehat{\chi}_m\} \subset \mathbb{P}_r(\widehat{T}_2)$ such that $||\widehat{\chi}_m||_{L^\infty(\widehat{T}_2)} = 1$ and $\lim_{m\to\infty} ||\widehat{\chi}_m||_{L^\infty(\widehat{T}_1)} = 0$. Since the unit ball in $(\mathbb{P}_r(\widehat{T}_2), ||\cdot||_{L^\infty(\widehat{T}_2)})$ is compact, we can extract a subsequence $\{\widehat{\chi}_{m'}\}$ that converges to some element $\widehat{\chi}_\infty$ of $\mathbb{P}_r(\widehat{T}_2)$. Obviously, $||\widehat{\chi}_\infty||_{L^\infty(\widehat{T}_1)} = 0$, which in turn implies that $\widehat{\chi}_\infty$ vanishes identically and contradicts the fact that $||\widehat{\chi}_\infty||_{L^\infty(\widehat{T}_2)} = 1$.

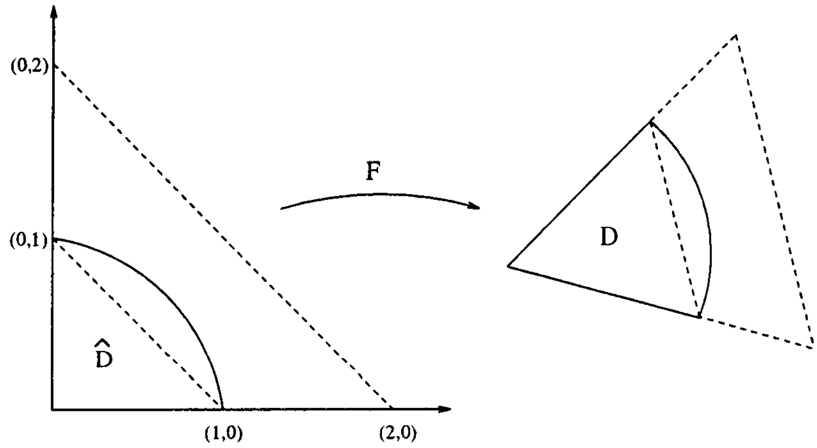


Figure 3.1: The affine map F.

All norms being equivalent on the finite dimensional space $\mathbb{P}_r(\widehat{T}_1)$, we have

$$||\widehat{\chi}||_{L^\infty(\widehat{T}_1)} \le c\,||\widehat{\chi}||_{L^p(\widehat{T}_1)}, \qquad 1 \le p \le \infty \tag{3.78}$$

for some constant c depending on r, $\widehat{T}_1$ and p. We shall show next that

$$\sup_{1\le p\le\infty} c(r, \widehat{T}_1, p) < \infty \tag{3.79}$$

Indeed, assuming the contrary we may find a sequence $\{\widehat{\chi}_m\} \subset \mathbb{P}_r(\widehat{T}_1)$ and a corresponding sequence $\{p_m\} \in [1,\infty]$ such that $||\widehat{\chi}_m||_{L^\infty(\widehat{T}_1)} = 1$ and $\inf_m ||\widehat{\chi}_m||_{L^{p_m}(\widehat{T}_1)} = 0$. Using Hölder's inequality we also have that $\inf_m ||\widehat{\chi}_m||_{L^1(\widehat{T}_1)} = 0$. Using a compactness argument as before, we arrive at a contradiction.

Finally, using Theorem 3.1.2 of [15] we obtain

$$||\widehat{\chi}||_{L^p(\widehat{T}_1)} \le ||\widehat{\chi}||_{L^p(\widehat{D})} \le ch^{-2/p}||\chi||_{L^p(D)} \tag{3.80}$$

where the constant c is independent of p and h but may depend on the "aspect ratio" of D, i.e., the ratio of the inscribed and circumscribed circles to D. The desired inverse inequality now follows from (3.78) and the estimate above. □

Remark. The above result, of course, simplifies greatly when D is a triangle.

The inverse inequality (3.43) now follows from Theorem 3.74 and Proposition 3.54 with the choice $p = |\ln \underline{h}_n|$.

Chapter 4

Implementation

In this chapter we address the practical issues of implementing a numerical method for the integration of the Nonlinear Schrödinger equation (henceforth referred to as the NLS equation), up to times close to the formation of singularities. We start by developing a code for the solution of an initial value problem for the NLS equation in a polygonal domain $\Omega \subset R^2$. At this stage, it is assumed that the solution does not experience a blow-up and that the starting mesh is sufficiently fine to produce accurate solutions. We discuss the computer representation of finite element meshes, the construction of the basis functions of our finite dimensional subspaces, and the solution of the resulting linear and nonlinear systems. To test our code, we verify experimentally the spatial convergence rates for the NLS equation by devising appropriate solutions and estimating the L^2–norms of the errors.

This implementation serves as the basis of an adaptive version of the code which allows for the automatic refinement of both the spatial and temporal meshes. The underlying refinement strategies amount to two computationally efficient and inexpensive 'tests' which are applied after each step in time. If the computed solution fails the temporal or spatial refinement tests then the time step is halved or the mesh is refined, and the solution is computed again. After describing the refinement strategy and the necessary modifications in the code, we proceed to test our method by following the evolution of several initial profiles which lead to blow-up in finite time. We consider both radially symmetric and non radially symmetric profiles, as well as profiles with several peaks, and compare our results with earlier numerical investigations of the blow-up in the NLS equation.

4.1 Finite Element Meshes and Basis Functions

Let us assume that $\mathcal{T}_h^0$ is an initial triangulation of the domain Ω, consisting of the triangles $\{\Omega_i^0\}_{i=1}^{d_0}$, and satisfying the assumptions stated in Chapter 2. We assume that all the book-keeping information necessary for the computer representation of this triangulation, e.g., enumeration of triangles, nodes and neighbors of each triangle, has been generated and stored in the form of an appropriate data base, from which it is easy to recover information about a specific triangle. Moreover, to facilitate the solution process and the construction of the basis functions of the associated finite element space, we shall assume that *metric* information about the current mesh, such as the location of each node, the lengths of the edges of each triangle and the areas of the triangles, has also been stored and can be readily made available.

With the triangulation $\mathcal{T}_h^0$, we associate the finite element space $V_h^0 \equiv V_h^{0,r}$, consisting of functions whose restrictions on the triangles Ω_i^0, $i = 1, \ldots, d_0$, are polynomials of degree $\leq r-1$. A construction of a basis of V_h^0 follows [15]. To this end, denote by $\hat{\Omega}$ the "reference triangle" with vertices $\hat{a}_1 = (1,0)$, $\hat{a}_2 = (0,1)$, $\hat{a}_3 = (0,0)$, and denote by $a_1 = (x_1, y_1)$, $a_2 = (x_2, y_2)$, $a_3 = (x_3, y_3)$ the vertices of Ω_i^0. There exists a unique, affine, invertible mapping

$$F_i : \hat{\mathbf{x}} \in R^2 \to F_i(\hat{\mathbf{x}}) = B_i \hat{\mathbf{x}} + b_i \tag{4.1}$$

satisfying $F_i(\hat{a}_j) = a_j, \quad 1 \leq j \leq 3$. In fact, it is easy to see that the matrix B_i and the vector b_i are given by

$$B_i = \begin{bmatrix} x_1 - x_3 & x_2 - x_3 \\ y_1 - y_3 & y_2 - y_3 \end{bmatrix}, \qquad b_i = \begin{bmatrix} x_3 \\ y_3 \end{bmatrix}$$

Furthermore, denote by $\mathbb{P}_r(\hat{\Omega})$ the space of polynomials functions defined over $\hat{\Omega}$ of degree $\leq r-1$, and assume $\{\hat{\phi}_l\}_{l=1}^{n_r}$ is a basis. The affine mapping (4.1) naturally induces a basis of $\{\tilde{\phi}_{i,l}\}_{l=1}^{n_r}$ of $\mathbb{P}_r(\Omega_i^0)$ by

$$\tilde{\phi}_{i,l} = \hat{\phi}_l \circ F_i, \qquad 1 \leq l \leq n_r \tag{4.2}$$

Then a basis of V_h^0 consists of the functions $\{\phi_{i,l}\}_{l=1}^{n_r}$, $1 \leq i \leq d_0$, defined by

$$\phi_{i,l}(\mathbf{x}) = \begin{cases} \tilde{\phi}_{i,l}(\mathbf{x}), & \mathbf{x} \in \Omega_i^0, \\ \mathbf{0} & \text{otherwise} \end{cases} \qquad 1 \leq l \leq n_r, \quad 1 \leq i \leq d_0 \tag{4.3}$$

It is shown in [15] that the dimension of $\mathbb{P}_r(\hat{\Omega})$ is $n_r = (r+1)(r+2)/2$. Since no continuity across interelement boundaries is enforced, the dimension of V_h^0 is then $N_r = d_0\, n_r$. As an example, we note that if $r = 2$, the basis functions of $\mathbb{P}_1(\hat{\Omega})$ may be chosen as

$$\lambda_1(\xi,\eta) = \xi, \qquad \lambda_2(\xi,\eta) = \eta, \qquad \lambda_3(\xi,\eta) = 1 - \xi - \eta,$$

These are precisely the functions that satisfy $\lambda_i(\hat{a}_j) = \delta_{ij}, \quad 1 \leq i,j \leq 3$. In turn, these functions can be used to construct a basis of $\mathbb{P}_r(\hat{\Omega})$ for $r \geq 2$. Indeed, it is easy to check that any polynomial $q(\mathbf{x})$ of degree ≤ 2 on $\hat{\Omega}$ may be written as

$$q = \sum_i \lambda_i(2\lambda_i - 1)q(\hat{a}_i) + \sum_{i<j} 4\lambda_i\lambda_j\, q(\hat{a}_{ij}),$$

where $\hat{a}_{ij} = \frac{1}{2}(\hat{a}_i + \hat{a}_j)$, $1 \leq i < j \leq 3$. From the practical point of view, it is clear from the above discussion that only the basis functions over the reference triangle need be computed. The basis functions over any triangle of $\mathcal{T}_h^0$ may then be determined from (4.1). In turn, this affine mapping is completely defined by the vertices of the triangle, which have already been stored.

4.2 Numerical solution of the fully discrete scheme

With the basis functions constructed in the previous section and with the triangulation $\mathcal{T}_h^0$ given, we focus our attention to the numerical solution of a fully discrete scheme for the initial and boundary value problem (1.2).

We discretize (1.2) in time by the *midpoint method*, or more precisely, the one-stage Gauss–Legendre method. Let $t_n = nk$, $n = 0, 1, \ldots, J$, where $t^* = Jk$. We seek $U^n \in V_h^0$ which approximates $u^n = u(\cdot, t^n)$ and such that

$$U^0 = \Pi_0 u^0 \tag{4.4}$$

where $\Pi_0 u^0$ denotes the L^2–projection of u^0 onto V_h^0. As stated already in Chapter 3, the approximation U^{n+1} is constructed from U^n by solving the following system for the intermediate unknown $U^{n,1}$

$$(U^{n,1}, \chi) + \frac{ik}{2} a_h^\gamma(U^{n,1}, \chi) = (U^n, \chi) + \frac{ik}{2}(f(U^{n,1}), \chi), \quad \forall\, \chi \in V_h^r,\ 0 \le n \le J-1, \tag{4.5}$$

and then setting

$$U^{n+1} = 2U^{n,1} - U^n. \tag{4.6}$$

At each time step we solve the nonlinear system represented by (4.5) using the so called Implicit–Explicit scheme, see [3]. Specifically, for $n \ge 0$, let $U_0^{n,1}$ be an initial guess for $U^{n,1}$. Then the iterates of the Implicit–Explicit scheme for (4.5) $U_{l+1}^{n,1}$, $l = 0, 1, \ldots,$ are defined by

$$(U_{l+1}^{n,1}, \chi) + \frac{ik}{2} a_h^\gamma(U_{l+1}^{n,1}, \chi) = (U^n, \chi) + \frac{ik}{2}(f(U_l^{n,1}), \chi), \quad \forall\, \chi \in V_h^r,\ 0 \le n \le J-1 \tag{4.7}$$

In practice, we performed only a small number of the Implicit–Explicit iterations and set $U^{n,1} = U_{l_n}^{n,1}$, where the integer l_n was equal to four for $n \le 1$ and to three for larger values of n. The starting values $U_0^{n,1}$ were computed by extrapolation from previous values as

$$U_0^{n,1} = \alpha_0 U^n + \alpha_1 U^{n-1} \tag{4.8}$$

The coefficients α_i are such that $U_0^{n,1}$ is the value at $t^n + k/2$ of the linear polynomial which interpolates the data U^{n-j} at the points t^{n-j}, $j = 0, 1$. For $n = 0$ we used $U^{n,1} =$

U^n and compensated for the reduced accuracy of the extrapolation by increasing the number of the Implicit–Explicit iterations to four. To obtain the explicit form of the linear system represented by (4.7) we write

$$U_l^{n,1} = \sum_{j=1}^{d_0} \sum_{m=1}^{n_r} c_{j,m}^l \phi_{j,m}, \qquad U_{l+1}^{n,1} = \sum_{j=1}^{d_0} \sum_{m=1}^{n_r} c_{j,m}^{l+1} \phi_{j,m}, \qquad U^n = \sum_{j=1}^{d_0} \sum_{m=1}^{n_r} c_{j,m} \phi_{j,m},$$

and set $\chi = \phi_{j,m}$, $m = 1, \ldots, n_r$, $j = 1, \ldots, d_0$, in (4.7). We thus obtain the system

$$\left[\begin{pmatrix} M^{(1,1)} & \cdots & M^{(1,d_0)} \\ \vdots & \ddots & \vdots \\ M^{(d_0,1)} & \cdots & M^{(d_0,d_0)} \end{pmatrix} + \frac{ik}{2} \begin{pmatrix} S^{(1,1)} & \cdots & S^{(1,d_0)} \\ \vdots & \ddots & \vdots \\ S^{(d_0,1)} & \cdots & S^{(d_0,d_0)} \end{pmatrix} \right] \begin{bmatrix} c_1^{l+1} \\ \vdots \\ c_{d_0}^{l+1} \end{bmatrix} \tag{4.9}$$

$$= \begin{pmatrix} M^{(1,1)} & \cdots & M^{(1,d_0)} \\ \vdots & \ddots & \vdots \\ M^{(d_0,1)} & \cdots & M^{(d_0,d_0)} \end{pmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_{d_0} \end{bmatrix} + \begin{bmatrix} f_1 \\ \vdots \\ f_{d_0} \end{bmatrix}$$

for the unknown vectors $c_j^{l+1} = \left[c_{j,1}^{l+1}, \ldots, c_{j,n_r}^{l+1} \right]^T$, $1 \le j \le d_0$. Here we have set,

$$f_j = [f_{j,m}, \ldots, f_{j,n_r}]^T, \qquad f_{j,m} = (f(U_l^{n,1}), \phi_{j,m}), \quad 1 \le m \le n_r$$

$$M_{l,m}^{(i,j)} = \int_\Omega \phi_{j,m} \phi_{i,l} \, d\mathbf{x}, \quad 1 \le i, j \le d_0$$

$$S_{l,m}^{(i,j)} = a^\gamma(\phi_{j,m}, \phi_{i,l}), \quad 1 \le i, j \le d_0$$

We may now summarize the steps involved in computing U^{n+1} from U^n in the following algorithm:

1. Form the matrices $M^{(i,j)}$, $S^{(i,j)}$ for $1 \le i, j \le d_0$.

2. Compute $U_0^{n,1}$ from (4.8).

3. For $l = 0, \ldots, l_n - 1$, solve (4.9) for c_j^{l+1}, $1 \le j \le d_0$.

4. Set $U^{n,1} = U_{l_n}^{n,1}$ and compute $U^{n+1} = 2U^{n,1} - U^n$.

Obviously the bulk of the computations are involved in steps 1 and 3. Fortunately, the matrices $M^{(i,j)}, S^{(i,j)}$ are independent of l and the time step k so they need only be computed once before starting the iterations. To shed some light into the structure of these matrices, let us denote by M the $d_0 \times d_0$ matrix whose (i,j)-entry is the $n_r \times n_r$ matrix $M^{(i,j)}$ and, similarly, denote by S the $d_0 \times d_0$ matrix whose (i,j)-entry is the $n_r \times n_r$ matrix $S^{(i,j)}$. We shall refer to M as the mass matrix and to S as the stiffness matrix. M is block diagonal, since the basis functions $\phi_{j,l}$, $l = 1, \ldots, n_r$ vanish identically outside the triangle Ω_j^0. Thus only the diagonal blocks $M^{(j,j)}$ need be computed and stored. Moreover, it is easy to see, using the affine mapping (4.1) that,

$$M_{l,m}^{(j,j)} = \int_{\Omega_j^0} \phi_{j,m}\phi_{j,l}\, d\mathbf{x} = J_j \int_{\hat{\Omega}} \hat{\phi}_m \hat{\phi}_l\, d\hat{\mathbf{x}} \tag{4.10}$$

where J_j denotes the Jacobian of the transformation F_j. Hence it suffices to compute $\hat{M}_{l,m} = \int_{\hat{\Omega}} \hat{\phi}_m \hat{\phi}_l\, d\hat{\mathbf{x}}$, $1 \leq m, l \leq n_r$. The elements of $M^{(j,j)}$ may then be found by multiplying the entries of $\hat{M}$ by J_j.

The stiffness matrix S is symmetric and positive definite since the bilinear form $a_h^\gamma(\cdot,\cdot)$ is symmetric and coercive for sufficiently large γ. Moreover, it follows from the definition of the bilinear form $a_h^\gamma(\cdot,\cdot)$, that the entry $S^{(i,j)}$ is nonzero if and only if the interface $\partial\Omega_{i,j}$ has positive one-dimensional measure. Hence, not only is S symmetric but sparse as well.

In the actual implementation of the scheme, we stored only the nonzero block-entries $S^{(i,j)}$ with $1 \leq j \leq i \leq d_0$, i.e., the nonzero entries of the *lower* triangle of S. The structure and bandwidth of S obviously depend on the enumeration of the triangles. Notice, however, that in the case of a conforming triangulation, each triangle has at most three neighbors so we expect the lower triangle of S to have at most three nonzero entries, excluding the diagonal blocks $S^{(i,i)}$. In the case of a general, non-conforming triangulation now, the number of neighbors of any given triangle may be larger than three, but in view of the mesh assumption (v) of Chapter 2, it cannot be arbitrarily large. In practice, the number of neighbors is, of course, a small finite number. In our implementation of an adaptive finite element code for the NLS we imposed the condition $\mu_3 = 2$, where μ_3 is the constant appearing in assumption (v) of Chapter 2, thus limiting the number of neighbors of each triangle to six. Figure 4.1 below, shows a typical structure of S for a uniform mesh of the domain $\Omega = [0,1]^2$, consisting of eight triangles. Each of the shaded squares represents a $n_r \times n_r$ matrix.

The integrals appearing in the definition of the bilinear form a_h^γ may be computed exactly, using numerical quadrature. For the integrals of the form $\int_{\Omega_i^0} \nabla\phi_{i,l}\nabla\phi_{i,m}\, d\mathbf{x}$ we use the affine mapping (4.1) to transform them onto integrals on the reference triangle and use a quadrature rule which is exact for polynomials of degree $\leq 2n_r$. The line integrals over the interfaces $\partial\Omega_{ij}$ or the boundary edges are also computed exactly by using a Gauss quadrature rule, exact for polynomials of one variable of degree $\leq 2n_r$.

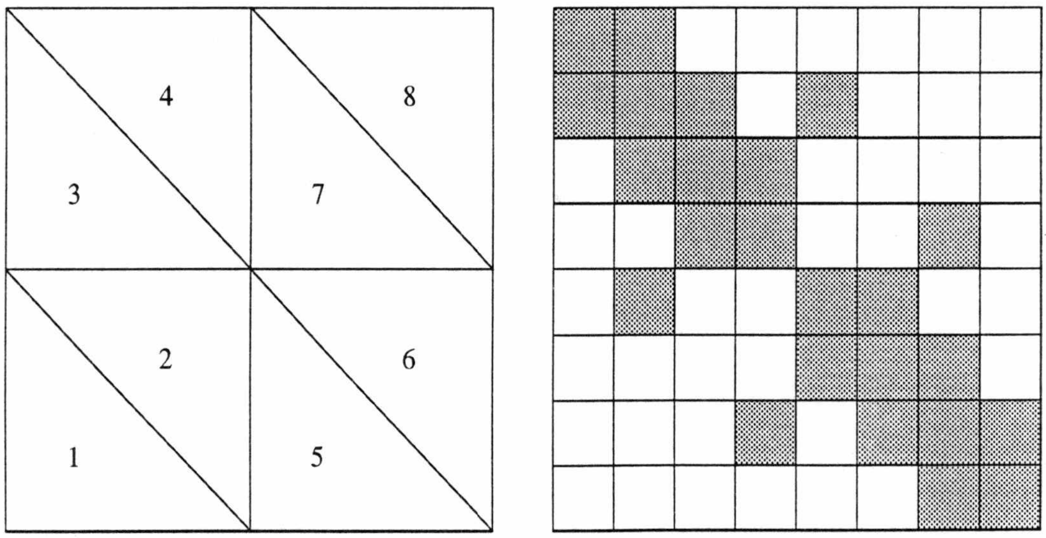


Figure 4.1: A triangulation of the domain $\Omega = [0,1]^2$ and the corresponding structure of the stiffness matrix S.

The sparsity of the matrices M and S is fully exploited in the code. To represent the mass matrix M we store only the $n_r \times n_r$ matrix $\hat{M}$ and the Jacobians J_l, $1 \leq l \leq d_0$, c.f. (4.10). For the stiffness matrix S we store the diagonal blocks $S^{(i,i)}$ $1 \leq i \leq d_0$ and the nonzero entries of the lower triangle of S. Clearly the number of nonzero entries in the lower triangle of S equals the number of interfaces $\partial\Omega_{ij}$. Thus, if NIE denotes the number of interfaces, the amount of storage required for the stiffness matrix is $(d_0 + \mathrm{NIE})\, n_r^2$. Since the diagonal blocks are symmetric, the storage may actually be reduced to $d_0 \frac{n_r(n_r+1)}{2} + \mathrm{NIE}\, n_r^2$. We chose to do that in our implementation.

Finally, the construction of the nonlinear term in the right-hand side of (4.7), amounts to computing the inner products $(f(U_l^{n,1}), \phi_{i,l})$ for $1 \leq l \leq n_r$ and all triangles Ω_i^0, $i = 1, \ldots, d_0$. This is done by mapping the triangle Ω_i^0 onto the reference triangle and using a quadrature rule exact for polynomials of degree ≤ 8.

For the solution of the complex, symmetric linear system (4.9) we used the Preconditioned Conjugate Gradient method, appropriately modified to handle complex matrices. The Jacobi and the Symmetric SOR matrices were used as preconditioners. Recall that if the symmetric matrix A is decomposed as

$$A = D + L + L^T$$

in its diagonal, lower, and upper triangular part, the Jacobi and Symmetric SOR matrices are defined as

$$P_{JAC} = D^{-1}, \qquad P_{SSOR} = (D + L)D^{-1}(D + L)^T$$

The Conjugate Gradient method involves one matrix-vector product, two inner products and three vector updates per iteration. The use of a preconditioner P, adds to the cost per iteration, the solution of a system of the form $Px = b$. If the Jacobi preconditioner is used then the matrix P is diagonal. In the case of preconditioning with the Symmetric SOR matrix, the solution of the system $Px = b$ involves solving two triangular systems and performing a matrix–vector product. Notice that in the case of the linear system (4.9) the diagonal blocks of the system matrix are $M^{(j,j)} +$

$\frac{ik}{2}S^{(j,j)}$. Both preconditioners necessitate that these matrices be factored. As long as the time step remains constant, this can be done once before starting stepping in time. In our implementation, we used the `LINPACK` subroutine `ZSPFA` to factor the diagonal blocks and the subroutine `ZSPSL` to solve systems involving these matrices.

We also took full advantage of the sparsity of the matrices M and S while performing the matrix-vector products. We multiplied only nonzero entries, thus reducing the operation count to two times the number of the nonzero blocks of $M + \frac{ik}{2}S$. Note that the size of the individual blocks is only n_r which is small compared with the dimension $N_r = d_0\, n_r$ of the system matrix $M + \frac{ik}{2}S$. We found that "inlining" the routines performing the matrix-vector products for the individual blocks contributed to a noticeable speed-up in the execution time of our code.

To test the performance of the preconditioners we constructed the system matrix $M + \frac{ik}{2}S$, corresponding to a uniform triangulation of the unit square $[0,1]^2$ with 288 triangles, and recorded the number of Conjugate Gradient iterations for one step of the Implicit–Explicit scheme (4.9). We used quadratic basis functions ($r = 3$) and $\gamma = 15$, resulting in a system of size 1728 by 1728. Results with different values of γ yielded entirely analogous results. The Conjugate Gradient iterations were stopped when the l_2-norm of the residual vector fell below 10^{-12}. Table 4.1 shows the number of iterations that were required, using both the Jacobi and the Symmetric SOR matrices as preconditioners, as well as the times required to solve the linear systems. The times were obtained on a SPARCstation 1 using the `etime` library routine. It can be seen from Table 4.1 that the number of Conjugate Gradient iterations is approximately halved when the Symmetric SOR matrix is used in place of the Jacobi matrix. However, the time needed to perform these iterations is not halved, since the Symmetric SOR preconditioner requires a larger overhead. We may also infer from Table 4.1 that the Jacobi preconditioned Conjugate Gradient method is highly competitive for large values of k^{-1}, but for smaller values, the Symmetric SOR preconditioned method is slightly faster despite the larger overhead.

We also studied the effect of the parameter γ on the number of Conjugate Gradient iterations. We performed the same experiments as above with $\gamma = 10, 15, 20, 25, 30$ and summarized the results in Table 4.2 and Table 4.3. Notice that γ has very little influence on the number of iterations for large values of k^{-1} but there is a significant effect for smaller values. This is, of course, explained by the fact that the entries of the stiffness matrix S are multiplied by the time step k and hence, if the time step is very small, the behavior of the system matrix $M + \frac{ik}{2}S$ is similar to that of M.

Table 4.1: Comparison of Jacobi and Symmetric SOR Preconditioners for $\gamma = 15$

	Number of iterations		Time (seconds)	
k^{-1}	Jacobi	SSOR	Jacobi	SSOR
10	142	66	9.54	8.43
20	129	61	8.71	7.75
40	113	52	7.75	6.74
80	95	42	6.51	5.50
100	82	41	5.70	5.48
200	66	29	4.64	3.97
400	47	21	3.46	3.03
800	35	15	2.70	2.31
1000	32	14	2.58	2.16

Table 4.2: Effect of the parameter γ on the number of Conjugate Gradient iterations with the Jacobi preconditioner.

	Number of iterations				
k^{-1}	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$	$\gamma = 25$	$\gamma = 30$
100	283	82	96	105	114
200	158	66	76	85	91
400	78	47	55	58	64
800	42	35	38	43	48
1600	26	26	30	33	35

Table 4.3: Effect of the parameter γ on the number of Conjugate Gradient iterations with the Symmetric SOR preconditioner.

	Number of iterations				
k^{-1}	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$	$\gamma = 25$	$\gamma = 30$
100	185	41	44	54	60
200	89	29	34	39	44
400	43	21	26	29	32
800	19	15	18	20	23
1600	11	12	14	15	17

To study the relationship between the mesh size h and the parameter γ, we performed the following experiments. We took $h^{-1} = 4, 8, 10, 12, 16, 20$, corresponding to uniform grids over the unit square $[0,1]^2$ with 32, 128, 200, 288, 512 and 800 triangles. We chose $k = 1/200$, and the same values of the parameter γ as before. The number of Conjugate Gradient iterations that were required to make the l_2-norm of the residual vector less than 10^{-12} are shown in Table 4.4 and Table 4.5. We observe that in both tables there is a significant difference between $\gamma = 10$ and $\gamma = 15$ for all values of h considered here. It also appears from Table 4.4 and Table 4.5 that 15 is the "optimal" value of the parameters γ, i.e. the value of γ which results in the least number of iterations. Notice, however, that although the number of iterations increase when $\gamma > 15$, the difference in the number of iterations from one value of γ to the next decreases with increasing h^{-1}. This motivated us to consider values of γ between 15 and 20. We found that, for both preconditioners and when $h^{-1} = 16$ or $h^{-1} = 20$, the least number of iterations is achieved for $\gamma = 17$. For the other values of h^{-1}, $\gamma = 15$ achieved the least number of iterations.

Table 4.4: Dependence of the number of Conjugate Gradient iterations with Jacobi preconditioning on the mesh size h and the parameter γ.

	Number of iterations				
h^{-1}	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$	$\gamma = 25$	$\gamma = 30$
4	35	34	37	39	46
8	72	49	59	68	72
10	114	59	68	72	81
12	158	66	76	85	91
16	261	84	84	93	99
20	363	86	98	109	121

Table 4.5: Dependence of the number of Conjugate Gradient iterations with Symmetric SOR preconditioning on the mesh size h and the parameter γ.

	Number of iterations				
h^{-1}	$\gamma = 10$	$\gamma = 15$	$\gamma = 20$	$\gamma = 25$	$\gamma = 30$
4	14	15	18	20	21
8	33	22	26	29	33
10	60	26	30	35	38
12	89	29	34	39	44
16	164	40	42	48	51
20	269	46	50	57	64

As a last test for our code we verified experimentally the spatial convergence rates for the NLS equation. We computed the solution of the problem

$$(4.11) \qquad \begin{cases} u_t = i\Delta u + i|u|^2 u, & \text{in}\, \Omega \times [0, 1/2] \\ u(x,0) = u^0(x), & \text{in}\, \Omega, \\ u = 0, & \text{on}\, \partial\Omega \times [0, 1/2] \end{cases}$$

where $\Omega = [0,1]^2$ and $u^0 = x(1-x)y(1-y)$. We used $k = 1/200$, as to make the error due to the temporal discretization small, $\gamma = 15$, and computed $U_{l+1}^{n,1}$, $l = 0, \ldots, l_n$ from (4.7). The Symmetric SOR–preconditioned Conjugate Gradient method was used as the linear system solver. To compensate for some loss of accuracy due to the reduced order of extrapolation when computing $U_0^{1,1}$, we chose $l_n = 4$ for $n = 1$ and $l_n = 3$ for $n > 1$. We used both linear ($r = 1$) and quadratic ($r = 2$) basis functions, and computed the convergence rate corresponding to different runs with mesh sizes h_1 and h_2 and errors E_1 and E_2 as $\log(E_1/E_2)/\log(h_1/h_2)$. The L^2-errors and the rates of convergence are summarized in the Table 4.6 below. The rates appear to be in a good overall agreement with the theoretical rate of convergence $r + 1$.

Table 4.6: L^2–Errors and Convergence rates for $r = 1$ and $r = 2$.

	$r=1$		$r=2$	
h^{-1}	Error	Rate	Error	Rate
8	0.133313(-2)		0.323075(-4)	
10	0.895737(-3)	1.782	0.132781(-4)	3.987
12	0.633751(-3)	1.897	0.896287(-5)	2.155
16	0.360361(-3)	1.962	0.398637(-5)	2.816
20	0.224667(-3)	2.117	0.175524(-5)	3.675

4.3 Implementation of a Fully Adaptive Method for the NLS

In this section we undertake the development of a method for the integration of the NLS equation up to times close to the formation of singularities. The scheme developed in the previous section serves as a basis, but it is appropriately modified to support a dynamically refined mesh and a variable time step. We begin by discussing the mesh refinement process and the underlying refinement strategy.

Mesh Refinement

The most commonly used process for refining a triangulation is *regular refinement.* The regular refinement of an individual triangle T consists of introducing new nodes on the midpoints of all edges and connecting these midpoints to form four new triangles similar to T. We shall refer to T as the "father" triangle and to the four new triangles T_1, T_2, T_3 and T_0 as the "sons" of T. This process is exhibited in Figure 4.2. It is clear that successive refinement of a conforming triangulation $\mathcal{T}_0$ yields a family of triangulations $\mathcal{T}_1, \mathcal{T}_2, \ldots$, which are all conforming and *regular* in the sense that the diameters of the triangles tend to zero but all angles stay bounded away from zero.

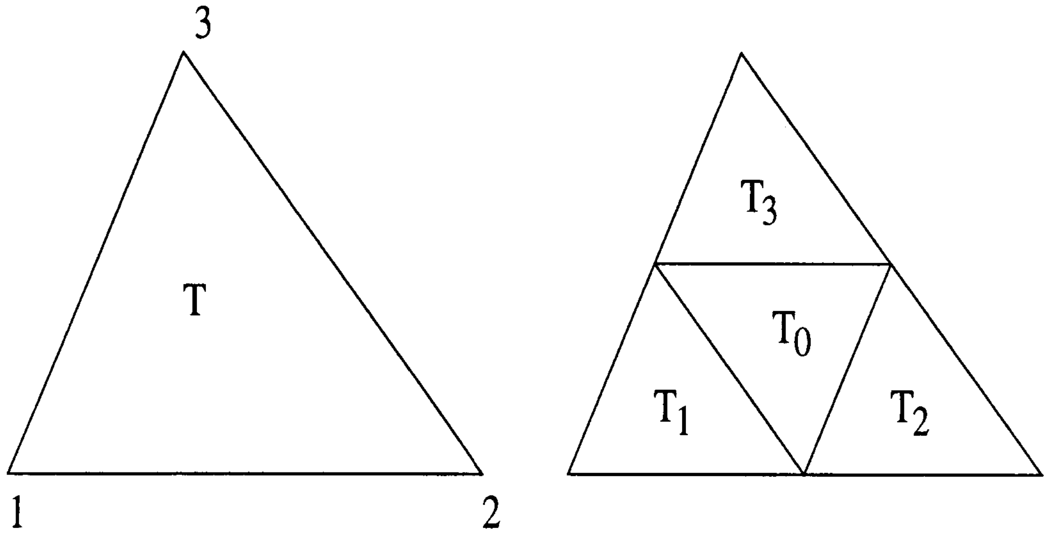


Figure 4.2: Regular Refinement of the triangle T

However, if conforming triangulations are required, regular refinement alone cannot support adaptivity, the reason being that triangles which are not refined will have non–conforming nodes. One way around this difficulty, is to refine the triangles with non–conforming nodes using *bisection* so that the resulting triangulation will be conforming. The process of refining a triangle using bisection amounts to forming two triangles by connecting one vertex with the midpoint of the opposite side. This process halves the angles of the triangles which are refined and may lead to extremely "thin" triangles. The bisecting edges are referred to in the literature as *green edges*, [8]. In many adaptive codes such as PLTMG and KASKADE the green edges are removed before the triangulation is refined further.

In our implementation of an adaptive method for the NLS, we use a slightly different approach. We refine the triangles of the mesh using regular refinement but allow only one non–conforming node per side. Figure 4.3 illustrates this idea. We have indicated the non–conforming nodes of the triangulation by dots. Notice that triangles R, S and T have one non–conforming node each. We should also point out that if the triangle Q in Figure 4.3 must be further refined, then its neighbors R and S will acquire two non–conforming nodes. Since this is not allowed, we refine the neighbors R and S *before* refining triangle Q. The refinement strategy explained above can be easily implemented by the recursive algorithm shown in Figure 4.4.

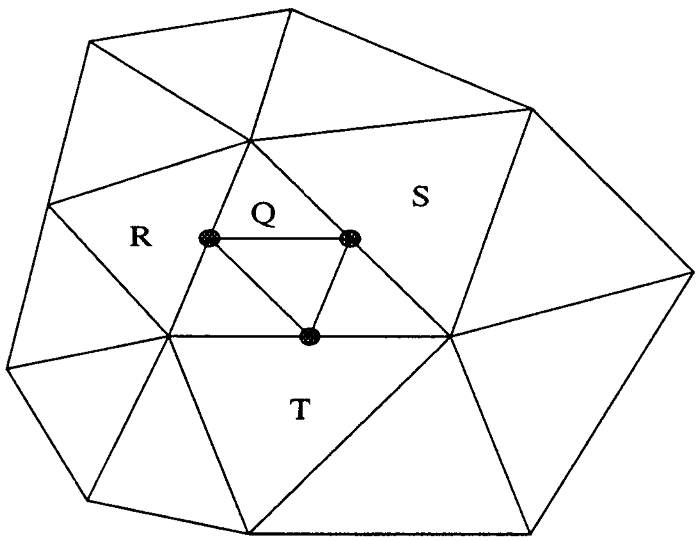


Figure 4.3: Regular refinement of the mesh with one non–conforming node per side.

```
procedure reftrng(i)
  for j in N(i) do
     if (j has a nonconforming node
         on the side shared with i) then
       reftrng(j)
     end;
  end;
  refdata(i);
end;
```

Figure 4.4: Local mesh refinement algorithm.

We assume here that `N(i)` is the set of neighbors of triangle `i` and that `refdata(i)` is a routine which modifies the data structures to reflect the regular refinement of the triangle `i`. This refinement approach has the benefit of simplifying the data structures necessary to represent such a non–conforming triangulation and allows for a smoother transition from large to smaller triangles. In addition, we avoid having to remove any green edges that should have been introduced otherwise. It is clear that this refinement process produces regular meshes, in the sense described earlier, and allows for local refinement of the mesh. As an example, let us consider the refinement of the mesh over an L–shaped region, near the re–entrant corner. It is well known, (cf. for example, [36]) that corner points pose particular problems in the numerical solution of boundary value problems. For example, the following boundary value problem for Laplace's equation $\Delta u = 0$ in the domain $\Omega = (-1,1)^2 \setminus (-1,0)^2$ and with boundary data $u = r^{2/3}\sin(\frac{2}{3}\phi)$, has a re–entrant corner singularity. It is shown in [35] that without local mesh refinement the error deteriorates globally to $O(h^{4/3})$. Figure 4.5 shows the mesh after two applications of the refinement strategy described above.

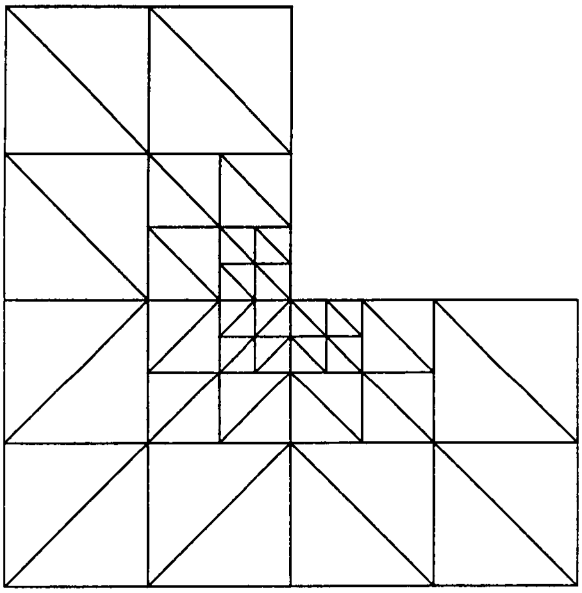


Figure 4.5: Local refinement of an L–shaped mesh

Obviously the mesh is regular and can be made arbitrarily fine in the neighborhood of the re–entrant corner. We should note that if no non–conforming nodes were allowed then refinement near the re–entrant corner would have resulted in the refinement of all triangles of the mesh.

Description of the Adaptive NLS code

With the adaptive mechanism in place, we now discuss the necessary modifications to the base scheme introduced in the previous section, to support dynamically refined meshes and variable time step. We shall assume that an initial triangulation $\mathcal{T}_h^0$, consisting of the triangles $\{\Omega_i^0\}_{i=1}^{d_0}$, is given. The adaptive mechanism generates a sequence of meshes $\mathcal{T}_h^1, \mathcal{T}_h^2, \ldots$, where each $\mathcal{T}_h^i$ is generated from $\mathcal{T}_h^{i-1}$ by appropriately refining some or all of its triangles. With each of these triangulations $\mathcal{T}_h^i$ we associate the finite space V_h^i, consisting of piecewise polynomial functions of degree $\leq r$ defined on $\mathcal{T}_h^i$. As in the previous section, no continuity along interelement boundaries is enforced. Moreover, since only regular refinement is used, the triangulations $\mathcal{T}_h^i$, $i = 0, 1, \ldots$, and the corresponding finite element spaces V_h^i, $i = 0, 1, \ldots$, are nested.

In our implementation of the adaptive code for the NLS, it was not necessary for us to keep track of the complete history of the refinement of the initial triangulation $\mathcal{T}_h^0$. For this reason, we agreed that if the triangle Ω_j^i of $\mathcal{T}_h^i$ must be refined then its "middle" son (c.f. Figure 4.2) acquires the number of its "father", i.e. j, while the other three triangles are given the numbers $d_i + 1, d_i + 2$ and $d_i + 3$. This simple convention allows us to actually use the same data structures developed for the non adaptive version of our code. We simply change the entry corresponding to Ω_j^i to reflect its refinement, and create new entries for its sons.

We also took additional steps to assure that the local refinement of the mesh will not necessitate a complete rebuilding of the mass and stiffness matrices M and S. Recall that both matrices depend on the underlying triangulation but the entries of the mass matrix can be computed from the corresponding entries of the mass matrix on the reference triangle and the Jacobian of the affine transformation (4.1), mapping the reference triangle onto the current triangle. For the stiffness matrix now, it is easy to see from the definition of the bilinear form $a_h^\gamma(\cdot,\cdot)$ that the refinement of a single triangle affects the entries of the stiffness matrix corresponding to this triangle and its neighbors. Since our refinement procedure implies that each triangle may acquire at most six neighbors, only a small number of the entries of the matrix S need be recomputed every time the mesh is refined.

For the time stepping scheme now, we implemented the one-stage Gauss–Legendre method with a variable time–step. This does not introduce any additional complexity to the algorithm but requires that the extrapolation coefficients α_0 and α_1 be recomputed when the time step is changed. The linear system solver required only slight modifications. We used the Conjugate Gradient method with the Symmetric SOR preconditioner and updated the routines performing the matrix-vector products to handle the changing mesh. Moreover, since the diagonal blocks of the system matrix depend on the triangulation and the time step, we updated their factorizations whenever the time step or the mesh changed.

It remains to develop a mechanism for deciding when the spatial or the temporal grids should be refined. The computational refinement tests employed in the code, have been successfully used in other contexts, see for example, [9], [10], [11] and [5]. This refinement mechanism is motivated by a local H^1–L^2 "inverse" property on each triangle Ω_i^m. We allow $||U^n||_{H^1(\Omega_i^m)}$ to grow by reducing the diameter of the triangle.

The refinement of the temporal grid is motivated by the need to control the second invariant associated with the NLS equation. We cut the time step in half whenever

$$\frac{|I_2(U^n) - I_2(U^{n+1})|}{||U^{n+1}||_{H^1(\Omega)}} > \mathrm{TOL}_k \tag{4.12}$$

where I_2 denotes the second invariant associated with the NLS equation. In our numerical experiments we used the value $\mathrm{TOL}_k = 5 \times 10^{-3}$, which controlled the second invariant in a satisfactory manner.

The flow chart in Figure 4.6 summarizes the steps taken by the adaptive NLS code and the sequence in which they are implemented. We have denoted by $\tilde{U}^{n+1}$ the solution at time t^{n+1} computed from (4.9), starting from U^n. It is on this vector that the spatial and temporal refinement tests are applied. If this temporary solution passes both tests then we declare $U^{n+1} = \tilde{U}^{n+1}$ and proceed to compute the solution at the next time level. If $\tilde{U}^{n+1}$ fails the temporal test then the time step is halved and the solution is computed again. If $\tilde{U}^{n+1}$ fails the spatial test then the underlying triangulation $\mathcal{T}_h^m$ is refined and the solution is computed again on the finer grid $\mathcal{T}_h^{m+1}$. Since the corresponding finite element spaces V_h^m and V_h^{m+1} are nested, the starting vector U^n is "transferred" onto the finer space by the natural injection of V_h^m into V_h^{m+1}.

The time stepping scheme was stopped when the size of the time step fell below a certain floor value or when the maximum amplitude of the peak exceeded a predefined value, or when the resources of our data structures were depleted. These possibilities are examined in the flowchart by the test "Is progress being made?". A negative answer to this question terminates the code while a positive answer causes the solution to be computed at the next time level.

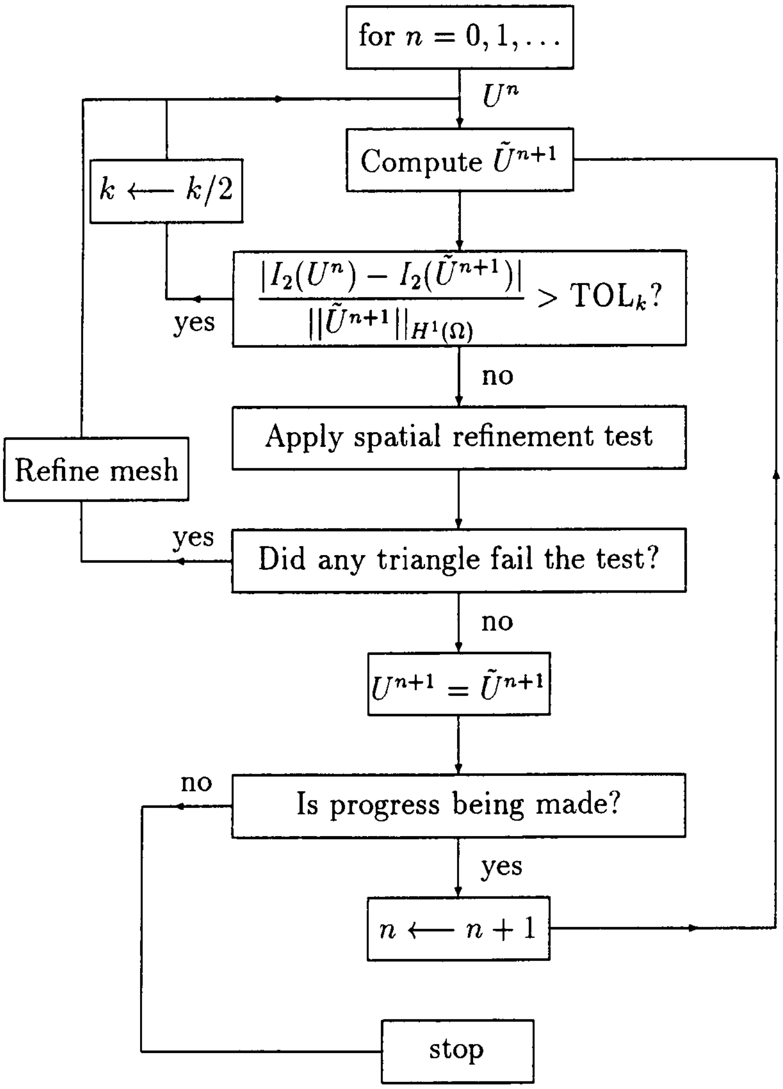


Figure 4.6: Flowchart for the adaptive NLS code.

Chapter 5

Numerical Experiments with the Adaptive NLS code

Having constructed a code with the adaptive mechanisms described in the previous chapter, we proceed to test the code by following the evolution of a class of initial profiles which proceed to blow–up in finite time. These are precisely the initial profiles for which the second invariant I_2, associated with the NLS equation is negative. Although most of the theoretical results on the blow–up of solutions of the NLS equation concern the initial value problem

$$\begin{cases} u_t = i\Delta u + if(u), & x \in R^d,\ t > 0, \\ u(x,0) = u^0(x), & x \in R^d, \end{cases} \tag{5.1}$$

there is both computational and theoretical evidence that blow–up also occurs for problems on bounded domains, cf, e.g. [41], [33]. In a collection of papers, Papanicolaou, C. Sulem, P.-L. Sulem, McLaughlin, and their collaborators ([41], [29], [30], [31], [27], [28]), studied the nature of the blow–up singularity by the method of dynamic rescaling. The method amounts to a time–dependent change of variables on the amplitude of the solution, on the space variable and on time such that the transformed equations have smooth solutions. The asymptotic behavior of the scale factors involved in the change of variables is then used to determine the nature of the singularity of the original problem. In [28] it was shown that in the critical case, $d = 2$, initial profiles with a single peak evolve into radially symmetric, self-similar singularities. In [5] fully adaptive, finite element schemes were constructed for studying the nature of the singularity, and the theoretically predicted blow-up rate

$$\left[\ln\ln\left(\frac{1}{t^*-t}\right)/(t^*-t)\right]^{\frac{1}{2}}$$

for the L^∞-norm of the solution u was confirmed. Here t^* denotes the blow-up time. The authors of the previously cited paper operated under the assumption that the initial profile and the solution of the initial value problem (5.1) are radially symmetric (isotropic case). After expressing the solution u in terms of the radial variable $\rho = (x_1^2 + x_2^2 + \ldots + x_d^2)^{1/2}$, they investigated, numerically, the formation of blow-up for solutions of an initial and boundary value problem approximating (5.1). The location of the blow-up was chosen as $\rho = 0$ and the adaptive mechanism was successful in refining the mesh near this point and controlling the time step as to integrate the equation for times very close to blow-up.

In our numerical experiments we made no assumption about the position of the singularity or the self-similarity of the solution. We approximated the solution of (5.1) by solving the NLS equation on a two-dimensional domain Ω and imposed zero Dirichlet conditions on the boundary $\partial\Omega$. In the literature (see for example, [29] and the references therein), most authors consider Gaussian initial profiles which do indeed go to zero at infinity. We chose the domain Ω large enough so that the restriction of the solution $u(t)$ on its boundary was negligibly small for $t \geq 0$. The starting approximation U^0 was computed as the L^2-projection of the initial profile u^0 on the finite element space V_h^0, and the approximations U^n for $n \geq 1$ were computed from (4.9). In all experiments reported in the sequel we used the Symmetric SOR matrix as a preconditioner for the Conjugate Gradient method, and performed four iterations of the Implicit–Explicit scheme per time step. The computational domain Ω was either the square $[0,2]^2$ or the unit square $[0,1]^2$. The parameter γ for the bilinear form $a_h^\gamma(\cdot,\cdot)$ was chosen as 15, while we used $\lambda = 25$ and a starting time step of size

4×10^{-4}. In all the experiments, the degree of the approximating functions, r, was equal to two.

5.1 Single Peak, Radially symmetric

The computational domain for this run was $\Omega = [0,2]^2$ and the starting grid was a uniform triangulation of Ω consisting of 512 triangles. The initial profile u^0, whose graph is shown in Figure 5.1, was chosen as $u^0(x,y) = 6\sqrt{2}\, e^{-45(x-1)^2-45(y-1)^2}$. The square of the L^2–norm for this profile was 2.51327, while the second invariant was -903.63827. We followed the evolution of the magnitude of the solution and the refinement of the spatial mesh and summarized the results in Figures 5.2–5.7. Our spatial refinement test refined the initial triangulation of Ω 91 times, resulting in a final grid containing 8990 triangles. Our temporal refinement strategy halved the initial time step 29 times and produced a final time step of $0.74505805969 \times 10^{-13}$. The L^∞–norm of the solution at the last time step was, approximately, 0.105880049×10^7.

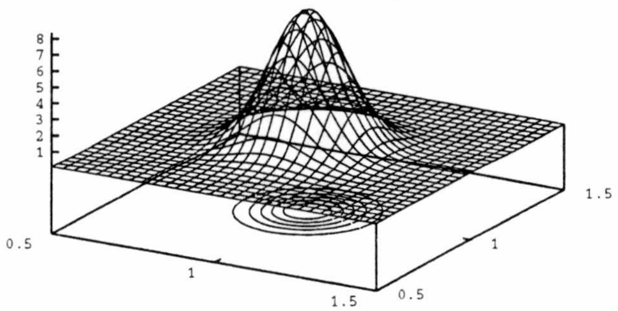


Figure 5.1: The initial profile u^0 for the first run.

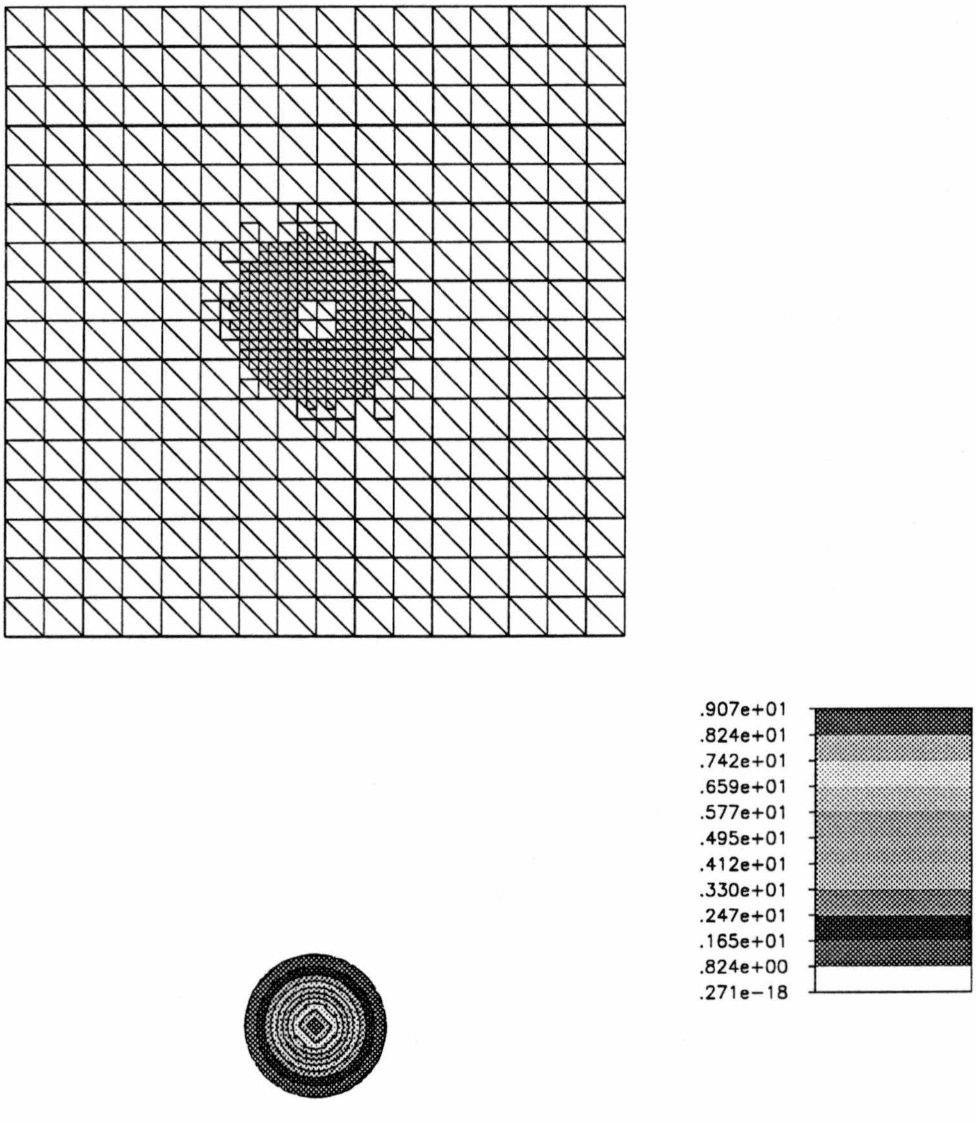


Figure 5.2: Experiment 1: Triangulation and contour plot of the solution at time 4×10^{-3}.

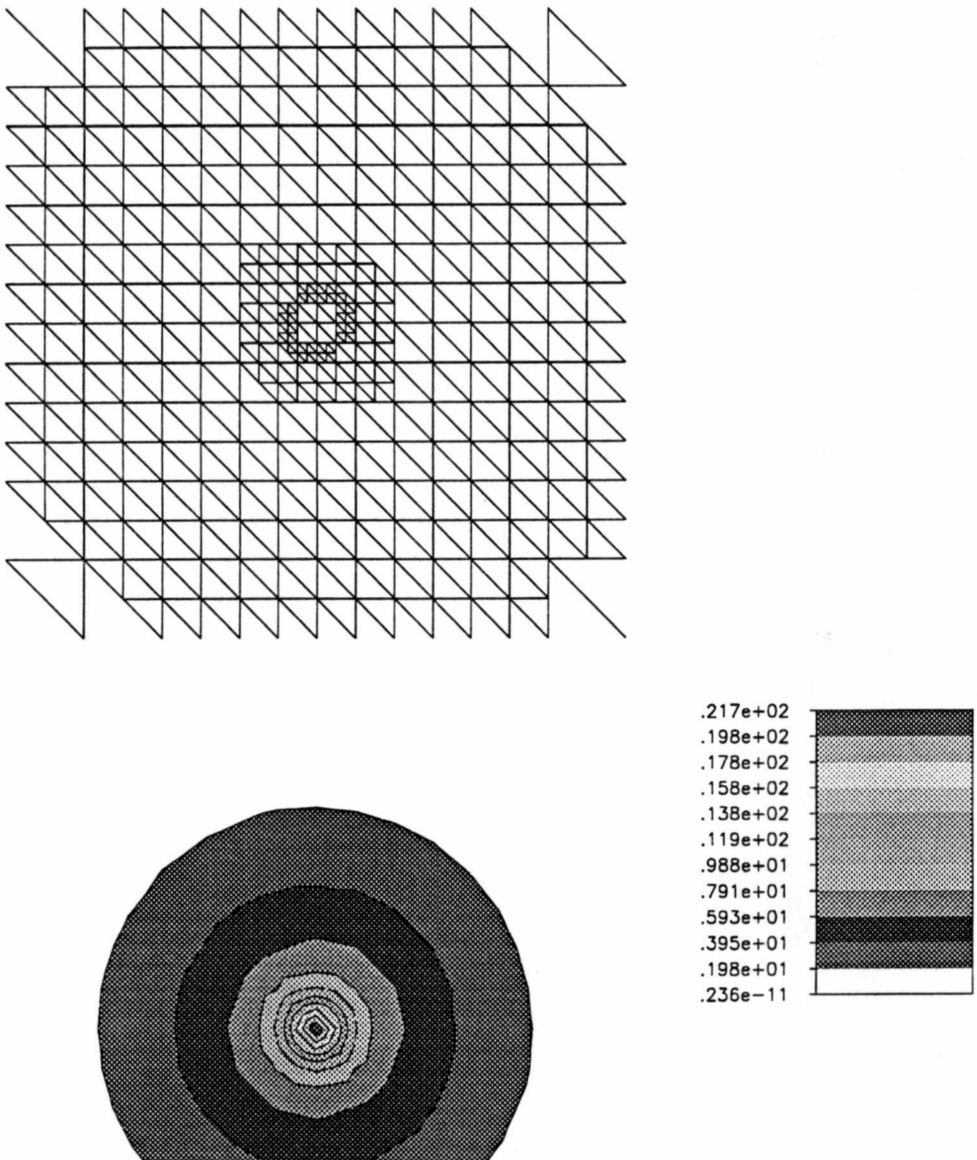


Figure 5.3: Experiment 1: Triangulation and contour plot of the solution at time 0.12×10^{-2}.

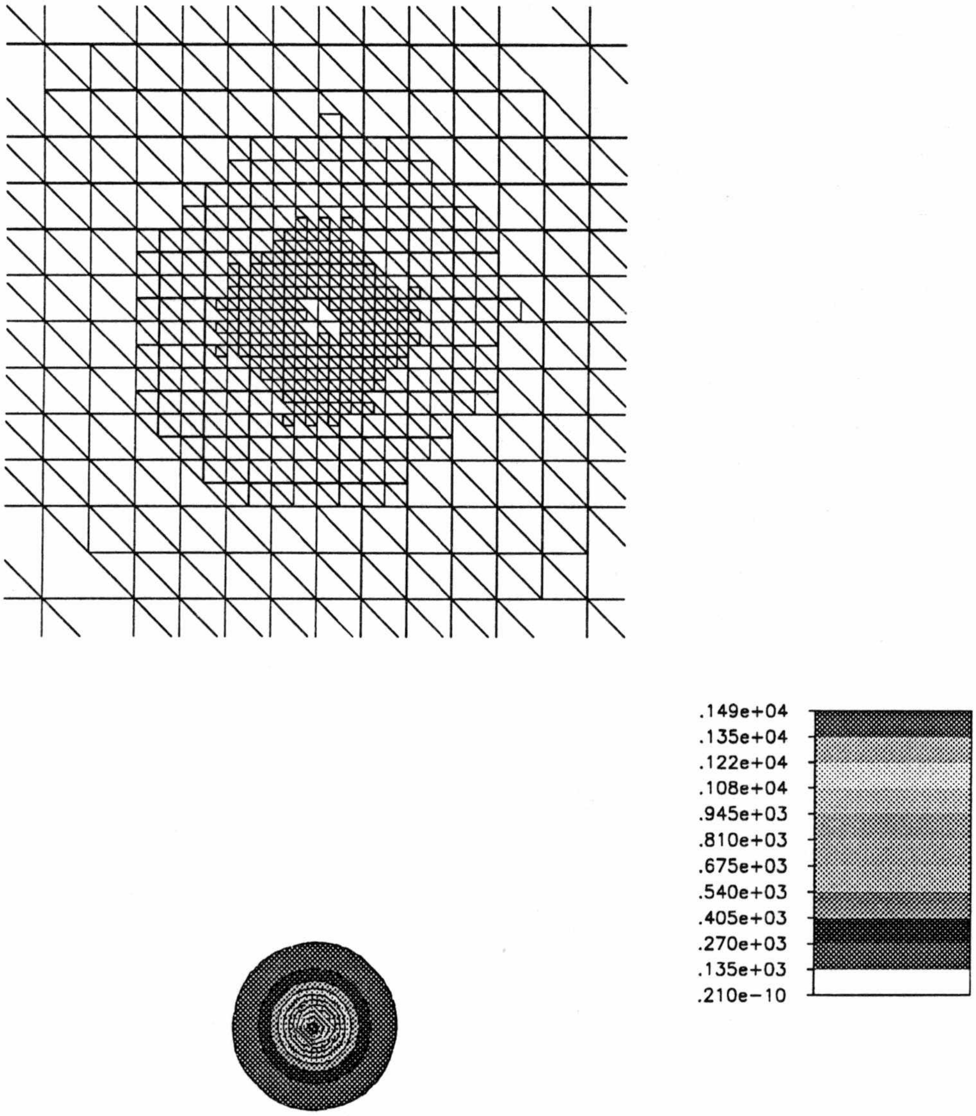


Figure 5.4: Experiment 1: Triangulation and contour plot of the solution at time $0.138234375 \times 10^{-2}$.

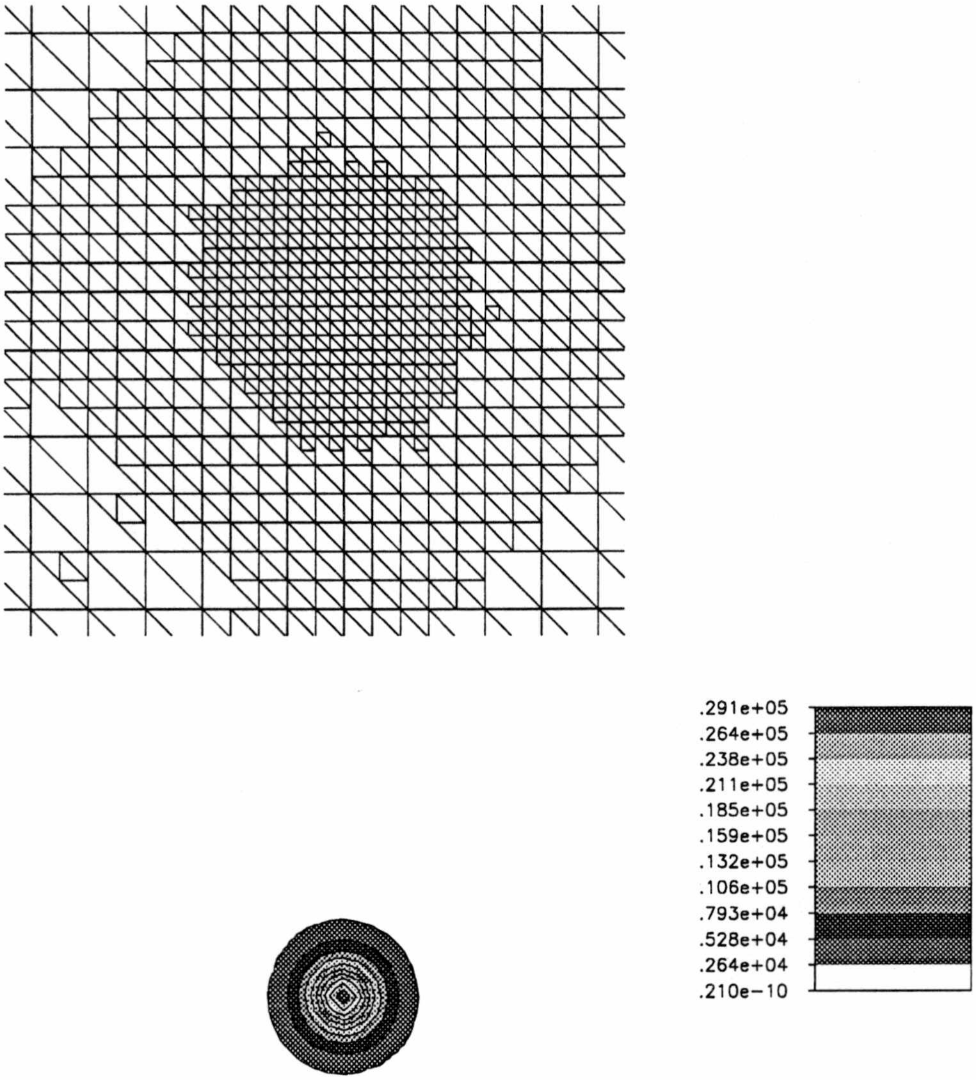


Figure 5.5: Experiment 1: Triangulation and contour plot of the solution at time $0.13824245 \times 10^{-2}$.

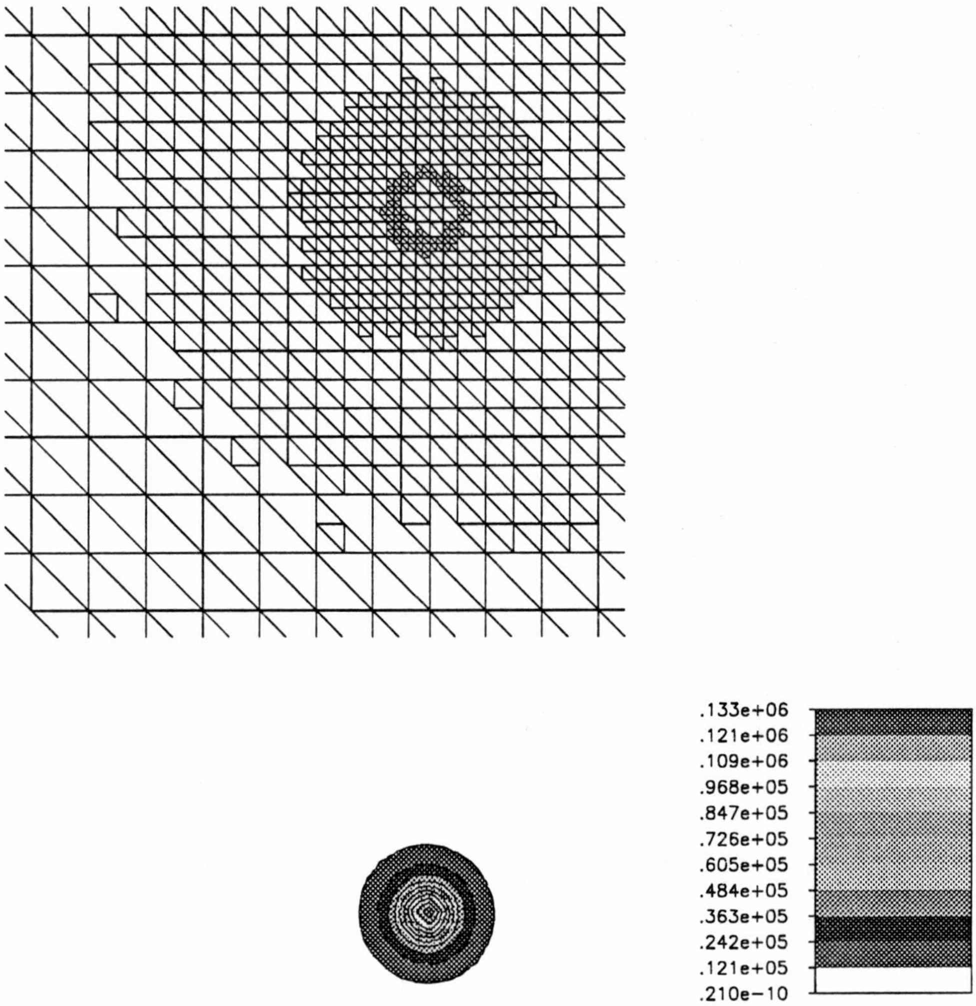
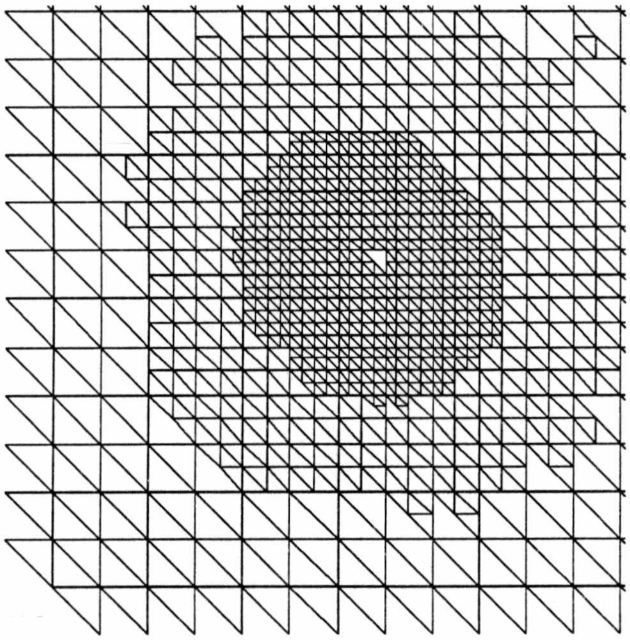


Figure 5.6: Experiment 1: Triangulation and contour plot of the solution at time $0.13824247 \times 10^{-2}$.



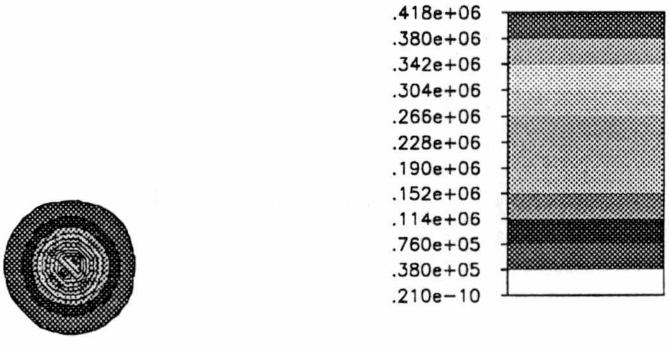


Figure 5.7: Experiment 1: Triangulation and contour plot of the solution at time $0.1382424764 \times 10^{-2}$.

5.2 Single Peak, Non–radially symmetric

The computational domain for this run was $\Omega = [0,2]^2$ and the starting grid was a uniform triangulation of Ω consisting of 512 triangles. The initial profile u^0, whose graph is shown in Figure 5.8, was chosen as

$$u^0(x,y) = 6\sqrt{2}\, e^{-45(x-1)^2 - 25(y-1)^2}$$

The square of the L^2–norm for this initial profile was approximately 3.37191, while the second invariant was -1280.17779. We followed the evolution of the magnitude of the solution and the refinement of the spatial mesh and summarized the results in Figures 5.9–5.13. Our spatial refinement test refined the initial triangulation of the domain Ω 119 times, resulting in a final triangulation containing 10808 triangles. Our temporal refinement strategy halved the initial time step 36 times and produced a final time step of $0.58207660913 \times 10^{-15}$. The L^∞–norm of the solution was, approximately, 0.102788379×10^8 at the end of the computation.

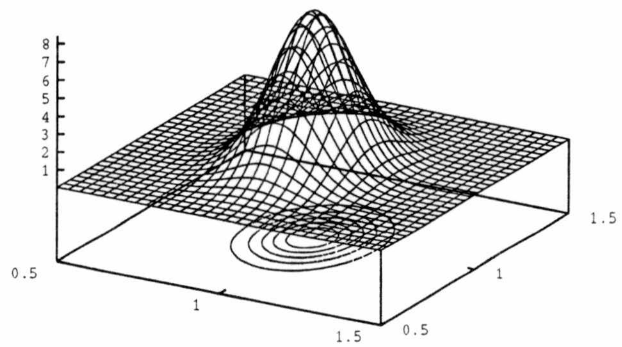


Figure 5.8: The initial profile u^0 for the second run.

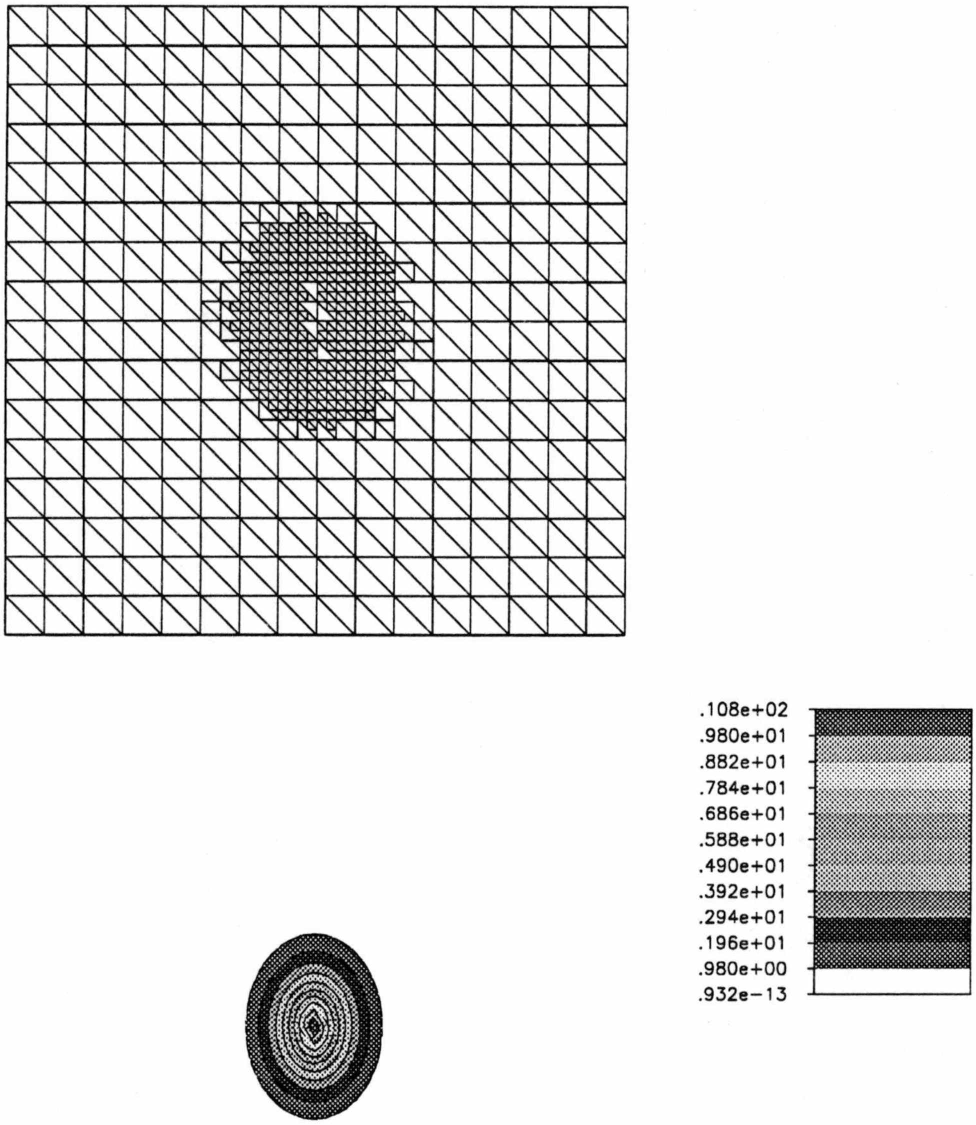


Figure 5.9: Experiment 2: Triangulation and contour plot of the solution at time 8×10^{-3}.

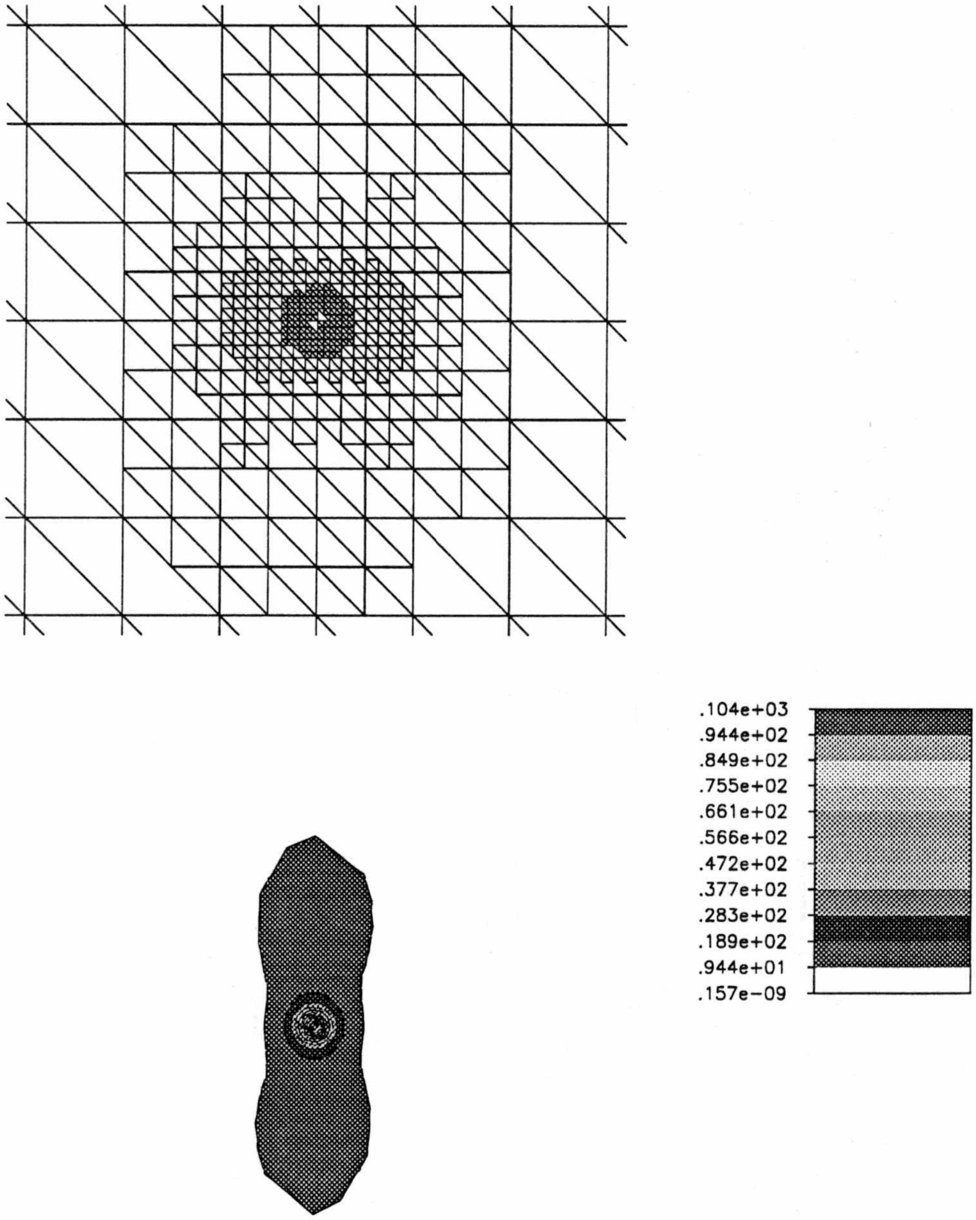


Figure 5.10: Experiment 2: Triangulation and contour plot of the solution at time 0.1485×10^{-2}.

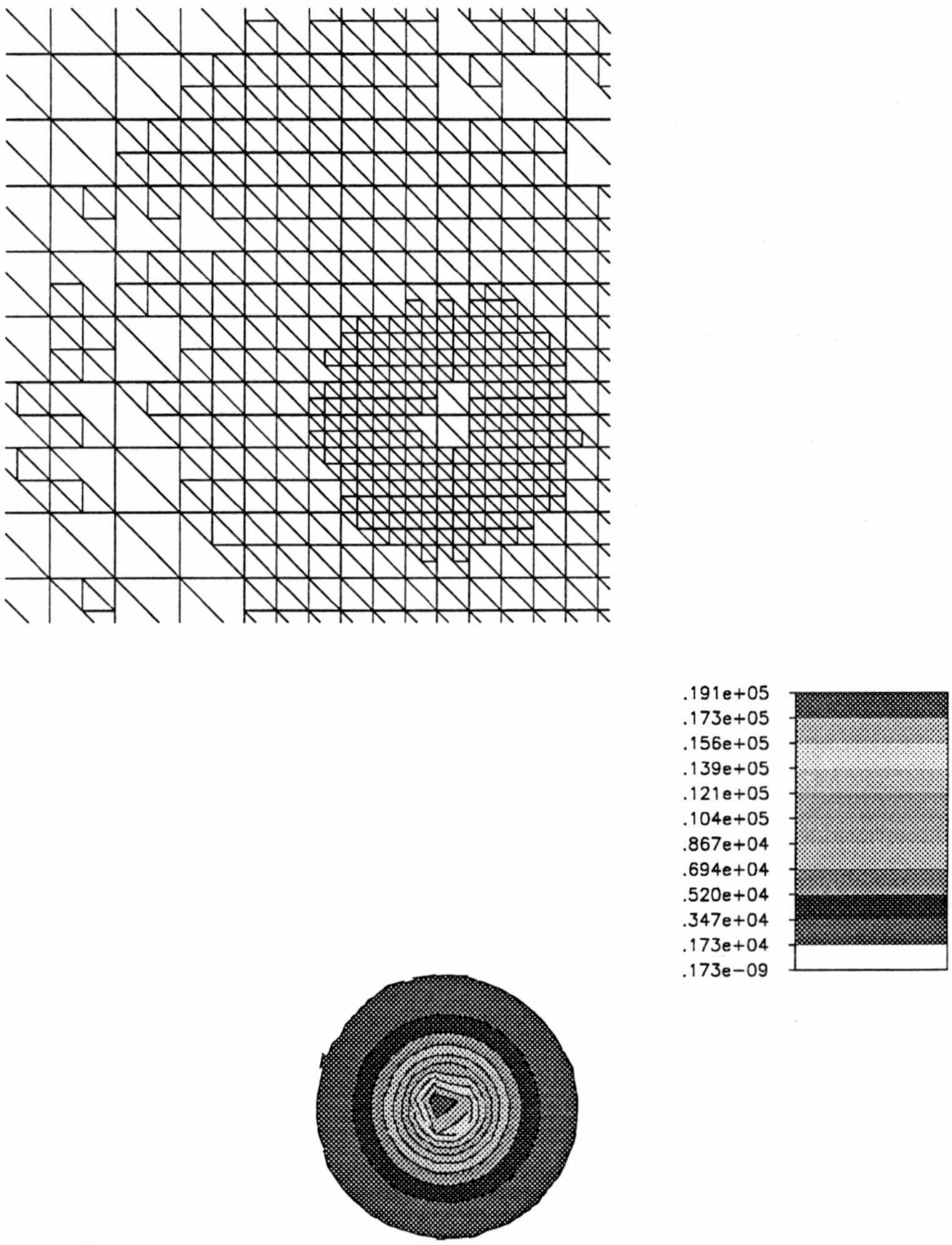


Figure 5.11: Experiment 2: Triangulation and contour plot of the solution at time $0.149596313 \times 10^{-2}$.

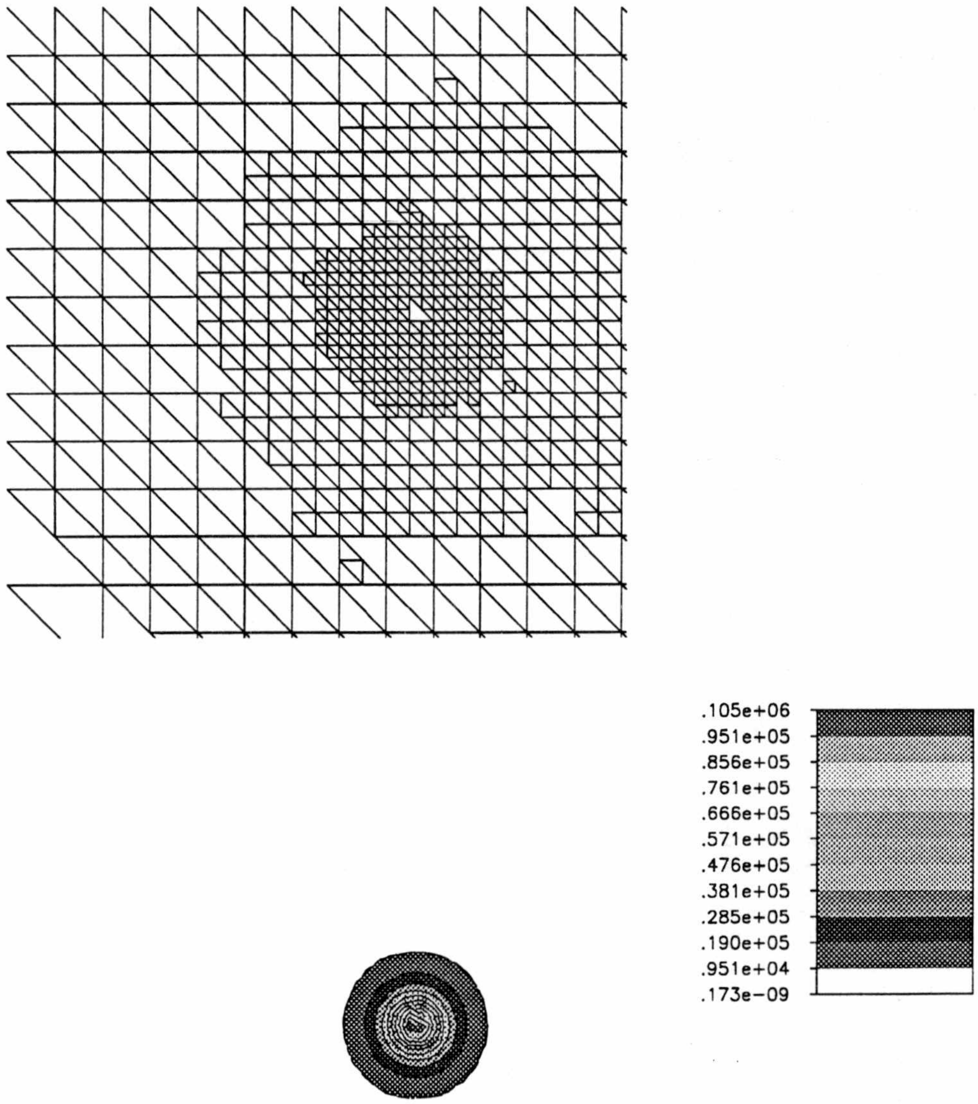


Figure 5.12: Experiment 2: Triangulation and contour plot of the solution at time $0.149596322 \times 10^{-2}$.

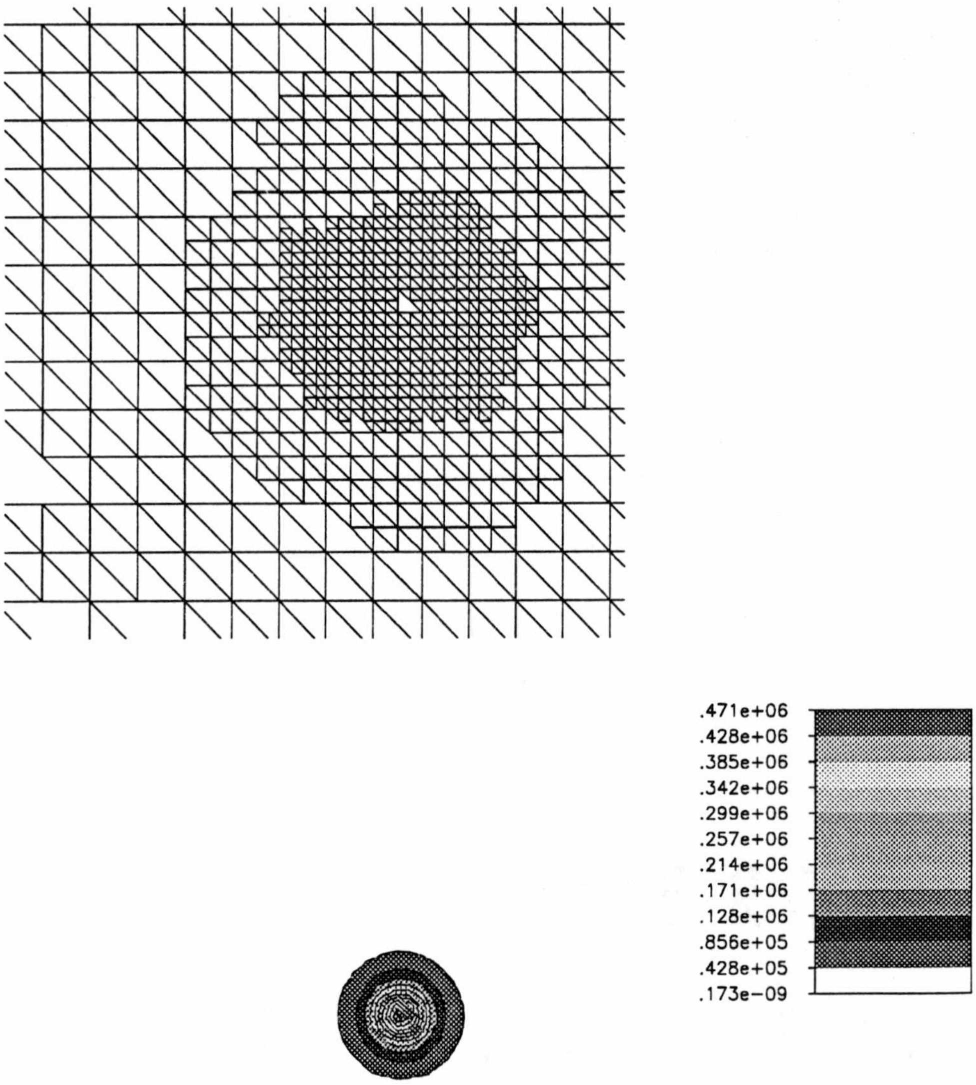


Figure 5.13: Experiment 2: Triangulation and contour plot of the solution at time $0.149596373 \times 10^{-2}$.

We can see from the contour plot of the solution in the preceding figures that the initially anisotropic profile, evolves very quickly into a conical profile which proceeds to blow up. (The apparent anisotropy in the contour plots near the point of maximum altitude is a result of insufficient resolution in the contouring procedure. The dimensions of the "window" over the point $(1,1)$ in the last contour plot were $10^{-5} \times 10^{-5}$).

5.3 "Twin peaks", Radially symmetric

The initial profile u^0 for this run consisted of two Gaussian peaks, each of maximum altitude $6\sqrt{2}$, placed at the points with coordinates $(0.8, 1.2)$ and $(1.2, 0.8)$. The peak centered at $(0.8, 1.2)$ was of the form $6\sqrt{2}\, e^{-45(x-0.8)^2-45(y-1.2)^2}$ and the peak centered at $(1.2, 0.8)$ was $6\sqrt{2}\, e^{-45(x-1.2)^2-45(y-0.8)^2}$. In the sequel, we shall refer to the former peak as the Northwest peak (NW) and to the latter one as the Southeast peak (SE). The surface plot of this initial profile is shown in Figure 5.14. The square of the L^2–norm for this profile was 5.03029 while the second invariant was, approximately, -1809.33113.

The computational domain for this experiment was once again the square $\Omega = [0,2]^2$, and the starting grid consisted of a uniform triangulation of Ω with 512 triangles. Our spatial refinement strategy refined the starting mesh 96 times, resulting in a final triangulation with 15716 triangles. The final time step was $0.7450580596 \times 10^{-13}$ and the solution achieved a maximum altitude of, approximately, $0.856556049 \times 10^{6}$. The evolution in time of the initial profile is shown in Figures 5.15–5.19. Although both peaks proceed to blow up, we only "zoom" into either the SE or the NW peak and show the contour plot and the corresponding triangulation in a neighborhood of the maximum altitude.

We also experimented with initial profiles consisting of two Gaussian peaks of different heights. As a first attempt, we positioned a Gaussian peak of height $6\sqrt{2}$

at the point with coordinates $(0.8, 1)$, and a second Gaussian peak of height four at the point with coordinates $(1.2, 1)$. The L^2-norm of this profile was 3.384541 and the second invariant was -1088.97186. The code refined the starting mesh 128 times, resulting in a final triangulation containing 11342 triangles. The temporal refinement mechanism halved the initial time step 36 times, resulting in a final time step of $0.5820766091 \times 10^{-15}$. The solution achieved an amplitude of, approximately, $0.100918927 \times 10^{8}$. Figures 5.20–5.23 show the contour plot of the amplitude of the solution and the underlying triangulations. The first three of the aforementioned figures show how the two peaks merge into a single, radially symmetric profile which proceeds to blow up.

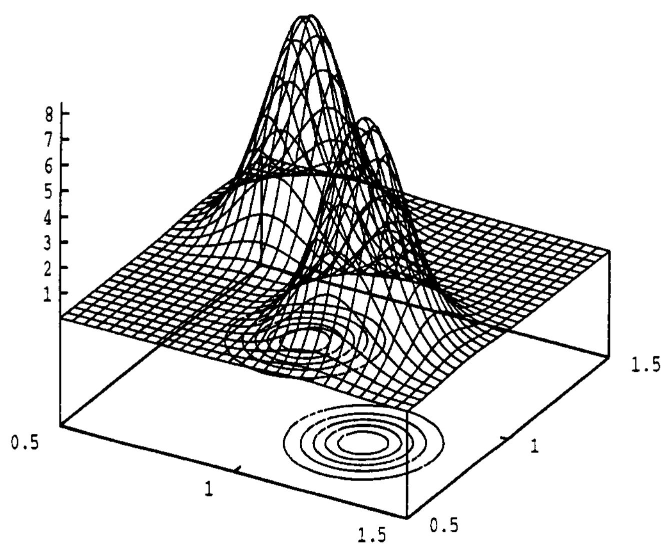


Figure 5.14: The initial profile u^0 for the third run.

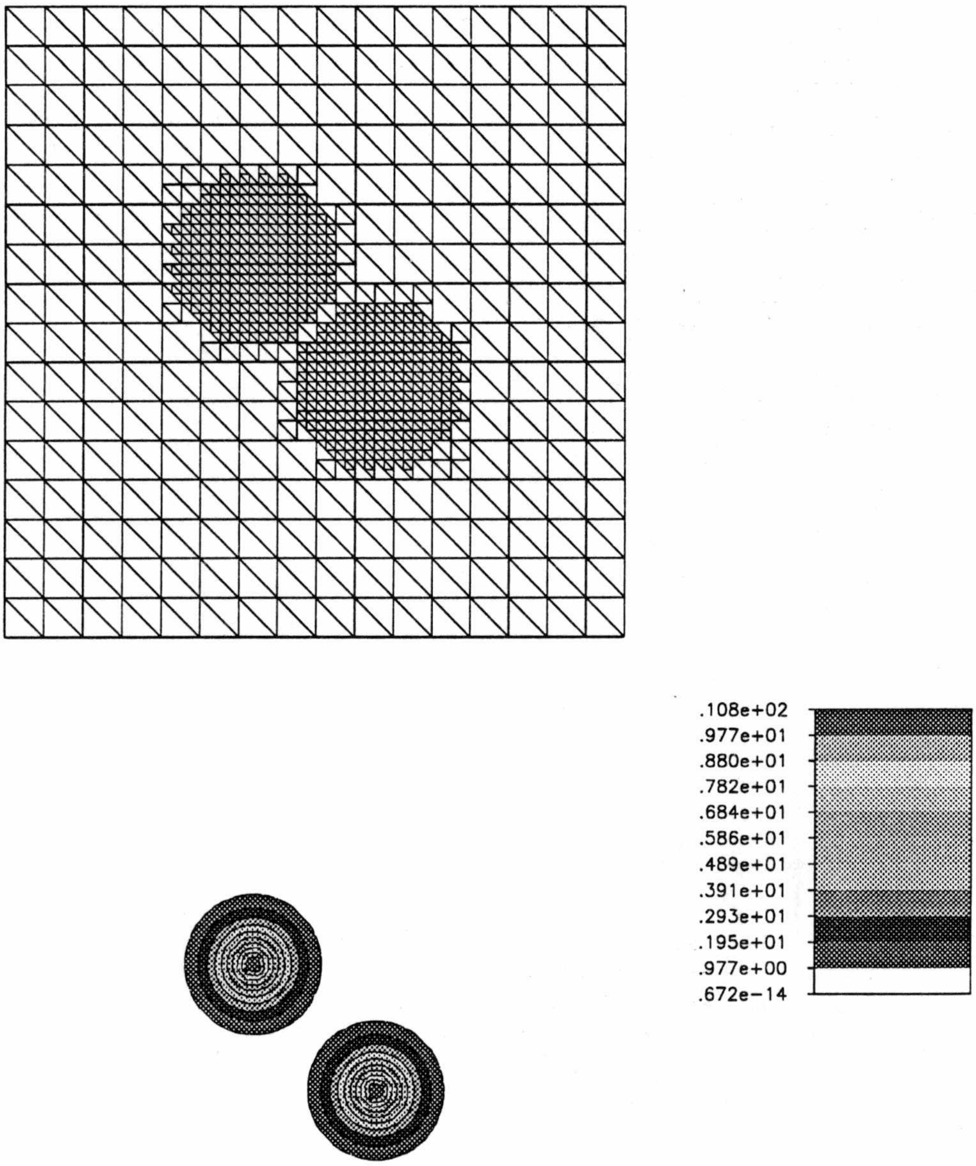
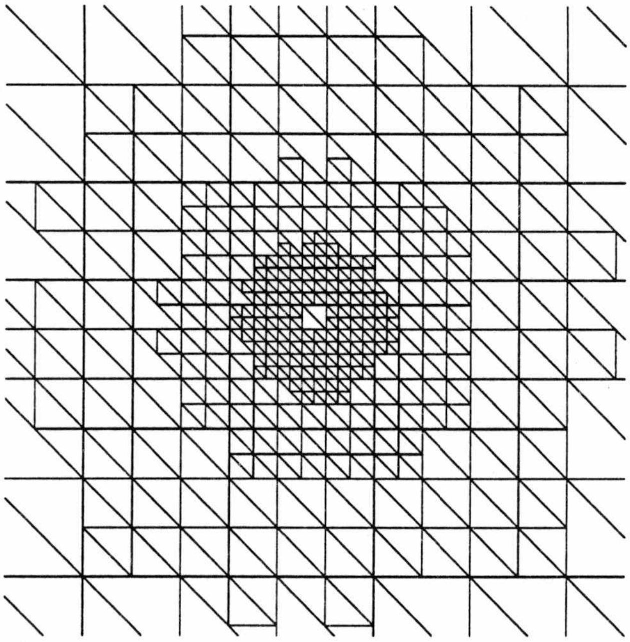


Figure 5.15: Experiment 3: Triangulation and contour plot of the solution at time 8×10^{-3}.



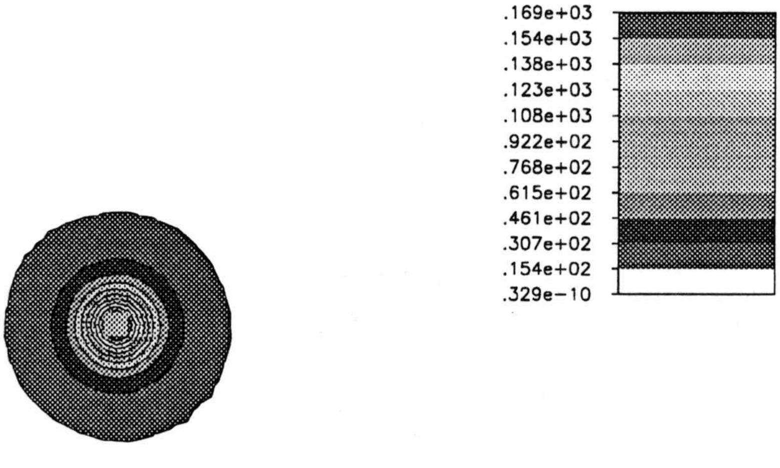


Figure 5.16: Experiment 3: Triangulation and contour plot of the solution at time 0.138×10^{-2}, NW peak.

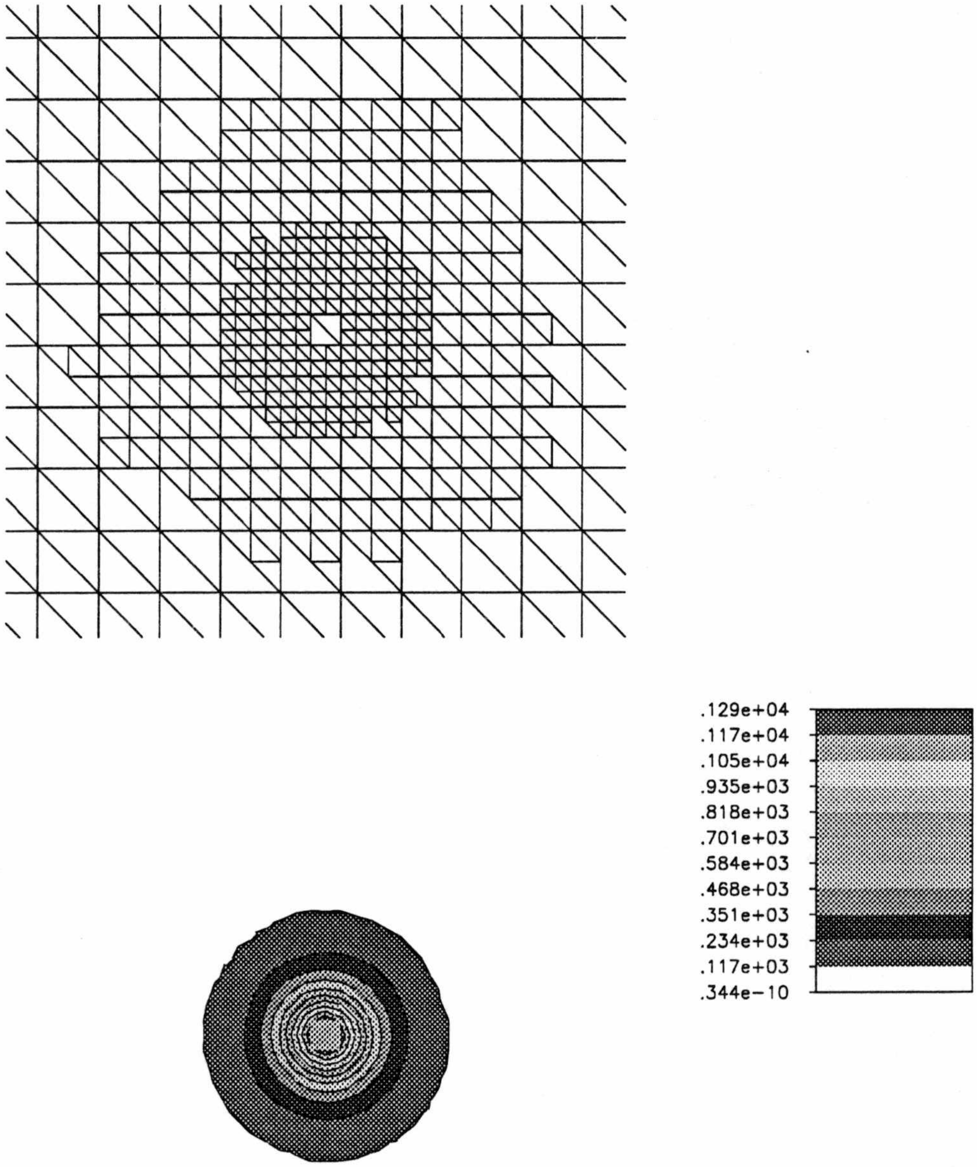


Figure 5.17: Experiment 3: Triangulation and contour plot of the solution at time 0.13839×10^{-2}, SE peak.

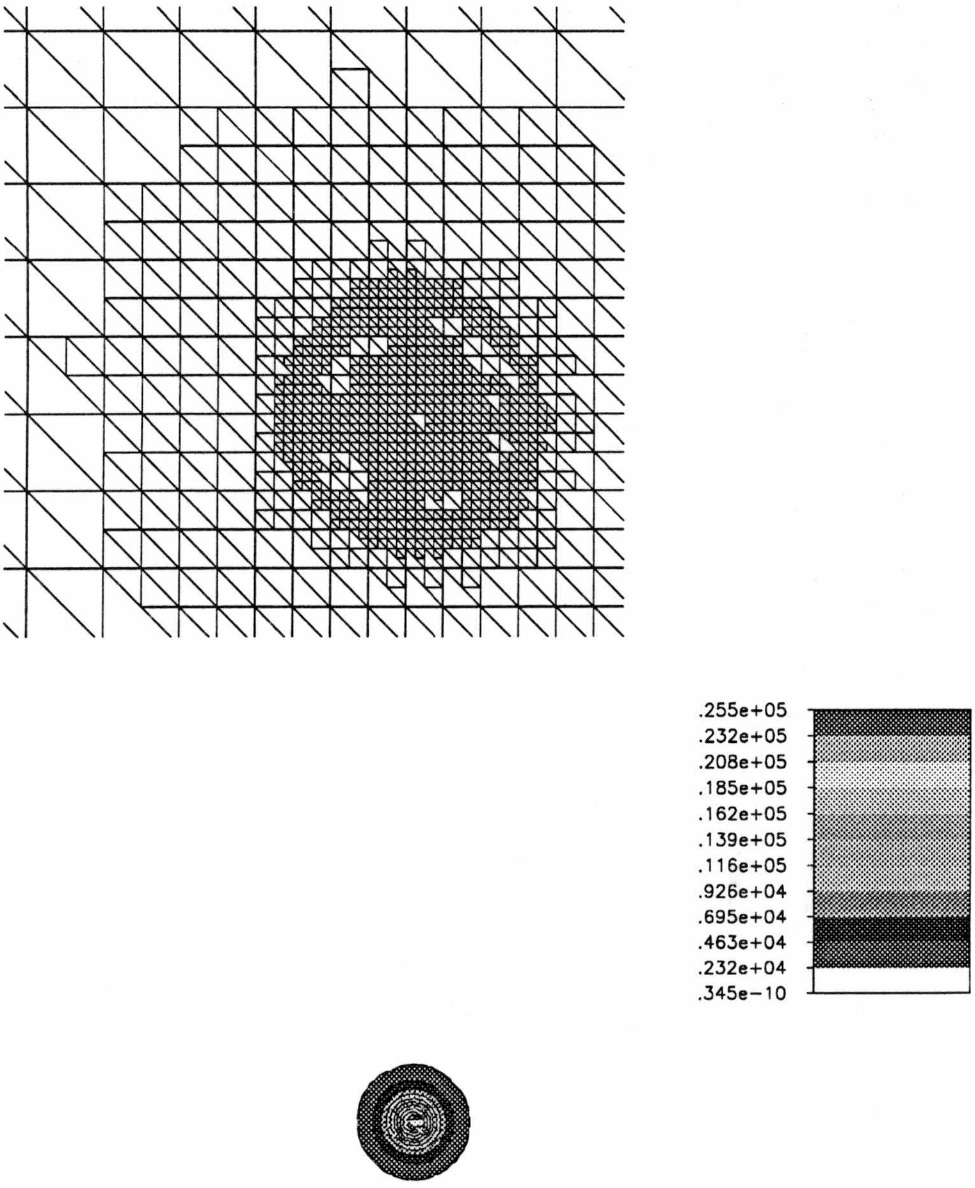


Figure 5.18: Experiment 3: Triangulation and contour plot of the solution at time 0.1383987×10^{-2}, SE peak.

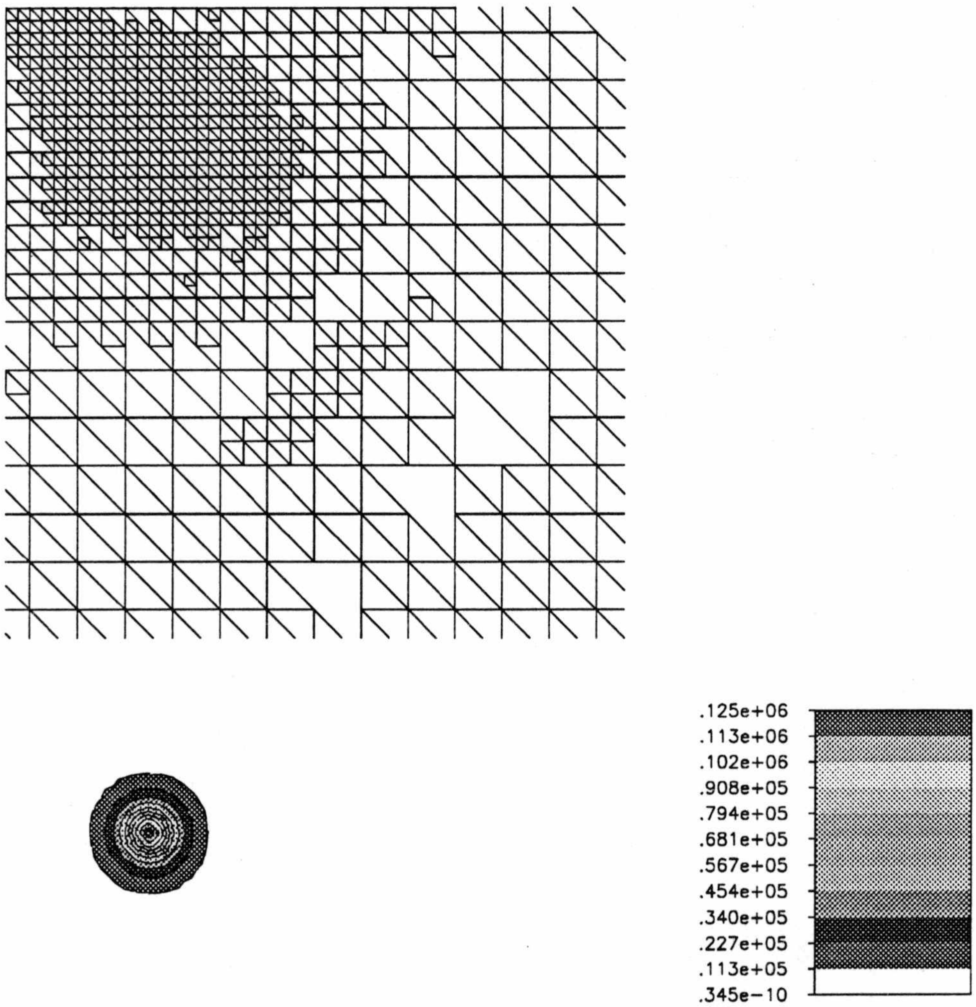
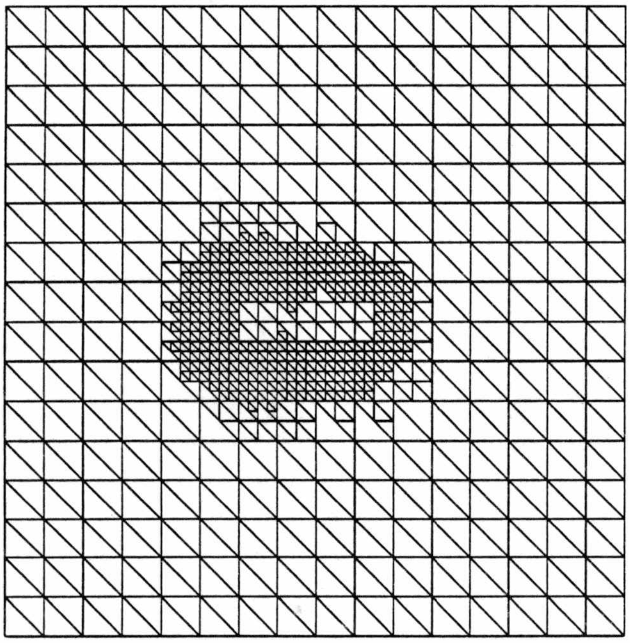


Figure 5.19: Experiment 3: Triangulation and contour plot of the solution at time $0.1383987917 \times 10^{-2}$, NW peak.



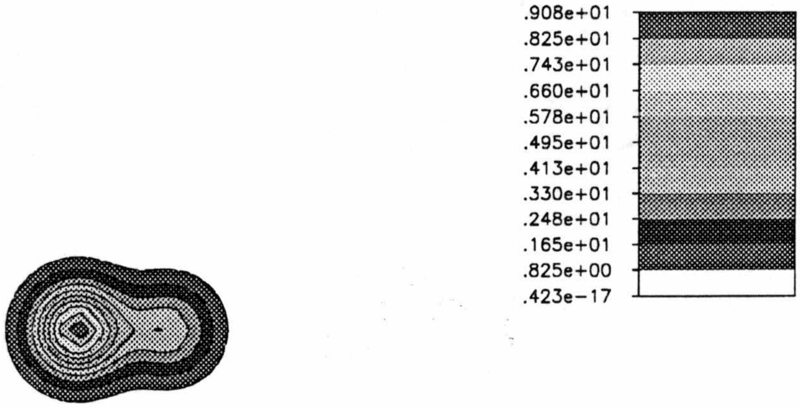


Figure 5.20: Two peaks of different heights: Triangulation and contour plot of the solution at time 0.4×10^{-3}.

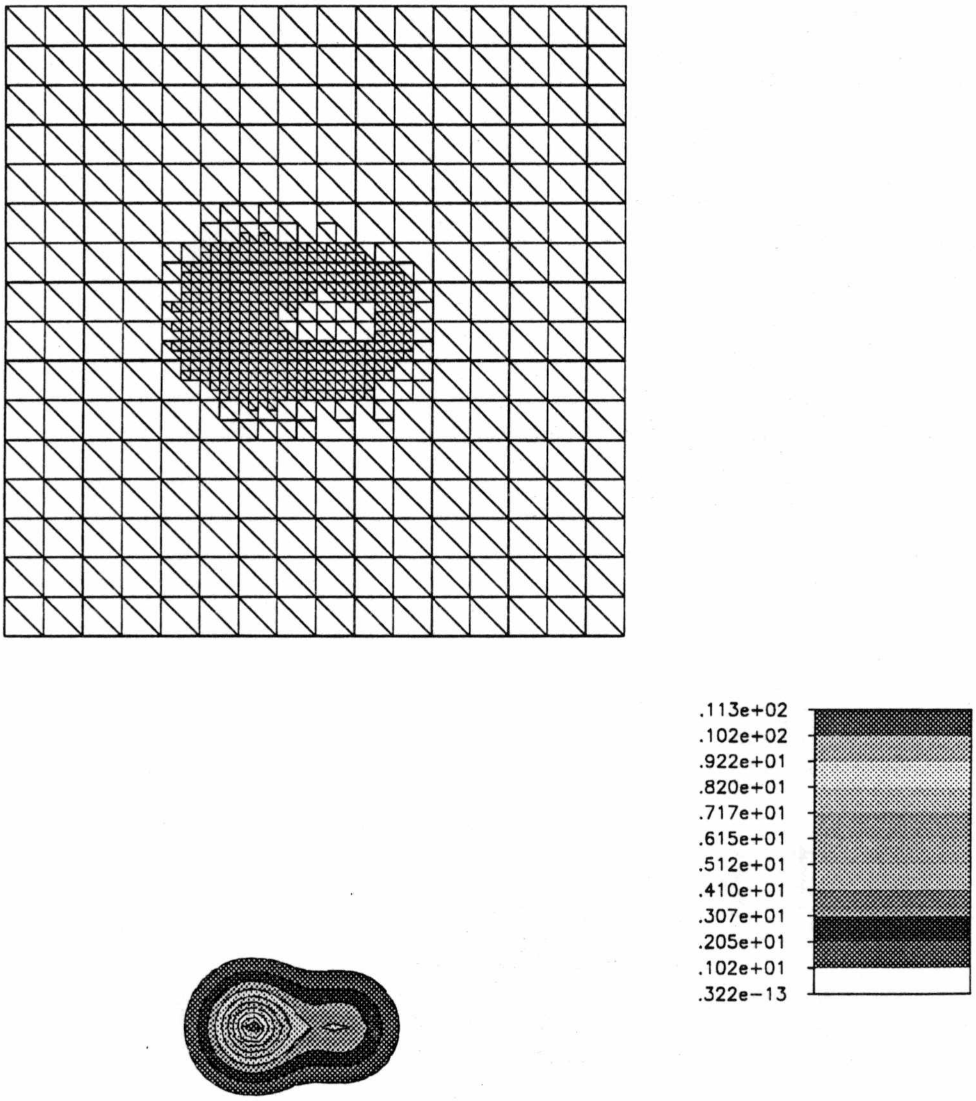


Figure 5.21: Two peaks of different heights: Triangulation and contour plot of the solution at time 0.8×10^{-3}.

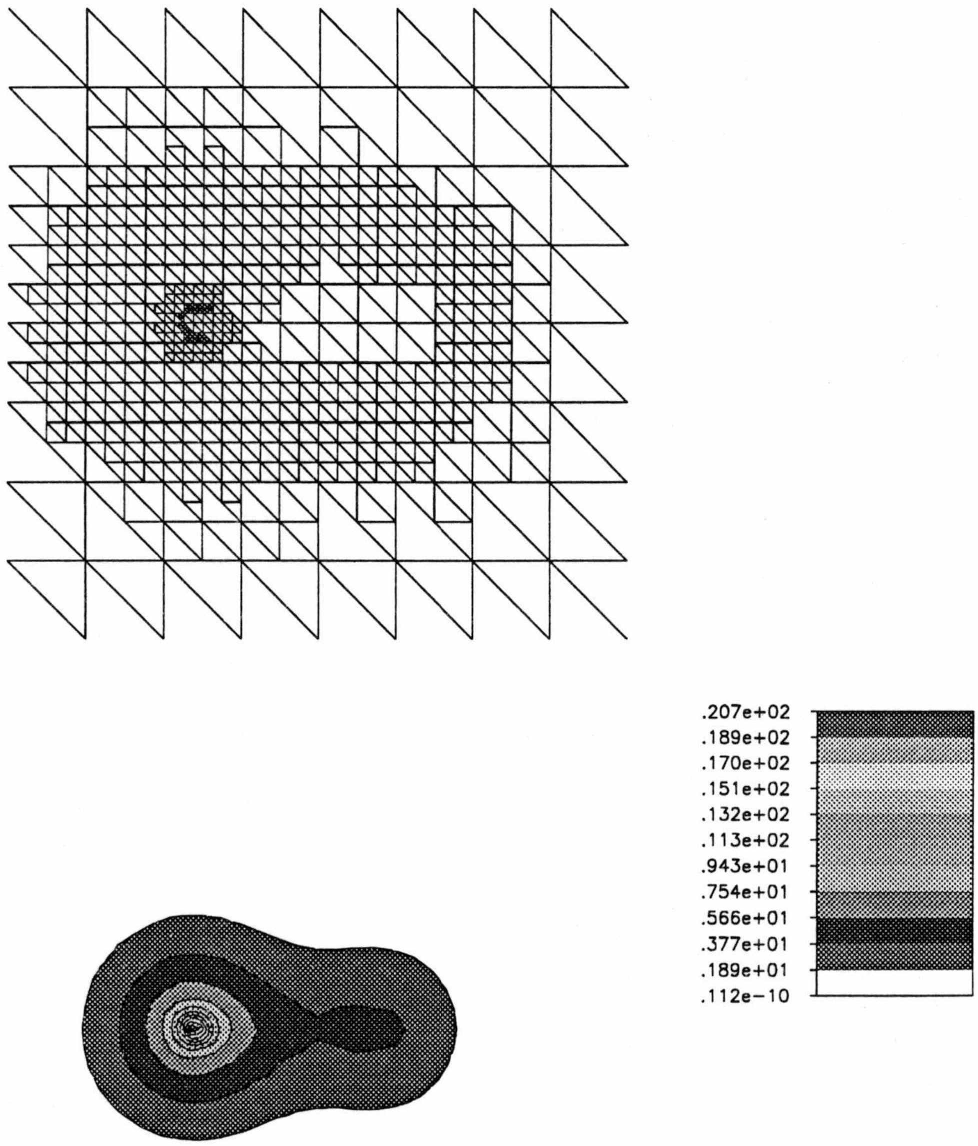


Figure 5.22: Two peaks of different heights: Triangulation and contour plot of the solution at time 0.12×10^{-2}.

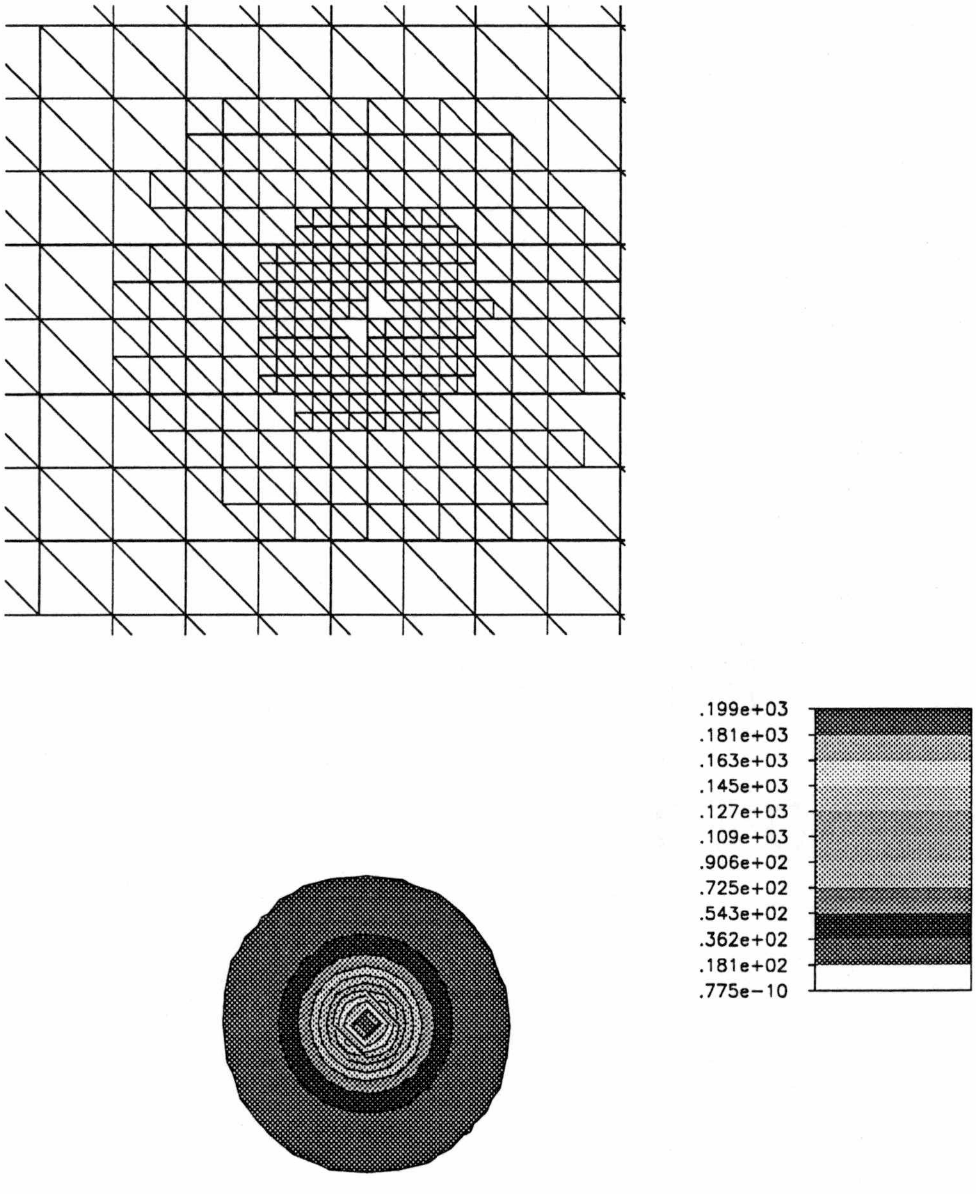


Figure 5.23: Two peaks of different heights: Triangulation and contour plot of the solution at time 0.138×10^{-3}.

5.4 "Twin peaks", Non–radially symmetric

The initial profile u^0 for this run, whose surface/contour plot is shown in Figure 5.24, was chosen as

$$u^0(x,y) = 6\sqrt{2}e^{-45(x-0.9)^2-45(y-1)^2} + 4e^{-25(x-1.1)^2-45(y-1)^2}$$

The square of the L^2–norm for this profile was 4.675273, while the second invariant was -2508.76769. The computational domain for this experiment was the square $\Omega = [0,2]^2$ and the starting grid consisted of a uniform triangulation of Ω with 512 triangles. Our spatial refinement test refined the initial triangulation of Ω 134 times, resulting in a final triangulation containing 12059 triangles. The temporal refinement test was employed 36 times and the final time step was $0.582076609 \times 10^{-15}$. The L^∞–norm of the solution at the end of the computation was, approximately, 0.100481428×10^8. The evolution of this initial profile is shown in Figures 5.25–5.30. We note that as times proceeds, the effect of the peak centered at $(1.1, 1)$ disappears and an apparently radially symmetric profile develops with a maximum whose position stabilizes close to the point $(1, 1)$. As the solution blows up, the initial anisotropy disappears, as shown in Figures 5.28–5.30, but at large distances the solution clears exhibits anisotropy (see Figure 5.27).

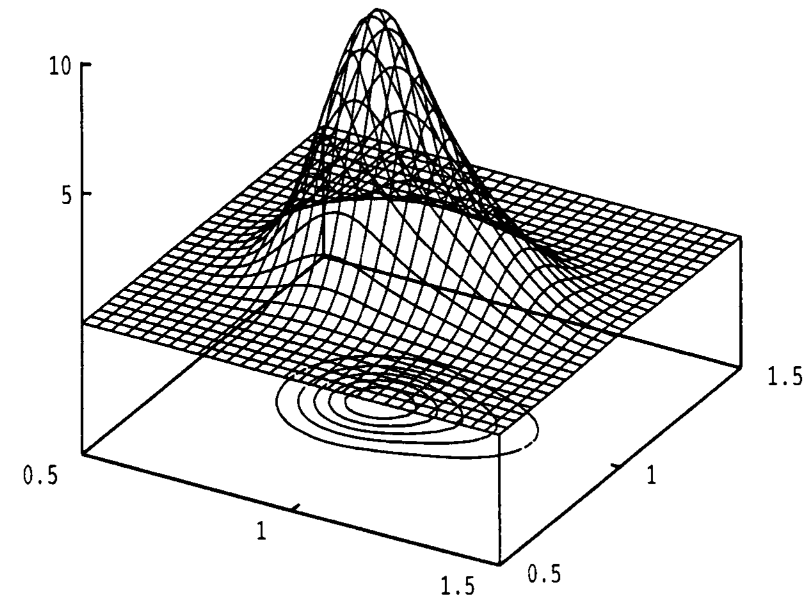


Figure 5.24: The initial profile u^0 for the fourth run.

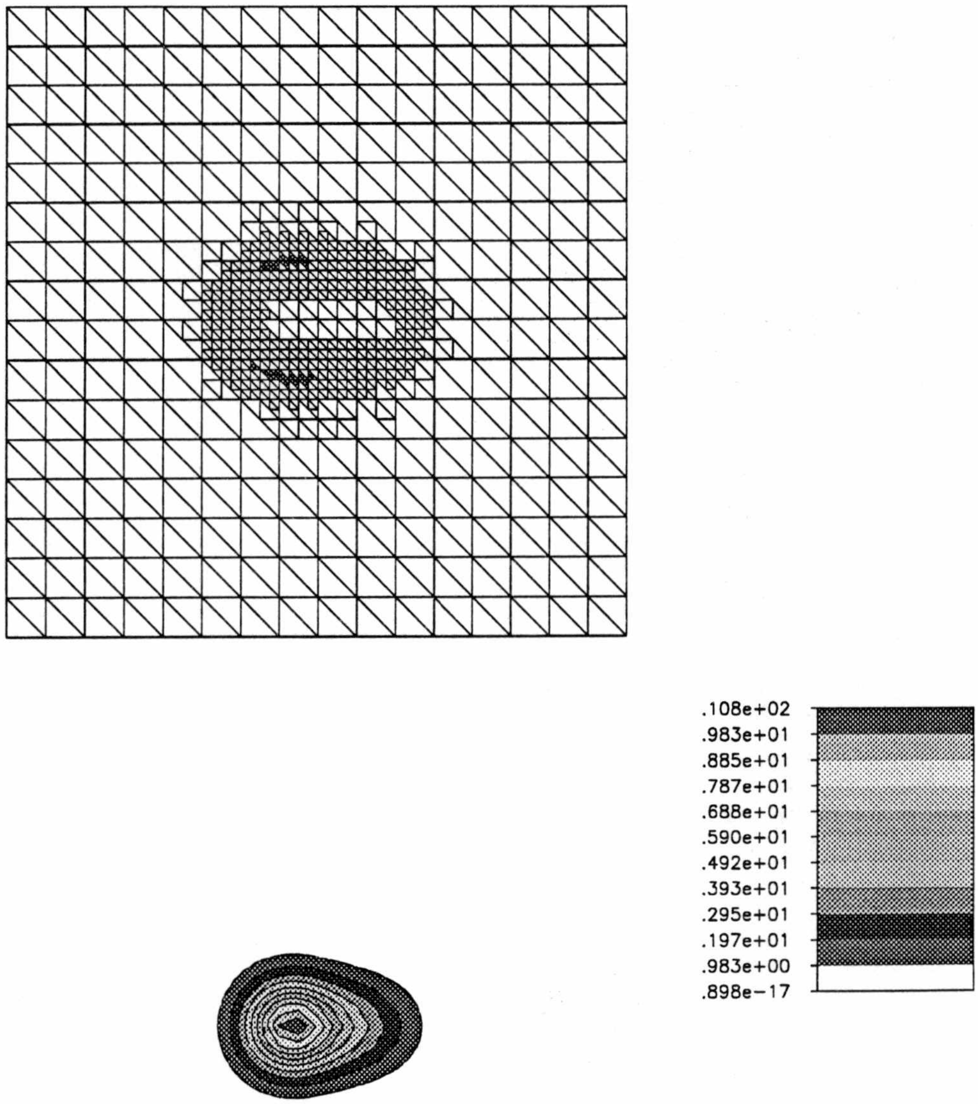


Figure 5.25: Experiment 4: Triangulation and contour plot of the solution at time 0.4×10^{-3}.

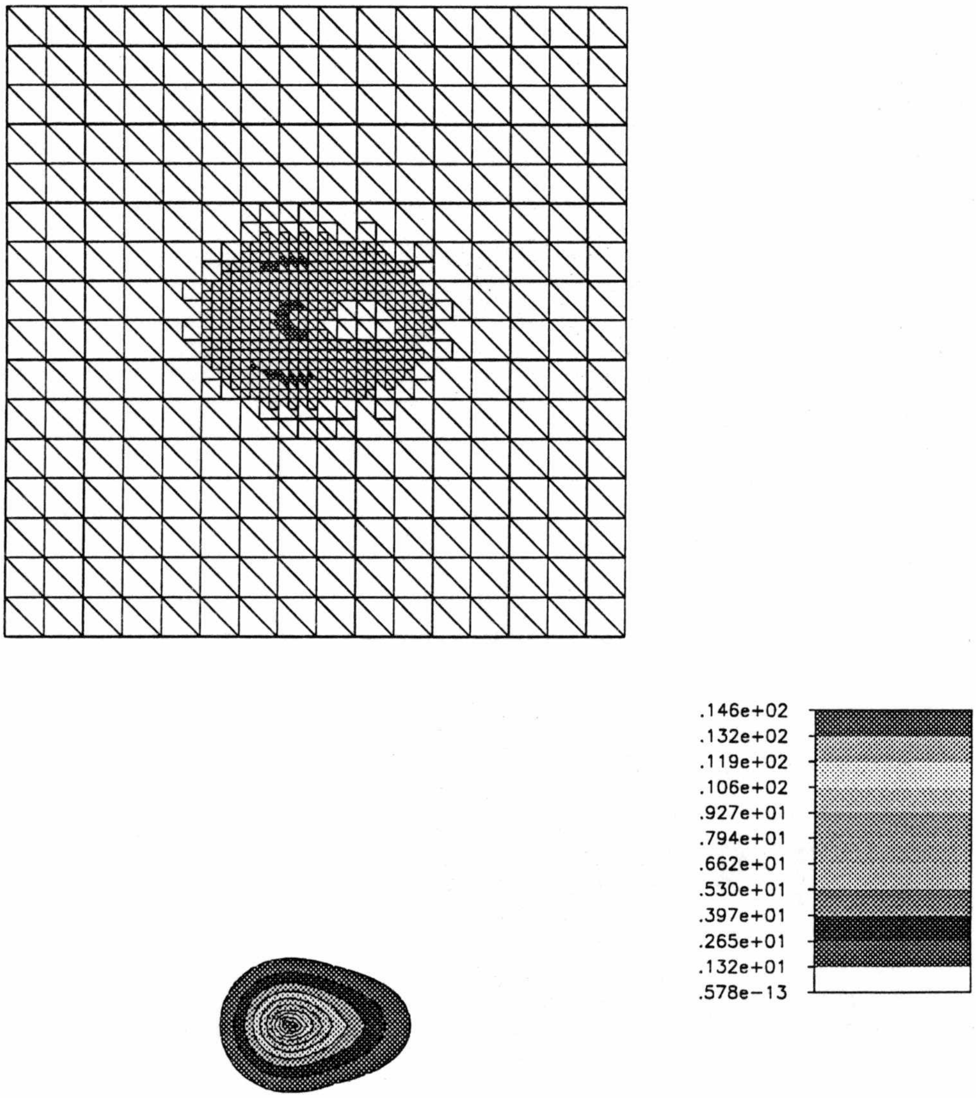


Figure 5.26: Experiment 4: Triangulation and contour plot of the solution at time 0.8×10^{-3}.

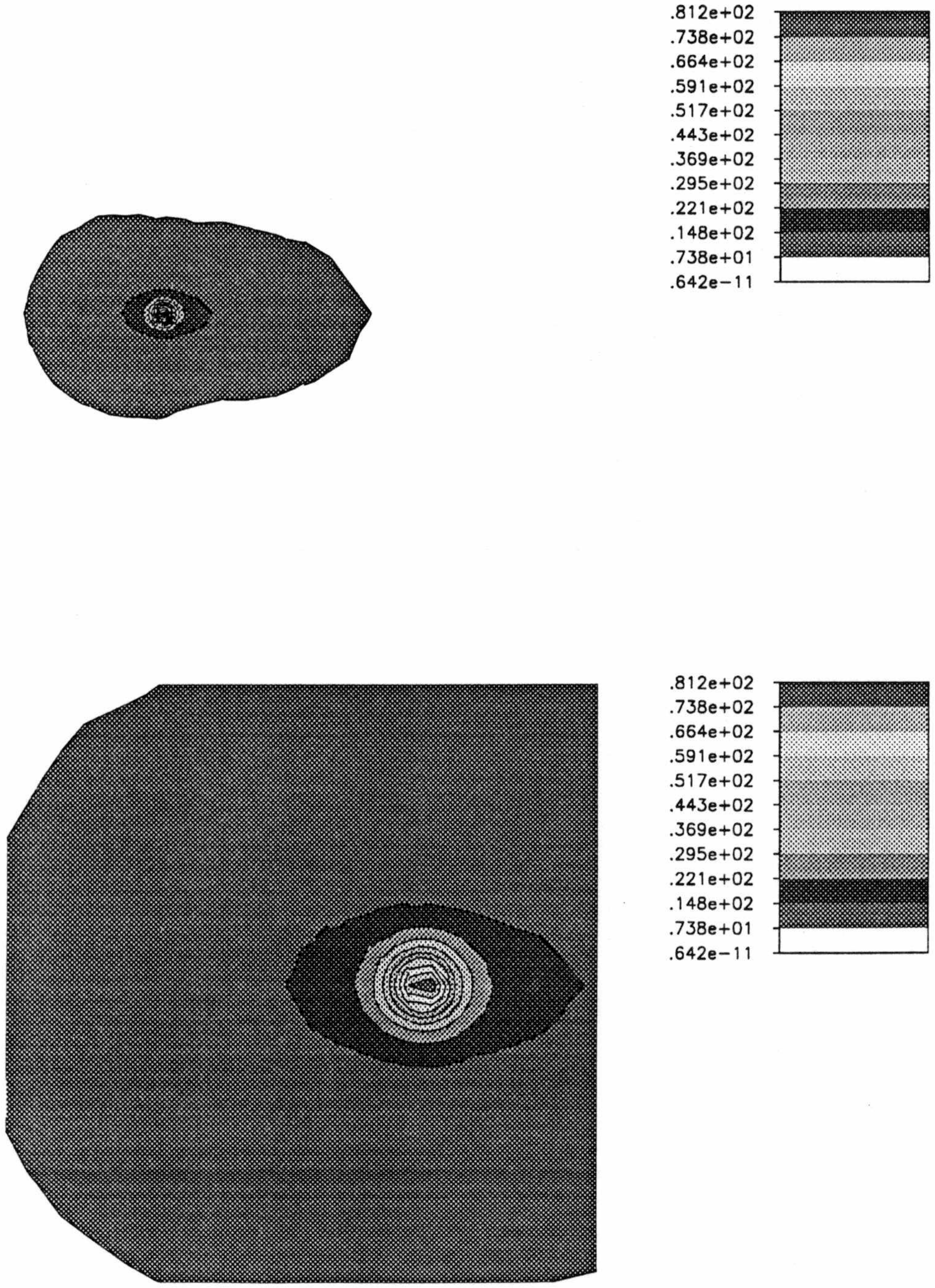


Figure 5.27: Experiment 4: Contour plots of the solution at time 0.11×10^{-2}. Global view (top) and detail around the point $(1, 1)$ (bottom).

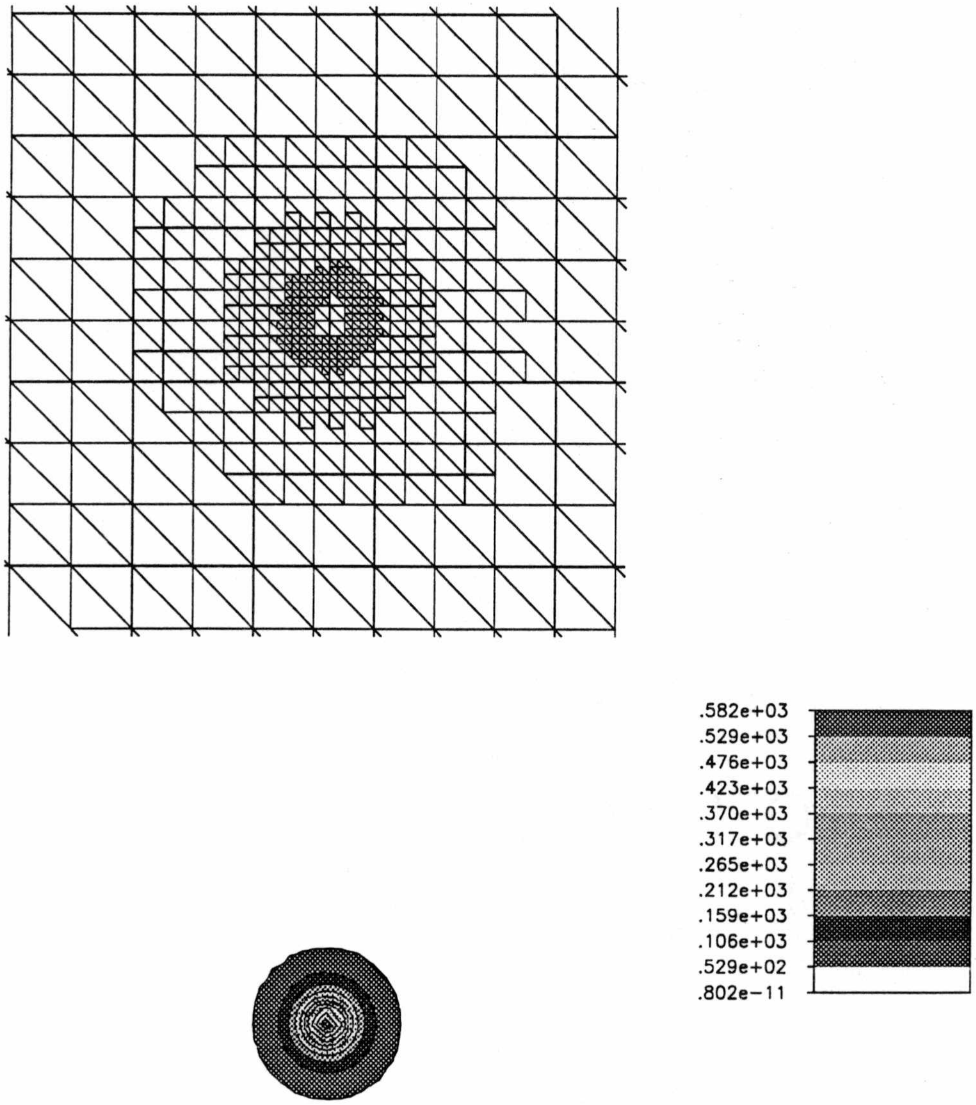


Figure 5.28: Experiment 4: Triangulation and contour plot of the solution at time $0.11153125 \times 10^{-2}$.

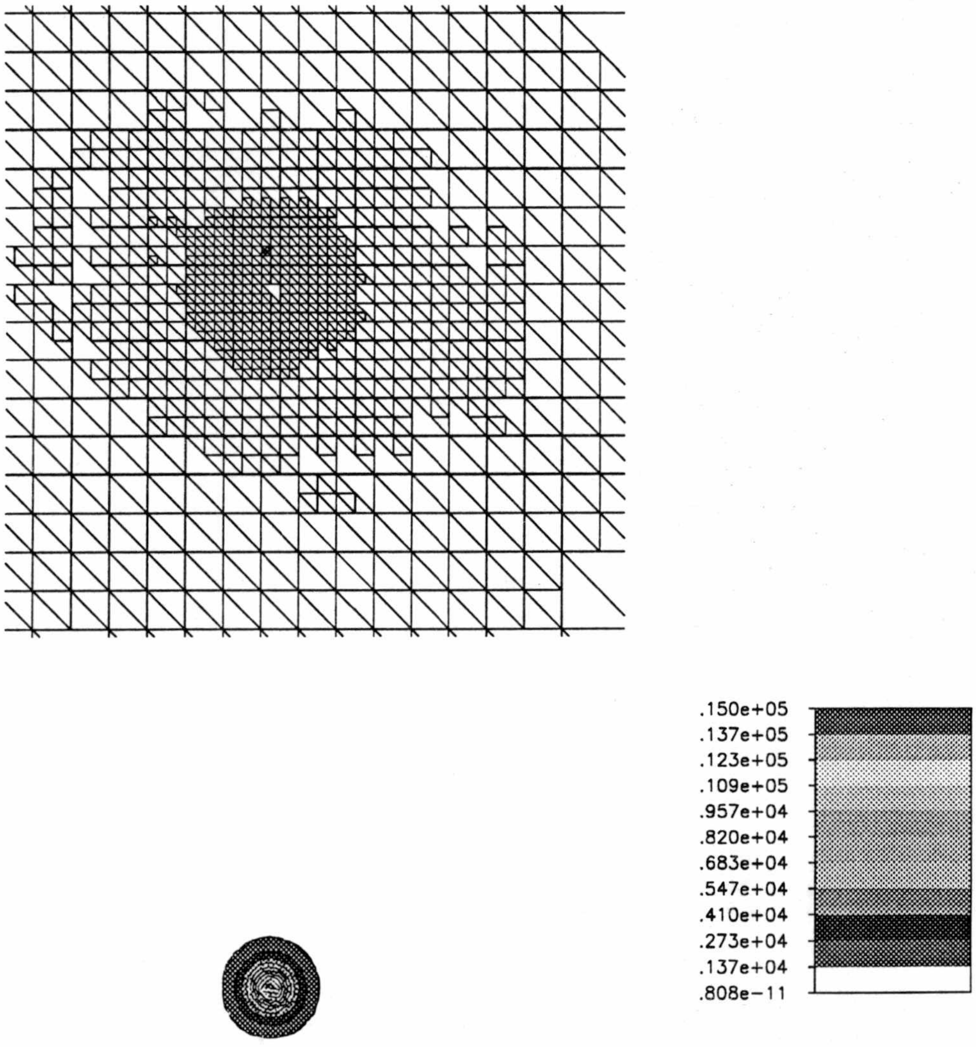


Figure 5.29: Experiment 4: Triangulation and contour plot of the solution at time $0.111579742 \times 10^{-2}$.

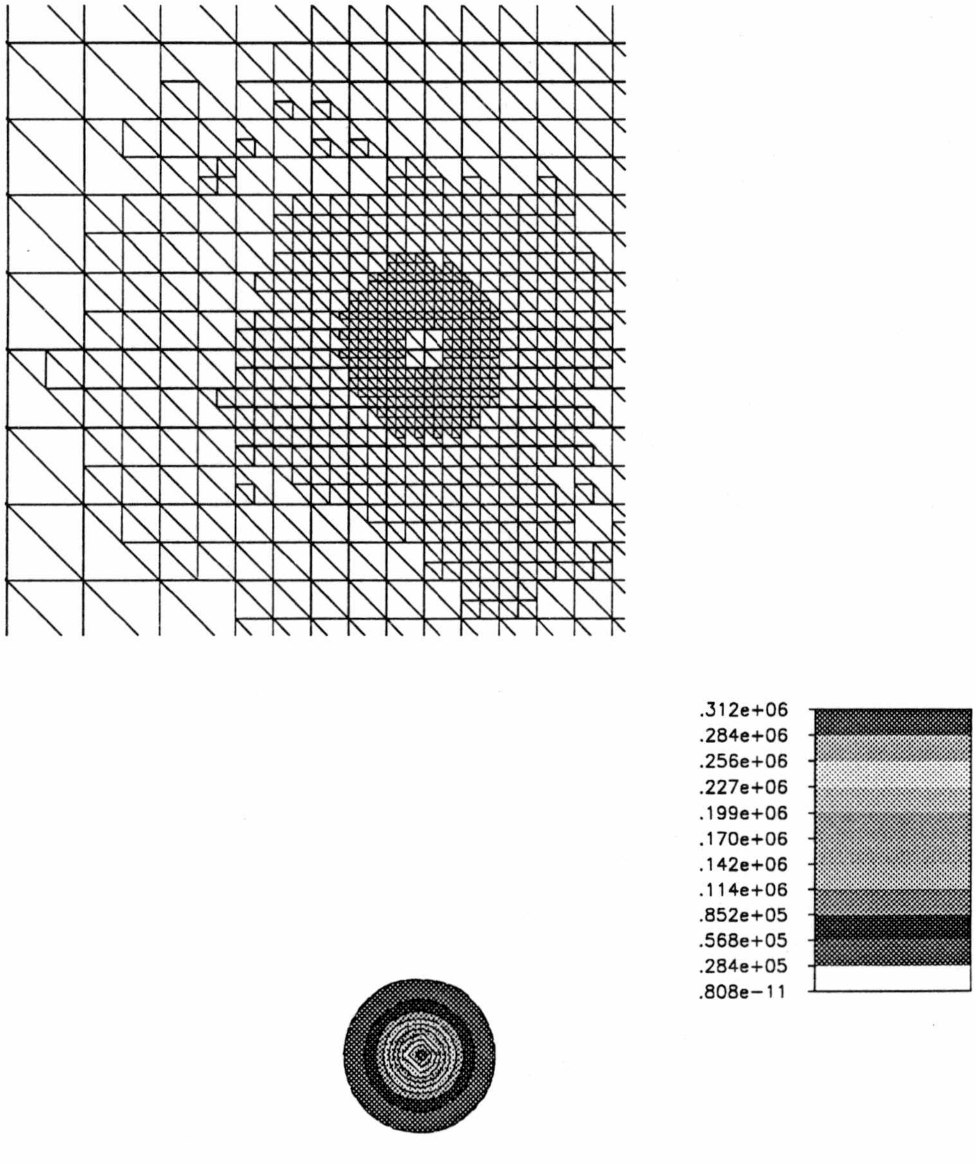


Figure 5.30: Experiment 4: Triangulation and contour plot of the solution at time $0.111579819 \times 10^{-2}$.

5.5 Three peaks, Radially symmetric

The initial profile u^0 for this experiment consisted of three Gaussian peaks positioned at the points with coordinates $(1, 1.2)$, $(0.8, 0.8)$ and $(1.2, 0.8)$. A surface/contour plot of this initial profile is shown in Figure 5.31. The square of the L^2–norm of the starting profile was 7.7888416 while the second invariant was -2847.78765. The computational domain for this experiment was the square $\Omega = [0, 2]^2$ and the starting grid a uniform triangulation of Ω with 512 triangles. Our refinement mechanism refined the starting triangulation 136 times, resulting in a final triangulation containing 14186 triangles. The starting time step was halved 36 times and the final time step was $0.582076609135 \times 10^{-15}$. It is interesting that, in contrast to the "Twin Peaks" experiment, only the Gaussian peak initially centered at $(1, 1.2)$ proceeded to blow up. The contour plot of the solution in a neighborhood of the point $(1, 1.2)$ and the corresponding triangulation are shown in Figures 5.32–5.35.

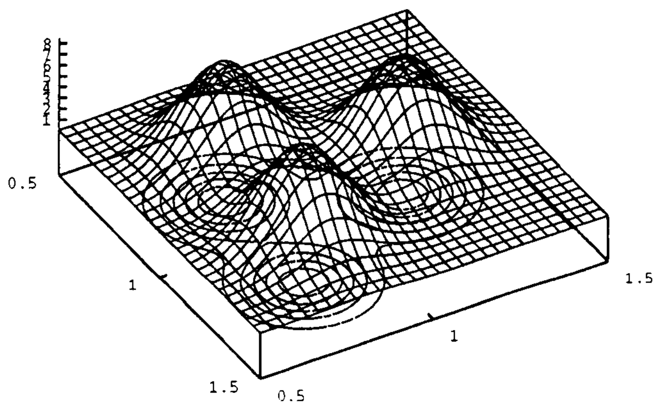


Figure 5.31: The initial profile u^0 for the fifth run.

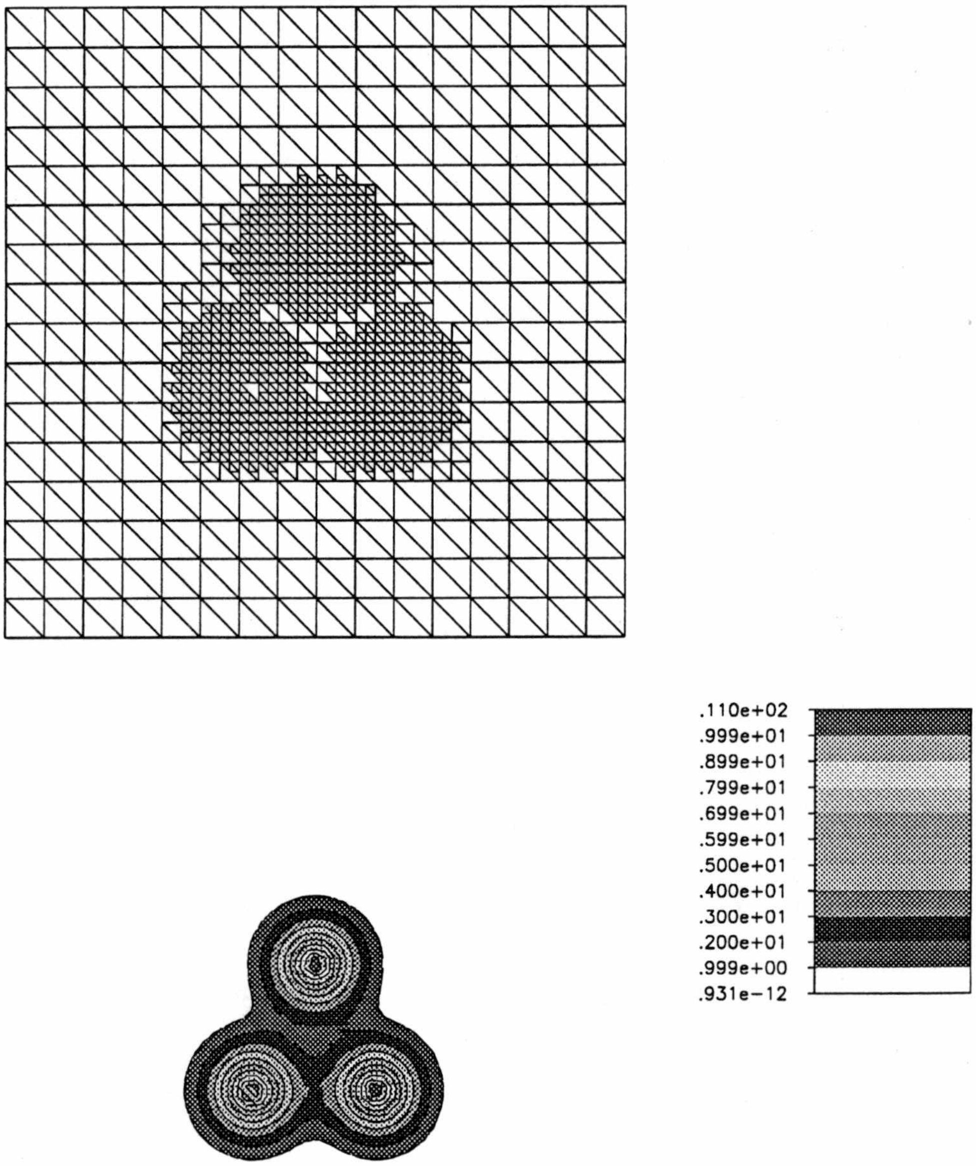


Figure 5.32: Experiment 5: Triangulation and contour plot of the solution at time 0.8×10^{-3}.

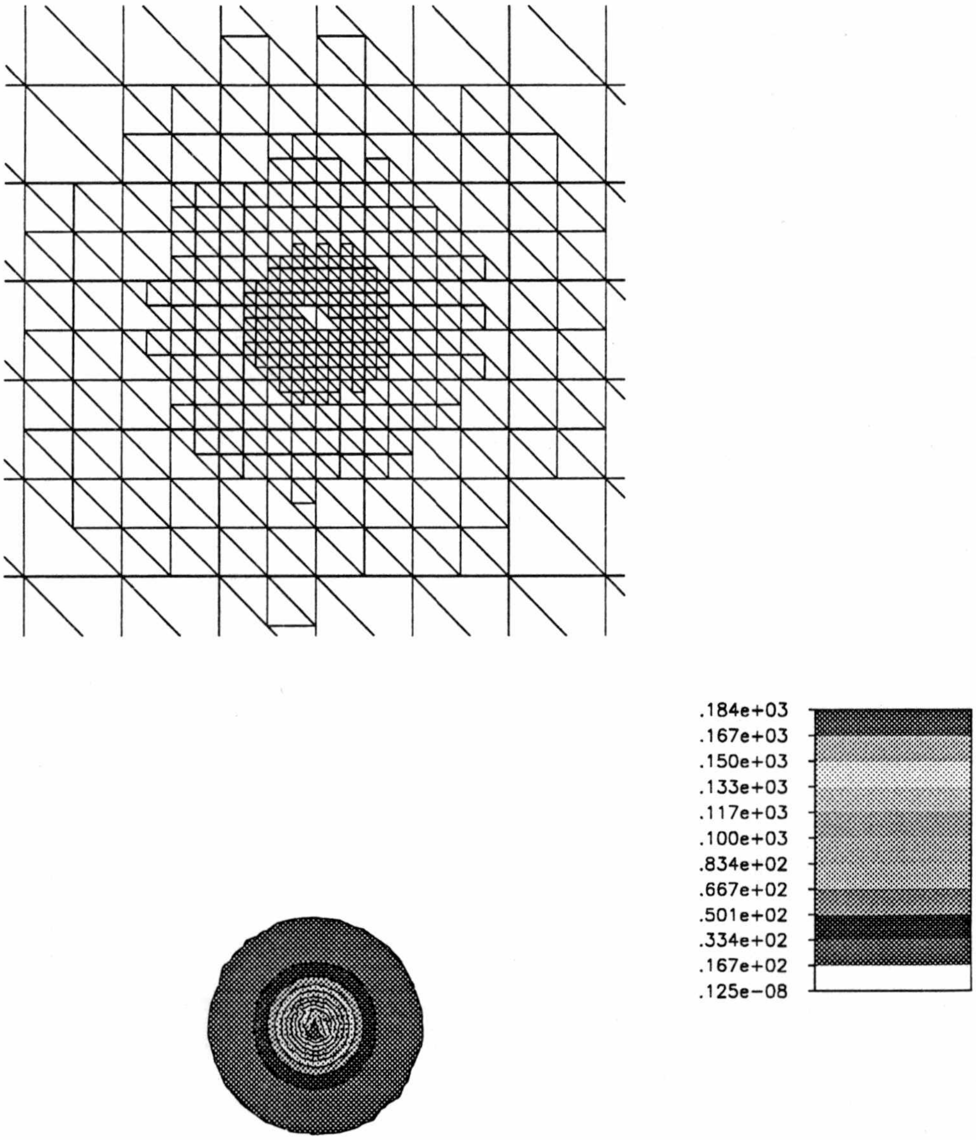


Figure 5.33: Experiment 5: Triangulation and contour plot of the solution at time 0.138125×10^{-3}.

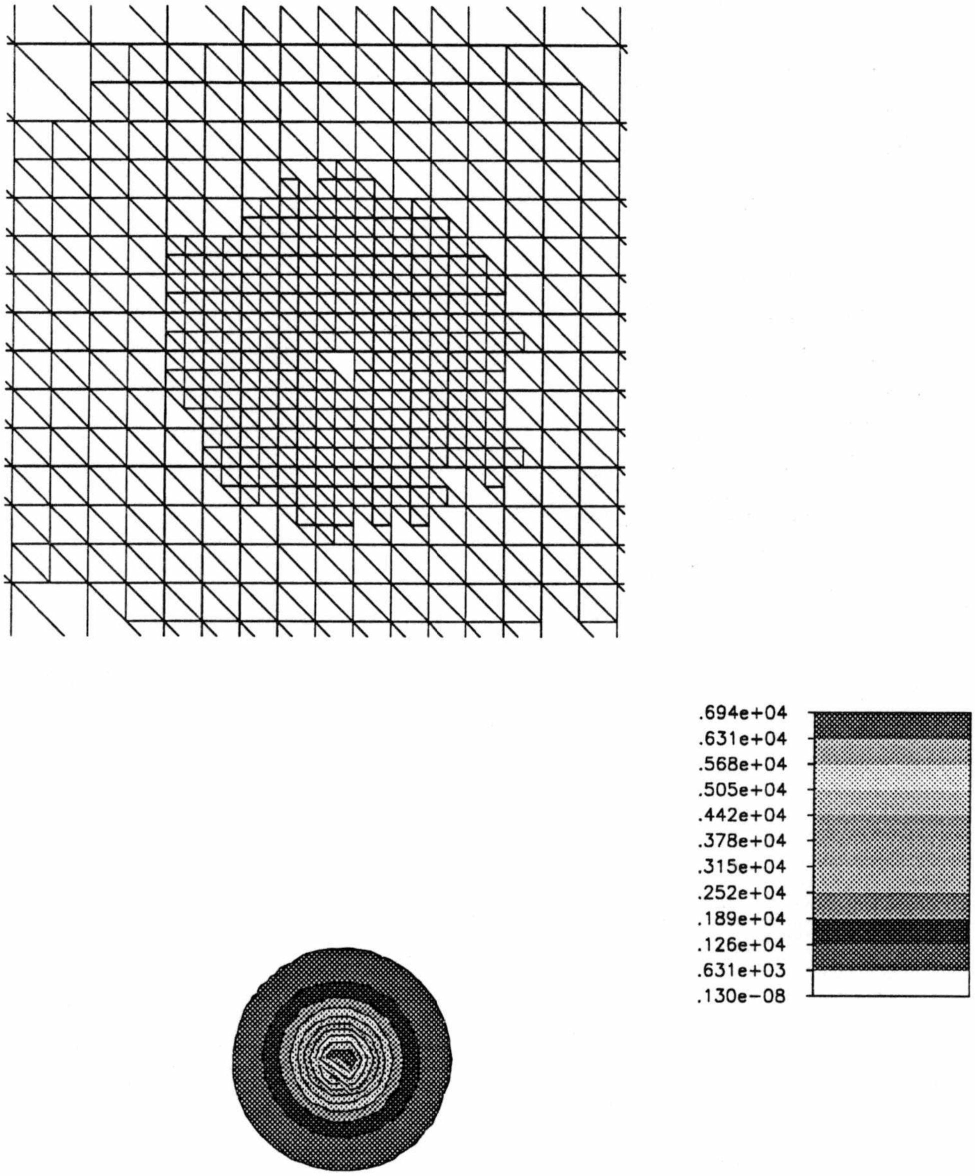


Figure 5.34: Experiment 5: Triangulation and contour plot of the solution at time $0.138495605 \times 10^{-2}$.

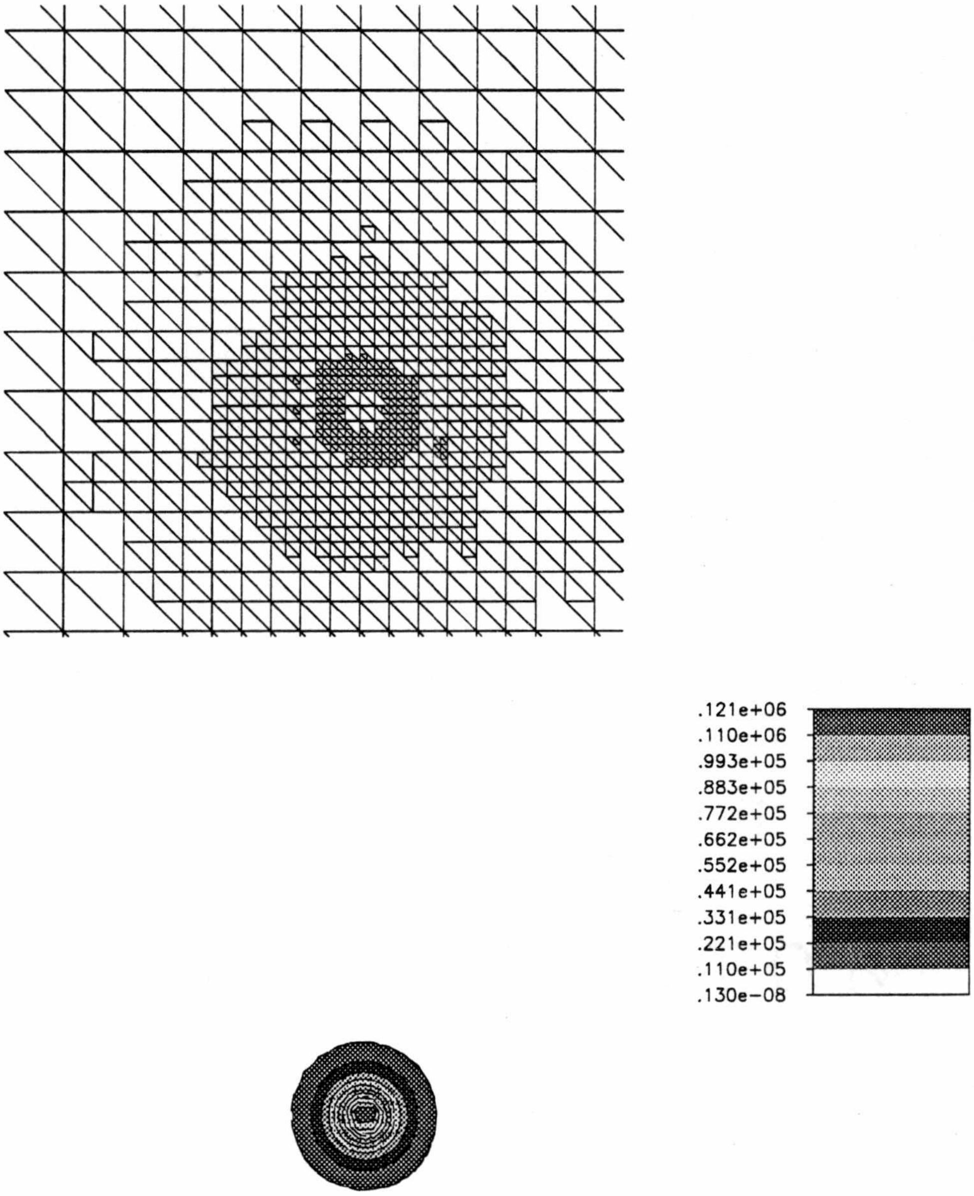


Figure 5.35: Experiment 5: Triangulation and contour plot of the solution at time $0.138495947 \times 10^{-2}$.

5.6 Four peaks, Radially symmetric

The initial profile for this experiment consisted of four Gaussian peaks positioned at the points with coordinates $(0.8, 1.2)$, $(0.8, 0.8)$, $(1.2, 0.8)$ and $(1.2, 1.2)$. Each of the four Gaussian peaks was of the form $6\sqrt{2}e^{-45(x-x_c)^2-45(y-y_c)^2}$, where (x_c, y_c) denotes any of the above four points. A surface/contour plot of this initial profile is shown in Figure 5.36. In the sequel, we shall refer to the peak positioned at $(0.8, 1.2)$ as the Northwest (NW) peak, while the other three peaks shall be referred to as the Southwest (SW), Southeast (SE) and Northeast (NE) peaks, respectively.

The square of the L^2–norm of this profile was 10.609972 and the second invariant was -3932.45936. The computational domain for this run was the square $\Omega = [0, 2]^2$ and the starting grid a uniform triangulation of Ω containing 512 triangles. Our temporal refinement mechanism halved the initial time step 35 times. The final time step was $0.11641532 \times 10^{-14}$. The spatial refinement test refined the starting grid 125 times, resulting in a final triangulation containing 16322 triangles. The maximum altitude achieved by the solution in this run was 0.63755503×10^{7}.

We followed the evolution of this initial profile in time and summarized the results in Figures 5.37–5.42. We observe from Figure 5.37 and Figure 5.38 that, initially, the four peaks evolve in a similar fashion. Figure 5.39, however, shows that the L^∞–norms attained by the SE and NW peaks are clearly smaller than the corresponding norms of the SW and NE peaks (see Figure 5.40). At later times, only the SW peak proceeds to blow up. Figure 5.41 shows the contour plot of the solution in a small neighborhood of the point $(0.8, 0.8)$. Figure 5.42 shows the contour plot of the solution and the underlying triangulation in a square centered at the point $(0.8, 0.8)$ and with side 3.125×10^{-6}.

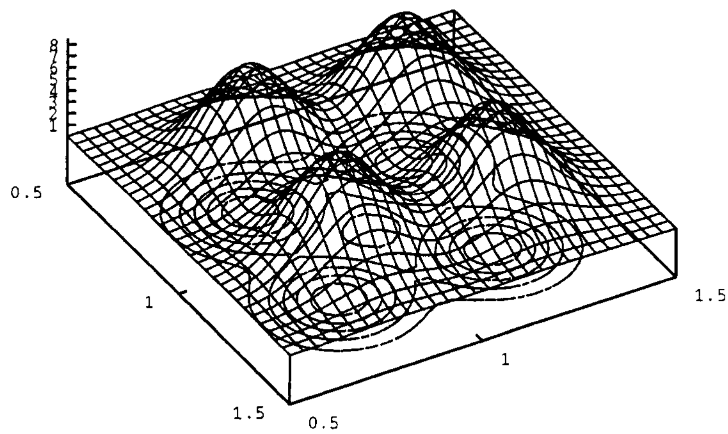


Figure 5.36: The initial profile u^0 for the sixth run.

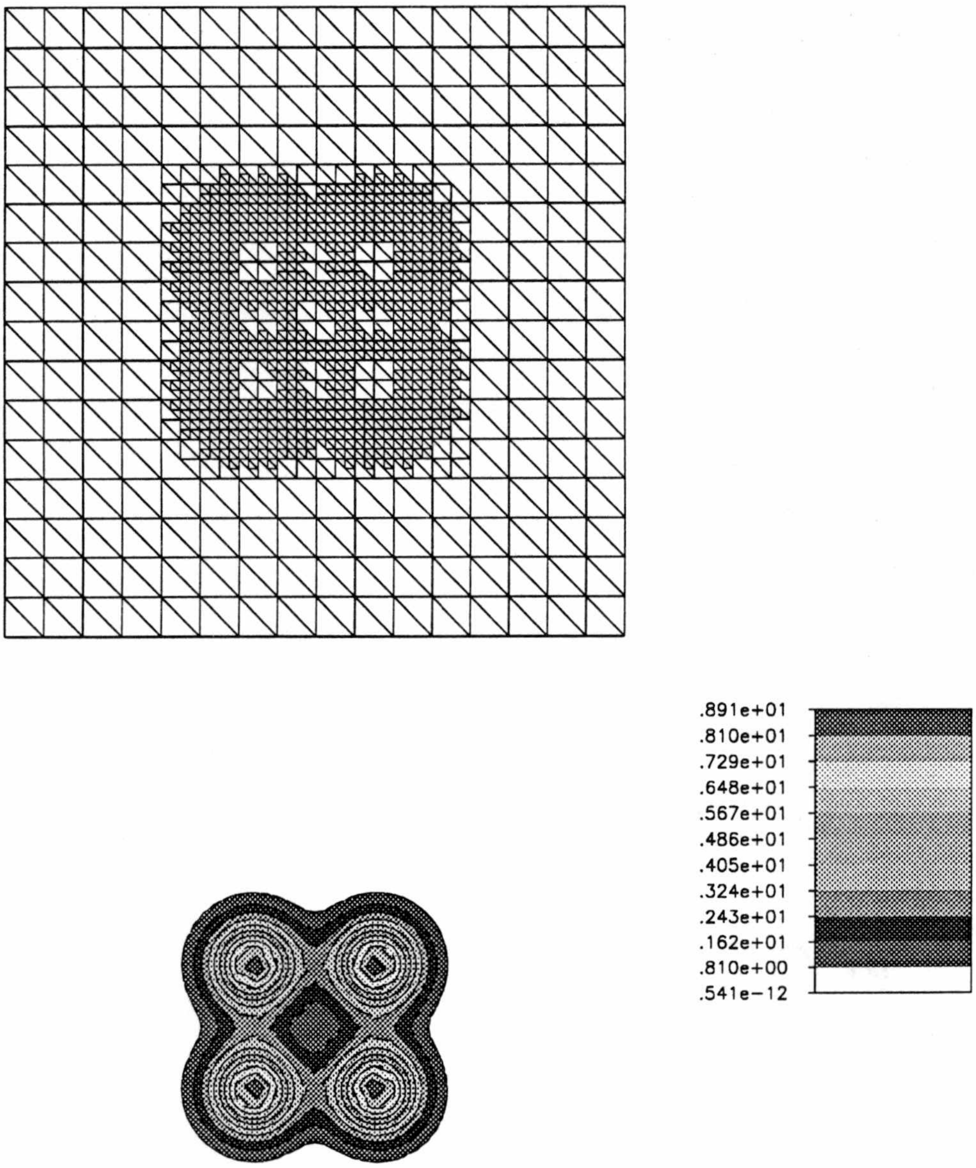
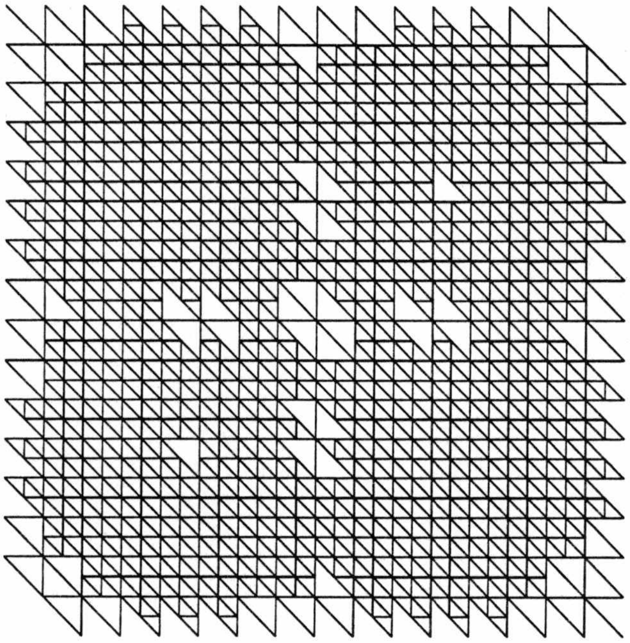


Figure 5.37: Experiment 6: Triangulation and contour plot of the solution at time 0.4×10^{-3}.



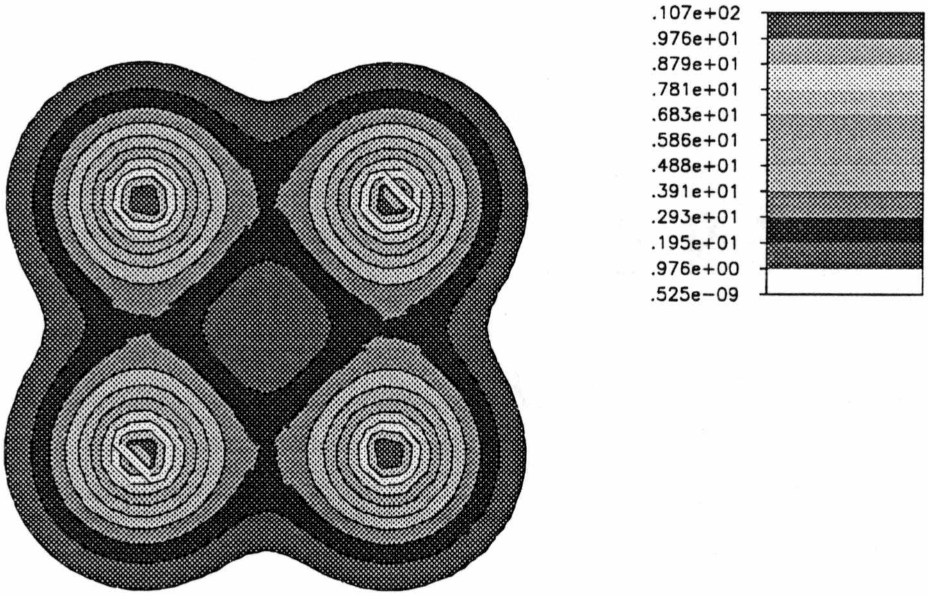


Figure 5.38: Experiment 6: Triangulation and contour plot of the solution at time 0.8×10^{-3}.

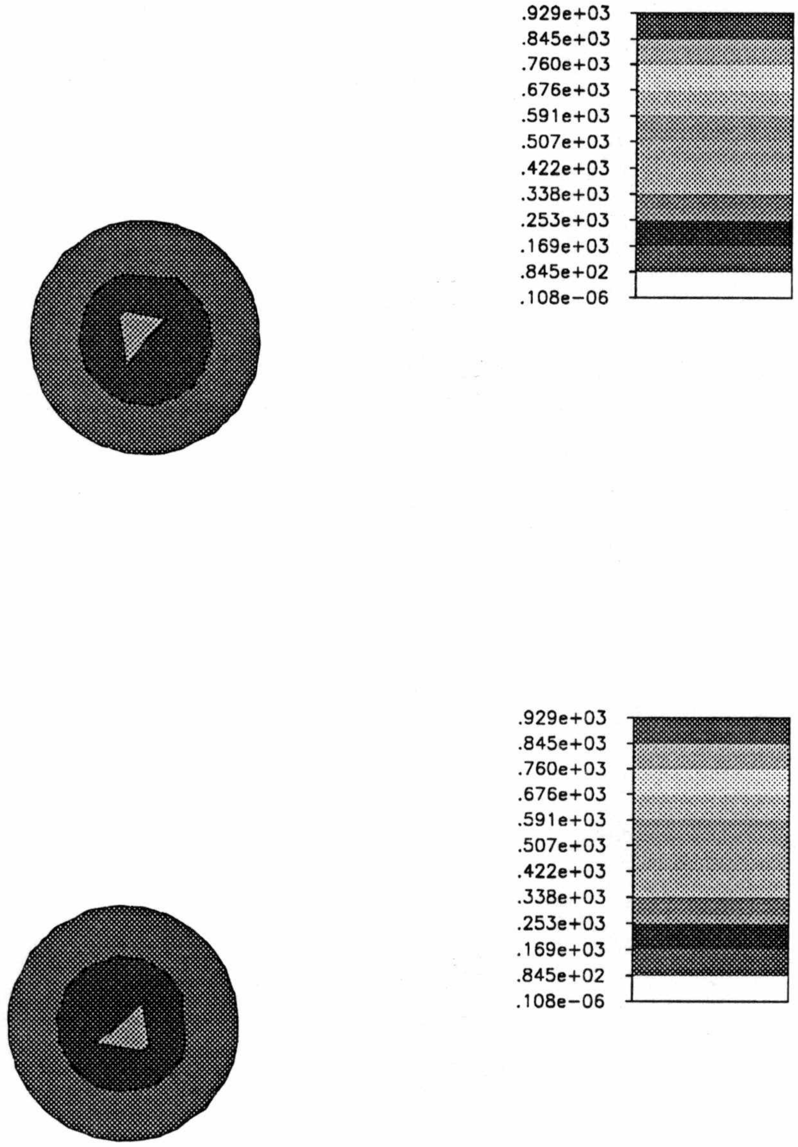


Figure 5.39: Experiment 6: Contour plots of the NW (top) and SE (bottom) peaks at time 0.138738×10^{-2}.

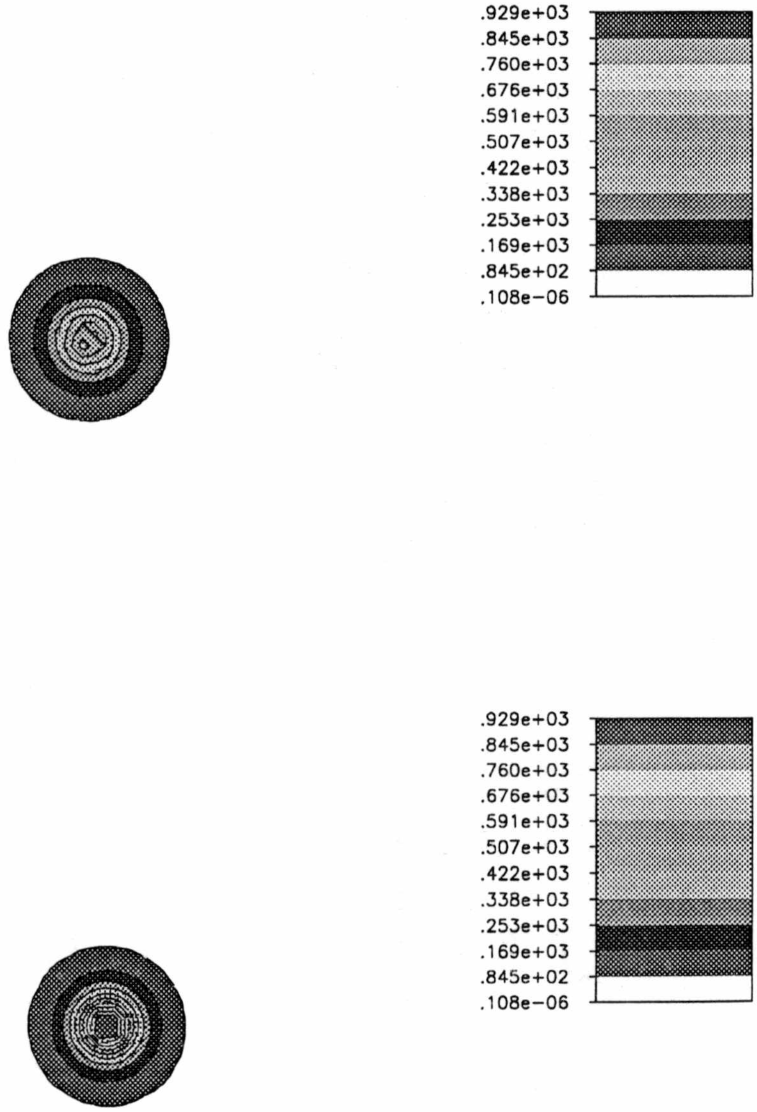
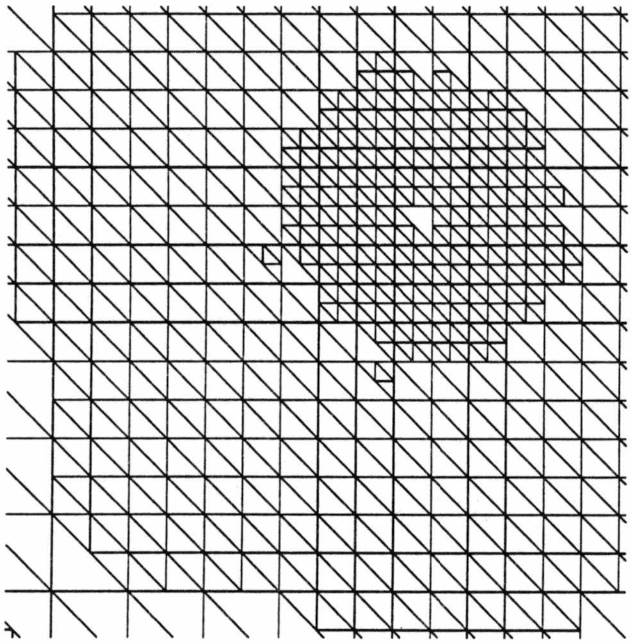


Figure 5.40: Experiment 6: Contour plots of the NE (top) and SW (bottom) peaks at time 0.138738×10^{-2}.



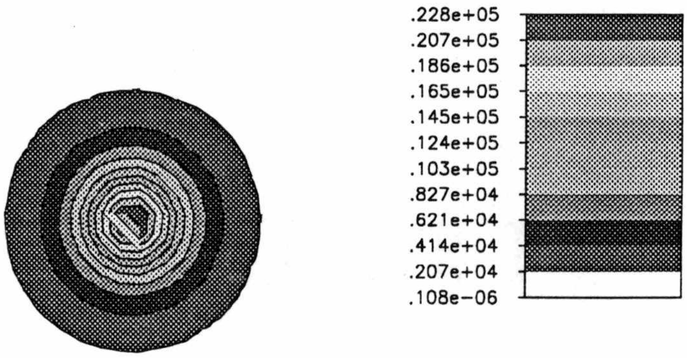


Figure 5.41: Experiment 6: Triangulation and contour plot of the SW peak at time 0.138754×10^{-2}.

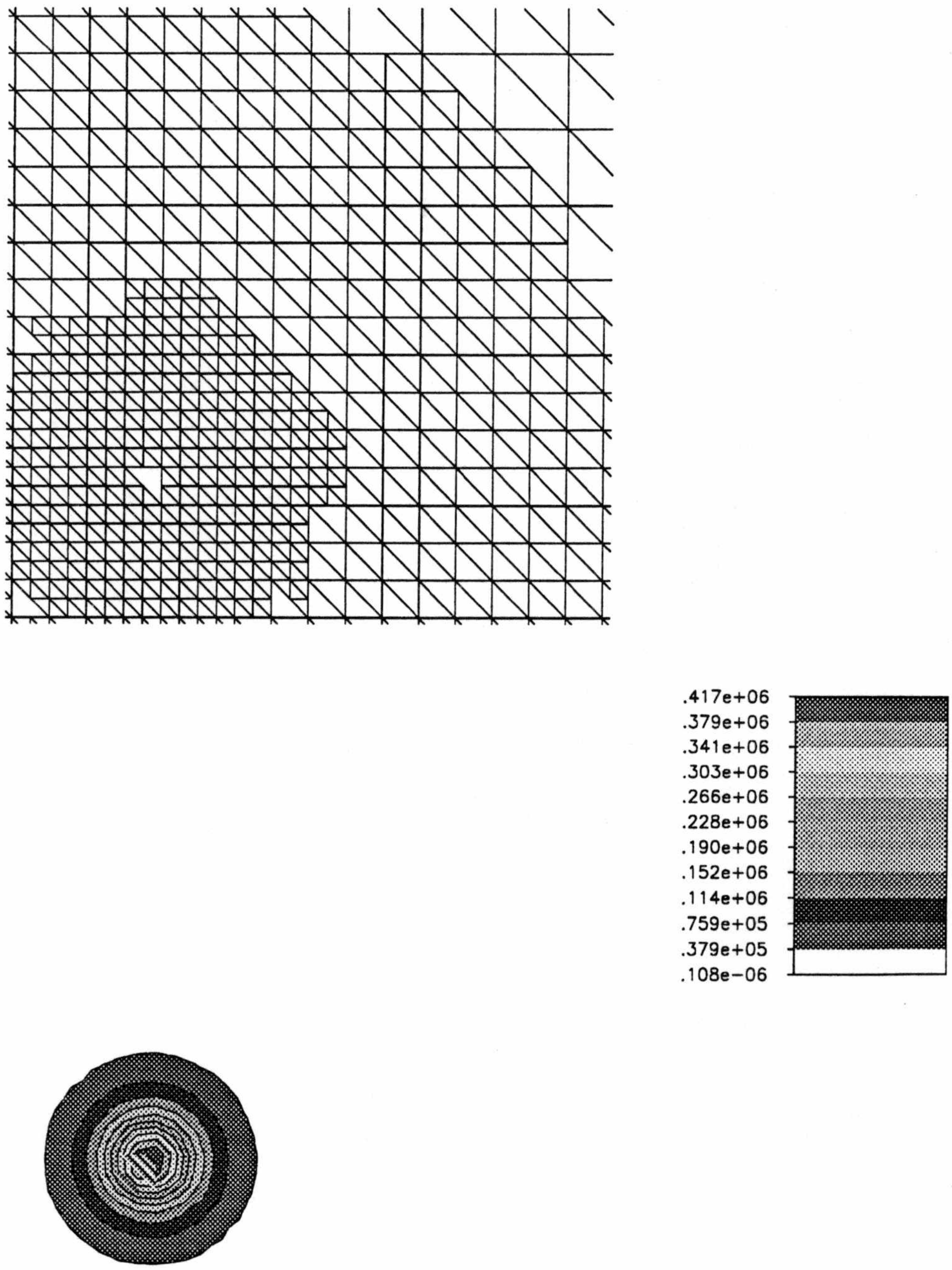


Figure 5.42: Experiment 6: Contour plots of the SW peak and the underlying triangulation at time $0.138754145 \times 10^{-2}$.

5.7 Blow–up Rates

In this section we attempt to estimate the rates with which various norms of the solution tend to infinity as the time t approaches the blow–up time t^*. We let $M(t)$ denote a quantity associated with the solution u at time t and denote by ρ its blow–up rate. We assumed that, cf. [5],

$$M(t) \cong \left[\ln\ln(\frac{1}{t^*-t})/(t^*-t)\right]^{\rho}$$

To estimate ρ, we compute the quantity $M(t)$ at two distinct times t_1 and t_2 near to, but less than t^* and compute the ratio

$$\frac{\ln(M(t_1)/M(t_2))}{\ln\left[\frac{\ln\ln\left(\frac{1}{t^*-t_1}\right)/(t^*-t_1)}{\ln\ln\left(\frac{1}{t^*-t_2}\right)/(t^*-t_2)}\right]} \tag{5.2}$$

Notice that the estimation of ρ from the above formula is rather delicate and requires that one computes the differences t^*-t_1 and t^*-t_2 carefully as to avoid cancellation. A systematic procedure for doing that is exhibited in [5] and [11]. In Table 5.1 we show the rates of blow–up for the L^∞ and L^4 norms of the solutions u of the Single Peak, Radially symmetric experiment (left) and the Single Peak, Non Radially symmetric one (right). We chose the blow–up times t^* as the times at which our code stopped when running the experiments, and computed the L^∞ and L^4 norms of the numerical solutions after the last six spatial grid refinements. The times at which these norms were computed were very close to the respective blow–up times t^* and therefore are not shown. Instead, under the column labeled Ref., we have indicated the difference of the last spatial grid refinement level form the current one. The numbers strongly suggest that the L^∞ norm of the solution blows up at the rate of $\frac{1}{2}$, while the L^4 norm blows up at the rate of $\frac{1}{4}$. Notice that the L^4 rates seem to approach the blow–up rate of $\frac{1}{4}$ faster than the L^∞ rates approach $\frac{1}{2}$. This is probably attributed to the fact that the L^4 norms of the numerical solution were computed, using numerical quadrature, *exactly*, while the L^∞ norms were approximated by the values of the

Table 5.1: Blow–up rates for the Single Peak, Radially Symmetric (left) and Non Radially Symmetric (right) experiments.

Ref.	L^∞	L^4
5	0.50830	0.25044
4	0.50352	0.25036
3	0.49951	0.24984
2	0.50112	0.25028
1	0.50111	0.25009

Ref.	L^∞	L^4
5	0.498013	0.246929
4	0.498015	0.246929
3	0.502399	0.249989
2	0.502398	0.249992
1	0.501026	0.250073

numerical solution at the integration points, as well as the vertices of each triangle. Recall that $|u|^2$, restricted on any particular triangle, is a polynomial of degree four which makes the exact determination of its absolute maximum difficult.

We next attempted to calculate the blow–up rates of the L^∞ and L^4 norms of the numerical solutions obtained in the experiments involving initial profiles with several peaks. We first considered the computation of the blow–up rates for the L^∞ and L^4 norms of the numerical solution of our fourth numerical experiment. Recall that the initial profile for this experiment consisted of two Gaussian peaks of heights $6\sqrt{2}$ and 4 positioned at the points $(0.9, 1)$, resp., $(1.1, 1)$, that very rapidly merged and formed a single, radially symmetric peak, which proceeded to blow up. As before, we computed all the relevant quantities at the times of the last four spatial grid refinements and calculated the blow–up rates from (5.2). The blow–up time t^* was chosen as the time of the last grid refinement. The blow–up rates for the L^∞ and L^4 norms of the numerical solution of this experiment are summarized in Table 5.2. Notice that the relatively small amplitude achieved by our numerical solution (of the order of $O(10^6)$) and, subsequently, the inexact estimation of the blow–up time t^*, resulted in blow–up rates which appear to be less robust than the rates reported earlier for the Single peak experiments. The computed rates for the L^∞ norm are, however, close to $\frac{1}{2}$ and the

Table 5.2: Blow–up rates for the Twin Peak, Non radially Symmetric experiment.

Ref.	L^∞	L^4
3	0.49885	0.25161
2	0.48710	0.24528
1	0.49743	0.24950

corresponding rates for the L^4 norm are close to $\frac{1}{4}$ as well. We next attempted to determine the blow–up rates for the L^∞ norm of the solution of our third numerical experiment, involving two Gaussian peaks of height $6\sqrt{2}$, positioned at the points with coordinates $(0.8, 1.2)$ and $(1.2, 0.8)$, cf. Section 5.3. There it was observed that the maximum amplitude of *both* initial peaks blows up as time approaches the blow-up time $t^* = 0.1383987930119 \times 10^{-2}$. The blow–up rates for the L^∞ norm of both peaks were almost identical, explained, of course, by the fact that both peaks achieved approximately the same L^∞ norm at the end of the computation. As with the previous experiment, all computed rates were close to $\frac{1}{2}$ but were not as robust as the rates computed for the experiments involving solutions that blow up at exactly one point.

5.8 Conclusions and future work

In this thesis we introduced an adaptive scheme for approximating the solution of the two-dimensional Nonlinear Schrödinger equation. The spatial discretization was based on nonconforming finite elements while the temporal discretization was achieved by a variable step Crank–Nicolson scheme. The theoretical convergence analysis of the fully discrete scheme was elaborated in Chapter 3. In order to simplify the analysis, a *modified scheme* was introduced in Section 3.1 whereby the nonlinear term $|u|^2u$ was replaced by $\gamma(|u|^2)u$. The function $\gamma : [0, \infty] \to [0, 4M^2]$, where $M = \sup_{[0,t^*]} ||u||_\infty + 1$, was chosen as a smooth function with bounded derivatives of order up to two, which coincided with $|u|^2$ for $|u| \leq M$. The existence of the approximations defined by the modified scheme was established in Section 3.2 and optimal order error estimates in $L^\infty(L^2)$ were shown in Section 3.3. The convergence of the original fully discrete scheme was carried out in Section 3.3.3. There it was shown that the solution of the modified scheme is also a solution of the original scheme by establishing the bound $||U^{n,1}||_\infty \leq M$ for the approximations $U^{n,1}$ generated by the modified scheme. This last result relied on the estimate

$$||v||_{L^p(\Omega)} \leq cp\,||v||_{1,h}, \qquad \forall\, v \in E_h$$

and the inverse inequality

$$||\chi||_{L^\infty(D)} \leq ch^{-2/p}||\chi||_{L^p(D)}, \qquad \forall \chi \in \mathbb{P}_r(D),$$

whose proofs appear in Section 3.4. The inverse inequality above is new and extends a similar result to the case of elements with curved boundaries.

Chapter 4 discussed the details of the implementation of the adaptive algorithm and the spatial and temporal mesh refinement mechanisms. In Chapter 5 several numerical experiments concerning the formation of singularities were performed with our adaptive code. Several profiles with either one or multiple peaks were used as initial data in our numerical experiments. Our method was able to follow the evolution of these profiles up to times very close to blow–up. Of particular interest were the experiments with anisotropic initial conditions and with several peaks. In the former case it was observed that the corresponding solutions go through an anisotropic phase before blowing up in a locally isotropic form. In the latter case, we were able to exhibit that the solution of the NLS equation with an initial profile consisting of two peaks of equal height blows up around two distinct points. The high amplitude achieved by our solutions and the apparent success of the mesh refinement mechanism motivated us to compute the blow–up rates for the L^∞ and L^4 norms of the solution. For both isotropic and anisotropic profiles the blow–up rates of Section 5.7 were found in good agreement with the theoretically predicted ones.

The computational results obtained in this work are quite encouraging. They establish our adaptive algorithm as a powerful tool in the numerical investigation of blow–up phenomena associated with the NLS equation. In the near future we plan to use this algorithm to study more thoroughly the evolution of anisotropic (Gaussian) profiles, as well as, profiles with several peaks. Moreover, we plan to use our mesh refinement mechanism to investigate blow–up phenomena associated with other equations. On the theoretical part, we plan to supplement this thesis by establishing convergence in the $||\cdot||_{1,h}$ norm and remove or improve the mesh conditions on the spatial and temporal discretization parameters h and k.

BIBLIOGRAPHY

Bibliography

[1] M. ABLOWITZ AND H. SEGUR, *Solitons and the Inverse Scattering Transform*, SIAM Studies in Applied Mathematics, Society for Industrial and Applied Mathematics, 1981.

[2] R. ADAMS, *Sobolev Spaces*, Academic Press, New York, 1970.

[3] G. D. AKRIVIS, V. A. DOUGALIS, AND O. A. KARAKASHIAN, *Solving the system of equations arising in the discretization of some nonlinear p.d.e.'s by Implicit Runge–Kutta methods.* To appear in RAIRO MMAN.

[4] ——, *On fully dicrete galerkin methods of second–order temporal accuracy for the nonlinear Schrödinger equation*, Numer. Math., 59 (1991), pp. 31–53.

[5] G. D. AKRIVIS, V. A. DOUGALIS, O. A. KARAKASHIAN, AND W. R. MCKINNEY, *Numerical approximation of blow–up of radially symmetric solution of the nonlinear Schrödinger equation.* In preparation.

[6] I. BABUŠKA AND J. E. OSBORN, *Analysis of finite element methods for second order boundary value problems with mesh dependent norms*, Numer. Math., 34 (1980), pp. 41–62.

[7] G. A. BAKER, W. N. JUREIDINI, AND O. A. KARAKASHIAN, *Piecewise solenoidal vector fields and the Stokes problem*, SIAM J. Numer. Anal., 27 (1990), pp. 1466–1485.

[8] R. E. BANK, *PLTMG: A software Package for Solving Elliptic Partial Differential Equations User's Guide 7.0*, vol. 15 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[9] J. L. BONA, V. A. DOUGALIS, O. A. KARAKASHIAN, AND W. R. MCKINNEY, *Fully-discrete methods with grid refinement for the generalized Korteweg-de Vries equation*, in Viscous profiles and numerical methods for shock waves, M. Shearer, ed., SIAM, Philadelphia, 1991, pp. 1–11.

[10] ——, *Computations of blow-up and decay for periodic solutions of the generalized Korteweg-de Vries equation*, Appl. Numer. Math., 10 (1992), pp. 335–355.

[11] ——, *Conservative, high-order numerical schemes for the generalized Korteweg-de Vries equation*, Phil. Trans. of Royal Society of London, A (1995), pp. 107–164.

[12] S. BRENNER AND L. R. SCOTT, *The Mathematical theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[13] F. E. BROWDER, *Existence and uniqueness theorems for solutions of nonlinear boundary value problems*, Applications of Partial Differential Equations, (R. Finn, ed.), American Mathematical Society, Providence, 1965.

[14] R. Y. CHIAO, E. GARMIRE, AND C. TOWNES, *Self-trapping of optical beams*, Phys. Rev. Lett., (1964), pp. 479–482.

[15] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[16] K. DEKKER AND J. G. VERWER, *Stability of Runge-Kutta methods for Stiff Nonlinear Differential equations*, North-Holland, Amsterdam, 1984.

[17] M. DELFOUR, M. FORTIN, AND G. PAYRE, *Finite-difference solutions of a non-linear Schrödiner equation*, J. Math. Phys., 44 (1981), pp. 277–288.

[18] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.

[19] ——, *Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[20] D. GILBARG AND N. TRUDINGER, *Partial Differential Equations of Second Order*, vol. 224 of Grundlehen der mathematischen Wissenschaften, Springer–Verlag, 2nd ed., 1983.

[21] J. GINIBRE AND G. VELO, *On a class of nonlinear Schrödinger equations i: the cauchy problem, general case*, J. Funct. Anal., 32 (1979), pp. 1–32.

[22] R. T. GLASSEY, *On the blowing up of solutions of the Cauchy problem for Nonlinear Schrödinger equations*, J. Math. Phys., 18 (1977), pp. 1974–1977.

[23] D. M. GRIFFITHS, A. R. MITCHELL, AND J. L. MORRIS, *A numerical study of the nonlinear Schrödinger equation*, Comput. Methods Appl. Mech. Engr., 45 (1984), pp. 177–215.

[24] A. HASEGAWA AND F. TAPPERT, *Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. I. anomalous dispersion*, Appl. Phys. Lett., 23 (1973), pp. 142–144.

[25] O. KARAKASHIAN AND C. MAKRIDAKIS, *A space–time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method.* Submitted.

[26] P. L. KELLEY, *Self-focusing of optical beams*, Phys. Rev. Lett., 5 (1965), pp. 1005–1012.

[27] M. J. LANDMAN, G. C. PAPANICOLAOU, C. SULEM, AND P.-L. SULEM, *Rate of blowup for solutions of the nonlinear Schrödinger equation at critical dimensions*, Phys. Rev. A, 38 (1988), pp. 3837–3843.

[28] M. J. LANDMAN, G. C. PAPANICOLAOU, C. SULEM, P.-L. SULEM, AND X. P. WANG, *Stability of isotropic singularities for the nonlinear Schrödinger equation*, Physica D, 47 (1991), pp. 393–415.

[29] D. W. MCLAUGHLIN, G. C. PAPANICOLAOU, C. SULEM, AND P.-L. SULEM, *Focusing singularity of the cubic Schrödinger equation*, Phys. Rev. A, 34 (1986), pp. 1200–1210.

[30] B. L. MESURIER, G. PAPANICOLAOU, C. SULEM, AND P.-L. SULEM, *The focusing singularity of the nonlinear Schrödinger equation*, in Directions in Partial Differential equations, M. G. Crandall, P. H. Rabinowitz, and R. E. Turner, eds., Academic Press, New York, 1987, pp. 159–201.

[31] ——, *Focusing and multifocusing solutions of the nonlinear Schrödinger equation*, Physica D, 32 (1988), pp. 210–226.

[32] A. C. NEWELL, *Solitons in Mathematics and Mathematical Physics*, vol. 48 of CBMS Applied Mathematics Series, Society for Industrial and Applied Mathematics, Philadelphia, 1988.

[33] T. OGAWA AND Y. TSUTSUMI, *Blow-up of solutions of for the Nonlinear Schrödinger Equation with quartic potential and periodic boundary condition*, in Functional Analytic Methods for Partial Differential Equations, A. Fujita, T. Ikebe, and S. T. Kuroda, eds., vol. 1450 of Lecture Notes in Math., Springer–Verlag, 1990, pp. 236–251.

[34] J. J. RASMUSSEN AND K. RYPDAL, *Blow-up in Nonlinear Schrödinger Equations-I, A General Review*, Physica Scripta, 33 (1986), pp. 481–497.

[35] U. RÜDE, *On the accurate computation of the solution of Laplace's and Poisson's equation*, in Multigrid Methods: Theory, Applications, Supercomputing: proceedings of the Third Copper Mountain Conference on Multigrid Methods, April 5–10, 1987, S. McCormick, ed., Marcel Dekker, New York, 1988.

[36] ——, *Mathematical and Computational Techniques for Multilevel Adaptive Methods*, vol. 13 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1993.

[37] J. M. SANS-SERNA, *Methods for the numerical solution of the nonlinear Schrödinger equation*, Math. Comp., 43 (1984), pp. 21–27.

[38] J. M. SANS-SERNA AND J. G. VERWER, *Conservative and nonconservative schemes for the solution of the nonlinear Schrödinger equation*, IMA J. Numer. Anal., 6 (1986), pp. 25–42.

[39] E. STEIN, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, NJ, 1970.

[40] W. A. STRAUSS, *The Nonlinear Schrödinger Equation*, in Contemporary developments in continuum mechanics and Partial Differential Equations, G. M. de la Penha and L. A. J. Medeiros, eds., North–Holland, Amsterdam, 1978, pp. 452–465.

[41] P.-L. SULEM, C. SULEM, AND A. PATERA, *Numerical simulation of singular solutions to the two–dimensional cubic schrödinger equation*, Comm. Pure Appl. Math., 37 (1984), pp. 755–778.

[42] V. I. TALANOV, *Self-focusing of wave beams in nonlinear media*, JETP Lett., 2 (1965), pp. 138–141.

[43] S. N. VLASOV, L. V. PISKUNOVA, AND V. I. TALANOV, *Structure of the field near the a singularity arising form self-focusing in a cubically nonlinear medium*, Soviet Physics JETP, 48 (1978), pp. 808–812.

[44] M. I. WEINSTEIN, *Nonlinear Schrödinger equation and sharp interpolation estimates*, Communs. in Math. Phys., 87 (1983), pp. 567–576.

[45] G. B. WHITHAM, *Linear and Nonlinear Waves*, J. Wiley, New York, 1974.

[46] D. WOOD, *The self-focusing singularity in the nonlinear Schrödinger equation*, Studies in Appl. Math., 71 (1984), pp. 103–115.

[47] V. E. ZAKHAROV, *Collapse of Langmuir waves*, JETP, 35 (1972), pp. 908–922.

[48] V. E. ZAKHAROV AND V. S. SYNAKH, *The nature of the self-focusing singularity*, Sov. Phys.–JETP, 41 (1976), pp. 465–468.

VITA

Michael Plexousakis was born in Heraklion, Greece, on October 12, 1965. He graduated from high school in 1983 and the fall of the same year he enrolled in the University of Crete at Heraklion, Greece, where he received a Bachelor of Science degree in Mathematics in 1987. During his undergraduate studies and a year after his graduation he worked as a research assistant in the Institute of Applied and Computational Mathematics at Heraklion, Greece. In August 1988 he was accepted as a graduate teaching associate in the Mathematics Department of the University of Tennessee, Knoxville. He received a Master of Science degree in June 1991 and in December 1996 he graduated with a Doctor of Philosophy degree with major in Mathematics.