

Li Tatsien

editor



Frontiers in
**Mathematical Analysis and
Numerical Methods**

In Memory of Jacques-Louis Lions

Frontiers in
**Mathematical Analysis and
Numerical Methods**

This page intentionally left blank

editor

Li Tatsien

Fudan University, China



Frontiers in
**Mathematical Analysis and
Numerical Methods**

In Memory of Jacques-Louis Lions

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: Suite 202, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

**FRONTIERS IN MATHEMATICAL ANALYSIS AND NUMERICAL METHODS
In Memory of Jacques-Louis Lions**

Copyright © 2004 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

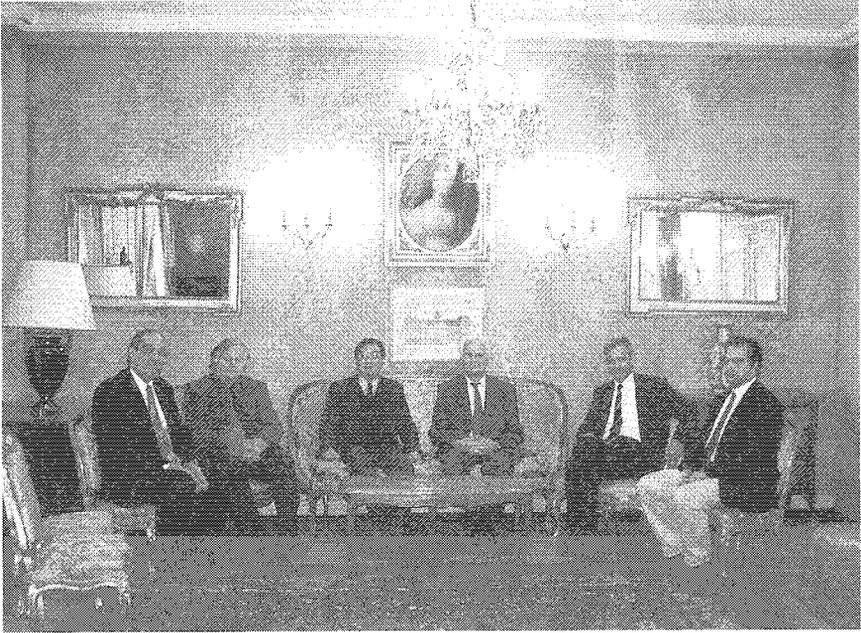
For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-238-941-5



Photo
COLLÈGE DE FRANCE
(J.-P. MARTIN)

Mention obligatoire



Taken at the Institut de France in Paris on May 29, 2000.

From left to right: A. Damlamian, R. Sénéor, Li Tatsien, J.-L. Lions, P. G. Ciarlet, F. Dubois.

PREFACE

About three years has passed away since the famous mathematician, a member of the Académie des Sciences de Paris, an oversea Member of the Chinese Academy of Sciences, Professor J.-L. Lions' death. As an outstanding mathematician, he made notable pioneer contributions to many fields of applied mathematics and enjoyed a worldwide high prestige and reputation. He held a series of high-level posts such as President of the IMU (International Mathematical Union), President of the Académie des Sciences de Paris. He played an important role and had a considerable influence in the international community of mathematics. The Journal *Chinese Annals of Mathematics* published a special issue (Vol. 23, Ser.B, No.2, 2002) dedicated to the memory of Professor J.-L. Lions, carrying papers contributed by his friends, students and colleagues. Afterwards the next several issues of the journal also published some papers of its kind. These papers involve many branches of mathematics and are of high quality and worth preserving. Here we collect them into a volume and offer it to the readers. Meanwhile a short essay of mine in memory of Professor J.-L. Lions is added as a supplement to the article of Professor P. G. Ciarlet, which is a brief account of Professor J.-L. Lions' life and achievements.

I would like to express my thanks to the faculty of the Editorial Office of *Chinese Annals of Mathematics*, especially to Professor Cai Zhijie and Professor Xue Mi, for their assistance in editing this book.

February 2004

LI Tatsien (LI Daqian)
Department of Mathematics
Fudan University
Shanghai 200433, China
dqli@fudan.edu.cn

This page intentionally left blank

CONTENTS

Preface	vii
Jacques-Louis LIONS, 1928–2001 <i>P. G. CIARLET</i>	1
Always Remembered <i>LI Tatsien (LI Daqian)</i>	19
Stable and Unstable Ideal Plane Flows <i>C. BARDOS, Y. GUO and W. STRAUSS</i>	25
Regularity Theory for Systems of Partial Differential Equations with Neumann Boundary Conditions <i>A. BENSOUSSAN and J. FREHSE</i>	47
Sensitivity of Darcy’s Law to Discontinuities <i>C. BERNARDI and O. PIRONNEAU</i>	69
Reiterated Homogenization of Degenerate Nonlinear Elliptic Equations <i>J. BYSTRÖM, J. ENGSTRÖM and P. WALL</i>	83
On the Connection in Finsler Space <i>S. S. CHERN</i>	97
Operator-Splitting Methods for the Simulation of Bingham Visco-Plastic Flow <i>E. J. DEAN and R. GLOWINSKI</i>	105

On the Uniqueness of the Weak Solutions of a Quasilinear Hyperbolic System with a Singular Source Term <i>J. P. DIAS and M. FIGUEIRA</i>	131
On the Classification of Initial Data for Nonlinear Wave Equations <i>GU Chaohao</i>	143
Regularity Results for Linear Elliptic Problems Related to the Primitive Equations <i>HU Changbing, R. TEMAM and M. ZIANE</i>	149
Local Exact Boundary Controllability for a Class of Quasilinear Hyperbolic Systems <i>LI Tatsien and RAO Bopeng</i>	171
On Nonlinear Differential Galois Theory <i>B. MALGRANGE</i>	185
Iterative Algorithms for Data Assimilation Problems <i>G. MARCHUK and V. SHUTYAEV</i>	197
Quadrilateral Mesh <i>MING Pingbing and SHI Zhongci</i>	207
On the Hyperbolic Obstacle Problem of First Order <i>J. F. RODRIGUES</i>	233
Instability of Traveling Waves of the Kuramoto-Sivashinsky Equation <i>W. STRAUSS and WANG Guanxiang</i>	253
Complexity of Large Time Behaviour of Evolution Equations with Bounded Data <i>J. L. VÁZQUEZ and E. ZUAZUA</i>	267

JACQUES-LOUIS LIONS, 1928–2001*

P. G. CIARLET

*Analyse Numérique, Université Pierre et Marie Curie, 4 Place
Jussieu, 75230 Paris, France. E-mail: pgc@ann.jussieu.fr*

Jacques-Louis Lions was born in the heart of Provence in the charming city of Grasse, much renowned for its perfume industry and its historical center.

In spite of his young age, he had the courage and determination to join the French Résistance at the end of 1943, as a soldier in the FFI (French Forces of the Interior). There he met Andrée, his wife and life-long companion.

Their son Pierre-Louis, who was born in 1956, would also be distinguished by mathematical talent. This gift earned him the highest mathematical distinction, the Fields Medal, awarded to him during the 1994 International Congress of Mathematicians in Zurich. His parents had the great joy of being present for this unique occasion.

At the early age of nineteen, Jacques-Louis Lions passed the entrance exam to the highly coveted Ecole Normale Supérieure de la rue d'Ulm. There he met Bernard Malgrange among others and at the end of their studies, they both decided to opt for a university carrier in mathematics (a rather uncommon choice at a time when most "Normaliens" would rather teach the famed "classes de Mathématiques Spéciales" in the lycées). They were then awarded a grant by CNRS (National Center for Scientific Research) to prepare their doctoral dissertations. And so they both went to

*This article is a translation adapted by the author from an article that appeared in French in the October 2001 issue of "MATAPLI", published here with the kind permission of its Editor-in-Chief, Brigitte LUCQUIN.

Nancy to work under the guidance of a prestigious thesis advisor, Laurent Schwartz, who had just received the Fields Medal for his theory of distributions in 1950.

After defending his thesis in 1954, Jacques-Louis Lions began his career “en province”, which is to say outside of the Paris region as was then customary, at the University of Nancy, where he held a professorship from 1954 to 1962.

Far from keeping him fully occupied, his remarkable mathematical achievements during that period left him enough time to envision the immense opportunities offered by Scientific Computing, which was then coming of age, with the manifold industrial applications that would henceforth become amenable. This constant quest for applications, which was to guide him all his life and to become one of the most exceptional aspects of his career, materialized in 1958 when he became scientific consultant for the SEMA (Society for Economics and Applied Mathematics), a society headed by Robert Lattès, who had entered the Ecole Normale Supérieure one year after him.

While this kind of inclination is common nowadays, it required lots of courage to follow it at the time. The applications of mathematics were then far from arousing the enthusiasm that they now generate!

After Nancy, Jacques-Louis Lions was named professor at the University of Paris, where he very quickly created a weekly seminar on “Numerical Analysis”, a discipline that was practically unheard of in France at that time. This seminar first met in the basement of the Institut Henri Poincaré, then in a dusty room of the Institut Blaise Pascal, which was situated rue du Maroc in the North of Paris.

When the University of Paris broke into thirteen distinct universities, he chose the sixth one, which was to be later named Université Pierre et Marie Curie. Two of his major initiatives there were to found the Laboratoire d’Analyse Numérique (after thirty years on the Jussieu campus, this department has been located rue du Chevaleret, near Place d’Italie, since 1999) and to create a DEA (Diplôme d’Etudes Approfondies, a set of advanced courses that a doctoral student has to pass before beginning a dissertation) specialized in Numerical Analysis. This DEA, from which a considerable number of applied mathematicians now holding positions in

universities, at CNRS, or in industry graduated over the years, was always highly regarded. As a tribute to his memory and as an expression of gratitude, the Laboratoire d'Analyse Numérique, which is currently headed by Yvon Maday, has just been renamed Laboratoire Jacques-Louis Lions.

In 1973, at the early age of forty-five, Jacques-Louis Lions had the highly unusual honor of being simultaneously named professor at the celebrated Collège de France and elected to the French Academy of Sciences. At the Collège de France, he held the Chair entitled "Mathematical Analysis of Systems and of their Control" for twenty-five years. His series of lectures, which in the tradition of the Collège had to be renewed each year, were always followed by vast audiences, attracting not only his own students but also students of his students!

The "Seminar of Applied Mathematics" that he organized there until 1998, first with Jean Leray, then with Haïm Brezis for many years, soon became an "institution within an institution". Indeed, countless applied mathematicians, either from Paris and its vicinity or French and foreign colleagues happening to be in Paris, gathered each Friday afternoon to hear prestigious lecturers, such as Stuart Antman, John Ball, Felix Browder, Ciprian Foias, Gu Chao-hao, Li Ta-tsien, Klaus Kirchgässner, Peter Lax, Andrew Majda, Louis Nirenberg, Olga Oleinik, Sergei Sobolev, Tosio Kato, Mark Vishik, and many others.

From 1966 to 1986, Jacques-Louis Lions was also part-time professor at the Ecole Polytechnique, where he created a course in numerical analysis from scratch that soon became a legend! Following the rule at the Ecole Polytechnique, he also wrote lecture notes, the contents of which were revolutionary for the time, at least in France. Indeed these notes constituted a kind of encyclopedia where, with his natural gift for teaching, Jacques-Louis Lions described and analyzed practically all that was then known about the numerical analysis of partial differential equations. Introductions to numerical optimization and numerical linear algebra were also presented in two separate chapters written by his first two doctoral students, Jean C ea and Pierre-Arnaud Raviart. A mystery remains about the first versions of these lecture notes. They were affectionately referred to as "the Diplodocus", even though no one including their author ever seemed to understand the reason for this!

But all these essentially academic activities, that would normally occupy all of one's time, did not take up all of his. Far from it!

From 1980 to 1984, he was also President of INRIA (National Institute for Research in Computer Science and Automatics), normally a full-time position! His leadership was of profound and lasting influence at INRIA: During his first weeks at the head of this institute, he used his incredible talents as an organizer to rejuvenate the organization and objectives, in particular by introducing the notion of project, gathering a clearly identified team around a well-defined objective on a specific theme.

During his four-year term as president, he strongly advocated the creation of start-up companies by researchers from the institute and he initialized its decentralization through the creation of similar institutes at Sophia-Antipolis and Rennes. Due to his personal prestige, the teams he was able to gather, and the numerous first-class international conferences that he organized there, he greatly contributed to the fame of INRIA.

From 1984 to 1992, he held another high-level, and also normally full-time, official position as President of CNES (National Center for Space Studies), where he continued and developed the action of his predecessor Hubert Curien (current President of the Academy of Sciences), who had just been named Minister of Research and Technology. There he used not only his eminent intellectual capacities but also his talents for intelligent persuasion to convince the French authorities that the orientations he advocated were well-founded. In this way, he played a major role in the conception of the French-American "Topex-Poseidon" space program for oceanography. Topex-Poseidon is also the name of the satellite that made it possible to at last understand "El Niño", a major event in climatology. His influence was likewise a decisive factor in the French-Russian negotiations that ultimately allowed Jean-Loup Chrétien and Michel Tognini to participate in manned space missions.

His presence at the Monday afternoon seances at the French Academy of Sciences remained rare for many years. But eventually, he gave new life to this noble "Compagnie" (as it is traditionally known among its members) when he became President in January 1997 for the customary two years. Immediately after Lions took office, President Jacques Chirac gave him the mission of supervizing the drafting of a document concerning the state of

the art worldwide in each of the following areas: "Access to knowledge for all and electronic processing of information, knowledge of our planet and ways of life, understanding life systems and improving health-care for all". He immediately began to work on this ambitious undertaking by creating and heading a "Committee 2000", whose task was to analyze the three questions and make proposals, under his ongoing close supervision. Remarkably, in spite of the scope of this project, he was able to meet the 2000 deadline he had set himself.

Indeed, he personally handed President Jacques Chirac the requested document during a ceremony held at the Elysée Palace on 25 January, 2000. He even succeeded in having all the Members and Corresponding Members of the Academy invited for the occasion, a first indeed!

However, Jacques-Louis Lions's actions during his presidency were not limited to the Committee 2000. His efforts were decisive in promoting the need for a profound reform in the status of the Academy. The principles of this reform have now been accepted. He also played a major role in the creation of an Academy of Technology, always desired but never achieved before. This academy was eventually created on 12 December, 2000.

As exemplified by his presidencies at INRIA and CNES, Lions was an exceptionally successful promoter of ever closer ties between academic research, too often seen as disconnected from the real world, and the more pragmatic industrial research. In this spirit, he headed scientific committees in major public utility companies, such as Météorologie Nationale, Gaz de France, France Telecom, or Electricité de France and he held high level advisory positions in major companies, such as Pechiney, Dassault Aviation, or Elf.

Jacques-Louis Lions's influence extended far beyond frontiers. Since the beginning of his career, he was an indefatigable traveler who, in addition to traditional venues in Europe or the Americas, very quickly added less canonical ones to his list of destinations. For instance, as early as 1957 he set out for a three-month visit to the Tata Institute of Fundamental Research in Bombay, at the time a genuinely adventurous trip! He enjoyed the splendor of the ancient Taj Mahal Hotel and the hospitality of Kollagunta Gopalaiyer Ramanathan, with whom he contributed to the creation of an applied branch of the Tata Institute on the campus of the Indian Institute

of Sciences in Bangalore twenty years later.

In 1966 he began a long series of visits in the former Soviet Union. Often invited by the USSR Academy of Sciences or by the Novosibirsk Institute for Computation, he initiated manifold scientific exchanges with eminent soviet mathematicians such as Guri Marchuk, Olga Oleinik, Lev Semenovitch Pontryagin, Ilia Vekua, Mark Visik, or Nicolay Nicolayevich Yanenko. One of his merits during this period, and not the least, was to contribute greatly to the dissemination of Soviet research in applied mathematics among Westerners.

A trip that left him with a lasting impression was the journey that he undertook in 1975 to Beijing, where he was received with great ceremony. He was in particular impressed there by the mathematical talents of Feng Kang, who had just independently rediscovered the finite element method. This was the first of a long series of trips to China, which later included three visits to Fudan University in Shanghai. In return he also invited the famous differential geometer Su Bu-chin to visit Paris in 1982 for signing an agreement between Fudan University, the INRIA, and the Ecole Polytechnique.

However, his international ventures were not limited to traditional scientific exchanges, as his talents as a lecturer, thesis adviser, and organizer produced many disciples throughout the countries he visited.

As early as the 1960s for instance, he was the adviser of Antonio Valle, the first in a long series of students from Spain and Portugal, who in turn set up numerical analysis departments at the Universities of Malaga, Sevilla, Santiago de Compostela, Lisboa, or at the Universidad Complutense de Madrid, modeled after the one he had created in Paris. In the same vein, he was in 1997 the main speaker in a European video-conference on Mathematics and the Environment, organized in Madrid by Jesus Ildefonso Diaz. He also chaired the Committee awarding the ten "Prizes for Young Mathematicians" during the Third European Congress of Mathematics held in Barcelona in 2000.

For many years, he chaired as well the Scientific Committee of the Istituto di Analisi Numerica del CNR of the University of Pavia, headed for several decades by Enrico Magenes, then by Franco Brezzi.

Together with Paul Germain, he represented France at the 1975 meet-

ing on “Functional Analysis and Mechanics” of the IUTAM (International Union for Theoretical and Applied Mechanics) held in Luminy, where the other representatives were Klaus Kirchgässner for West Germany, Sir James Lighthill for the United Kingdom, and William Prager for the United States.

His intelligent proselytizing was not limited to Europe, however. In China for instance, he was one of the driving forces behind the creation in 1997 of the LIAMA (French-Chinese Laboratory of Computer Science, Automatics, and Applied Mathematics), an offspring of INRIA and the Chinese Academy of Sciences, housed since then in Beijing by the Institute of Automatics of the academy. He likewise played a major role in the creation in 1998 of the ISFMA (Chinese-French Institute of Applied Mathematics), splendidly housed by the Department of Mathematics of Fudan University in Shanghai, thanks to the tireless efforts of his Director Li Ta-tsien. As President of the Scientific Committee of this institute, he attended the opening ceremony that marked its creation. He was Honorary Editor of “Chinese Annals of Mathematics”. He also established and kept close scientific ties with colleagues in Hong Kong, notably with Roderick Wong, Dean at City University of Hong Kong.

Jacques-Louis Lions was also President of the IMU (International Mathematical Union) from 1991 to 1995. During a meeting of this organization in Rio de Janeiro on 6 May, 1992, he proposed that the year 2000 be baptized “World Mathematical Year”. This proposal, which was later supported by UNESCO, turned out to be a genuine success story that significantly contributed to the improvement of the image of mathematics in the general public and helped encourage mathematical research in developing countries.

While he served as “Past-President” of the IMU from 1995 to 1999, the decision was taken that Beijing would host the twenty-third International Congress of Mathematicians in 2002, the second in Asia after that of Kyoto in 1990.

He likewise was a constant supporter of the initiatives of the Third World Academy of Sciences (TWAS), either directly or through colleagues from his group. Particularly noteworthy in this respect were his undertakings for the progress of mathematical research in Africa.

The mathematical works of Jacques-Louis Lions are immense. Alone or in collaboration, he wrote more than twenty books, most of which have

become classics often translated into several foreign languages, as well as more than five hundred papers. The different themes of his work are briefly described below, in an order that approximately follows their chronology.

He had, and will continue to have for a long time, a considerable influence on mathematics and their applications, not only through his own work, but also through that of the School he created and constantly kept in touch with. This School, which numbers some fifty initial students and scores of students of students, etc., has acquired a widespread fame over the years not only in university circles, but also in industry, an accurate indication that the directions of research he envisioned and promoted were highly relevant.

If a single title were to be attached to Jacques-Louis Lions's mathematical works, it might be with a fair degree of accuracy that of "Partial differential equations in all their aspects: Existence, uniqueness, regularity, control, homogenization, numerical analysis, etc., and the applications they model, such as fluid and solid mechanics, oceanography, climatology, etc.".

Jacques-Louis Lions produced his first mathematical works in 1951. At the same time two major books were published, one by Laurent Schwartz on the theory of distributions and one by Sergei Sobolev on their applications to mathematical physics, as well as a founding paper by John von Neumann and Robert Richtmyer on the numerical approximation of nonlinear hyperbolic problems arising in hydrodynamics.

Inspired by these works, Lions's first objectives were to undertake a systematic study of linear and nonlinear boundary value problems, notably by constantly using the theory of distributions, and then to find ways to numerically approximate their solutions.

He began in 1954 a series of collaborations and lasting friendships with eminent Italian mathematicians, such as Enrico Magenes, Guido Stampacchia, Ennio de Giorgi, or Giovanni Prodi (brother of the current President of the European Union). One such collaboration resulted in an exhaustive analysis of boundary value problems posed in fractional Sobolev spaces, thanks in particular to the theory of interpolation between Banach spaces that he initiated with Jack Peetre in 1961. This analysis is the object of the celebrated three-volume treatise "Non-Homogeneous Boundary Value Problems and Applications" (1968-1970) that he wrote with Enrico Ma-

genes. From 1965 to 1967, he also developed with Guido Stampacchia the foundations of the theory of variational inequalities, as they appear for instance in unilateral problems in elasticity.

His inclination for applications led him to propose a particularly elegant proof of Korn's inequality, based notably on a fundamental result in distribution theory known as "Lions's lemma" (although several other results of his bear the same name!). He further developed applications of the theory of variational equations or inequalities to mechanics by mathematically analyzing Bingham fluids, friction, viscoelasticity, or plasticity models. These applications constitute the substance of another well-known book, "Inequalities in Mechanics and Physics" (1972) that he wrote with Georges Duvaut.

He was equally interested in the numerical simulation of these problems, at a time when it was realized that the applicability of finite difference methods had reached its limits. For instance, these methods do not perform well when the problems to be approximated have rapidly varying coefficients or are posed on domains with complicated geometries. On the other hand, the finite element method, already familiar to engineers in handling these types of difficulties, remained essentially unknown to mathematicians.

With a remarkable intuition, Jacques-Louis Lions immediately foresaw that it was preferable to discretize the variational, or weak, formulations of partial differential equation problems rather than the partial differential equations themselves. Accordingly, he quickly pointed out to his group of colleagues and students the interest of studying and analyzing Galerkin methods in general and finite element methods in particular. A quite productive period ensued, to which he himself contributed with another classic, "Numerical Analysis of Variational Inequalities" (1976), co-authored with Roland Glowinski and Raymond Trémoières.

His book "Some Methods for Solving Nonlinear Boundary Value Problems" (1969) is a major contribution to the theory of nonlinear partial differential equations, which remains even today a substantial source of inspiration (it is unfortunate that this book was never translated into English). In this work, Jacques-Louis Lions introduced and systematically analyzed the so-called compactness methods, which play a key role in the existence theory for the Navier-Stokes and von Kármán equations, the monotony

methods he had developed with Jean Leray, and the regularization and penalty methods, which can for instance be applied to the Schrödinger or Korteweg-de Vries equations. For the most part, the results found in this book are either due to himself or to his students, in particular Haïm Brezis and Luc Tartar.

Most of the works mentioned so far, together with many generalizations they led to, were to be assembled in the monumental treatise "Mathematical Analysis and Numerical Methods for Sciences and Technology" (1984-1985), conceived and edited by Jacques-Louis Lions and Robert Dautray. This work, which comprises almost four thousand pages, is justly regarded as the modern counterpart of the celebrated Courant-Hilbert.

His ongoing interest for problems with small parameters led him to write "Singular Perturbations in Boundary Value Problems and in Optimal Control" (1973), a book where he laid down the foundations of the asymptotic analysis of such problems. The methods and notions that he then introduced and analyzed, such as a priori estimates, stiff problems, boundary layers, multiple scales, and so on, were subsequently recognized as fundamental for many applications. For instance, they later played a major role in the mathematical modeling of elastic structures or "multi-structures", made of plates, rods, or shells.

Another field where small parameters naturally arise is the modeling of composite materials, of constant use in the aerospace industry for instance. Their asymptotic analysis, which is a special case of what became known as homogenization theory, was abundantly developed and illustrated by applications in another seminal work, "Asymptotic Analysis for Periodic Structures" (1978), which he wrote with Alain Bensoussan and George Papanicolaou. In this book, a substantial number of essentially empirical formulas used in the modeling of periodic structures were rigorously justified for the first time, thanks notably to the compensated compactness method due to his students Francois Murat and Luc Tartar and to the oscillating test-functions method of Luc Tartar.

A fundamental work by Lev Semenovitch Pontryagin about the optimal control of systems governed by ordinary differential equations (the objective was to control the trajectories of artificial satellites) immediately attracted his attention in 1958. Through the contacts he already had at that time

with the engineering community, he became quickly convinced that the next step was to extend optimal control to distributed systems, i.e., systems whose state is governed by partial differential equations. The inclination that he then developed for the optimal control of such systems was to always remain at the center of his mathematical interests.

A pionner as always, he began by laying down the foundations of a general theory in yet another celebrated book, "Optimal Control of Systems Governed by Partial Differential Equations" (1968), where he notably introduced an infinite-dimensional version of the Riccati equation.

In two books co-authored with Alain Bensoussan, "Applications of Variational Inequalities to Stochastic Control" (1978) and "Impulse Control and Variational Inequalities" (1983), he continued his investigations by considering in particular the optimal control of systems that are not necessarily well-posed or that have multiple states.

After having so thoroughly analyzed the main aspects of optimal control theory, Jacques-Louis Lions shifted his interests to the study of "controllability", a discipline that basically seeks to answer the following type of question: Given a system in an arbitrary initial state, the question is to devise a way of acting on it in such a fashion that its solution reaches a given final state in a finite time, for instance by imposing adequate boundary conditions.

During the prestigious "John von Neumann Lecture" that he gave at the SIAM Congress in Boston in 1986, he presented for the first time his now famous "HUM" (Hilbert Uniqueness Method) for the exact controllability of linear time-dependent equations. He chose this particular terminology to emphasize the fact that the feasibility of such controllability is related in an essential way to the uniqueness of the solution to the adjoint problem, typically obtained by the Holmgren or the Carleman theorem.

This lecture was the starting point of numerous works by himself or his School. In particular, he began by publishing no less than three books on the subject in the same year, "Exact Controllability, Perturbations and Stabilization of Distributed Systems" in two volumes (1988) and, with John Lagnese, "Modeling, Analysis, and Control of Thin Plates" (1988), which contains an abundance of applications to the theory of elastic plates. In 1995, he also established with Enrique Zuazua the generic character of

the controllability of the three-dimensional Stokes equations: If there is no approximate controllability for a given open set, it is always possible to find another arbitrarily close open set for which this type of controllability holds.

These works constituted yet another mark of his constant interest for real-life applications. He also had the concern of proposing numerically feasible approximation methods. These were the theme of a long article co-authored with Roland Glowinski, an article so long that its nearly three hundred pages took up two consecutive issues of "Acta Numerica" (1994–1995).

Even though the last works of Jacques-Louis Lions were in different areas, they continued to be partly influenced by the methodology he had developed for questions of controllability.

In 1990, he started to express his great interest in climatology in "El Planeta Tierra". In this book, which was directly published in Spanish, he described in a masterly fashion and in a remarkably accessible style the most important problems originating in this science, such as modeling, numerical simulation, sensitivity to initial conditions, etc. From 1994 to 1998, his last courses at the Collège de France were all about these subjects.

The models found in climatology include complex systems of partial differential equations such as those of Navier-Stokes or of thermodynamics. But these systems had never been seriously analyzed from the mathematical viewpoint, although they had been blithely used in a massive way since the 1960s for numerical simulation in weather forecasting.

In spite of the "truly diabolic complexity" (as he was fond of saying) of the combination of partial differential equations, boundary conditions, transmission conditions, nonlinearities, physical assumptions, etc., that enter these models, Jacques-Louis Lions, together with Roger Temam and Shouhong Wang, was able to study questions of existence and uniqueness of solutions, to establish the existence of attractors, and to propose numerical methods. He even succeeded in teaching these works on a blackboard, a pedagogical tour de force!

In 1995, he began with Evariste Sanchez-Palencia another series of works where they developed the theory of sensitive problems, exemplified by the boundary value problems that appear in the theory of linearly elastic membrane shells. In such problems, which in a sense constitute the antithesis

of well-posed problems, arbitrarily small, yet arbitrarily smooth changes in the right-hand sides of the equations may induce “sudden” changes on the properties of their solutions. It is perhaps no coincidence that the analysis of such problems relies in particular on uniqueness theorems that bear resemblance with those needed in the HUM.

In his last works, Jacques-Louis Lions returned to the numerical analysis of parallel algorithms and domain decomposition methods, in a long series of Notes aux Comptes Rendus de l'Académie des Sciences most often co-authored with Olivier Pironneau and published from 1997 to 2001. In fact, these kinds of topics had been on his mind for a long time. As early as the 1980s, he had already been an ardent advocate of installing a parallel computer at INRIA. The main idea in these Notes is to introduce parallelism in the continuous problem rather than in the discrete one, an approach that is in fact quite general, since it applies equally well to any problem that is approximated by an iterative method, such as a fractional step method, a decomposition method into sub-problems in optimization theory, a domain decomposition method, and so forth.

One can only be impressed by these immense works, whether by the quality, diversity, and novelty of the mathematics used or by the permanent quest for new applications that were previously believed to be inaccessible.

Like John von Neumann, whom he deeply admired, Jacques-Louis Lions was a visionary, who very quickly understood that the availability of ever-increasing computational power could revolutionize the modeling of phenomena and thereby improve our knowledge and mastery of the physical world, provided however that the required mathematics were simultaneously created and developed. He admirably contributed to this latter task.

Jacques-Louis Lions justly received numerous honors. Although he always remained modest about them, their list is truly astonishing: He was Commandor of the French Legion of Honor and Great Officer of the French National Order of Merit, a distinction he received at the hands of President Chirac on 23 February, 1999. He was a member of twenty-two foreign academies and Honoris Causa Doctor of nineteen universities, he received

the most prestigious prizes, and he delivered the most coveted lectures.

He was in particular awarded three Prizes by the French Academy of Sciences, the John von Neumann Prize in 1986, the Harvey Prize from the Technion in 1991, and the Lagrange Prize at the ICIAM meeting in Edinburgh in 1999. Jacques-Louis Lions was proud to have had the rare honor of having shaken the hand of Emperor Akihito when he received the highly prestigious Japan Prize in 1991. It was the climax of a perfectly and meticulously organized week which had particularly impressed him!

In particular, he was an invited speaker three times at an International Congress of Mathematicians, in 1958, 1970, and 1974. He gave the John von Neumann Lecture at the SIAM meeting in Boston in 1986. He was plenary speaker at the ICIAM Congress in Hamburg and at the SIAM Congress in Philadelphia in 1995, and he held the Galileo Chair at the University of Pisa in 1996. He also had the extremely rare honor for a scientist, especially for a mathematician, of speaking before a parliament! More specifically, he delivered an address entitled "Will it ever be possible to describe, understand, and control the inanimate and animate world by means of the languages of mathematics and informatics?" in front of the Cortes who had specially gathered in Madrid on 21 January, 2000 for the occasion, as part of the celebrations of the "World Mathematical Year 2000".

He belonged to the most famous academies, such as the USSR Academy of Sciences and the American Academy of Arts and Sciences to which he was elected in 1982 and 1986. In 1996, he was simultaneously elected to the Royal Society of the United Kingdom, to the National Academy of Sciences of the USA, and to the Third World Academy of Sciences, just before being simultaneously elected in 1998 to the Chinese Academy of Sciences and to the Accademia Nazionale dei Lincei in Rome. The Pontifical Academy of Sciences, to which he was elected in 1990, seemed to be particularly dear to him. It may perhaps be not so well-known, but it is an unusually rare honor to become one of its members!

Who was the man behind all these endeavors? What I knew of him convinced me that both his human and professional qualities were truly exceptional.

Jacques-Louis Lions wrote abundantly, very rapidly, and with an astonishing ease not only mathematics but also countless letters, which consti-

tuted his favorite means of communication. He was a master with the fax, which he used with an amazing efficiency! For instance it was not uncommon for each one of his various collaborators at any given time to receive four or five faxes per week, sometimes of up to thirty pages if their contents were mathematical.

Even though he wrote abundantly, he confessed that he did not usually keep track of his manifold letters, perhaps because he could rely entirely on his memory, or perhaps because he preferred to spare himself a herculean archival task! Let us hope that his correspondents had the good idea to keep his letters, which could thus be later compiled.

By any measure, his abilities were astounding. For instance, he once told me that it only took him a few weeks to write the several hundred pages of the “Diplodocus” lecture notes mentioned earlier. Likewise his indifference to lack of sleep or to the most extreme jetlags and his freshness after lengthy flights were always a subject of astonishment among his travel companions.

As John Ball put it so well, “Jacques-Louis Lions was a man of considerable personal magnetism and charm, whose charisma, brilliance as a teacher, and accessibility attracted others to work with him”. And indeed it was obvious that Jacques-Louis Lions had an ample share of charisma, even if charisma is not easy to define in a rigorous manner! He was also incredibly open and displayed such amiable and simple manners that any one of his current students or collaborators felt they were at the center of his attention.

He was also very brave in the face of physical danger and suffering. Even when the pain became unbearable, he never complained, keeping on the contrary his compassion for others.

All those who met him will cherish the memory of his warm personality, the vision that he so well conveyed, and his profound intelligence.

Books by Jacques-Louis Lions

- [1] Lectures on Elliptic Partial Differential Equations, Tata Institute, Bombay, 1957.
- [2] Équations Différentielles Opérationnelles et Problèmes aux Limites, Springer-Verlag, Berlin, 1961.

- [3] Problèmes aux Limites dans les Équations aux Dérivées Partielles, Presses de l'Université de Montréal, Montréal, 1962.
- [4] Méthodes de Quasi-Réversibilité et Applications (with R. Lattès), Dunod, Paris, 1967;
English translation; Quasi-Reversibility Methods and Applications, American Elsevier, New York, 1969;
Russian translation: Mir, Moscow, 1970.
- [5] Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles, Dunod, Paris, 1968;
English translation: Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1970;
Russian translation: Mir, Moscow, 1972.
- [6] Problèmes aux Limites Non Homogènes et Applications (with E. Magenes), 3 volumes, Dunod, Paris, 1968–1970;
Russian translation: Mir, Moscow, 1971;
English translation: Non-Homogeneous Boundary Value Problems and Applications, 3 volumes, Springer-Verlag, Berlin, 1972–1973.
- [7] Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires, Dunod, Paris, 1969;
Russian translation: Mir, Moscow, 1972.
- [8] Les Inéquations en Mécanique et en Physique (with G. Duvaut), Dunod, Paris, 1972;
English translation: Inequalities in Mechanics and Physics, Springer-Verlag, Berlin, 1976;
Russian translation: Nauka, Moscow, 1980.
- [9] Some Aspects of the Optimal Control of Distributed Parameter Systems, CBMS Conference No. 6, SIAM, Philadelphia, 1972.
- [10] Sur le Contrôle Optimal des Systèmes Distribués, Monographie No. 20, Enseignement Mathématique, Genève, 1973;
Russian translation: UMN 28, 1973.
- [11] Perturbations Singulières dans les Problèmes aux Limites et en Contrôle Optimal, Lecture Notes in Mathematics, Vol. 323, Springer-Verlag, Berlin, 1973.
- [12] Analyse Numérique des Inéquations Variationnelles (with R. Glowinski and R. Trémollières), 2 volumes, Dunod, Paris, 1976;

- English translation: *Numerical Analysis of Variational Inequalities*, North Holland, Amsterdam, 1981.
- [13] *Sur quelques Questions d'Analyse, de Mécanique et de Contrôle Optimal*, Presses de l'Université de Montréal, Montréal, 1976.
- [14] *Applications des Inéquations Variationnelles au Contrôle Stochastique* (with A. Bensoussan), Dunod, Paris, 1978;
English translation: *Applications of Variational Inequalities to Stochastic Control*, North-Holland, 1982.
- [15] *Asymptotic Analysis for Periodic Structures* (with A. Bensoussan and G. Papanicolaou), North-Holland, Amsterdam, 1978.
- [16] *Some Methods in the Mathematical Analysis of Systems and their Control*, Science Press, Beijing and Gordon & Breach, New York, 1981.
- [17] *Contrôle Impulsionnel et Inéquations Variationnelles* (with A. Bensoussan), Gauthier-Villars, Paris, 1982;
English translation: *Impulse Control and Variational Inequalities*, Heyden & Son, Philadelphia, 1984.
- [18] *Contrôle des Systèmes Distribués Singuliers*, Gauthier-Villars, Paris, 1983; Russian translation: Nauka, Moscow, 1987.
- [19] *Analyse Mathématique et Calcul Numérique pour les Sciences et les Techniques* (with R. Dautray), 3 volumes, Masson, Paris, 1984–1988;
English translation: *Mathematical Analysis and Numerical Methods for Science and Technology*, 9 volumes, Springer-Verlag, Berlin, 1988–1993.
- [20] *Contrôlabilité Exacte, Perturbations et Stabilisation de Systèmes Distribués*, 2 volumes, Masson, Paris, 1988.
- [21] *Modelling, Analysis and Control of Thin Plates* (with J. Lagnese), Masson, Paris, 1988.
- [22] *El Planeta Tierra*, Instituto de España, Espasa-Calpe, Madrid, 1990.
- [23] *Sentinelles pour les Systèmes Distribués à Données Incomplètes*, Masson, Paris, 1992.
- [24] *Models for the Coupled Atmosphere and Ocean* (with R. Temam and S. Wang), *Computational Mechanics Advances*, Vol. 1, North-Holland, Amsterdam, 1993.

Series Edited by Jacques-Louis Lions

- [1] *Studies in Mathematics and its Applications* (with H. Fujita, H. B. Keller and G. Papanicolaou), 30 volumes, North-Holland, Amsterdam, 1976–2001.
- [2] *Nonlinear Partial Differential Equations and their Applications. Collège de France Seminar* (with H. Brezis), 12 volumes, Pitman-Longman, 1981–1994.
- [3] *Mathématiques Appliquées pour la Maîtrise* (with P.G. Ciarlet), 21 volumes, Masson, Paris, 1982–1997.
- [4] *Recherches en Mathématiques Appliquées* (with P.G. Ciarlet), 42 volumes, Masson, Paris, 1985–1997.
- [5] *Handbook of Numerical Analysis* (with P.G. Ciarlet), 7 volumes, North-Holland, Amsterdam, 1990–2001.
- [6] *Nonlinear Partial Differential Equations and their Applications. Collège de France Seminar* (with D. Cioranescu), 2 volumes, Pitman, 1998–2001.

Books Dedicated to Jacques-Louis Lions

- [1] *Analyse Mathématique et Applications. Contributions en l'honneur de Jacques-Louis Lions*, Gauthier-Villars, Montrouge, 1988.
- [2] *Frontiers in Pure and Applied Mathematics. A collection of papers dedicated to Jacques-Louis Lions on the occasion of his sixtieth birthday* (R. Dautray, Editor), North-Holland, Amsterdam, 1991.
- [3] *Équations aux Dérivées Partielles et Applications. Articles dédiés à Jacques-Louis Lions*, Gauthier-Villars, Paris, 1998.

ALWAYS REMEMBERED

LI TATSIEN (LI DAQIAN)

*Department of Mathematics, Fudan University,
Shanghai 200433, China. E-mail: dqli@fudan.edu.cn*

On the desk in my office at the Institut Sino-Français de Mathématiques Appliquées (ISFMA), there is a carefully framed photo, which was taken when I saw Professor J.-L. Lions for the last time. That was the afternoon of May 29, 2000, and I was attending the meeting of the (French) Scientific Committee of the ISFMA presided by Prof. J.-L. Lions at the Institut de France in Paris. The participants were in high spirits, and after the meeting a group photo was taken in a quaint hall. I was sitting next to him for that photo and could not imagine that this would be our last time together!

Prof. J.-L. Lions in the photo seems to be always watching me with his flashing eyes, urging me never to forget his concerns and instructions, urging me never to forget those unforgettable past matters which greatly influenced me, urging me to move forward and do my best, and urging me to complete his great unfinished work of developing applied mathematics.

In my memory, the figure of Prof. J.-L. Lions was always tall and strong, and his face was still glowing with health and radiating vigour. Perhaps due to the aesthetic disposition of French nation, he was unwilling to have his sick look seen by his relatives and friends. The evening of March 25, 2001, soon after I arrived in Paris from Shanghai, I was shocked to learn that Prof. J.-L. Lions was seriously ill and hospitalized. Next morning I hurried to Collège de France, where I confirmed the news and had more details. I was deeply concerned about his sickness. I left messages several times hoping that I would be allowed to visit him in the hospital. He had been a strong man and I hoped he could recover from his illness and regain his energy. Unfortunately, when I finished my visit to Strasbourg

and Clermont-Ferrand in May and returned to Paris to leave for China, I heard that he was more seriously ill. Soon after I returned to Shanghai the bad news of his death came to me through different channels. I didn't expect that he would pass away so soon!

Prof. J.-L. Lions, a master of applied mathematics who enjoyed high prestige and reputation and acquired great achievements, and a respected and beloved teacher of mine, left us in such a hurry. At the moment of grief the fond memories of our time together freshened in my mind, stirring me for a long time.

Prof. J.-L. Lions was my advisor during my stay at Collège de France as a visiting scholar from January, 1979 to April, 1981. The first time we met, he gave me a copy of his monograph *Quelques méthodes de résolution des problèmes aux limites non linéaires*. I spent over half one year in studying this book deliberately gained an overall understanding to the method of solving nonlinear partial differential equations by using the theory of Sobolev space. It was a rigorous training in the basic skill of Lions School. During that period Prof. J.-L. Lions proposed three problems to me. I chose the most difficult one and wrote a paper on it. He was pleased by my paper and encouraged me to publish it. But I realized that at that time he just tested my ability, so I did not submit the paper to any journal. In the second half of 1979, he was lecturing at Collège de France every Friday morning. In one class he introduced a new kind of function space for solving pointwise control problems and gave a conjecture that the function space should be independent of the concrete forms of the elliptic operators. Due to my hard study during the first half of year, I proved the conjecture next day and immediately sent it to him. I was surprised that he addressed my proof in details as a theorem in the next class. It inspired me very much and also taught me the way he exchanged thoughts and communicated with his audience through his lectures. With his encouragement and advice, I successfully published several papers on pointwise control problems and related limit behaviors. These results were added into his monograph *Some methods in the mathematical analysis of systems and their control* published in 1981 by Scientific Press, Beijing, and he also asked me to write two appendices for this book.

During the period of my stay at Collège de France over two years, I spent

most of my time and energy in the study of quasilinear hyperbolic systems. My work related to his research fields might be only what was mentioned above. It is not significant and I wish I have done more. However, living and working near Prof. J.-L. Lions, I was imperceptibly influenced by his working style and thinking method, I was very impressed by many talented disciples of Prof. J.-L. Lions, and I could see the prosperity and strength of Lions School. It was really of benefit to me. Since then each time I visited France Prof. J.-L. Lions always welcomed me and gave me advice and we became closer and closer. As a result, my understanding of applied mathematics was greatly improved. I further realized his working style and thinking method. I respect him not only for his academic achievement but also for his great personality. It will no doubt impact me forever.

As a master in applied mathematics, Prof. J.-L. Lions had kept close contact with some important organizations in practical field as their academic consultant for long time. He proposed many new mathematical concepts, methods and theories just from urgent needs of practical applications. They not only played important instructive roles in solving practical problems but also greatly enriched the contents of applied mathematics. His success shows that combining theory with practice is a broad way for the development of applied mathematics. His effective working mode is a valuable legacy to us.

I remember that as a professor of Collège de France, each year Prof. Lions always gave a new course which addressed his latest research. He also published a new monograph almost every year. Even in May, 2000, one year before his death, I followed his lecture at Université Pierre et Marie Curie on the control problems of Navier-Stokes equations. It was believable that what he bore in mind constantly even in his last moment probably was various mathematical problems which he never gave up. Such a talented and diligent mathematician who devoted his entire life to mathematics is respectable that we would learn from him and keep him in mind forever.

Prof. J.-L. Lions was a very busy person during his lifetime, holding several important posts at home and abroad concurrently; nevertheless, he managed his life elegantly. During the days when I was a visiting scholar at Collège de France, every Friday in the afternoon before and after the seminar he presided there were always so many persons waiting for him,

who wanted to discuss with him or consult him. Though my office was next to his, our talk often could not last over five minutes each time since he was so busy. But such a talk was always concise and effective. Each time I handed him my research work or problems, two or three days later I would receive a long letter from him, giving me his advice and instructive opinions. Having known him for more than twenty years, I was deeply impressed with his efficient working style and I urged myself to learn from him.

It is especially worth mentioning that Prof. J.-L. Lions was full of enthusiasm about the development of mathematics in developing countries, reflecting an honest mathematician's wide vision and broad mind. During the period that he took office of the Secretary-General of the International Mathematical Union (IMU), he made great efforts to regain seat of the People's Republic of China in IMU, resulting in a satisfactory settlement at the end. As the President and the past President of IMU, he again spared no effort to help China meet with success in applying for hosting the International Congress of Mathematicians (ICM) at Beijing in 2002. When he visited Shanghai in 1981, he was awarded the title of Honorary Professor of Fudan University. Since then, he was always concerned about the establishment of the mathematical disciplines in Fudan University. He invited Professor Su Buchin, the President of Fudan University, to visit Paris in 1982 and facilitated a cooperative agreement among Fudan University, Ecole Polytechnique and Institut National de Recherche en Informatique et en Automatique (INRIA). As a result, the academic exchange and cooperation of Fudan University with France in mathematical fields became more frequently. With his effort, supported by the Presidents of China and France, the Institut Sino-Français de Mathématiques Appliquées (ISFMA) was established in Shanghai in February, 1998. He was the French Chairman of the Scientific Committee of the Institute. He came to Shanghai to attend the opening ceremony with a group of French mathematicians, and gave a scientific lecture there. The journal *Chinese Annals of Mathematics*, whose past Editor-in-Chief was Prof. Su Buchin, reorganized its Editorial Committee in 1999. Prof. J.-L. Lions kindly accepted the invitation to be an Honorary Editor of the journal. He was also the Honorary Editor-in-Chief of the *Series in Contemporary Applied Mathematics* published by the Higher

Education Press, Beijing, from 2000 onwards. When Macao returned to China, Prof. J.-L. Lions helped to organize the conference “Mathematics and its Role in Civilization”, which added a piece of bright landscape for new-returned Macao. After he took office of the President of the Académie des Sciences de Paris, his first outside visit was to the Academia Sinica in Beijing, and he was elected as an oversea member of the Chinese Academy of Sciences in 1998, unique member from France at that time. We could say without any exaggeration that Prof. J.-L. Lions was an truly devoted friend of Chinese mathematicians and he won popular love and respect in Chinese mathematical community.

Today, due to the efforts and encouragement of Prof. J.-L. Lions in his lifetime:

— The International Congress of Mathematicians (ICM2002) was successfully held in Beijing in August 2002, and the General Assembly of the IMU was also smoothly held in Shanghai before ICM2002.

— The ISFMA operates smoothly, and is playing a more important role in promoting the exchanges between Chinese and French applied mathematicians. Efforts have been and will be made to realize his ideas in developing applied mathematics.

— The journal *Chinese Annals of Mathematics* has become more influential in recent years. Hopefully the journal would become an international mathematical journal of great influence in near future. The second issue of *Chinese Annals of Mathematics Ser. B* (2002) as a special issue and the next several issues published articles contributed by his friends, students and colleagues, dedicated to the memory of him.

— The theory of exact boundary controllability, which he initiated and made important contributions to, has got rather important progress in the case of general quasilinear hyperbolic systems of first order.

.....

When Prof. J.-L. Lions came to Shanghai to attend the opening ceremony of the ISFMA in 1998, he was deeply moved by Shanghai’s great changes. At the reception hosted by Shanghai’s Vice-Mayor, he said: “This is my third visit to Shanghai, and I have seen three different cities in three visits !” After then I have invited him to visit Shanghai again, in particular, I hoped he could come to Shanghai during the period of the General

Assembly of the IMU in 2002. I would, as host, show him the new changes in Shanghai in recent years. Regretfully, it has been now impossible.

STABLE AND UNSTABLE IDEAL PLANE FLOWS

C. BARDOS

Université Denis Diderot and LAN Université Pierre et Marie Curie, Paris, France. E-mail: bardos@math.jussieu.fr

Y. GUO

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912, USA. E-mail: guoy@cfm.brown.edu

W. STRAUSS

Department of Mathematics, Brown University, Providence, Rhode Island 02912, USA. E-mail: wstrauss@math.brown.edu

Abstract

The authors investigate the stability of a steady ideal plane flow in an arbitrary domain in terms of the L^2 norm of the vorticity. Linear stability implies nonlinear instability provided the growth rate of the linearized system exceeds the Liapunov exponent of the flow. In contrast, a maximizer of the entropy subject to constant energy and mass is stable. This implies the stability of certain solutions of the mean field equation.

Keywords Stable ideal plane flows, Unstable ideal plane flows, Liapunov exponent

2000 MR Subject Classification 35L50, 35B35

§1. Introduction

We consider solutions $u = (u_1(x_1, x_2, t), u_2(x_1, x_2, t))$ of the incompressible two-dimensional Euler equation in a bounded domain $\Omega \subset \mathbb{R}^2$ with a smooth impermeable boundary $\partial\Omega$:

$$\partial_t u + u \cdot \nabla u = -\nabla p \text{ in } \mathbb{R}_t \times \Omega, \quad u \cdot \vec{n} = 0 \text{ on } \mathbb{R}_t \times \partial\Omega. \quad (1.1)$$

\vec{n} denotes the outward normal to the boundary. The vorticity

$$\omega = \nabla \wedge u \equiv \partial_{x_1} u_2 - \partial_{x_2} u_1$$

is then transported by the flow according to the equation

$$\partial_t \omega + u \cdot \nabla \omega = 0 \text{ in } \mathbb{R}_t \times \Omega. \quad (1.2)$$

We define the operator curl^{-1} by the formula

$$\text{curl}^{-1} \omega = \nabla \wedge \Psi, \text{ with } -\Delta \Psi = \omega \text{ in } \Omega, \Psi = 0 \text{ on } \partial\Omega. \quad (1.3)$$

If the domain Ω is simply connected, then $u = \text{curl}^{-1} \omega$. If it is not simply connected, but $u = \text{curl}^{-1} \omega$ at time $t = 0$, then $u = \text{curl}^{-1} \omega$ at all times. This is the only case that we consider and therefore the Euler equation is equivalent to the equation

$$\partial_t \omega + \text{curl}^{-1}(\omega) \cdot \nabla \omega = 0. \quad (1.4)$$

If $u_0(x)$ is a stationary solution of the Euler equation (1.4), then the vorticity $\omega_0(x)$ satisfies the equation

$$0 = u_0 \cdot \nabla \omega_0 = \nabla \Psi_0 \wedge \nabla \omega_0, \quad (1.5)$$

which says that the level lines of Ψ_0 and Ω_0 coincide. This condition is satisfied in particular for any solution of the nonlinear elliptic equation

$$-\Delta \Psi = f(\Psi) \text{ in } \Omega, \quad \Psi = 0 \text{ on } \partial\Omega. \quad (1.6)$$

The present article is composed of two parts. The first concerns the stability of solutions of the mean field equation which was introduced to the subject by Onsager [15]:

$$-\Delta \Psi = C e^{-\beta \Psi} \text{ in } \Omega, \quad C > 0. \quad (1.7)$$

The convexity of the entropy functional

$$S(\omega) = \int_{\Omega} \omega \log \omega dx \quad (1.8)$$

is used in conjunction with the tools developed in [3] and [4]. In a standard normalization, the stability is proven for $\beta > -8\pi$, extending previous results obtained by Arnold's method [1] for β negative but small in absolute value. The case of $-\beta$ large corresponds to large energy and seems to be the relevant one in the formation of coherent structures.

Our first stability theorem is as follows.

Theorem 1.1. *Consider the variational problem*

$$S(A, E) = \inf_{\omega} \int_{\Omega} \omega \log \omega dx \quad (1.9)$$

subject to the mass and energy constraints

$$\omega \geq 0, \quad \int_{\Omega} \omega dx = A, \quad \frac{1}{2} \int_{\Omega} |\operatorname{curl}^{-1} \omega|^2 dx = E. \quad (1.10)$$

Assume that the minimizer μ , which satisfies

$$\int_{\Omega} \mu \log \mu dx = S(A, E) \quad (1.11)$$

and always exists [3], is unique. Consider the family F of initial data defined as the nonnegative functions that belong to a Holder space $C^{0,\alpha}$ for some $\alpha > 0$. Then for all $\epsilon > 0$ there exists $\delta > 0$ such that, for the solutions $\omega(t)$ of the Euler equation (1.4) with $\omega(0) \in F$, the implication

$$\|\omega(0) - \mu\|_{L^2(\Omega)} \leq \delta \quad \Rightarrow \quad \sup_{t \in \mathbb{R}} \|\omega(t) - \mu\|_{L^2(\Omega)} \leq \epsilon \quad (1.12)$$

holds.

The proof of this theorem and some variants to handle the case where the minimizer is not unique will be given in Section 2.

In Section 3 it is shown how to deduce nonlinear instability from linearized instability. The method follows closely the program initiated by the second and third authors of this paper (in collaboration with others) [6, 7, 10], where it was observed that the notion of instability is a robust property. In particular, using a perturbation method, nonlinear instabilities can be deduced from instabilities of the linearized operator.

In the present situation the linearized equation is

$$(\partial_t + u_0 \cdot \nabla) \tilde{\omega} + \tilde{u} \cdot \nabla \omega_0 = 0. \quad (1.13)$$

Solutions of (1.13) are described by the group of operators e^{tA_0} with generator

$$A_0 \tilde{\omega} = -u_0 \cdot \nabla \tilde{\omega} - \operatorname{curl}^{-1} \omega \cdot \nabla \omega_0. \quad (1.14)$$

This generator is the sum of an advection term which generates an isometry group in any $L^p(\Omega)$ ($1 \leq p < \infty$), denoted by $e^{-tu_0 \cdot \nabla}$, and a compact perturbation. It follows that the part $\Sigma(A_0)$ of the spectrum of A_0 outside the imaginary axis is purely discrete and that if $\Sigma(A_0) \neq \emptyset$, the type Λ of

the semi group e^{-tA_0} in any L^p space is given by

$$\Lambda = \sup_{\lambda \in \Sigma(A_0)} \operatorname{Re} \lambda. \quad (1.15)$$

Linear instability corresponds to the case when $\Lambda > 0$ (that is, $\Sigma(A_0) \neq \emptyset$). We denote by σ the Liapunov exponent of the autonomous flow associated to u_0 . Our instability theorem is as follows.

Theorem 1.2 (From Linear to Nonlinear Instability). *Given a steady flow $u_0 \in C^3(\Omega)$, consider the linearized equation*

$$\partial_t \tilde{\omega} + u_0 \cdot \nabla \tilde{\omega} + \operatorname{curl}^{-1} \tilde{\omega} \cdot \nabla \omega_0 = 0. \quad (1.16)$$

Assume that the type Λ of the semigroup and the Liapunov exponent σ of the flow generated by the vector field u_0 satisfy the inequality

$$\Lambda > \sigma. \quad (1.17)$$

Then for any $p > 2$ there exist positive constants C, ϵ_0, δ_0 and a family of solutions of the nonlinear Euler equation $\{\omega_\delta, 0 < \delta \leq \delta_0\}$ which satisfy both

$$\|\omega_\delta(0) - \omega_0\|_{W^{1,p}} \leq \delta \quad (1.18)$$

and

$$\sup_{0 < t \leq C|\log \delta|} \|\omega_\delta(t) - \omega_0\|_{L^2} \geq \epsilon_0. \quad (1.19)$$

In two space variables the use of the vorticity equation forces the discreteness of the spectrum of A_0 off the imaginary axis. In contrast, Friedlander and Vishik [5] consider a different linearized operator. They linearize the Euler equation for the velocity to get the operator

$$B_0 \tilde{u} = -(u_0 \cdot \nabla) \tilde{u} - (\tilde{u} \cdot \nabla) u_0 - \nabla q$$

on the space $\{\tilde{u} : \tilde{u} \in L^2(\Omega), \nabla \cdot \tilde{u} = 0\}$. They prove [5, 18] that the essential spectral radius of e^{tB_0} is equal to the maximal growth rate of the bicharacteristic-amplitude equations

$$\begin{aligned} \dot{x} &= u_0(x), \quad \dot{\xi} = -(\partial_x u_0)^T \xi, \\ \dot{b} &= -(\partial_x u_0) b + 2(\xi \cdot \partial_x u_0) b \xi / |\xi|^2. \end{aligned}$$

Thus the essential spectra of A_0 and B_0 are quite different.

In [6], it is proven that certain flows are nonlinearly unstable in H^s for $s > 2$, by making use of the point spectrum of B_0 . In [8], Grenier

proves similar results in the space $\{u \in L^2\}$. Here we prove the same kind of theorem for the L^2 norm of the vorticity. This is the very norm for which we obtain stability results. Furthermore, in contrast to [6] and [8], our results are valid (i) with no geometric hypothesis on the shape of the domain Ω and (ii) with less regularity assumed on the steady flow.

Two basic examples are shear flow and simple rotating flow. In both cases, the exponent σ vanishes. Indeed, shear flow is

$$u_0 = \begin{pmatrix} u(x_2) \\ 0 \end{pmatrix} \quad \text{with} \quad X_0(t, x_1, x_2) = \begin{pmatrix} u(x_2)t + x_1 \\ x_2 \end{pmatrix}.$$

Since the flow grows only linearly in time, $\sigma = 0$. Simple rotating flow in a disk is

$$u_0 = \begin{pmatrix} u(r) \sin \theta \\ -u(r) \cos \theta \end{pmatrix} \quad \text{with} \quad X_0(t, x_1, x_2) = \begin{pmatrix} r \cos(\theta - tu(r)/r) \\ r \sin(\theta - tu(r)/r) \end{pmatrix}$$

for which σ is also clearly equal to 0.

There exist classical examples [6] of shear or rotating flows going back to Rayleigh for which the linearized instability is proven. The previous comments imply that for these cases the hypothesis (1.17) is satisfied.

§2. Stability of Solutions of the Mean Field Equation

The solutions of the mean field equation

$$-\Delta \Psi = A \frac{e^{-\beta \Psi}}{\int_{\Omega} e^{-\beta \Psi} dx} \quad (2.1)$$

are invariant under the Euler flow. They are also closely related to the minimization of the functional

$$S(\omega) = \int_{\omega} H(\omega) dx \quad \text{with} \quad H(\omega) = \omega \ln \omega \quad (2.2)$$

subject to the constraints

$$\omega \geq 0, \quad A = \int \omega dx, \quad E = \frac{1}{2} \int_{\Omega} |\text{curl}^{-1} \omega|^2 dx. \quad (2.3)$$

The constants A and E are positive while β is the Lagrange multiplier of the energy constraint. More precisely, starting from the convexity of the function $\omega \rightarrow H(\omega) \geq e^{-1}$ the following facts are proven in [3, 4, 11] and are restated in the next two theorems.

Theorem 2.1. (i) *Denote*

$$P(E, A) = \left\{ \omega \geq 0 \text{ a.e. in } \Omega, \int_{\Omega} \omega(x) dx = 1, \frac{1}{2} \int_{\Omega} |\operatorname{curl}^{-1} \omega|^2 dx = E \right\}. \quad (2.4)$$

Then for any pair (E, A) of positive constants there is at least one ω_{\min} which solves the Microcanonical Variation Principle

$$S(E, A) \equiv \inf_{\omega \in P(E, A)} S(\omega) = S(\omega_{\min}). \quad (2.5)$$

Furthermore, any nonnegative solution ω of (2.5) is strictly positive and there exists at least one β such that $\omega = -\Delta\Psi$ with Ψ solving the corresponding β -mean field equation (2.1).

(ii) *Conversely, any solution of the β -mean field equation (2.1) is a solution of (2.5) with a given mass and energy $E(\beta)$.*

With appropriate scalings, the following renormalizations are assumed for the domain Ω and for the solution of the mean field equation:

$$|\Omega| = \operatorname{meas} \Omega = 1 \text{ and } A = \int_{\Omega} \omega(x) dx = 1. \quad (2.6)$$

The set of minimizers in (2.5) is denoted by $\mathcal{M}(E, A)$. In particular, $\mathcal{M}(E, 1)$, $P(E, 1)$ and $S(E, 1)$ are denoted by $\mathcal{M}(E)$, $P(E)$ and $S(E)$.

Theorem 2.2. (i) *For any $\beta > -8\pi$ there is at least one solution of (2.1); this solution is always unique for $\beta > 0$.*

(ii) *This solution is also unique for $\beta > -8\pi$ if the domain Ω is simply connected. In this setting the mapping $\beta \mapsto E(\beta)$ is well defined. It is also strictly decreasing and is onto from the interval $(-8\pi, \infty)$ to the interval $(0, E_c)$, where*

$$E_c = \lim_{\beta \rightarrow -8\pi^+} E(\beta).$$

(iii) *For starshaped domains there exists a number $\beta_c \leq -8\pi$ such that the mean field equation has no solution for $\beta < \beta_c$. In particular if Ω is a disk, $E_c = \infty$ and $\beta_c = -8\pi$.*

In fact, the situation of the disk corresponds to the ideal case where for any energy $E \in (0, \infty)$ any minimizer corresponds to a unique temperature $\beta \in (-8\pi, \infty)$. As a consequence, the minimizer is uniquely defined. In general, the situation turns out to be more complicated, either when the domain is not simply connected (no uniqueness of the solution of the mean

field equation for given β), or when $E_c < \infty$ and $\beta_c < -8\pi$ in which case several values of β may correspond to the same energy.

For our purposes observe that for a simply connected domain a solution of the mean field equation

$$-\Delta\Psi_\beta = \frac{e^{-\beta\Psi}}{\int_\Omega e^{-\beta\Psi_\beta} dx}, \quad \beta - 8\pi \quad (2.7)$$

provides a minimizer of $S(\omega_\beta)$ with the constraint $E(\omega) = E(\beta) \in (0, E_c)$. Furthermore if $\omega_{\beta'}$ is another minimizer with the same energy,

$$\text{either } \beta' = \beta \text{ or } \beta' < -8\pi. \quad (2.8)$$

Now we prove the stability theorem stated in the introduction.

Proof of Theorem 1.1. The hypothesis $\omega(0) \in F$ implies (cf. [20, 21]) that the corresponding solution is well defined and smooth. In particular one has for $\omega(t)$ and $u(t) = \text{curl}^{-1}\omega(t)$ the invariance

$$E(t) = \frac{1}{2} \int_\Omega |u(x, t)|^2 dx = \frac{1}{2} \int_\Omega |u(x, 0)|^2 dx = E(0), \quad (2.9)$$

and for any continuous function f

$$\int_\Omega f(\omega(x, t)) dx = \int_\Omega f(\omega(x, 0)) dx. \quad (2.10)$$

By contradiction, it is easy to see that the statement of the theorem is equivalent to the following one. For any sequence of initial data $\omega_n(0) \in F$ and for any sequence of times t_n , the limit

$$\lim_{n \rightarrow \infty} \|\omega_n(0) - \mu\|_{L^2(\Omega)} = 0 \quad (2.11)$$

implies the existence of a subsequence n_j such that

$$\lim_{n \rightarrow \infty} \|\omega_{n_j}(t_{n_j}) - \mu\|_{L^2(\Omega)} = 0. \quad (2.12)$$

So we assume (2.11).

We claim that

$$\lim_{n \rightarrow \infty} \int_\Omega H(\omega_n(x, 0)) dx = \int_\Omega H(\mu) dx \equiv S(\mu). \quad (2.13)$$

To prove (2.13), the strict positivity of μ as stated in Theorem 2.1 is used and a positive constant η_0 such that $\mu(x) \geq 2\eta_0$ in Ω is introduced. Let $0 \leq \eta \leq \min(\eta_0, e^{-1})$ and observe the inclusion $\{x \in \Omega \mid |\omega_n(x, 0)| < \eta\} \subset$

$\{x \in \Omega \mid |\omega_n(x, 0) - \mu(x)| > \eta\}$. Then we have

$$\begin{aligned}
& \left| \int_{\Omega} H(\omega(x, 0)) dx - S(\mu) \right| \\
& \leq \int_{\Omega \cap \{\omega_n(x, 0) \geq \eta\}} |H(\omega_n(x, 0)) - H(\mu(x))| dx \\
& \quad + \int_{\Omega \cap \{\omega_n(x, 0) < \eta\}} (|H(\omega_n(x, 0))| + |H(\mu(x))|) dx \\
& \leq \frac{2}{\eta} \int |\omega_n(x, 0) - \mu(x)| dx + \eta |\log \eta| \\
& \quad + \|H(\mu)\|_{L^\infty(\Omega)} \int_{|\omega_n(x, 0) - \mu(x)| \geq \eta} dx. \tag{2.14}
\end{aligned}$$

To complete the proof of the claim, choose η to make $\eta |\log \eta|$ less than $\epsilon/2$ and then choose n large enough to ensure, with the strong L^2 convergence of $\omega_n(0)$ to μ , that the sum of the two other terms is less than $\epsilon/2$.

Then there exists a subsequence n_j denoted below by n such that $\omega_n(t_n)$ converges weakly in $L^2(\Omega)$ to a function $\nu(x)$. With the notation $u_n(t_n) = \text{curl}^{-1} \omega_n(t_n)$, $u_\nu = \text{curl}^{-1} \nu$, the relations (2.9), (2.10) and the “entropic convergence lemma” (Proposition 3.1 of [2]) one has, for $\nu \in L^2(\Omega)$, $\nu \geq 0$, the following properties:

$$\int_{\Omega} \nu(x) dx = \lim \int_{\Omega} \omega_n(x, t_n) dx = \lim \int_{\Omega} \omega_n(x, 0) dx = 1, \tag{2.15}$$

$$\frac{1}{2} \int_{\Omega} |u_\nu(x)|^2 dx = \lim \int_{\Omega} |u_n(x, t_n)|^2 dx = E, \tag{2.16}$$

$$\int_{\Omega} H(\nu(x)) dx \leq \lim \int_{\Omega} H(\omega_n(x, t_n)) dx = H(\mu). \tag{2.17}$$

Thus ν is a minimizer. The hypothesis concerning the uniqueness of the minimizer implies the relation $\nu = \mu$ and the strong $L^2(\Omega)$ convergence because

$$\begin{aligned}
\int_{\Omega} |\nu(x)|^2 dx & \leq \lim \int_{\Omega} |\omega_n(x, t_n)|^2 dx = \lim \int_{\Omega} |\omega_n(x, 0)|^2 dx \\
& = \int_{\Omega} |\mu(x)|^2 dx = \int_{\Omega} |\nu(x)|^2 dx. \tag{2.18}
\end{aligned}$$

This proves Theorem 1.1.

Since the minimizer may not be unique, we consider several extensions of the above theorem. We have the following variant.

Theorem 2.3. *No assumption is made on the uniqueness of the minimizer. On the other hand, consider for a given energy E , a minimizer $\mu \in P(E)$. Let $F' \subset F$ be the subset of initial data uniformly bounded in $L^\infty(\Omega)$ by a fixed constant k . Then for all $\epsilon > 0$ there exists $\delta > 0$ such that, for the solutions $\omega(t)$ of the Euler equation (1.4) with $\omega(0) \in F'$, the assertion*

$$\|\omega(0) - \mu\|_{L^2(\Omega)} \leq \delta \quad \Rightarrow \quad \sup_{t \in \mathbb{R}} \inf_{\nu \in \mathcal{M}(E)} \|\omega(t) - \nu\|_{L^2(\Omega)} \leq \epsilon \quad (2.19)$$

holds.

Proof. The proof follows the preceding one and leads with no major modification to the extraction of a subsequence such that $\omega_n(t_n)$ converges weakly in $L^2(\Omega)$ to a minimizer $\nu \in \mathcal{M}(E)$. Since the uniqueness of this minimizer is not assumed, the relation (2.18) is no longer valid and the convexity of the entropy H (with the uniform boundedness) is used instead. In fact, by a Taylor expansion we have

$$\begin{aligned} & \int_{\Omega} (H(\omega_n(x, t_n)) - H(\nu(x, t))) dx \\ &= \int_{\Omega} (1 + \log \nu(x, t)) (\omega_n(x, t_n) - \nu(x, t)) dx \\ & \quad + \int_{\Omega} \frac{1}{2(\overline{\omega(x, t)})} (\omega_n(x, t_n) - \nu(x, t))^2 dx \end{aligned} \quad (2.20)$$

with $\overline{\omega(x, t)}$ between $\omega_n(x, t_n)$ and $\nu(x)$. Using the uniform boundedness of $\omega_n(x, 0)$ and the weak L^2 convergence of $\omega_n(t_n)$, we have for $0 < \alpha$ small enough the inequality

$$\alpha \lim \int_{\Omega} (\omega_n(x, t_n) - \nu(x, t))^2 dx \leq \lim \int_{\Omega} (H(\omega_n(x, t_n)) - H(\nu(x, t))) dx = 0, \quad (2.21)$$

which proves the strong convergence.

If the domain is simply connected, there is, as recalled above in Theorem 2.2, a one-to-one correspondence between the solutions of the mean field equation for $\beta > -8\pi$ and the minimizers with energy $E > E_c$. This does not seem to exclude in general the existence of other minimizers for the same energy which solve a mean field equation with $\beta < -8\pi$. In accordance with this observation we give the following theorem.

Theorem 2.4. *Assume that the open set Ω is simply connected and consider a solution Ψ^* of the mean field equation*

$$-\Delta \Psi^* = \frac{e^{-\beta \Psi^*}}{\int_{\Omega} e^{-\beta \Psi^*} dx}, \quad \Psi^* = 0 \quad \text{on } \partial\Omega. \quad (2.22)$$

Denote by $\mu = -\Delta \Psi^*$ the corresponding vorticity with energy $E(\mu)$ and consider the same set of initial data F' as in Theorem 2.3. Then for all $\epsilon > 0$ there exists $\delta > 0$ such that, for the solutions $\omega(t)$ of the Euler equation (1.4) with $\omega(0) \in F'$, the assertion

$$\|\omega(0) - \mu\|_{L^2(\Omega)} \leq \delta \Rightarrow \sup_{t \in \mathbb{R}} \|\omega(t) - \mu\|_{L^2(\Omega)} \leq \epsilon \quad (2.23)$$

holds.

Proof. According to Theorem 3.2 ii) of [4], one has

$$\mathcal{M}(E(\mu)) = \{\mu\} \cup \mathcal{M}^*, \quad (2.24)$$

where any element $\nu \in \mathcal{M}^*$ is a solution of the mean field equation with a temperature $\beta(\nu) < -8\pi$. Observe by contradiction that the $L^2(\Omega)$ -distance d between μ and \mathcal{M}^* is strictly positive. By Theorem 2.3, for all $\epsilon > 0$ there exists $\delta > 0$ such that the relation

$$\|\omega(0) - \mu\|_{L^2(\Omega)} \leq \delta \quad (2.25)$$

implies the relation

$$\sup_{t \in \mathbb{R}} \inf_{\nu \in \mathcal{M}(E(\mu))} \|\omega(t) - \nu\|_{L^2(\Omega)} \leq \epsilon. \quad (2.26)$$

We observe that for $\omega(0) \in F'$ we have $\omega(\cdot) \in C(\mathbb{R}_t; L^2(\Omega))$. Choosing $\epsilon < \frac{d}{2}$ and $\delta < \frac{d}{2}$ and using the triangle inequality, we conclude that

$$\sup_{t \in \mathbb{R}} \inf_{\nu \in \mathcal{M}(E(\mu))} \|\omega(t) - \nu\|_{L^2(\Omega)} = \sup_{t \in \mathbb{R}} \|\omega(t) - \mu\|_{L^2(\Omega)}. \quad (2.27)$$

This completes the proof.

The previous stability results involve, as in most of the contributions in the subject (cf. for instance [1, 13, 19]), the L^2 norm of the vorticity. Finally we conclude this section by proving, using the notion of entropic convergence, a weaker result under a weaker hypothesis.

Theorem 2.5. *The family of initial data F is defined as in Theorem 1.1. For some $p > 1$, $F_p \subset F$ denotes the subset of initial data uniformly bounded in $L^p(\Omega)$; that is, there exists $k < \infty$ such that*

$$\forall \omega(0) \in F_p, \quad \int_{\Omega} |\omega(0, x)|^p dx \leq k. \quad (2.28)$$

Then for all (A, E) and $\epsilon > 0$ there exists $\delta > 0$ such that

$$\begin{aligned} & \left| \int_{\Omega} \omega(0, x) dx - A \right| + \left| \frac{1}{2} \int_{\Omega} |(\operatorname{curl}^{-1} \omega(0))(x)|^2 dx - E \right| \\ & + \left| \int_{\Omega} H(\omega(0, x)) dx - S(A, E) \right| \leq \delta \end{aligned} \quad (2.29)$$

implies

$$\sup_{t \in \mathbb{R}} \inf_{\nu \in \mathcal{M}(A, E)} \|\operatorname{curl}^{-1} \omega(t) - \operatorname{curl}^{-1} \nu\|_{L^2(\Omega)} \leq \epsilon. \quad (2.30)$$

Proof. By (2.25) we extract from any sequence $\omega_n(t_n)$ a subsequence, still denoted by $\omega_n(t_n)$, which converges weakly in $L^p(\Omega)$ to a function $\nu \in L^p(\Omega)$. The inequality (2.29) and the compactness of the Sobolev imbedding $W^{1,p}(\Omega) \subset L^2(\Omega)$ imply the following limits:

$$A = \lim \int_{\Omega} \omega_n(t_n, x) dx = \int_{\Omega} \nu(x) dx, \quad (2.31)$$

$$E = \lim \frac{1}{2} \int_{\Omega} |(\operatorname{curl}^{-1} \omega_n(t_n))(x)|^2 dx = \frac{1}{2} \int_{\Omega} |\operatorname{curl}^{-1} \nu|^2 dx. \quad (2.32)$$

Then by the entropic convergence we have

$$\int_{\Omega} H(\nu) dx \leq \liminf \int_{\Omega} H(\omega_n(t_n)) dx = \lim \int_{\Omega} H(\omega_n(0)) dx = S(A, E). \quad (2.33)$$

Therefore ν is a minimizer, the equality (2.32) implies the $L^2(\Omega)$ strong convergence of the sequence $\operatorname{curl}^{-1} \omega_n(t_n)$, and the proof is complete.

§3. From Linear to Nonlinear Instability

This section is devoted to the proof of Theorem 1.2. First we recall some facts about the Liapunov exponent. The classical Liapunov exponent σ for the flow $X_0(t, x)$ induced by u_0 ,

$$\frac{\partial X_0}{\partial t} = u_0(X_0), \quad X_0(0, x) = x, \quad (3.1)$$

is defined by

$$\sigma = \sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \left| \frac{\partial X_0}{\partial x} \right|, \quad (3.2)$$

where $\frac{\partial X_0}{\partial x}$ denotes the 2×2 matrix $(\partial X_0^i / \partial x^j)$.

Proposition 3.1. *The Liapunov exponent can also be defined as*

$$\sigma = \lim_{t \rightarrow \infty} \frac{1}{t} \sup_x \log \left| \frac{\partial X_0}{\partial x} \right|. \quad (3.3)$$

This fact might exist in the theory of dynamical systems but in the absence of a reference a short proof is given. Clearly

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sup_x \log \left| \frac{\partial X_0}{\partial x} \right| \geq \sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \left| \frac{\partial X_0}{\partial x} \right|. \quad (3.4)$$

To prove the converse we follow the argument of [7]. By definition of σ , for every pair $\epsilon > 0$ and $x \in \Omega$, there exists a "time" $T[x] > 0$ such that

$$t \geq T[x] \Rightarrow \left| \frac{\partial X_0(t, x)}{\partial x} \right| < e^{t(\sigma + \epsilon)}.$$

By the continuity of $\partial_x X_0$ and the boundary condition, for all $x \in \bar{\Omega}$ there is a neighborhood $B_x \subset \bar{\Omega}$ such that

$$\left| \left(\frac{\partial X_0}{\partial x} \right) (T[x], x) \right| < e^{T[x](\sigma + \epsilon)}.$$

We introduce a finite covering of $\bar{\Omega}$ by such open sets: $\bar{\Omega} = B_{x_1} \cup B_{x_2} \cup \dots \cup B_{x_N}$ with $N < \infty$. Denote $B_i = B_{x_i}, T_i = T[x_i]$ for $1 \leq i \leq N$. Given $x \in \bar{\Omega}$, choose i_1 such that $x \in B_{i_1}$ and then choose the sequence

$$y_1 = X_0(T_{i_1}, x) \in B_{i_2},$$

$$y_2 = X_0(T_{i_1} + T_{i_2}, x) \in B_{i_3},$$

.....

$$y_1 = X_0(T_{i_1} + \dots + T_{i_k}, x) \in B_{i_{k+1}} \quad (1 \leq k < \infty).$$

Now

$$y_2 = X_0(T_{i_1} + T_{i_2}, x) = X_0(T_{i_2}, y_1) = X_0(T_{i_2}, X_0(T_{i_1}, x)). \quad (3.5)$$

Therefore

$$\frac{\partial y_2}{\partial x} = \left(\frac{\partial X_0}{\partial x} \right) (T_{i_2}, y_1) \cdot \left(\frac{\partial X_0}{\partial x} \right) (T_{i_1}, x), \quad (3.6)$$

so that

$$\left| \frac{\partial [X_0(T_{i_1} + T_{i_2}, x)]}{\partial x} \right| \leq e^{[T_{i_1} + T_{i_2}](\sigma + \epsilon)}. \quad (3.7)$$

Similarly for any x and k we have the inequality

$$\left| \frac{\partial [X_0(T_{i_1} + \dots + T_{i_k}, x)]}{\partial x} \right| \leq e^{[T_{i_1} + \dots + T_{i_k}](\sigma + \epsilon)}. \quad (3.8)$$

Now, given t and x , we choose

$$S_k = T_{i_1} + \dots + T_{i_k} \leq t < T_{i_1} + \dots + T_{i_k} + T_{i_{k+1}}, \quad (3.9)$$

and denote by X'_0 the derivative of

$$X_0(t, x) = X_0(t - S_k, X_0(S_k, x))$$

with respect to x . By the chain rule and estimate (3.8) we obtain, for any pair $(t, x) \in [0, \infty) \times \bar{\Omega}$,

$$\begin{aligned} \left| \frac{\partial [X_0(t, x)]}{\partial x} \right| &\leq \sup_{y \in \bar{\Omega}} |X_0'(t - S_k, y)| \cdot |X_0'(S_k, x)| \\ &\leq \left[\sup_{y \in \bar{\Omega}, 0 \leq s \leq \max_i T_i} |X_0'(s, y)| \right] \cdot e^{(\sigma + \epsilon)S_k} \\ &\leq C e^{(\sigma + \epsilon)S_k} \leq C' e^{(\sigma + \epsilon)t}, \end{aligned} \quad (3.10)$$

which implies (3.3).

The following key lemma states that if a velocity field $v(t, x)$ is close enough to $u_0(x)$ in C^1 , then their corresponding flows are close together in a sufficiently short time interval where $\eta e^{(t-s)\mu}$ is small.

Lemma 3.1. *Let $u_0(x)$ be a steady C^1 solution. Let $X_0(t, s, x) = X_0(t - s, x)$ be its classical flow with Liapunov exponent σ . Let $v(t, x) \in C^1$ be another vector field defined on $\mathbb{R}_t \times \bar{\Omega}$ that is incompressible and tangent to the boundary*

$$\nabla \cdot v = 0 \text{ in } \Omega \text{ and } v \cdot \vec{n} = 0 \text{ on } \partial\Omega. \quad (3.11)$$

Denote by $X(t, s, x)$ its flow

$$\frac{\partial X}{\partial t} = v(t, X), \quad X(s, s, x) = x.$$

Let $0 < \epsilon$ and $\mu > \sigma + \epsilon$. Then there exist positive constants C_1, C_2 and θ_0 with the following property.

For any positive constant η the estimate

$$\|v(t, \cdot) - u_0(\cdot)\|_{C^1(\Omega)} \leq \eta e^{(t-s)\mu} \quad (3.12)$$

for $s \leq t \leq s + S_\eta$ with

$$S_\eta \equiv \frac{1}{\mu} \ln \frac{\theta_0}{\eta} \quad (3.13)$$

implies for $s \leq t \leq s + S_\eta$ the a priori estimates

$$|X(t, s, x) - X_0(t, s, x)| \leq C_1 \eta e^{(t-s)\mu}, \quad (3.14)$$

$$\left| \frac{\partial (X(t, s, x) - X_0(t, s, x))}{\partial x} \right| \leq C_2 \theta_0 e^{(t-s)(\sigma + \epsilon)}. \quad (3.15)$$

Proof. When there is no risk of confusion, the arguments (t, s, x) in X and X_0 will be omitted. The difference $X - X_0$ satisfies

$$\frac{\partial (X - X_0)}{\partial t} = v(t, X) - u_0(X_0); \quad (X - X_0)(s, s, x) = 0.$$

Thus

$$\begin{aligned} & \left[\frac{\partial}{\partial t} - \frac{\partial u_0}{\partial X}(X_0) \right] (X - X_0) \\ &= \left[-\frac{\partial u_0}{\partial X}(X_0) \right] (X - X_0) + [u_0(X) - u_0(X_0)] + [v(t, X) - u_0(X)] \equiv g. \end{aligned} \quad (3.16)$$

By the Taylor expansion and (3.12), we have

$$|g| \leq \frac{1}{2} \left| -\frac{\partial^2 u_0}{\partial X^2}(\bar{X})(X - X_0)^2 \right| + \eta e^{(t-s)\mu} \leq C|X - X_0|^2 + \eta e^{(t-s)\mu}.$$

Let

$$T^* = \sup\{t_1 : |X(t_1) - X_0(t_1)| \leq C_1 \eta e^{(t_1-s)\mu}\} \quad (3.17)$$

with C_1 to be determined. In the interval $[s, T^*]$, we have from (3.2) and (3.16) the estimate

$$\begin{aligned} & |X(t) - X_0(t)| \\ & \leq C_\epsilon \int_s^t e^{(\sigma+\epsilon)(t-\tau)} [C_1 \eta e^{(\tau-s)\mu}]^2 d\tau + C'_\epsilon \eta \int_s^t e^{(\sigma+\epsilon)(t-\tau)} e^{(\tau-s)\mu} d\tau \\ & \leq C_\epsilon e^{(\sigma+\epsilon)(t-s)} C_1^2 \eta^2 \int_0^{t-s} e^{(\sigma+\epsilon)\rho} e^{2\mu\rho} d\rho + C''_\epsilon \eta e^{\mu(t-s)} \\ & \leq C_1^2 C'_\epsilon [\eta e^{\mu(t-s)}]^2 + C''_\epsilon [\eta e^{\mu(t-s)}]. \end{aligned}$$

Putting $t = T^*$, we have from (3.17)

$$C_1 [\eta e^{\mu(T^*-s)}] = |X(T^*) - X_0(T^*)| \leq C_1^2 C'_\epsilon [\eta e^{\mu(T^*-s)}]^2 + C''_\epsilon [\eta e^{\mu(T^*-s)}].$$

Hence

$$\eta e^{\mu(T^*-s)} \geq \frac{1}{C_1 C'_\epsilon} - \frac{C''_\epsilon}{C_1^2 C'_\epsilon} > \theta_0$$

if we choose

$$C_1 = 2C''_\epsilon \quad \text{and} \quad 0 < \theta_0 < \frac{1}{4C'_\epsilon C''_\epsilon}.$$

By (3.13) we have

$$S_\eta < T^* - s$$

and we deduce from (3.17) that (3.14) is valid.

Next we define $Y = \frac{\partial}{\partial x}(X - X_0)$, where $\frac{\partial}{\partial x} = \frac{\partial}{\partial x_1}$ or $\frac{\partial}{\partial x_2}$. It satisfies

$$\begin{aligned} \frac{\partial Y}{\partial t} &= \frac{\partial}{\partial x} [v(t, X(t, s, x)) - u_0(X_0(t, s, x))] = \frac{\partial v}{\partial X} \frac{\partial X}{\partial x} - \frac{\partial u_0}{\partial X} \frac{\partial X_0}{\partial x} \\ &= \frac{\partial u_0}{\partial X} Y + \frac{\partial [v - u_0]}{\partial X} \left\{ \frac{\partial X_0}{\partial x} + Y \right\} \equiv \frac{\partial u_0}{\partial X} Y + h. \end{aligned} \quad (3.18)$$

By assumption (3.12) and (3.14), for $0 \leq t - s \leq S_\eta$,

$$\begin{aligned} \frac{\partial[v - u_0]}{\partial X} &= \frac{\partial v(t, X)}{\partial X} - \frac{\partial u_0(X_0)}{\partial X} \\ &= \left\{ \frac{\partial v(t, X)}{\partial X} - \frac{\partial u_0(X)}{\partial X} \right\} + \left\{ \frac{\partial u_0(X)}{\partial X} - \frac{\partial u_0(X_0)}{\partial X} \right\} = O(\theta_0). \end{aligned}$$

Therefore by (3.13) we have

$$|h| \leq C\theta_0[1 + |Y|].$$

By (3.18) we can now estimate

$$\begin{aligned} |Y(t)| &\leq C \int_s^t e^{(t-\tau)(\sigma+\epsilon/2)} |h(\tau)| d\tau \leq C\theta_0 \int_s^t e^{(t-\tau)(\sigma+\epsilon/2)} [1 + |Y(\tau)|] d\tau \\ &\leq C\theta_0 e^{(t-s)(\sigma+\epsilon/2)} \left\{ 1 + \int_s^t e^{(s-\tau)(\sigma+\epsilon/2)} |Y(\tau)| d\tau \right\}. \end{aligned}$$

It thus follows, for $t - s \leq S_\eta$ and θ_0 sufficiently small, that

$$|Y(t)| \leq C\theta_0 e^{(t-s)(\sigma+\epsilon)}.$$

Lemma 3.2. *Given u_0 and v as in Lemma 3.1, let $\tilde{\omega}$ solve the linear equation*

$$(\partial_t + v \cdot \nabla)\tilde{\omega} = 0.$$

Then for any $1 \leq p \leq \infty$ and

$$0 \leq t \leq S_\eta = \frac{1}{\mu} \log \frac{\theta_0}{\eta} \quad (3.19)$$

($0 \leq t < \infty$ if $v \equiv u_0$), we have

$$\|\tilde{\omega}(t)\|_{W^{1,p}} \leq C_\epsilon e^{(\sigma+\epsilon)t} \|\tilde{\omega}(0)\|_{W^{1,p}}.$$

Proof. Denoting $\Gamma(x) = \tilde{\omega}(0, x)$, we have

$$\tilde{\omega}(t, x) = \Gamma(X(t, 0, x)).$$

Clearly $\|\tilde{\omega}(t)\|_{L^p} = \|\Gamma\|_{L^p}$ and

$$\left| \frac{\partial \tilde{\omega}}{\partial x} \right| = \left| \frac{\partial \Gamma}{\partial X} \cdot \frac{\partial X}{\partial x} \right| \leq C e^{(\sigma+\epsilon)t} \left| \frac{\partial \Gamma}{\partial X} \right|$$

by Lemma 3.1. Hence

$$\int_\Omega \left| \frac{\partial \tilde{\omega}}{\partial x} \right|^p dx \leq C e^{p(\sigma+\epsilon)t} \int_\Omega \left| \frac{\partial \Gamma}{\partial X} \right|^p dx$$

since the Jacobian equals one. This proves the desired estimate on the first derivatives.

Now we prove that the eigenfunctions are smooth.

Lemma 3.3. *Any eigenfunction $\phi_\lambda \in L^2(\Omega)$ corresponding to an eigenvalue λ with $\operatorname{Re} \lambda > \sigma$ belongs to $W^{1,p}(\Omega)$ for all $p < \infty$.*

Proof. The eigenfunction satisfies

$$(\lambda + u_0 \cdot \nabla)\phi_\lambda + \operatorname{curl}^{-1}\phi_\lambda \cdot \nabla\omega_0 = 0, \quad (3.20)$$

from which we deduce the relation

$$\phi_\lambda = \int_0^\infty e^{-\lambda\tau} e^{-\tau u_0 \cdot \nabla} (\operatorname{curl}^{-1}\phi_\lambda \cdot \nabla\omega_0) d\tau. \quad (3.21)$$

By assumption, ϕ_λ is in $L^2(\Omega)$ so that $\operatorname{curl}^{-1}\phi_\lambda \cdot \nabla\omega_0$ is in $H^1(\Omega)$. By Lemma 3.1 in the simple case $v = u_0$ and $0 \leq t < \infty$, we have

$$e^{-t u_0 \cdot \nabla} : W^{1,p}(\Omega) \rightarrow W^{1,p}(\Omega)$$

with norm $O(e^{(\sigma+\epsilon)t})$ for any $\epsilon > 0$. Therefore the assumption $\operatorname{Re} \lambda > \sigma$ implies that ϕ_λ is in $H^1(\Omega)$. Now this implies that $\operatorname{curl}^{-1}\phi_\lambda \cdot \nabla\omega_0$ belongs to $H^2(\Omega) \subset W^{1,p}(\Omega)$ for $p < \infty$. Using once again the estimate

$$\|e^{-t u_0 \cdot \nabla}\|_{W^{1,p}} \leq O(e^{(\sigma+\epsilon)t}) \text{ and } \operatorname{Re} \lambda > \sigma, \quad (3.22)$$

we conclude the proof of the lemma.

The main ingredient in the proof of instability is a bootstrap lemma (Theorem 3.2). Before giving this lemma it is worthwhile to compare it with classical results on the 2D Euler equation. It is known that for smooth initial data the solution of the Euler equation remains as smooth as the data. For instance, for any $W^{1,p}$ norm we have the crude estimate

$$\|\omega(t)\|_{W^{1,p}} \leq \|\omega(0)\|_{W^{1,p}} \exp\left(C \int_0^t \|\nabla u(s)\|_{L^\infty(\Omega)} ds\right). \quad (3.23)$$

However due to the fact that curl^{-1} is not continuous from L^∞ to $W^{1,\infty}$, time dependent estimates on $\|\nabla u(t)\|_{L^\infty(\Omega)}$ are subtle [20]. The same observation applies to the difference $\tilde{\omega}(t) = \omega(t) - \omega_0$ where ω_0 is a stationary solution and $\omega(t)$ a perturbation. The equation for $\tilde{\omega}$ is

$$\partial_t \tilde{\omega} + u \nabla \tilde{\omega} + \operatorname{curl}^{-1}(\tilde{\omega}) \cdot \nabla \omega_0 = 0. \quad (3.24)$$

Applying the operator D (derivative with respect to the first or the second spatial variable) to the equation (3.24), multiplying this equation by $(D\tilde{\omega})^{p-1}$, integrating over Ω and using the Gronwall lemma, we deduce for

$\tilde{\omega}$ the inequality

$$\|\tilde{\omega}(t)\|_{W^{1,p}} \leq \|\tilde{\omega}(0)\|_{W^{1,p}} \exp\left(C \int_0^t (\|\nabla \tilde{u}(s)\|_{L^\infty(\Omega)} + 1) ds\right). \quad (3.25)$$

By the same argument involving a nonlinear Gronwall lemma,

$$\|\nabla \tilde{u}(t)\|_{L^\infty(\Omega)}$$

can be estimated, but only for a short time which depends on the initial data. More precisely, for

$$e^{Ct} \|\tilde{\omega}(0)\|_{W^{1,p}} \leq D, \quad (3.26)$$

we have

$$\|\nabla \tilde{u}(t)\|_{L^\infty(\Omega)} \leq \frac{E \|\tilde{\omega}(0)\|_{W^{1,p}}}{D - e^{Ct} \|\tilde{\omega}(0)\|_{W^{1,p}}} \quad (3.27)$$

with convenient constants C , D and E . Combining the formulas (3.25) and (3.27), we easily obtain the following result.

Theorem 3.1. *Given any stationary solution ω_0 of the 2D Euler equation in a bounded domain Ω , there exist positive constants C, C' and θ which depend only on Ω and ω_0 such that for any perturbation*

$$\tilde{\omega}(t) = \omega(t) - \omega_0$$

we have the implication

$$|t| \leq \frac{1}{C} \log\left(\frac{\theta}{\|\tilde{\omega}(0)\|_{W^{1,p}}}\right) \Rightarrow \|\tilde{\omega}(t)\|_{W^{1,p}} \leq \|\tilde{\omega}(0)\|_{W^{1,p}} e^{C't}. \quad (3.28)$$

In contrast to the sharp, but global, Wolibner estimate [20], the above estimates are only local in time and the constants are not “sharp”. The bootstrap lemma states that under convenient hypotheses we can sharpen the constants in (3.28).

Theorem 3.2 (Bootstrap Lemma). *Denote by u_0 a steady flow with classical Liapunov exponent σ as in (3.2). Let μ be any real number strictly greater than σ . Let $T > 0$ and $p > 2$. Then there exist positive constants θ, C_1, C_2 , and C_3 with the following property. Let $\omega(t, x) \in C(\mathbb{R}_t; W^{1,p})$ be any solution of the nonlinear Euler equation which satisfies the initial estimate*

$$\|\omega(0) - \omega_0\|_{W^{1,p}(\Omega)} \leq C_1 \delta \quad (3.29)$$

and the L^2 estimate

$$\|\omega(t) - \omega_0\|_{L^2(\Omega)} \leq C_2 \delta e^{\mu t} \text{ in } [0, T]. \quad (3.30)$$

Let

$$T_\delta = \frac{1}{\mu} \log \frac{\theta}{\delta}.$$

Then in the time interval $0 \leq t \leq \min\{T, T_\delta\}$, the solution also satisfies the $W^{1,p}$ estimate

$$\|\omega(t) - \omega_0\|_{W^{1,p}(\Omega)} \leq C_3 \delta e^{\mu t}. \quad (3.31)$$

Proof. Introduce the notation

$$\tilde{\omega}(t) = \omega(t) - \omega_0 \quad \text{and} \quad \tilde{u}(t) = u(t) - u_0 = \operatorname{curl}^{-1} \tilde{\omega}(t).$$

Given $\eta > 0$, let

$$S_\eta = \frac{1}{\mu} \log \frac{\theta_0}{\eta} \quad \text{and} \quad S = \sup\{t : \|\tilde{u}(t)\|_{C^1} \leq \eta e^{\mu t}\}. \quad (3.32)$$

Observe that $(\tilde{\omega}, \tilde{u})$ solves the equation

$$(\partial_t + u \cdot \nabla) \tilde{\omega} = -\tilde{u} \cdot \nabla \omega_0 \quad (3.33)$$

with initial data

$$\|\tilde{\omega}(0)\|_{W^{1,p}(\Omega)} \leq C_1 \delta. \quad (3.34)$$

By the Duhamel Principle and Lemma 3.2, this implies, for $0 \leq t \leq \min\{T, S_\eta, S\}$, the estimate

$$\|\tilde{\omega}(t)\|_{W^{1,p}} \leq C \delta e^{(\sigma+\epsilon)t} + C \int_0^t e^{(\sigma+\epsilon)(t-\tau)} \|\tilde{u}(\tau) \cdot \nabla \omega_0\|_{W^{1,p}} d\tau. \quad (3.35)$$

The norm on the right is estimated by

$$\begin{aligned} \|\tilde{u}(\tau) \cdot \nabla \omega_0\|_{W^{1,p}} &\leq C \|\tilde{u}\|_{W^{1,p}} \leq C \|\tilde{\omega}\|_{L^p} \leq \gamma \|\tilde{\omega}\|_{W^{1,p}} + C_\gamma \|\tilde{\omega}\|_{L^2} \\ &\leq \gamma \|\tilde{\omega}\|_{W^{1,p}} + C_\gamma \delta e^{\mu\tau} \end{aligned}$$

by (3.30), where γ is arbitrarily small. This is placed into the integral inequality (3.35) as

$$\|\tilde{\omega}(t)\|_{W^{1,p}} \leq C \delta e^{(\sigma+\epsilon)t} + C_\gamma \delta e^{\mu t} + C_\gamma \int_0^t e^{(\sigma+\epsilon)(t-\tau)} \|\tilde{\omega}(\tau)\|_{W^{1,p}} d\tau.$$

Multiplying by $e^{-\mu t}$, we obtain

$$\begin{aligned} e^{-\mu t} \|\tilde{\omega}(t)\|_{W^{1,p}} &\leq C \delta + C_\gamma \left\{ \sup_{0 \leq \tau \leq t} e^{-\mu\tau} \|\tilde{\omega}(\tau)\|_{W^{1,p}} \right\} \int_0^t e^{(\sigma+\epsilon-\mu)s} ds \\ &\leq C \delta + C_\gamma \left\{ \sup_{0 \leq \tau \leq t} e^{-\mu\tau} \|\tilde{\omega}(\tau)\|_{W^{1,p}} \right\}. \end{aligned}$$

Hence for $\gamma = 1/2C$ and for $0 \leq t \leq \min\{T, S_\eta, S\}$, we have

$$e^{-\mu t} \|\tilde{\omega}(t)\|_{W^{1,p}} \leq 2C\delta. \quad (3.36)$$

Thus, since $p > 2$, we have for $0 \leq t \leq \min\{T, S_\eta, S\}$,

$$\|\tilde{u}(t)\|_{C^1} \leq C\|\tilde{u}(t)\|_{W^{2,p}} \leq C\|\tilde{\omega}(t)\|_{W^{1,p}} \leq C_0\delta e^{\mu t} \quad (3.37)$$

for some constant C_0 . Now we choose $\eta = 2C_0\delta$, and $\theta = \theta_0/2C_0$. It follows from the definition of S that $S > S_\eta$. Hence

$$\|\tilde{\omega}(t)\|_{W^{1,p}} \leq 2C\delta e^{\mu t} \quad (3.38)$$

for $0 \leq t \leq \min\{T, S_\eta\}$. Noticing that $S_\eta = T_\delta$, we deduce (3.31).

Proof of the Instability Theorem 1.2. We now return to the nonlinear equation

$$(\partial_t + u_0 \cdot \nabla)\tilde{\omega} + (\tilde{u} \cdot \nabla)\omega_0 = -(\tilde{u} \cdot \nabla)\tilde{\omega} \quad (3.39)$$

satisfied by the perturbation $\tilde{\omega} = \omega - \omega_0$. It takes the Duhamel form

$$\tilde{\omega}(t) = e^{tA_0}\tilde{\omega}(0) - \int_0^t e^{(t-\tau)A_0}(\tilde{u} \cdot \nabla)\tilde{\omega}(\tau)d\tau. \quad (3.40)$$

Assuming that $\Lambda > \sigma$, by compactness (see (1.15)) there is at least one eigenvalue λ such that $\text{Re } \lambda = \Lambda$. If $\lambda = \Lambda$ is a real eigenvalue, we can choose $\tilde{\omega}(0, x) = \delta\phi_\lambda(x)$, where δ is small and ϕ_λ is the eigenfunction

$$A_0\phi_\lambda = \lambda\phi_\lambda. \quad (3.41)$$

However for the sake of generality the case where Λ is not an eigenvalue is considered below and then we choose for initial data the function $\tilde{\omega}(0, x) = \delta\mathcal{I}m\phi_\lambda(x)$. Taking the L^2 norm in (3.40) gives

$$\|\tilde{\omega}(t) - \delta e^{tA_0}\mathcal{I}m\phi_\lambda\|_{L^2} \leq C_\nu \int_0^t e^{(t-\tau)\nu} \|(\tilde{u} \cdot \nabla)\tilde{\omega}\|_{L^2} d\tau \quad (3.42)$$

for any $\Lambda < \nu < 2\Lambda$. Define

$$T = \sup \left\{ s : \|\tilde{\omega}(t) - \delta e^{tA_0}\mathcal{I}m\phi_\lambda\|_{L^2} \leq \frac{\delta}{2} \|e^{tA_0}\mathcal{I}m\phi_\lambda\|_{L^2}, \quad \forall t \in [0, s] \right\}. \quad (3.43)$$

Then for $0 \leq t \leq T$ we have

$$\|\tilde{\omega}(t)\|_{L^2} \leq C_2\delta e^{\mu t}, \quad (3.44)$$

where $\mu = \text{Re } \lambda = \Lambda$. Hence from Theorem 3.2 with $\mu = \Lambda > \sigma$, we crudely estimate the nonlinear term as

$$\|(\tilde{u} \cdot \nabla)\tilde{\omega}\|_{L^2} \leq \|\tilde{u}\|_{L^\infty} \|\nabla\tilde{\omega}\|_{L^2} \leq C\|\tilde{\omega}\|_{W^{1,p}}^2 \leq C\{C_3\delta e^{\mu t}\}^2 \quad (3.45)$$

provided $t \leq \min\{T_\delta, T\}$. We shall prove that $T_\delta \leq T$ for δ small. If not, notice that

$$\|e^{tA_0} \delta \mathcal{I}m\phi_\lambda\|_{L^2} \geq c_0 \delta e^{\mu t}. \quad (3.46)$$

Therefore we have for $0 \leq t \leq \min\{T_\delta, T\}$ the estimate

$$\begin{aligned} \|\tilde{\omega}(t) - \delta e^{tA_0} \mathcal{I}m\phi_\lambda\|_{L^2} &\leq C \int_0^t e^{(t-\tau)\nu} C_3 \{\delta e^{\mu\tau}\}^2 d\tau \leq C \{\delta e^{\mu t}\}^2 \\ &\leq \left\{ \frac{C}{c_0} \delta e^{\mu t} \right\} \|\delta e^{tA_0} \mathcal{I}m\phi_\lambda\|_{L^2} \leq C' \theta \|\delta e^{tA_0} \mathcal{I}m\phi_\lambda\|_{L^2} \\ &\leq \frac{\delta}{4} \|e^{tA_0} \mathcal{I}m\phi_\lambda\|_{L^2} \end{aligned} \quad (3.47)$$

by choosing θ small. If $T < T_\delta$, we choose $t = T$ above to obtain

$$\|\tilde{\omega}(T) - \delta e^{TA_0} \mathcal{I}m\phi_\lambda\|_{L^2} \leq \frac{\delta}{4} \|e^{TA_0} \mathcal{I}m\phi_\lambda\|_{L^2}, \quad (3.48)$$

which contradicts the definition of T . Therefore $T_\delta \leq T$ and we can put $t = T_\delta$ to obtain

$$\|\tilde{\omega}(T_\delta)\|_{L^2} \geq \frac{3}{4} \|e^{T_\delta A_0} \delta \mathcal{I}m\phi_\lambda\|_{L^2} \geq \frac{3}{4} c_0 \theta \equiv \epsilon_0 > 0. \quad (3.49)$$

§4. Conclusion and Acknowledgments

It has become increasingly apparent that the notions of stability and instability are dependent on the norms. Therefore our purpose has been to analyze both of them in the same norm.

The analysis of the stability could be also done for other types of minimizers. Consider a convex function $\omega \rightarrow G(\omega)$ and the minimizer of the functional

$$\int_\omega G(\omega) dx$$

under the constraints of mass equal to 1 and energy given. Formally such a minimizer is a solution of a generalized mean field equation

$$-\Delta \Psi = C (G')^{-1}(-\beta \Psi), \quad (4.1)$$

which ought to be studied in the same way as the standard mean field equation. In this direction some related stability results can be found in [19].

However it is the mean field equation itself or its generalization given in [16, 14, 12] which seems really pertinent for the description of coherent

structures. While the result of Section 2 may explain why these structures persist, it does not explain why they appear in the first place. In fact, these stability results are time-reversible, so that in order to come close to a solution of the mean field equation one has to start close to this solution. The stationary solutions do not behave like attractors but like centers of a dynamical system. However, justification of their frequent appearance could be found in one of the following possibilities.

- These solutions should be the most probable ones in terms of a convenient probability measure to be defined on the configuration space.

- They might be produced by the conjunction of several circumstances, such as the fact that the initial vorticity is bounded in L^∞ and a family of solutions converges merely weakly to a nontrivial Young measure, the fact that in mean time the solutions are limits of solutions of the Navier-Stokes equations with viscosity tending to zero, and the fact that only the ω limit set for $t \rightarrow \infty$ is important.

This contribution is part of a research program which was initiated during the visit of C. Bardos to Brown University in the spring of 2000 and he wishes to thank Brown University for its support and warm hospitality during this period.

Any specialist of the field reading this paper will guess that we had the chance to benefit from some helpful discussions with several colleagues, including S. Friedlander, G. Haller, M. Kiessling, M. Vishik and G. Wolansky and we would like to thank all of them.

And finally we would like to thank Professor Li Ta-tsien who gave us the opportunity to publish this article in this special issue dedicated to the memory of Jacques-Louis Lions. He played a fundamental role in the scientific training of two of us, having been the thesis advisor of the first author and one of the first collaborators of the third author during his postdoctoral year in Paris.

REFERENCES

- [1] Arnold, V. I. & Khesin, B. A., Topological methods in hydrodynamics, Applied Mathematical Sciences, **125**, Springer, 1997.
- [2] Bardos, C., Golse, F. & Levermore, D., Fluid dynamics limits of kinetic equations II: convergence proofs for the Boltzmann equation, *Comm. on Pure and Appl. Math.*, **46**(1993), 667–753.

- [3] Caglioti, E., Lions, P. L., Marchioro, C. & Pulvirenti, M., A special class of stationary flows for two dimensional Euler equation, a statistical mechanics description, Part I, *Comm. Math. Phys.*, **143**(1992), 501–525.
- [4] Caglioti, E., Lions, P. L., Marchioro, C. & Pulvirenti, M., A special class of stationary flows for two dimensional Euler equation, a statistical mechanics description, Part II, *Comm. Math. Phys.*, **174**(1995), 229–260.
- [5] Friedlander, S. & Vishik, M., Dynamo theory methods for hydrodynamic stability, *J. Math. Pures Appl.*, **72**(1993), 145–180.
- [6] Friedlander, S., Strauss, M. & Vishik, M., Nonlinear instability in an ideal fluid, *Ann. I. H. P. (Anal Nonlin.)*, **14**:2(1997), 187–209.
- [7] Friedlander, S., Strauss, W. & Vishik, M., Robustness of instability for the two-dimensional Euler equation, *SIAM J. Math. Anal.*, **30**(1999), 1343–1354.
- [8] Grenier, E., On the nonlinear instability of Euler and Prandtl equations, *Comm. Pure Appl. Math.*, **53**(2000), 1067–1091.
- [9] Guo, Y. & Rein, G., Isotropic steady states in galactic dynamics, *Comm. Math. Phys.*, **219**(2001), 607–629.
- [10] Guo, Y. & Strauss, W., Instability of periodic BGK equilibria, *Comm. Pure Appl. Math.*, **48**(1995), 861–894.
- [11] Kiessling, M., Statistical mechanics of classical particles with logarithmic interaction, *Comm. Pure Appl. Math.*, **47**(1993), 27–56.
- [12] Lynden-Bell, D., Statistical mechanics of violent relaxation in stellar systems, *Mont. Not. R. Astr. Soc.*, **136**(1967), 101–121.
- [13] Marchioro, C. & Pulvirenti, M., Mathematical theory of incompressible nonviscous fluids, Applied Mathematical Sciences Vol 96, Springer Verlag, New York, Berlin, Heidelberg, 1998.
- [14] Miller, J., Peter B. Weichman & Cross, C., Statistical mechanics, Euler’s equation and Jupiter’s red spot, *Phys. Rev. A*, **45**:4(1992), 2328–2358.
- [15] Onsager, L., Statistical hydrodynamics, *Supplemento al Nuovo Cimento*, **6**(1949), 279–287.
- [16] Robert, R. & Sommeria, J., Statistical equilibrium states for two dimensional flows, *J. Fluid Mech.* **229**(1991), 291–310.
- [17] Turkington, B., Statistical equilibrium measures and coherent states in two-dimensional turbulence, *Comm. Pure Appl. Math.*, **52**(1999), 781–809.
- [18] Vishik, M., Spectrum of small oscillations of an ideal fluid and Lyapunov exponents, *J. Math. Pures Appl.*, **75**(1996), 531–557.
- [19] Wolansky, G. & Ghil, M., Nonlinear stability for saddle solutions of ideal flows and symmetry breaking, *Comm. Math. Phys.*, **192**(1998), 713–736.
- [20] Wolibner, W., Un théorème sur l’existence du mouvement plan d’un fluide parfait, homogène et incompressible, pendant un temps infiniment long, *Mat. Z.*, **37**(1933), 698–726.
- [21] Yudovich, V. I., Nonstationary flow of an ideal incompressible liquid, *Zh. Vych. Mat.*, **3**(1963), 1032–1066.

REGULARITY THEORY FOR SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS WITH NEUMANN BOUNDARY CONDITIONS

A. BENSOUSSAN

*University Paris Dauphine and CNES (Centre National d'Etudes
Spatiales), 2, Place Maurice Quentin, 75001, Paris, France*

J. FREHSE

Institut für Angewandte Mathematik, Universität Bonn, Bonn, Germany

Abstract

The objective of this paper is to consider the theory of regularity of systems of partial differential equations with Neumann boundary conditions. It complements previous works of the authors for the Dirichlet case. This type of problem is motivated by stochastic differential games. The Neumann case corresponds to stochastic differential equations with reflection on boundary of the domain.

Keywords Regularity theory, Neumann boundary conditions, Dirichlet problem

2000 MR Subject Classification 35J55, 35B65

§1. Introduction

We consider here a system of nonlinear P.D.E. with Neumann boundary conditions. It is the counterpart of the Dirichlet problem considered by the authors in several publications (see in particular the book [1]). Although a natural counterpart of the Dirichlet case, the results of this paper are presented here in a simplified and self-contained manner.

§2. Setting of the Problem

2.1. Notation and Assumptions

We consider a smooth, bounded domain Ω of R^n . In particular, the boundary will be representable by local charts, so that in a sense the Neumann problem can be reduced to the Dirichlet problem. We consider the operator

$$A = -\operatorname{div} a D, \quad (2.1)$$

where $a(x) \equiv a_{ij}(x)$ satisfies

$$\begin{aligned} a_{ij}(x) &= a_{ji}(x) \quad \text{is Lipschitz continuous on } \bar{\Omega}, \\ a_{ji}(x)\xi_i\xi_j &\geq \alpha_0|\xi|^2, \quad \forall \xi \in R^n, \alpha_0 > 0. \end{aligned} \quad (2.2)$$

We now consider "Hamiltonians", namely functions $H_\nu(x, p)$, $\nu = 1, \dots, N$, $p = p_1, \dots, p_N$, with $p_\nu \in R^n$, with the following properties

$$H_\nu(x, p) \quad \text{are Caratheodory functions,} \quad (2.3)$$

$$\sum_\nu H_\nu(x, p) \geq -\lambda - \lambda^0 \left| \sum_\nu p_\nu \right|^2, \quad (2.4)$$

$$H_\nu(x, p) \leq \lambda_\nu + \lambda_\nu^0 |p_\nu|^2, \quad (2.5)$$

$$H_\nu(x, p) = Q(x, p)p_\nu + H_\nu^0(x, p), \quad (2.6)$$

with

$$|Q(x, p)| \leq k + K|p|, \quad (2.7)$$

$$|H_\nu^0(x, p)| \leq k_\nu + K_\nu \sum_{\mu \leq \nu} |p_\mu|^2. \quad (2.8)$$

Let α be a positive constant, we are interested in the system

$$\begin{aligned} Au_\nu + \alpha u_\nu &= H_\nu(x, Du), \\ \frac{\partial u_\nu}{\partial n_A} \Big|_{\partial\Omega} &= 0, \end{aligned} \quad (2.9)$$

which is written in the variational form

$$\int_\Omega a D u_\nu D v dx + \alpha \int_\Omega u_\nu v dx = \int_\Omega H_\nu(x, Du) v dx, \quad \forall v \in H^1(\Omega) \cap L^\infty(\Omega). \quad (2.10)$$

2.2. Statement of the Main Result

Theorem 2.1. *We assume (2.2) to (2.8). Then there exists a solution of (2.10) which belongs to $(W^{2,s}(\Omega))^N$, $\forall 2 \leq s < \infty$.*

Remark 2.1. The assumptions are subject to some flexibility, in the sense that we can combine in a linear manner the equations, to achieve the desired structure (see [1] for details).

Remark 2.2. The assumptions (2.6), (2.7) and (2.8) are restrictive only for $\nu = 1, \dots, N - 1$. A general quadratic growth can be assumed by $H_N(x, p)$, and it is always possible to define $H_N^0(x, p)$ by the relation

$$H_N^0(x, p) = H_N(x, p) - Q(x, p)p_N$$

and (2.8) will be verified.

§3. The $H^1(\Omega) \cap L^\infty(\Omega)$ Theory

3.1. Approximation

The problem (2.10) makes sense for solutions in $H^1(\Omega) \cap L^\infty(\Omega)$. This is what we are looking for to begin with. We consider as an approximation the problem with Hamiltonians

$$H_\nu^\varepsilon(x, p) = \frac{H_\nu(x, p)}{1 + \varepsilon|H(x, p)|}, \quad (3.1)$$

where $H(x, p)$ stands for the vector H_1, \dots, H_N . Note that all the assumptions (2.3)–(2.8) are satisfied for the hamiltonian $H_\nu^\varepsilon(x, p)$ with the same constants, provided we define

$$Q^\varepsilon(x, p) = \frac{Q(x, p)}{1 + \varepsilon|H(x, p)|}, \quad (3.2)$$

$$H_\nu^{0, \varepsilon}(x, p) = \frac{H_\nu^0(x, p)}{1 + \varepsilon|H(x, p)|}. \quad (3.3)$$

Moreover one has

$$|H^\varepsilon(x, p)| \leq \frac{1}{\varepsilon}, \quad \forall x, p. \quad (3.4)$$

Therefore, by application of Schauder's fixed point theorem, one can obtain the existence of a solution $u^\varepsilon = (u_\nu^\varepsilon)$ in $(W^{2,s}(\Omega))^N$ of the problem

$$\int_\Omega a D u_\nu^\varepsilon D v dx + \alpha \int_\Omega u_\nu^\varepsilon v dx = \int_\Omega H_\nu^\varepsilon(x, D u^\varepsilon) v dx, \quad \forall v \in H^1(\Omega) \cap L^\infty(\Omega). \quad (3.5)$$

3.2. L^∞ Estimate

Since the estimate are exactly the same for (3.5) and (2.10), we shall to simplify the notation consider only (2.10), with a priori estimates, assuming

a solution in $(W^{2,s}(\Omega))$ exists. Let us write

$$\tilde{u} = \sum_{\nu} u_{\nu},$$

and summing up (2.10), we get

$$\int_{\Omega} aD\tilde{u}_{\nu}Dvdx + \alpha \int_{\Omega} \tilde{u}_{\nu}vdx = \int_{\Omega} \sum_{\nu} H_{\nu}(x, Du)vdx. \quad (3.6)$$

Consider the function

$$E = \exp \gamma(\tilde{u} + L)^-, \quad (3.7)$$

where $\gamma, L > 0$ will be defined later. We test (3.6) with $v = 1 - E$, which is a negative function in $H^1 \cap L^{\infty}$. We get

$$\gamma \int_{\Omega} aD\tilde{u}.D\tilde{u}E\mathbf{1}_{\{\tilde{u}+L<0\}}dx + \alpha \int_{\Omega} \tilde{u}(1-E)dx = \int_{\Omega} \sum_{\nu} H_{\nu}(x, u)(1-E)dx$$

and using the assumption (2.4), together with the fact that $1 - E < 0$, yields

$$\leq \int_{\Omega} (-\lambda - \lambda^0|D\tilde{u}|^2)(1-E)dx.$$

Therefore, also using (2.2) we get

$$\begin{aligned} & \alpha_0\gamma \int_{\Omega} |D\tilde{u}|^2E\mathbf{1}_{\{\tilde{u}+L<0\}}dx + \alpha \int_{\Omega} (\tilde{u} + L)(1-E)dx \\ & \leq \int_{\Omega} (\alpha L - \lambda)(1-E)dx + \lambda^0 \int_{\Omega} |D\tilde{u}|^2E\mathbf{1}_{\{\tilde{u}+L<0\}}dx. \end{aligned}$$

Choosing

$$\gamma = \frac{\lambda^0}{\alpha_0}, \quad L = \frac{\lambda}{\alpha}$$

yields

$$\int_{\Omega} (\tilde{u} + L)(1-E)dx \leq 0.$$

Since clearly $(\tilde{u} + L)(1 - E) \geq 0$, we deduce

$$(\tilde{u} + L)(1 - E) = 0 \quad \text{a.e.}, \text{ hence } \tilde{u} + L \geq 0 \quad \text{a.e.}$$

Therefore we have shown

$$\tilde{u} \geq -\frac{\lambda}{\alpha} \quad \text{a.e.} \quad (3.8)$$

We next consider in (2.10) the test function $E_{\nu} - 1$, with

$$E_{\nu} = \exp \gamma_{\nu}(u_{\nu} - l_{\nu})^+$$

for convenient positive constants γ_ν, l_ν . Note that $E_\nu - 1 \geq 0$. Making use of the assumption (2.5), it follows that

$$\begin{aligned} & \gamma_\nu \int_{\Omega} a D u_\nu E_\nu \mathbf{1}_{\{u_\nu > l_\nu\}} dx + \alpha \int_{\Omega} u_\nu (E_\nu - 1) dx \\ & \leq \int_{\Omega} \lambda_\nu (E_\nu - 1) dx + \lambda_\nu^0 \int_{\Omega} |D u_\nu|^2 E_\nu \mathbf{1}_{\{u_\nu > l_\nu\}} dx. \end{aligned}$$

Thus taking

$$\gamma_\nu = \frac{\lambda_\nu^0}{\alpha_0}, \quad l_\nu = \frac{\lambda_\nu}{\alpha},$$

we deduce again

$$(u_\nu - l_\nu)(E_\nu - 1) = 0 \quad \text{a.c.}$$

which implies

$$u_\nu \leq \frac{\lambda_\nu}{\alpha}. \quad (3.9)$$

Continuing (3.8) and (3.9) we obtain

$$u_\nu \geq \frac{-\left(\lambda + \sum_{\mu \neq \nu} \lambda_\mu\right)}{\alpha}. \quad (3.10)$$

So we have proven

$$\|u_\nu\|_\infty \leq \frac{\zeta_\nu}{\alpha} = \frac{1}{\alpha} \text{Max}\left(\lambda_\nu, \lambda + \sum_{\mu \neq \nu} \lambda_\mu\right). \quad (3.11)$$

Remark 3.1. The estimate (3.11) is the same as for the solution of the Dirichlet problem. This is due to the presence of the zero order term αu_ν , and to the fact that for the Dirichlet as well as for the Neumann problem, the maximum and the minimum of the function u_ν do not take place at the boundary of the domain.

3.3. H^1 Estimate

The H^1 estimate is done by using the special structure (2.6), (2.7), (2.8). Recall that

$$|u_\nu(x)| \leq \frac{\zeta_\nu}{\alpha}. \quad (3.12)$$

We introduce the function

$$\beta(s) = e^s - s - 1, \quad F = \prod_{\nu=1}^N \exp \beta(\gamma_\nu u_\nu),$$

where γ_ν is a positive constant to be defined later. We test (2.10) with $v = F\gamma_\nu\beta'(\gamma_\nu u_\nu) \in H^1 \cap L^\infty$. So we get (as in the Dirichlet case)

$$\begin{aligned} & \sum_\nu \int_\Omega \gamma_\nu^2 a Du_\nu \cdot Du_\nu e^{\gamma_\nu u_\nu} F dx + \int_\Omega a \frac{DF \cdot DF}{F} dx \\ &= \int_\Omega Q DF dx + \int_\Omega \sum_\nu \gamma_\nu (H_\nu^0(Du) - \alpha u_\nu) F (e^{\gamma_\nu u_\nu} - 1) dx, \end{aligned}$$

hence also

$$\begin{aligned} & \sum_\nu \int_\Omega \gamma_\nu^2 a Du_\nu \cdot Du_\nu e^{\gamma_\nu u_\nu} F dx \\ & \leq \int_\Omega F \frac{a^{-1} Q \cdot Q}{4} dx + \int_\Omega \sum_\nu \gamma_\nu (H_\nu^0(Du) - \alpha u_\nu) F (e^{\gamma_\nu u_\nu} - 1) dx. \end{aligned} \quad (3.13)$$

Introduce next the function

$$X = \prod_{\nu=1}^N (\exp \beta(\gamma_\nu u_\nu) + \exp \beta(-\gamma_\nu u_\nu)) \quad (3.14)$$

and the related quantities

$$\begin{aligned} X_\nu &= X \frac{e^{\gamma_\nu u_\nu} \exp \beta(\gamma_\nu u_\nu) + e^{-\gamma_\nu u_\nu} \exp \beta(-\gamma_\nu u_\nu)}{\exp \beta(\gamma_\nu u_\nu) + \exp \beta(-\gamma_\nu u_\nu)}, \\ \tilde{X}_\nu &= X \frac{(e^{\gamma_\nu u_\nu} - 1) \exp \beta(\gamma_\nu u_\nu) - (e^{-\gamma_\nu u_\nu} - 1) \exp \beta(-\gamma_\nu u_\nu)}{\exp \beta(\gamma_\nu u_\nu) + \exp \beta(-\gamma_\nu u_\nu)} \end{aligned} \quad (3.15)$$

and we have the inequalities

$$2^N \leq X \leq X_\nu \leq X e^{\frac{\gamma_\nu \zeta_\nu}{\alpha}}, \quad |\tilde{X}_\nu| \leq X_\nu.$$

Applying the relation (3.13) with γ_ν changed into $-\gamma_\nu$, and summing up the 2^N relations obtained in this way, we get the inequality

$$\begin{aligned} & \sum_\nu \int_\Omega \gamma_\nu^2 a Du_\nu \cdot Du_\nu X_\nu dx \\ & \leq \int_\Omega X \frac{a^{-1} Q \cdot Q}{4} dx + \int_\Omega \sum_\nu \gamma_\nu (H_\nu^0(Du) - \alpha u_\nu) \tilde{X}_\nu dx \end{aligned} \quad (3.16)$$

and using the assumption (2.8) yields

$$\begin{aligned} & \leq \int_\Omega X \frac{a^{-1} Q \cdot Q}{4} dx + \sum_\nu \int_\Omega \gamma_\nu (k_\nu + \zeta_\nu + K_\nu |Du_\nu|^2) X_\nu dx \\ & \quad + \sum_\nu \int_\Omega |Du_\nu|^2 \sum_{\mu > \nu} \gamma_\mu K_\mu X_\mu dx. \end{aligned}$$

Use (2.7) to note that

$$\frac{a^{-1}Q \cdot Q}{4} \leq \frac{1}{2\alpha_0}(k^2 + K^2|Du|^2),$$

hence it follows from (3.16) also

$$\begin{aligned} & \alpha_0 \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 |Du_{\nu}|^2 X_{\nu} dx \\ & \leq \int_{\Omega} \left(\frac{1}{2\alpha_0} k^2 X + \sum_{\nu} \gamma_{\nu} (k_{\nu} + \zeta_{\nu}) X_{\nu} \right) dx \\ & \quad + \int_{\Omega} \sum_{\nu} |Du_{\nu}|^2 \left(\frac{K^2}{2\alpha_0} X_{\nu} + \gamma_{\nu} K_{\nu} X_{\nu} + \sum_{\mu > \nu} \gamma_{\mu} K_{\mu} X_{\mu} \right) dx, \end{aligned}$$

but also

$$X_{\mu} \leq X_{\nu} e^{\frac{\gamma_{\mu} \zeta_{\mu}}{\alpha}}, \quad \forall \mu, \nu,$$

hence

$$\begin{aligned} & \sum_{\nu} \int_{\Omega} |Du_{\nu}|^2 X_{\nu} \left(\alpha_0 \gamma_{\nu}^2 - \frac{K^2}{2\alpha_0} - \gamma_{\nu} K_{\nu} - \sum_{\mu > \nu} \gamma_{\mu} K_{\mu} e^{\frac{\gamma_{\mu} \zeta_{\mu}}{\alpha}} \right) dx \\ & \leq \int_{\Omega} \left(\frac{1}{2\alpha_0} k^2 X + \sum_{\nu} \gamma_{\nu} (k_{\nu} + \zeta_{\nu}) X_{\nu} \right) dx. \end{aligned} \quad (3.17)$$

Choosing the constants γ_{ν} so that

$$\alpha_0 \gamma_{\nu}^2 - \frac{K^2}{2\alpha_0} - \gamma_{\nu} K_{\nu} - \sum_{\mu > \nu} \gamma_{\mu} K_{\mu} e^{\frac{\gamma_{\mu} \zeta_{\mu}}{\alpha}} > 0, \quad \forall \nu \quad (3.18)$$

and recalling that $X_{\nu} \geq 2^N$, we obtain

$$\int_{\Omega} |Du|^2 dx \leq K_0. \quad (3.19)$$

3.4. Convergence

Since the preceding estimates hold for u^{ε} , we have also

$$\|u^{\varepsilon}\|_{\infty} \leq \zeta, \quad (3.20)$$

$$\|u^{\varepsilon}\|_{(H^1(\Omega))^N} \leq K_1. \quad (3.21)$$

So we can extract a subsequence, still denoted by u^{ε} , such that

$$u^{\varepsilon} \rightarrow u_{\nu} \text{ in } H^1(\Omega) \text{ weakly and } L^{\infty}(\Omega) \text{ weak star, also a.e.} \quad (3.22)$$

We want to prove that

$$u_{\nu}^{\varepsilon} \rightarrow u_{\nu} \text{ in } H^1(\Omega) \text{ strongly.} \quad (3.23)$$

We perform a calculation close to that of Section 3.3, with this time

$$F = \prod_{\nu=1}^N \exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})),$$

and test (3.5) with $v = F\gamma_{\nu}\beta'(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu}))$. We obtain

$$\begin{aligned} & \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D(u_{\nu}^{\varepsilon} - u_{\nu}) \cdot D(u_{\nu}^{\varepsilon} - u_{\nu}) e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} F dx + \int_{\Omega} a \frac{DF \cdot DF}{F} dx \\ &= \int_{\Omega} \left(Q^{\varepsilon} - \gamma_{\nu} a D u_{\nu} (e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} - 1) \right) DF dx \\ & \quad + \int_{\Omega} F (H_{\nu}^{0,\varepsilon} - c u_{\nu}^{\varepsilon} + Q^{\varepsilon} D u_{\nu}) \gamma_{\nu} (e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} - 1) dx \\ & \quad - \int_{\Omega} F a D u_{\nu} \cdot D(u_{\nu}^{\varepsilon} - u_{\nu}) \gamma_{\nu}^2 e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} dx \\ &= \text{I} + \text{II} + \text{III}. \end{aligned} \tag{3.24}$$

Introduce the function

$$h(s) = \sum_{\nu} \gamma_{\nu}^2 (e^{|\gamma_{\nu} s_{\nu}|} - 1)^2, \tag{3.25}$$

where s represents the vector (s_1, \dots, s_N) . Note

$$|a(x)| \leq M, \quad |Q^{\varepsilon}| \leq k + K|Du| + K|Du^{\varepsilon} - Du|.$$

Moreover

$$DF = F \sum_{\nu=1}^N \gamma_{\nu} (e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} - 1) D(u_{\nu}^{\varepsilon} - u_{\nu}),$$

hence

$$|DF| \leq F |D(u^{\varepsilon} - u)| h^{1/2}(u^{\varepsilon} - u). \tag{3.26}$$

Therefore considering the integrals on the right hand side of (3.24), we have

$$\begin{aligned} |\text{II}| &\leq \int_{\Omega} (K|Du^{\varepsilon} - Du| + M|Du| h^{1/2}(u^{\varepsilon} - u)) |DF| dx \\ &\quad + \int_{\Omega} F (k + K|Du|) |D(u^{\varepsilon} - u)| h^{1/2}(u^{\varepsilon} - u) dx, \end{aligned} \tag{3.27}$$

so

$$\begin{aligned} \int_{\Omega} a \frac{DF \cdot DF}{F} dx &\geq \text{I} - \int_{\Omega} F \left[\left(\frac{K^2}{2\alpha_0} + \frac{1}{2} \right) |D(u^{\varepsilon} - u)|^2 \right. \\ &\quad \left. + h(u^{\varepsilon} - u) \left(k^2 + K^2 + \frac{M^2}{2\alpha_0} \right) |Du|^2 \right] dx, \end{aligned} \tag{3.28}$$

$$\begin{aligned}
& \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D(u_{\nu}^{\varepsilon} - u_{\nu}) e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} F dx \\
& \leq \int_{\Omega} F \left[\left(\frac{K^2}{2\alpha_0} + \frac{1}{2} \right) |D(u^{\varepsilon} - u)|^2 \right. \\
& \quad \left. + h(u^{\varepsilon} - u) \left(k^2 + \left(K^2 + \frac{M^2}{2\alpha_0} \right) |Du|^2 \right) \right] dx + \text{II} + \text{III}. \tag{3.29}
\end{aligned}$$

Introducing the quantities analogous to (3.14), (3.15), namely

$$\begin{aligned}
X^{\varepsilon} &= \prod_{\nu=1}^N (\exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu}))), \\
X_{\nu}^{\varepsilon} &= X \{ e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} \exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \\
& \quad + e^{-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} \exp \beta(-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \} \\
& \quad / \{ \exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \}, \\
\tilde{X}_{\nu}^{\varepsilon} &= X \{ (e^{\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} - 1) \exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \\
& \quad - (e^{-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})} - 1) \exp \beta(-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \} \\
& \quad / \{ \exp \beta(\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})) \}, \tag{3.31}
\end{aligned}$$

we have of course

$$2^N \leq X^{\varepsilon} \leq X_{\nu}^{\varepsilon} \leq X^{\varepsilon} e^{2\frac{\gamma_{\nu} \xi_{\nu}}{\alpha_0}}, \quad |\tilde{X}_{\nu}^{\varepsilon}| \leq X_{\nu} \tag{3.32}$$

and also

$$|\tilde{X}_{\nu}^{\varepsilon}| \leq (e^{|\gamma_{\nu}(u_{\nu}^{\varepsilon} - u_{\nu})|} - 1) X^{\varepsilon}. \tag{3.33}$$

From (3.30), we deduce, writing the 2^N inequalities corresponding to all changes of γ_{ν} into $-\gamma_{\nu}$,

$$\begin{aligned}
& \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D(u_{\nu}^{\varepsilon} - u_{\nu}) D(u_{\nu}^{\varepsilon} - u_{\nu}) X_{\nu}^{\varepsilon} dx \\
& \leq \int_{\Omega} X^{\varepsilon} \left[\left(\frac{K^2}{2\alpha_0} + \frac{1}{2} \right) |D(u^{\varepsilon} - u)|^2 \right. \\
& \quad \left. + h(u^{\varepsilon} - u) \left(k^2 + \left(K^2 + \frac{M^2}{2\alpha_0} \right) |Du|^2 \right) \right] dx \\
& \quad + \sum_{\nu} \int_{\Omega} (H_{\nu}^{0,\varepsilon} - cu_{\nu}^{\varepsilon} + Q^{\varepsilon} Du_{\nu}) \gamma_{\nu} \tilde{X}_{\nu}^{\varepsilon} dx \\
& \quad - \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 X_{\nu}^{\varepsilon} D(u_{\nu}^{\varepsilon} - u_{\nu}) Du_{\nu} dx = \text{I}' + \text{II}' + \text{III}'. \tag{3.34}
\end{aligned}$$

We first check that

$$|H_\nu^{0,\varepsilon} - cu_\nu^\varepsilon| \leq k_\nu + \zeta_\nu + 2K_\nu |Du| + 2K_\nu \sum_{\mu \leq \nu} |D(u_\mu^\varepsilon - u_\mu)|^2,$$

hence using (3.33) and (3.34) we have

$$\begin{aligned} & \left| \sum_\nu (H_\nu^{0,\varepsilon} - cu_\nu^\varepsilon) \gamma_\nu \tilde{X}_\nu^\varepsilon \right| \\ & \leq X^\varepsilon \gamma_\nu (k_\nu + \zeta_\nu + 2K_\nu |Du|) (e^{|\gamma_\nu(u_\nu^\varepsilon - u_\nu)|} - 1) \\ & \quad + 2K_\nu \gamma_\nu X_\nu^\varepsilon \sum_{\mu \leq \nu} |D(u_\mu^\varepsilon - u_\mu)|^2 \\ & \leq X^\varepsilon h^{1/2}(u^\varepsilon - u) \left(\left(\sum_\nu (k_\nu + \zeta_\nu)^2 \right)^{1/2} + 2 \left(\sum_\nu K_\nu^2 \right)^{1/2} |Du| \right) \\ & \quad + 2 \sum_\nu k_\nu \gamma_\nu X_\nu^\varepsilon |D(u_\nu^\varepsilon - u_\nu)|^2 + 2 \sum_\nu |D(u_\nu^\varepsilon - u_\nu)|^2 \sum_{\mu > \nu} \gamma_\mu K_\mu X_\mu^\varepsilon \end{aligned}$$

and

$$\left| \sum_\nu Q^\varepsilon Du_\nu \gamma_\nu \tilde{X}_\nu^\varepsilon \right| \leq X^\varepsilon (k + K |Du^\varepsilon|) |Du| h^{1/2}(u^\varepsilon - u).$$

Collecting results, we deduce from (3.36)

$$\begin{aligned} & \sum_\nu \int_\Omega |D(u_\nu^\varepsilon - u_\nu)|^2 \left(\alpha_0 \gamma_\nu^2 - \frac{K^2}{2\alpha_0} - \frac{1}{2} - 2K_\nu \gamma_\nu \right. \\ & \quad \left. - 2 \sum_{\mu > \nu} \gamma_\mu K_\mu e^{\frac{2\zeta_\mu \gamma_\mu}{\alpha}} \right) X_\nu^\varepsilon dx \\ & \leq \int_\Omega X^\varepsilon h(u^\varepsilon - u) \left(k^2 + \left(K^2 + \frac{M}{2\alpha_0} \right) |Du|^2 \right) dx \\ & \quad + \int_\Omega X^\varepsilon h^{1/2}(u^\varepsilon - u) \left(\left(\sum_\nu (k_\nu + \zeta_\nu)^2 \right)^{1/2} + 2 \left(\sum_\nu K_\nu^2 \right)^{1/2} |Du| \right) dx \\ & \quad + \int_\Omega X^\varepsilon h^{1/2}(u^\varepsilon - u) |Du| (k + K |Du^\varepsilon|) dx \\ & \quad - \sum_\nu \int_\Omega \gamma_\nu^2 X_\nu^2 D(u_\nu^\varepsilon - u_\nu) Du_\nu dx \end{aligned} \tag{3.35}$$

as $\varepsilon \rightarrow 0$, $h(u^\varepsilon - u) \rightarrow 0$ pointwise, $X^\varepsilon \rightarrow 2^N$, $X_\nu^\varepsilon \rightarrow 2^N$. In view of the weak convergence in L^2 of Du_ν^ε to Du_ν , and the L^∞ bounds, it is easy to convince oneself that the right hand side of (3.37) tends to 0. Choosing the γ_ν so that

$$\alpha_0 \gamma_\nu^2 - \frac{K^2}{2\alpha_0} - \frac{1}{2} - 2K_\nu \gamma_\nu - 2 \sum_{\mu > \nu} \gamma_\mu K_\mu e^{\frac{2\zeta_\mu \gamma_\mu}{\alpha}} > 0$$

and recalling the $X_\nu^\varepsilon \geq 2^N$, we deduce immediately from (3.37) that

$$\int_{\Omega} |D(u^\varepsilon - u)|^2 dx \rightarrow 0.$$

Hence we have proven (3.23). We may assume that

$$Du^\varepsilon \rightarrow Du \quad \text{a.e.},$$

thus

$$H_\nu^\varepsilon(x, Du^\varepsilon) \rightarrow H_\nu(x, Du) \quad \text{a.e.}$$

Moreover since

$$|H^\varepsilon(x, Du^\varepsilon)| \leq C(1 + |Du^\varepsilon|^2),$$

the convergence of Du^ε in L^2 implies the equi-integrability of $H_\nu^\varepsilon(x, Du^\varepsilon)$.

Hence from Vitali's theorem

$$H_\nu^\varepsilon(x, Du^\varepsilon) \rightarrow H_\nu(x, Du) \quad \text{in } L^1.$$

Therefore we can pass to the limit in (3.5), showing that u is an $H^1(\Omega) \cap L^\infty(\Omega)$ solution of (2.10).

§4. Regularity

4.1. An Inequality

We associate to u_ν a constant c_ν which is arbitrary, provided that

$$|c_\nu| \leq \frac{\zeta_\nu}{\alpha}, \tag{4.1}$$

and consider as in Section 3.3, the function

$$F = \prod_{\nu=1}^N \exp \beta(\gamma_\nu(u_\nu - c_\nu)).$$

Let ψ be in $C^1(\bar{\Omega})$, $\psi \geq 0$, and test (2.10) with

$$v = F \gamma_\nu \beta'(\gamma_\nu(u_\nu - c_\nu)) \psi.$$

We obtain

$$\begin{aligned} & \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D u_{\nu} D u_{\nu} e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} F \psi dx + \int_{\Omega} a \frac{DF \cdot DF}{F} \psi dx + \int_{\Omega} a DF \cdot D \psi dx \\ &= \int_{\Omega} Q \cdot DF \psi dx + \sum_{\nu} \int_{\Omega} \gamma_{\nu} (H_{\nu}^0(Du) - \alpha u_{\nu}) F (e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} - 1) \psi dx, \end{aligned}$$

hence instead of (3.13)

$$\begin{aligned} & \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D u_{\nu} D u_{\nu} e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} F \psi dx + \int_{\Omega} a DF \cdot D \psi dx \\ & \leq \int_{\Omega} F \frac{a^{-1} Q \cdot Q}{4} \psi dx + \sum_{\nu} \int_{\Omega} \gamma_{\nu} (H_{\nu}^0(Du) - \alpha u_{\nu}) F (e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} - 1) \psi dx. \end{aligned} \quad (4.2)$$

Introduce the function

$$X = \prod_{\nu=1}^N (\exp \beta(\gamma_{\nu}(u_{\nu} - c_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu} - c_{\nu}))) \quad (4.3)$$

and the related quantities

$$\begin{aligned} X_{\nu} &= X \{ e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} \exp \beta(\gamma_{\nu}(u_{\nu} - c_{\nu})) \\ & \quad + e^{-\gamma_{\nu}(u_{\nu}-c_{\nu})} \exp \beta(-\gamma_{\nu}(u_{\nu} - c_{\nu})) \} \\ & \quad / \{ \exp \beta(\gamma_{\nu}(u_{\nu} - c_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu} - c_{\nu})) \}, \\ \tilde{X}_{\nu} &= X \{ (e^{\gamma_{\nu}(u_{\nu}-c_{\nu})} - 1) \exp \beta(\gamma_{\nu}(u_{\nu} - c_{\nu})) \\ & \quad - (e^{-\gamma_{\nu}(u_{\nu}-c_{\nu})} - 1) \exp \beta(-\gamma_{\nu}(u_{\nu} - c_{\nu})) \} \\ & \quad / \{ \exp \beta(\gamma_{\nu}(u_{\nu} - c_{\nu})) + \exp \beta(-\gamma_{\nu}(u_{\nu} - c_{\nu})) \} \end{aligned} \quad (4.4)$$

again with 2^N relations with the possible choices of γ_{ν} and $-\gamma_{\nu}$, and adding up, we get

$$\begin{aligned} & \sum_{\nu} \int_{\Omega} \gamma_{\nu}^2 a D u_{\nu} D u_{\nu} X_{\nu} \psi dx + \int_{\Omega} a DX \cdot D \psi dx \\ & \leq \int_{\Omega} X \frac{a^{-1} Q \cdot Q}{4} \psi dx + \sum_{\nu} \int_{\Omega} \gamma_{\nu} (H_{\nu}^0(Du) - \alpha u_{\nu}) \tilde{X}_{\nu} \psi dx. \end{aligned}$$

Performing as in Section 3.3 for all terms where ψ arises, and picking the constants γ_{ν} as there, we obtain

$$k_0 \int_{\Omega} |Du|^2 \psi dx + \int_{\Omega} a DX \cdot D \psi dx \leq K_0 \int_{\Omega} \psi dx. \quad (4.5)$$

Note that

$$DX = \sum_{\nu} \gamma_{\nu} \tilde{X}_{\nu} Du_{\nu} \quad (4.6)$$

and thus

$$|DX| \leq C|u - c||Du|, \quad (4.7)$$

where c stands for the vector of constants (c_1, \dots, c_N) .

4.2. $W^{1,p}$ Regularity for $2 \leq p < p + \varepsilon$

Consider balls $B_R(x_0)$ of center x_0 and radius R , such that $|\Omega \cap B_R(x_0)| > 0$. By the smoothness of the boundary, we have (denote $\Gamma = \partial\Omega$)

$$|\Omega \cap B_R(x_0)| \geq c_0 R^n \quad |(R^n - \Omega) \cap B_R(x_0)| \geq c_1 R^n, \text{ if } x_0 \in \Gamma. \quad (4.8)$$

To the function u_{ν} and to a ball $B - R(x_0)$ we associate the constant

$$c_{\nu}^R = \begin{cases} \frac{1}{|B_{2R}|} \int_{B_{2R}(x_0)} u_{\nu} dx, & \text{if } B_{2R}(x_0) \subset \Omega, \\ \frac{1}{|B_{4R}(x'_0) \cap \Omega|} \int_{B_{4R}(x'_0) \cap \Omega} u_{\nu} dx, & \text{if } B_{2R}(x_0) \cap (R^n - \Omega) \neq \emptyset, \\ \text{where } x'_0 \in \Gamma \cap B_{2R}(x_0). \end{cases} \quad (4.9)$$

We shall use the Poincaré's inequality

$$\left(\int_{\Omega \cap B_{2R}(x_0)} |u_{\nu} - u_{\nu}^R|^{\lambda} dx \right)^{\frac{1}{\lambda}} \leq cR^{n \left(\frac{1}{\lambda} - \frac{1}{\mu} \right) + 1} \left(\int_{\Omega \cap B_{6R}(x_0)} |Du_{\nu}|^{\mu} dx \right)^{\frac{1}{\mu}} \quad (4.10)$$

with $\lambda, \mu \geq 1, n \left(\frac{1}{\lambda} - \frac{1}{\mu} \right) + 1 \geq 0$. Let τ be a cut off function

$$\tau = \begin{cases} 1 & \text{on } B_1(0), \\ 0 & \text{outside } B_2(0) \end{cases}$$

and $0 \leq \tau \leq 1, \tau \in C^{\infty}$. We denote

$$\tau_R(x) = \tau \left(\frac{x - x_0}{R} \right).$$

We take in (4.5) $\psi = \tau_R^2$, and $c_{\nu} = c_{\nu}^R$. So from (4.7) we have

$$|DX| \leq C|u - c^R||Du|$$

and thus we deduce from (4.7) the inequality

$$\int_{B_R(x_0) \cap \Omega} |Du|^2 dx \leq C \int_{B_{2R}(x_0) \cap \Omega} |Du| \frac{|u - c^R|}{R} dx + CR^n. \quad (4.11)$$

From Hölder's inequality and Poincaré's inequality, we have

$$\begin{aligned}
& \int_{B_{2R} \cap \Omega} |Du| \frac{|u - c^R|}{R} dx \\
& \leq \frac{C}{R} \left(\int_{B_{2R} \cap \Omega} |Du|^{\frac{2n}{n+1}} dx \right)^{\frac{n+1}{2n}} \left(\int_{B_{2R} \cap \Omega} |u - c^R|^{\frac{2n}{n-1}} dx \right)^{\frac{n-1}{2n}} \\
& \leq \frac{C}{R} \left(\int_{B_{6R} \cap \Omega} |Du|^{\frac{2n}{n+1}} dx \right)^{\frac{n+1}{2n}}.
\end{aligned}$$

So set $z = |Du|^{\frac{2n}{n+1}} \mathbf{1}_\Omega$, then we have the inequality

$$\int_{B_R} z^{\frac{n+1}{n}} dx \leq C \left(\int_{B_{6R}} z dx \right)^{\frac{n+1}{n}} + C, \quad (4.12)$$

where

$$\int_{B_R} = \frac{1}{|B_R|} \int_{B_R}.$$

This is the reverse Hölder's inequality, implying Gehring's result, namely $z^{\frac{n+1}{n}} + \varepsilon$ is integrable for some positive ε , hence $u \in W^{1,p}(\Omega)$, for $2 \leq p < p + \varepsilon'$.

4.3. Interior C^δ Regularity

Let Ω_1 be an open subdomain of Ω , with $\bar{\Omega}_1 \subset \Omega$. We want first to prove the following estimate

$$\int_{\Omega_1} |Du|^2 |x - x_0|^{2-n} dx \leq C_{\Omega_1}, \quad \forall x_0 \in \bar{\Omega}_1. \quad (4.13)$$

We consider the Green function $G = G^{x_0}$, solution of

$$\begin{aligned}
-\operatorname{div}(a(x)DG^{x_0}) &= \delta(x - x_0), \\
G^{x_0}|_{\partial\Omega} &= 0.
\end{aligned} \quad (4.14)$$

Note that

$$G^{x_0} \in L^q(\Omega) \cap W_0^{1,r}(\Omega), \quad 1 \leq q < \frac{n}{n-2}, \quad 1 \leq r < \frac{n}{n-1}. \quad (4.15)$$

Moreover

$$c_0 |x - x_0|^{2-n} \leq G^{x_0} \leq c_1 |x - x_0|^{2-n}, \quad \forall x \in \Omega_1, \quad (4.16)$$

where the constants c_0, c_1 depend only on Ω_1 . We just take in (4.5) $\psi =$

$G^{x_0^1}$, and from the definition of the Green function, one has

$$\int_{\Omega} aDX.DG^{x_0} dx = X(x_0) > 0,$$

hence

$$k_0 \int_{\Omega} |Du|^2 G^{x_0} dx \leq K_0 \int_{\Omega} G^{x_0} dx \leq C.$$

Using (4.16), the estimate (4.13) follows immediately. We want to prove C^δ regularity in Ω_1 . We shall use Morrey's result, namely

$$\sup_{\substack{x,y \in \Omega_1 \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|^\delta} \leq C \left(\sup_{\substack{x_0 \in \Omega_1 \\ B_R(x_0) \subset \Omega_1}} \frac{\int_{B_R(x_0)} |Du|^2 dx}{R^{n-2+2\delta}} \right)^{1/2}. \quad (4.17)$$

So we want to show that

$$\int_{B_R(x_0)} |Du|^2 dx \leq KR^{n-2+2\delta}, \quad \forall x_0 \in \Omega_1, B_R(x_0) \subset \Omega_1; \quad (4.18)$$

of course, the constant K will depend on Ω . To check (4.18), we shall prove the inequality

$$\int_{B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx \leq C \int_{B_{\sigma R}(x_0) - B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx + CR^\beta \quad (4.19)$$

with $\sigma > 2, \beta > 0, B_{\sigma R}(x_0) \subset \Omega_1$. If (4.19) holds, then we can rely on the hole filling technique of Widman [2]. Filling the hole, we deduce from (4.19)

$$\int_{B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx \leq \theta \int_{B_{\sigma R}(x_0)} |Du|^2 |x - x_0|^{2-n} dx + CR^\beta \quad (4.20)$$

with $\theta < 1$. Set (x_0 being fixed in Ω_1), choosing $2\delta < \beta$,

$$\varphi(R) = \frac{\int_{B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx}{R^{2\delta}}, \quad R \leq R_0,$$

where

$$R_0 = \sup\{R \mid B_R(x_0) \subset \Omega_1\},$$

we may write from (4.20)

$$\varphi(R) \leq \mu\varphi(\sigma R) + C, \quad R \leq \frac{R_0}{\sigma},$$

¹This is formal, since G^{x_0} is not in C^1 . One proceeds first with an approximation of the Green function. We skip this step

with $\mu = \theta\sigma^{2\delta} < 1$. Since $\varphi(R_0) < \infty$, as a consequence of (4.13), we get $\varphi(R) \leq C, \forall R \leq R_0$. So we have shown that

$$\int_{B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx \leq KR^{2\delta}, \quad \forall x_0 \in \Omega_1, B_R(x_0) \subset \Omega_1,$$

which implies (4.18).

Proof of (4.13). We apply (4.5) with

$$\psi = G^{x_0} \tau_R^2, \quad x_0 \in \Omega_1, \quad B_{2R}(x_0) \subset \Omega_1, \quad (4.21)$$

$$c_\nu = c_\nu^R = \frac{1}{|B_{2R} - B_{R/2}|} \int_{B_{2R}(x_0) - B_{R/2}(x_0)} u_\nu dx. \quad (4.22)$$

We obtain first

$$k_0 \int_{\Omega} |Du|^2 G^{x_0} \tau_R^2 \geq C \int_{B_R(x_0)} |Du|^2 |x - x_0|^{2-n} dx, \quad (4.23)$$

$$K \int_{\Omega} G^{x_0} \tau_R^2 \leq CR^{\frac{n}{\delta}}. \quad (4.24)$$

Next

$$D\psi = DG^{x_0} \tau_R^2 + 2G^{x_0} \tau_R D\tau_R.$$

Consider

$$I = 2 \int_{\Omega} aDXD\tau_R G^{x_0} \tau_R dx,$$

we have from (4.7)

$$\begin{aligned} |I| &\leq \int_{B_{2R} - B_R} |Du| \frac{|u - c^R|}{R} G^{x_0} dx \\ &\leq C \int_{B_{2R} - B_R} |Du|^2 |x - x_0|^{2-n} dx + C \int_{B_{2R} - B_R} \frac{|u - c^R|^2}{R^2} |x - x_0|^{2-n} dx. \end{aligned}$$

But, using Poincaré's inequality we have

$$\begin{aligned} \int_{B_{2R} - B_R} \frac{|u - c^R|^2}{R^2} |x - x_0|^{2-n} dx &\leq \frac{C}{R^n} \int_{B_{2R} - B_R} |u - c^R|^2 dx \\ &\leq \frac{C}{R^{n-2}} \int_{B_{2R} - B_{R/2}} |Du|^2 dx \\ &\leq C \int_{B_{2R} - B_{R/2}} |Du|^2 |x - x_0|^{2-n} dx, \end{aligned}$$

hence

$$|I| \leq C \int_{B_{2R} - B_{R/2}} |Du|^2 |x - x_0|^{2-n} dx. \quad (4.25)$$

Consider next the term

$$\begin{aligned}
\Pi &= \int_{\Omega} aDXDG^{x_0}\tau_R^2 dx \\
&= \int_{\Omega} aDG^{x_0}D((X-2^N)\tau_R^2)dx - 2 \int_{\Omega} aDG^{x_0}D\tau_R(X-2^N)\tau_R dx \\
&\geq -2 \int_{\Omega} aDG^{x_0}D\tau_R(X-2^N)\tau_R dx \\
&\geq -C \int_{B_{2R}-B_R} |DG| \frac{|u-c^R|^2}{R} \tau_R dx,
\end{aligned}$$

where we have used the estimate

$$|X-2^N| \leq C|u-c^R|^2. \quad (4.26)$$

We estimate Π from below as follows

$$\Pi \geq -C \int_{B_{2R}-B_R} G \frac{|u-c^R|^2}{R^2} dx - C \int_{B_{2R}-B_R} G^{-1}|DG|^2|u-c^R|^2\tau_R^2 dx. \quad (4.27)$$

The first term on the right hand side is estimated by the right hand side of (4.21). So there remains to estimate the term

$$\text{III} = \int_{B_{2R}-B_R} G^{-1}|DG|^2|u-c^R|^2\tau_R^2 dx.$$

Now, we introduce a new cut off function ξ , satisfying

$$\xi = \begin{cases} 0 & \text{for } |x| \leq \frac{1}{2}, \\ \tau & \text{for } |x| > 1, \end{cases}$$

and we set

$$\xi_R(x) = \xi\left(\frac{x-x_0}{R}\right),$$

thus

$$\xi_R = \tau_R \text{ on } B_{2R} - B_R.$$

We first test the Green function equation (4.14) with $G^{-1/2}|u-c^R|^2\xi_R^2$ which vanishes on the boundary of Ω , since $B_{2R} \subset \Omega_1$. It also vanishes on x_0 , hence

$$\frac{1}{2} \int_{\Omega} aDGDGG^{-3/2}|u-c^R|^2\xi_R^2 dx = \int_{\Omega} aD(|u-c^R|^2\xi_R^2).DGG^{-1/2} dx. \quad (4.28)$$

On the other hand, taking in (2.10) $v = (u_\nu - c_\nu^R)G^{1/2}\xi_R^2$ yields, summing up in ν ,

$$\begin{aligned} & \sum_\nu \int_\Omega a D u_\nu D_\nu G^{1/2} \xi_R^2 dx + \frac{1}{4} \int_\Omega a D (|u - c^R|^2 \xi_R^2) D G G^{-1/2} dx \\ & - \frac{1}{2} \int_\Omega a D G . D \xi_R |u - c^R|^2 G^{-1/2} \xi_R dx + \int_\Omega a D |u - c^R|^2 . D \xi_R G^{1/2} \xi_R dx \\ & = \int_\Omega (H_\nu(x, Du) - \alpha u_\nu) (u_\nu - c_\nu^R) G^{1/2} \xi_R^2 dx. \end{aligned}$$

So we deduce, taking into account the quadratic growth of H_ν ,

$$\begin{aligned} & \int_\Omega a D (|u - c^R|^2 \xi_R^2) D G G^{-1/2} dx \leq 2 \int_\Omega a D G . D \xi_R |u - c^R|^2 G^{-1/2} \xi_R dx \\ & + C R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} |Du|^2 dx + C R^{n(1-\frac{1}{2q})} \\ & + C R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} \frac{|u - c^R|^2}{R^2} dx. \end{aligned} \quad (4.29)$$

Furthermore

$$\begin{aligned} & \int_\Omega a D G . D \xi_R |u - c^R|^2 G^{-1/2} \xi_R dx \\ & \leq C \delta \int_\Omega a D G . D G G^{-3/2} |u - c^R|^2 \xi_R^2 dx \\ & + \frac{C}{\delta} R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} \frac{|u - c^R|^2}{R^2} dx, \quad \forall \delta, \end{aligned}$$

and combining this estimate in (4.29), (4.28) for δ sufficiently small yields

$$\begin{aligned} & \int_\Omega a D G D G G^{-3/2} |u - c^R|^2 \xi_R^2 dx \\ & \leq C R^{n(1-\frac{1}{2q})} + C R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} |Du|^2 dx \\ & + C R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} \frac{|u - c^R|^2}{R^2} dx, \end{aligned}$$

and using Poincaré's inequality, we obtain

$$\int_\Omega a D G D G G^{-3/2} |u - c^R|^2 \xi_R^2 dx \leq C R^{n(1-\frac{1}{2q})} + C R^{\frac{2-n}{2}} \int_{B_{2R} - B_{R/2}} |Du|^2 dx. \quad (4.30)$$

Going back to the definition of III, and recalling that $\xi_R = \tau_R$ on $B_{2R} - B_R$,

we get

$$\text{III} \leq CR^{\frac{2-n}{n}} \int_{\Omega} G^{-3/2} a D G D G |u - c^R|^2 \xi_R^2 dx,$$

and from (4.30) it follows that

$$\begin{aligned} \text{III} &\leq CR^{2-n} \int_{B_{2R}-B_{R/2}} |Du|^2 dx + CR^{1+\frac{n}{2q'}} \\ &\leq C \int_{B_{2R}-B_{R/2}} |Du|^2 |x - x_0|^{2-n} dx + CR^{1+\frac{n}{2q'}}. \end{aligned}$$

Finally from (4.27) we obtain

$$\text{II} \geq -C \int_{B_{2R}-B_{R/2}} |Du|^2 |x - x_0|^{2-n} dx - CR^{1+\frac{n}{2q'}}.$$

Collecting results in the application of (4.5) and changing R into $2R$, we obtain (4.19) with $\sigma = 4$, and $\beta = \frac{n}{q'}$ since $\frac{n}{q'} < 1 + \frac{n}{2q'}$ and $\beta < 2$.

4.4. C^δ Regularity on the Boundary

By local maps representation of the boundary, the problem amounts to the following. Consider a sufficiently small ball B centered on the boundary, and a diffeomorphism ψ from B onto $D \subset R^n$, such that

$$\begin{aligned} \Omega^+ &= \psi(B \cap \Omega) \subset \{y \in R^n \mid y_n > 0\}, \\ \Gamma' &= \psi(B \cap \Gamma) \subset \{y \in R^n \mid y_n = 0\}, \end{aligned}$$

ψ, ψ^{-1} sufficiently smooth. Define next

$$\begin{aligned} \Omega^- &= \{y \mid y_n < 0, (y_1, \dots, y_{n-1}, -y_n) \in \Omega^+\}, \\ \Omega' &= \Omega^+ \cup \Omega^- \cup \Gamma'. \end{aligned}$$

If z is defined on Ω , set

$$z'(y) = z(\psi^{-1}(y)), \quad y \in \Omega^+ \cup \Gamma'$$

and define z' on Ω^- by reflection, namely

$$z'(y_1, \dots, y_{n-1}, y_n) = z'(y_1, \dots, y_{n-1}, -y_n) \text{ if } y_n < 0.$$

Applying the procedure to the function u_ν , we have to prove

$$u'_\nu(y) \in C^\delta(\Omega'). \quad (4.31)$$

We begin to reduce the variational problem (2.10) to a problem on $B \cap \Omega$.

Consider in (2.10) functions which vanish in $\Omega - \Omega \cap B$, we obtain

$$\begin{aligned} \int_{\Omega \cap B} a D u_\nu D v d x + \alpha \int_{\Omega \cap B} u_\nu v d x &= \int_{\Omega \cap B} H_\nu(x, D u) v d x, \\ \forall v \in H^1(\Omega \cap B) \cap L^\infty(\Omega \cap B), v|_{\partial B \cap \Omega} &= 0. \end{aligned} \quad (4.32)$$

We then perform the change of coordinates $x = \psi^{-1}(y)$, denote

$$J_\psi(x) = \text{matrix} \left(\frac{\partial \psi_i}{\partial x_j} \right),$$

and set

$$\begin{aligned} a'(y) &= \frac{J_\psi(\psi^{-1}(y)) a(\psi^{-1}(y)) J_\psi^*(\psi^{-1}(y))}{|\det J_\psi(\psi^{-1}(y))|}, \\ H'_\nu(y, p) &= \frac{H_\nu(\psi^{-1}(y), J_\psi^*(\psi^{-1}(y)) p)}{|\det J_\psi(\psi^{-1}(y))|}, \\ a'_0(y) &= \frac{\alpha}{|\det J_\psi(\psi^{-1}(y))|}. \end{aligned}$$

The variational problem (4.32) reads

$$\begin{aligned} \int_{\Omega^+} a'(y) D u'_\nu D v' d y + \int_{\Omega^+} a'_0(y) u'_\nu v' d y &= \int_{\Omega^+} H'_\nu(y, D u') v' d y, \\ \forall v' \in H^1(\Omega^+) \cap L^\infty(\Omega^+), v'|_{\partial \Omega^+ - \Gamma'} &= 0, \\ u'_\nu \in H^1(\Omega^+) \cap L^\infty(\Omega^+), & \end{aligned} \quad (4.33)$$

$$a'(y) \xi \cdot \xi \geq \alpha'_0 |\xi|^2, \quad a_0(y) \geq \bar{a}_0. \quad (4.34)$$

Moreover the Hamiltonian $H'_\nu(y, p)$ verify the special structure assumption, namely

$$H'_\nu(y, p) = Q'(y, p) p_\nu + H_\nu^0(y, p) \quad (4.35)$$

with

$$Q'(y, p) = \frac{J_\psi(\psi^{-1}(y)) Q(\psi^{-1}(y), p)}{|\det J_\psi(\psi^{-1}(y))|}, \quad H_\nu^0(y, p) = \frac{H_\nu^0(\psi^{-1}(y), p)}{|\det J_\psi(\psi^{-1}(y))|}$$

and we have

$$|Q'(y, p)| \leq k + K|p|, \quad (4.36)$$

$$|H_\nu^0(y, p)| \leq k_\nu + K_\nu \sum_{\mu \leq \nu} |p_\mu|^2. \quad (4.37)$$

We now proceed with the reflection procedure.

Write $y = (y', y_n)$ where $y' = (y_1, \dots, y_{n-1})$ and define for $y_n < 0$,

$$\begin{aligned} a'_{ii}(y', y_n) &= a'_{ii}(y', -y_n), \quad \forall i, \\ a'_{ij}(y', y_n) &= a'_{ij}(y', -y_n), \quad \forall i, j \ i \neq j, \ i, j \neq n, \\ a'_{in}(y', y_n) &= -a'_{in}(y', -y_n), \quad \forall i \neq n, \\ a'_0(y', y_n) &= a_0(y', -y_n). \end{aligned}$$

Let also $p_\nu = (p'_\nu, p_{\nu n})$, where p'_ν stands for the components $p_{\nu 1}, \dots, p_{\nu(n-1)}$. Then, we write for $y_n < 0$,

$$H'_\nu(y', y_n; p'_1, p_{1n}; \dots; p'_N, p_{Nn}) = H'_\nu(y', -y_n; p'_1, -p_{1n}; \dots; p'_N, -p_{Nn}).$$

Therefore the functions u'_ν extended by reflection appear to be solutions of the problem

$$\begin{aligned} \int_{\Omega'} a'(y) Du'_\nu Dv' dy + \int_{\Omega'} a'_0(y) u'_\nu v' dy &= \int_{\Omega} H'_\nu(y, Du') v' dy, \\ \forall v' \in H_0^1(\Omega') \cap L^\infty(\Omega'), \quad u'_\nu \in H^1(\Omega') \cap L^\infty(\Omega') \end{aligned} \quad (4.38)$$

and our objective is to prove (4.31).

If we drop the prime symbol, and use x instead of y , our problem amounts to the following. Consider the variational problem

$$\begin{aligned} \int_{\Omega} a Du_\nu Dv dx + \int_{\Omega} a_0 u_\nu v dx &= \int_{\Omega} H_\nu(x, Du) v dx, \\ \forall v \in H_0^1(\Omega) \cap L^\infty(\Omega), \quad u_\nu \in H^1(\Omega) \cap L^\infty(\Omega). \end{aligned}$$

We know that the solution is $C^\delta(\Omega)$.

It is similar to the interior C^δ regularity of (2.10) considered in §4. However we have to be careful to consider only test functions, which vanish on the boundary. Nevertheless, thanks to the special structure on H_ν we derive again (4.5), for any ψ in $C^1(\bar{\Omega})$, $\psi \geq 0$, $\psi|_{\partial\Omega} = 0$. In the proof of interior C^δ regularity, we use only test functions which vanish on the boundary of Ω , hence the proof carries over, and the result follows. •

4.5. End of Proof of Theorem 2.1

We know that we have a solution of (2.10) which belongs to $W^{1,p}(\Omega) \cap C^\delta(\bar{\Omega})$, $2 \leq p < 2 + \varepsilon$. We can now rely on the linear theory and Miranda-Nirenberg interpolation theorem [3]. Indeed $H_\nu(x, Du)$ belongs to $L^{\frac{p_0}{2}}$ with $p_0 > 2$. Hence from the linear theory $u_\nu \in W^{2, \frac{p_0}{2}}$ and since $u_\nu \in C^\delta$, we

have also

$$u_\nu \in W^{1,p_1} \text{ with } \frac{1}{p_1} = \frac{1}{p_0} - \frac{\delta}{2n},$$

provided $p_0 < \frac{2n}{\delta}$, and thus $p_1 > p_0$. After a finite number of steps we get $p_i \geq \frac{2n}{\delta}$, and it follows that $u_\nu \in W^{1,p}(\Omega)$, $p > 2n$, and from the linear theory again $u_\nu \in W^{2,s}(\Omega)$, $s > n$. From Sobolev embedding theorem, $u_\nu \in W^{1,r}(\Omega)$, $\forall s$ and thus from the linear theory again $u_\nu \in W^{2,s}(\Omega)$, $\forall s$.

REFERENCES

- [1] Bensoussan, A. & Frehse, J., *Nonlinear partial differential equations and applications*, Springer-Verlag, New York, 2001.
- [2] Widman, K. O., Hölder continuity of solutions of elliptic equations, *Manuscr. Math.*, **5**(1971), 299–308.
- [3] Nirenberg, L., An extended interpolation inequality, *Ann. Sc. Norm. Sup. Pisa*, **cl. Sc. IV. ser. 20**(1966), 733–737.

SENSITIVITY OF DARCY'S LAW TO DISCONTINUITIES

C. BERNARDI

*Laboratoire Jacques-Louis Lions, Université Paris
VI-CNRS, France. E-mail: bernardi@ann.jussieu.fr*

O. PIRONNEAU

*Laboratoire Jacques-Louis Lions, Université Paris
VI-IUF, France. E-mail: pironneau@ann.jussieu.fr*

Abstract

The authors investigate the sensitivity of hydrostatic pressure of flows through porous media with respect to the position of the soil layers. Indeed, these induce discontinuities of the porosity which is a piecewise constant coefficient κ of the partial differential equation satisfied by the pressure u and it leads to the computation of the derivative of u with respect to changes in position of discontinuity surface of κ .

The analysis relies on a mixed formulation of the problem. Preliminary numerical simulations are given to illustrate the theory. An application to a simple inverse problem is also given.

Keywords Partial differential equations, Flow through porous media,
Sensitivity, Darcy's equation, Topological optimization
2000 MR Subject Classification 35Q35

§0. Introduction

The hydrostatic pressure u in a porous medium of porosity κ can be found, in simple situations, by solving

$$-\nabla \cdot (\kappa \nabla u) = f.$$

When the porous medium has layers of different materials, κ is smooth except at the interface between layers where it is discontinuous. For ground identification from surface data it is important to compute the sensitivity of u with respect to the discontinuities of κ . This is similar to topological

optimization (cf. [2]), but here we can make an explicit use of the fact that κ is piecewise regular and use the topology of the layers. Continuity in κ with respect to the L^∞ norm has been studied earlier (cf. [4]), but such a formalism does not allow changes in the position of the discontinuities of κ .

Problem Statement. More precisely, consider a bounded open set Ω of \mathcal{R}^2 and a family of closed curves $a \mapsto \Sigma(a)$ strictly inside Ω and function of a scalar parameter a separating Ω into two non overlapping sets $\Omega_i(a)$, $i = 1, 2$, with Lipschitz continuous boundaries:

$$\bar{\Omega} = \bar{\Omega}_1(a) \cup \bar{\Omega}_2(a), \quad \Omega_1(a) \cap \Omega_2(a) = \emptyset, \quad \bar{\Sigma}(a) = \bar{\Omega}_1(a) \cap \bar{\Omega}_2(a).$$

Note that we do not allow $\Sigma(a)$ to touch $\partial\Omega$. This is for mathematical convenience only and it is conjectured that the following results apply also to the general case both in \mathcal{R}^2 and \mathcal{R}^3 .

Let $\kappa(a)$ be piecewise constant and equal to κ_i on $\Omega_i(a)$ and consider

$$-\nabla \cdot (\kappa(a)\nabla u(a)) = f \tag{0.1}$$

with Neumann or Dirichlet condition on $\Gamma = \partial\Omega$. We wish to compute the derivative u' of $u(a)$ with respect to a at $a = 0$.

Again the following analysis is not hard to generalize to the case of a piecewise differentiable κ with similar discontinuity surface $\Sigma(a)$, and it is for the sake of clarity that we present the piecewise constant case only.

The Traditional Approach Fails. Denoting by $I_{\Omega_2(a)}$ the characteristic function of $\Omega_2(a)$ and by $[\kappa]$ the jump $\kappa_2 - \kappa_1$, we have

$$\kappa(a) = \kappa_1 + [\kappa]I_{\Omega_2(a)}.$$

So it is tempting to write that

$$\int_{\Omega} \kappa(a)\nabla u'\nabla w = - \int_{\Omega} \kappa'\nabla u\nabla w = - \int_{\Sigma} [\kappa]\nabla u\nabla w$$

but the last integral makes no sense because ∇u is discontinuous across $\Sigma(a)$.

Similarly an ‘‘optimal shape design’’ approach seems to lead nowhere.

Indeed with the notation of [6], we assume that $\Sigma(a)$ depends on a via a given function α of $\mathcal{C}^1(\Sigma)$ and the following local variation in the ‘‘direction’’ α around $\Sigma(0)$ taken as a reference curve:

$$\Sigma(a) = \{x + a\alpha(x)n(x) : x \in \Sigma(0)\}, \tag{0.2}$$

where n is the unit normal to $\Sigma(0)$ which points outside Ω_1 .

For clarity we shall drop the parameter a when it is equal to 0 and write Σ instead of $\Sigma(0)$ whenever non ambiguous.

Recall that in smooth situations ($f \in C^1(\Omega)$, see [6, Chapter 3] for details), the following form of the mean value theorem is shown

$$\int_{\Omega_i(a)} f - \int_{\Omega_i(0)} f = -(-1)^i a \int_{\Sigma} f \alpha + o(|a|), \quad i = 1, 2.$$

Therefore with the shorthand notation $\int_{\delta\Omega_i} f = \int_{\Omega_i(a)} f - \int_{\Omega_i(0)} f$, a formal differentiation of (0.1) would lead to

$$\begin{aligned} 0 &= \delta \left(\int_{\Omega} \kappa(a) \nabla u(a) \nabla w - \int_{\Omega} f w \right) \\ &= \int_{\delta\Omega_1} \kappa_1 \nabla u \nabla w + \int_{\delta\Omega_2} \kappa_2 \nabla u \nabla w + \int_{\Omega} \kappa(a) \nabla \delta u \nabla w + \text{higher order terms} \\ &= a \int_{\Sigma} \left(\kappa_1 \frac{\partial u}{\partial n} - \kappa_2 \frac{\partial u}{\partial n} \right) + \int_{\Omega} \kappa(a) \nabla \delta u \nabla w + \text{higher order terms.} \end{aligned} \tag{0.3}$$

However the jump of $\kappa \frac{\partial u}{\partial n}$ across Σ is zero so this calculation indicates that u' would be zero. The example below shows that it is not the case. Hence the ‘‘higher order terms’’ are wrongly called so.

An Example. Consider the same problem on $(0, L)$ with $\Omega_1(a) = (0, a)$ and $\Omega_2(a) = (a, L)$, $0 < a < L$:

$$\nabla \cdot \kappa \nabla u = 0, \quad u(0) = 0, \quad u(L) = b.$$

Let $H(x)$ be the Heaviside function and

$$\kappa(x) = \kappa_1 + (\kappa_2 - \kappa_1)H(x - a), \quad 0 < a < L.$$

There is an analytical solution to the PDE

$$u = \begin{cases} a_1 x, & x < a, \\ a_2(x - L) + b, & x > a, \end{cases}$$

and the continuity of u and $\kappa \partial_n u$ requires

$$a_1 a = a_2(a - L) + b, \quad \kappa_1 a_1 = \kappa_2 a_2 \Rightarrow \quad a_2 = \frac{b \kappa_1}{L \kappa_1 + a[\kappa]}, \quad a_1 = \frac{b \kappa_2}{L \kappa_1 + a[\kappa]}.$$

So the derivative u' of u with respect to a , at $a \neq 0$, is

$$u'(a) = \begin{cases} -x b[\kappa] \kappa_2 (L \kappa_1 + a[\kappa])^{-2}, & x < a, \\ -(x - L) b[\kappa] \kappa_1 (L \kappa_1 + a[\kappa])^{-2}, & x > a. \end{cases}$$

Notice that u' is discontinuous at a :

$$[u']_{x=a} = \frac{b[\kappa](a[\kappa] + L\kappa_1)}{(L\kappa_1 + a[\kappa])^2} = \frac{b[\kappa]}{L\kappa_1 + a[\kappa]} \neq 0 \quad \text{if } b \neq 0.$$

Outline. The paper is organized as follows.

- First we recall the mixed formulation of the problem and we guess the result by giving a heuristic argument for the derivative.
- Then in Sections 2 and 3 we prove existence, uniqueness and regularity for the solution of the mixed problem. The conjecture is proved in Section 4.
- A simple numerical illustration is given in Section 5; a Raviart-Thomas element is used, the derivative is compared with the finite difference approximation. Finally a simple inverse problem is analyzed numerically in Section 6; the curve of discontinuity of κ is recovered by least square (optimal control) from observations on a set distant from the curve.

§1. The Conjecture

From now on we work with the curves $\Sigma(a)$ introduced in (0.2). For the sake of clarity and without loss of generality, let us consider a slightly different problem, for a given function f and a given vector valued function F :

$$\nabla \cdot (\kappa \nabla u) = f - \nabla \cdot (\kappa F) \quad \text{in } \Omega, \quad \kappa \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = -\kappa F \cdot n. \quad (1.1)$$

Let

$$\begin{aligned} H(\text{div}, \Omega) &= \{W \in L^2(\Omega)^d : \nabla \cdot W \in L^2(\Omega)\}, \\ X &= \{V \in H(\text{div}, \Omega) : V \cdot n = 0 \quad \text{on } \partial\Omega\}, \\ L_0^2(\Omega) &= \left\{ w \in L^2(\Omega) : \int_{\Omega} w = 0 \right\}. \end{aligned}$$

Let F belong to $L^2(\Omega)^d$. Consider the mixed formulation with $U = \kappa(\nabla u + F)$.

Find $(U, u) \in X \times L_0^2(\Omega)$ with

$$\begin{cases} \forall w \in L_0^2(\Omega), \int_{\Omega} (\nabla \cdot U)w = \int_{\Omega} fw, \\ \forall W \in X, \int_{\Omega} \frac{1}{\kappa} U \cdot W + \int_{\Omega} u \nabla \cdot W = \int_{\Omega} F \cdot W. \end{cases} \quad (1.2)$$

This is a slight departure from the standard mixed formulation (see [3]) in that we have divided by κ .

Next we observe that with a piecewise constant κ ,

$$\frac{1}{\kappa(a)}(x) = \frac{1}{\kappa_1} + \left[\frac{1}{\kappa}\right]I_{\Omega_2}, \quad \text{therefore} \quad \frac{d}{da}\left(\frac{1}{\kappa}\right) = \alpha\delta_{\Sigma}\left[\frac{1}{\kappa}\right]$$

with δ_{Σ} defined by

$$\forall f \in H^{1/2+\epsilon}(\Omega), \quad \int_{\Omega} f\delta_{\Sigma} = \int_{\Sigma} f.$$

So differentiating (1.2) yields

$$\begin{aligned} \int_{\Omega} (\nabla \cdot U')w &= 0, \\ \int_{\Omega} \left(\frac{1}{\kappa}U' \cdot W + u'\nabla \cdot W\right) &= - \int_{\Sigma} \alpha\left[\frac{1}{\kappa}U\right] \cdot W, \end{aligned} \quad (1.3)$$

assuming that the trace of $U \cdot W$ on Σ exists. Note, however, that if s, n denote a tangent vector and a normal vector to Σ , $U \cdot n/\kappa$ jumps across Σ but $U \cdot s/\kappa = \partial u/\partial s$ does not because $[u]_{\Sigma} = 0$. Hence

$$\left[\frac{1}{\kappa}U\right] \cdot W = \left[\frac{1}{\kappa}U\right] \cdot nW \cdot n = \left[\frac{1}{\kappa}\right]U \cdot nW \cdot n,$$

and so it seems that $U \cdot n \in L^2(\Sigma)$ could be sufficient.

Conjecture 1.1. *The derivative $(V, v) \equiv (U', u')$ of (U, u) solution of (1.2) is given by:*

Find $(V, v) \in X \times L^2_0(\Omega)$ with

$$\begin{cases} \forall w \in L^2_0(\Omega), & \int_{\Omega} (\nabla \cdot V)w = 0, \\ \forall W \in X, & \int_{\Omega} \left(\frac{1}{\kappa}V \cdot W + v\nabla \cdot W\right) = \int_{\Sigma} gW \cdot n, \end{cases} \quad (1.4)$$

with

$$g = -\alpha\left[\frac{1}{\kappa}\right]U \cdot n.$$

Remark 1.1. In the distribution sense (1.4) is

$$\nabla \cdot V = 0, \quad \frac{1}{\kappa}V - \nabla v = \delta_{\Sigma}gn, \quad V \cdot n|_{\Gamma} = 0.$$

Note then that it is not possible to eliminate V and write a single equation for v because $\kappa\delta_{\Sigma}$ has no meaning on account of the discontinuity of κ on Σ .

Remark 1.2. It is easy to see that the conjecture is true for the one-dimensional example described in the introduction.

$$\begin{aligned}
U' &= \frac{d}{da} \left(\kappa \frac{du}{dx} \right) = -b\kappa_1\kappa_2[\kappa](L\kappa_1 + a[\kappa])^{-2} \\
\Rightarrow \int_0^L U' \frac{dw}{dx} &= U' \int_0^L \frac{dw}{dx} = 0, \\
&\int_0^L \left(\frac{1}{\kappa} U' W + u' \nabla \cdot W \right) \\
&= -b\kappa_1\kappa_2[\kappa](L\kappa_1 + a[\kappa])^{-2} \left(\frac{1}{\kappa_1} \int_0^a W + \frac{1}{\kappa_2} \int_a^L W \right) \\
&\quad + b\kappa_2[\kappa](L\kappa_1 + a[\kappa])^{-2} \int_0^a W + b\kappa_1[\kappa](L\kappa_1 + a[\kappa])^{-2} \int_a^L W \\
&\quad + u' W \Big|_0^a + u' W \Big|_a^L \\
&= W(a) \frac{b[\kappa]}{a[\kappa] + L\kappa_1} = -W(a) \left[\frac{1}{\kappa} \right] \kappa \frac{du}{dx}.
\end{aligned}$$

§2. Existence and Uniqueness

Assume $\Gamma = \partial\Omega$ of class $C^{1,1}$. Let Σ be a smooth closed curve inside Ω .

Let $\kappa > 0$ be a piecewise constant function discontinuous across Σ only. Then, provided that $W \mapsto \int_{\Sigma} gW \cdot n$ is continuous on X , an existence and uniqueness result can be shown by adapting the proof of Theorem 2.1 in [3] to the case $\kappa \neq 1$. It is an application of the inf-sup lemma, namely, all bilinear forms being continuous,

$$\|V\|_{H(\operatorname{div}, \Omega)} + \frac{\beta^*}{2} \|v\|_0 \leq C \|g\|_{H^{1/2}(\Sigma)},$$

where β^* is the inf-sup constant of $(W, w) \mapsto (\nabla \cdot W, w)$:

$$\beta^* = \inf_{w \in L_0^2(\Omega)} \sup_{W \in X} \frac{(\nabla \cdot W, w)}{\|v\|_{H(\operatorname{div}, \Omega)} \|w\|_0}.$$

§3. Regularity

For the sake of clarity, we assume for what follows that $F = 0$. Recall also that we have assumed that $\Gamma \cap \Sigma = \emptyset$.

In (1.4) $U \cdot n \Big|_{\Sigma}$ appears in an integral. We need to show that the integral exists. Functions of V have their normal component traces $V \cdot n$ on Σ in $H^{-1/2}(\Sigma)$. So we need to show that $U \cdot n \in H^{1/2}(\Sigma)$.

Proposition 3.1. *If Σ is regular and f is in $H^1(\Omega)$, then $U \cdot n$ belongs to $H^{1/2}(\Sigma)$.*

For clarity the proof is given in dimension 2. Assume that Σ is sufficiently regular so that in a neighborhood \mathcal{O} of Σ we can define a coordinate system σ, ν in which the equation of Σ is $\nu = 0, n$ the normal to Σ , is tangent to the curves $\sigma = \text{constant}$ and σ is its curvilinear abscissa.

Problem (1.1) in variational form writes

$$\forall w \in H^1(\Omega), \quad \int_{\mathcal{O}} \kappa \nabla u \nabla w = \int_{\Omega} f w.$$

By taking $\text{supp } w' \subset \mathcal{O}, w = \frac{\partial w'}{\partial \sigma}$ and by integrating by parts in σ we find that

$$\forall w' \in H_0^1(\mathcal{O}), \quad \int_{\mathcal{O}} \kappa \nabla \frac{\partial u}{\partial \sigma} \nabla w' = \int_{\mathcal{O}} \frac{\partial f}{\partial \sigma} w' \tag{3.1}$$

because κ is not a function of σ . This shows that if f is regular all partial derivatives in σ of u are in $H^1(\mathcal{O})$. Therefore $U \cdot \vec{s}$ has the same regularity.

Now $\frac{\partial U \cdot \vec{s}}{\partial \sigma} = \kappa \frac{\partial^2 u}{\partial \sigma^2}$ belongs to $L^2(\mathcal{O})$. So by $\nabla \cdot U = 0$ we see that $\frac{\partial U \cdot n}{\partial \nu} |_{\mathcal{O}_i}$ is in $L^2(\mathcal{O}_i)$ for any open set $\mathcal{O}_i \in \mathcal{O}$ not intersecting Σ .

Similarly, $U \cdot \vec{s} \in H^1(\mathcal{O})$ implies $\frac{\partial}{\partial \nu} U \cdot \vec{s} \in L^2(\mathcal{O})$, i.e. $\frac{\partial}{\partial \nu} \kappa \frac{\partial u}{\partial \sigma} \in L^2(\mathcal{O})$. Therefore $\frac{\partial U \cdot n}{\partial \sigma} = \kappa \frac{\partial^2 u}{\partial \nu \partial \sigma}$ is in $L^2(\mathcal{O} \setminus \Sigma)$. Hence U is in $H^1(\Omega \setminus \Sigma)$.

Corollary 3.1. *If Σ is regular and F belongs to $W^{2,\infty}(\Omega)$ then U is continuous in $\Omega \setminus \Sigma$ and U is in $L^\infty(\Omega)$.*

This is because (3.1) shows that $\frac{\partial u}{\partial \sigma}$ satisfies a partial differential equation of the same type as the one of u . So by the proposition above $\kappa \nabla \frac{\partial u}{\partial \sigma} \cdot n$ is in $H^{1/2}(\Sigma)$. Therefore $U \cdot n$ is in $H^{3/2}(\Sigma)$, hence it is also continuous and bounded. By the maximum principle (in $\Omega \setminus \Sigma, \Delta U$ is bounded), U is bounded everywhere.

§4. Differentiability

4.1. Continuity

In order to study the changes $\delta u, \delta U$ of u, U , when $a \rightarrow 0$, let

$$\eta = \kappa^{-1}, \quad \delta u = u(a) - u(0), \quad \delta U = U(a) - U(0), \quad \delta \eta = \eta(a) - \eta(0).$$

(1) The first equation in (1.4) is easy to establish because

$$\int_{\Omega} (\nabla \cdot U) w = 0, \quad \int_{\Omega} (\nabla \cdot (U + \delta U)) w = 0 \Rightarrow \int_{\Omega} (\nabla \cdot \delta U) w = 0. \tag{4.1}$$

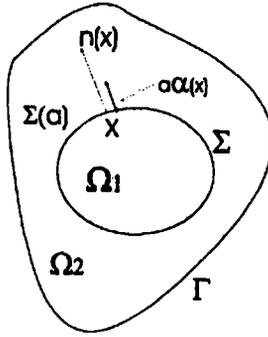


Fig. 1. The coefficient κ is constant on Ω_1 and constant on Ω_2 and discontinuous across Σ . When Σ becomes $\Sigma(a)$, the distance from $\Sigma(a)$ to Σ is $a\|\alpha\|_\infty$, the solution of the partial differential equation changes. We wish to find the derivative with respect to a for a given function α .

(2) For the second equation we have

$$\int_{\Omega} ((\eta + \delta\eta)(U + \delta U) \cdot W - \eta U \cdot W + \delta u \nabla \cdot W) = 0,$$

which is also

$$\int_{\Omega} ((\eta + \delta\eta)\delta U \cdot W + \delta u \nabla \cdot W) = - \int_{\Omega} \delta\eta U \cdot W. \quad (4.2)$$

Take $W = \delta U$ in (4.2), and use $w = \delta u$ in (4.1), then with $\kappa_{\max} = \max(\kappa_1, \kappa_2)$, $\kappa_{\min} = \min(\kappa_1, \kappa_2)$

$$\begin{aligned} \int_{\Omega} (\eta + \delta\eta)|\delta U|^2 &= - \int_{\Omega} \delta\eta U \cdot \delta U \leq \|\delta U\|_0 \|U\|_\infty \|\delta\eta\|_0 \\ &\Rightarrow \frac{1}{\kappa_{\max}} \|\delta U\|_0 \leq \|U\|_\infty \|\delta\eta\|_0. \end{aligned}$$

The support $S(a)$ of $\delta\eta$ is a thin strip around Σ of width αa , and $\eta + \delta\eta$ is either κ_{\max}^{-1} or κ_{\min}^{-1} ; therefore

$$\|\delta\eta\|_0 \leq C_1 \left| \left[\frac{1}{\kappa} \right] \right| \sqrt{a\|\alpha\|_\infty}$$

for some constant C_1 . So we have the following result.

Proposition 4.1. *If $U \in L^\infty(\Omega)$, the change $\{\delta u, \delta U\}$ in $\{u, U\}$ due to a is bounded in $L^2_0(\Omega) \times H(\operatorname{div}, \Omega)$, verifies $\nabla \cdot \delta U = 0$ and*

$$\|\delta u\|_0 + \|\delta U\|_{H(\operatorname{div}, \Omega)} \leq C_2 \sqrt{\|\alpha\|_\infty a} \|U\|_\infty \frac{|\kappa|}{\kappa_{\min}}.$$

4.2. Differentiability

Lemma 4.1. *Let \mathcal{O} be a neighborhood of Σ containing the support $S(a)$ of $\eta(a) - \eta(0)$ but sufficiently thin near Σ so as to be able to extend n in \mathcal{O} . If U, W are continuous in $\mathcal{O} \setminus \Sigma$ and $U \cdot n, W \cdot n$ are continuous in \mathcal{O} then, for a small enough,*

$$\lim_{a \rightarrow 0} \left| \int_{S(a)} \frac{\delta(\eta U)}{a} \cdot W - \int_{\Sigma} \alpha \left[\frac{1}{\kappa} \right] U \cdot n W \cdot n \right| = 0. \quad (4.3)$$

Proof. We assume that W is smooth; the result can be extended later by density. Consider the case $a > 0, \alpha \geq 0$ everywhere. Then (see Fig. 1), assuming that n is oriented from $\Omega_1(a)$ to $\Omega_2(a)$, and denoting by $U_i(x, a)$ (resp. $(U_i(\sigma, n, a))$) the value of U at $x \in \Omega_i$ (resp. σ, n for the value a of the parameter), we have

$$\begin{aligned} & \int_{S(a)} \frac{\delta(\eta U)}{a} \cdot W \\ &= \int_{S(a)} \frac{\eta_2 U_2(x, a) - \eta_1 U_1(x, 0)}{a} \cdot W(x) dx \\ &= \int_{\Sigma} d\sigma \int_0^{a\alpha(\sigma)} \frac{\eta_2 U_2(\sigma, n, 0) - \eta_1 U_1(\sigma, n, 0)}{a} \cdot W(\sigma, n) dn + O(\sqrt{a}), \end{aligned} \quad (4.4)$$

where we have used the continuity with respect to a (Proposition 4.1), and where $S(a)$ is the support of $\delta\eta$:

$$S(a) = \{(\sigma, n) : n \in [0, a\alpha(\sigma)]\}.$$

By the mean value theorem, when g is continuous, there exists $n(\sigma)$ such that

$$\int_{S(a)} g = \int_{\Sigma} d\sigma \int_0^{a\alpha(\sigma)} g(\sigma, n) dn = \int_{\Sigma} a\alpha(\sigma) g(\sigma, n(\sigma)) d\sigma.$$

Therefore (4.4) becomes

$$\int_{S(a)} \delta \left(\frac{\eta}{a} U \right) \cdot W = \int_{\Sigma} \alpha(\sigma) [\eta U(x, 0)] \cdot W(x) d\sigma(x) + O(\sqrt{a}).$$

Naturally the same argument applies when $\alpha < 0$ and also to the general case by cutting Σ into pieces on which α does not change sign.

Finally (4.3) is found because

$$\eta_2 U_2 \cdot s - \eta_1 U_1 \cdot s = \left[\frac{\partial u}{\partial s} \right] = 0.$$

So with $\{v_\delta, V_\delta\} = \left\{ \frac{\delta u}{a}, \frac{\delta U}{a} \right\}$, (4.1) gives

$$\forall w \in L_0^2(\Omega), \quad \int_{\Omega} (\nabla \cdot V_\delta) w = 0. \quad (4.5)$$

Let \mathcal{O} be a thin strip around Σ containing $\Sigma(a)$ for all a under consideration below. After division by a Equation (4.2) reads

$$\forall W \in H(\text{div}, \Omega), \quad \int_{\Omega \setminus \mathcal{O}} (\eta V_\delta \cdot W + u_\delta \nabla \cdot W) + \int_{\mathcal{O}} \frac{\delta(\eta U)}{a} \cdot W = 0.$$

Accordingly by Lemma 4.1 there exists a weakly converging subsequence of v_δ, V_δ and the limit v, V satisfies

$$\forall W \in H(\text{div}, \Omega), \quad \int_{\Omega} (\eta V \cdot W + v \nabla \cdot W) + \left[\frac{1}{\kappa} \right] \int_{\Sigma} \alpha U \cdot n W \cdot n = 0.$$

Remark 4.1. With sufficient regularity we have in fact proved that

$$\left| \int_{S(a)} \delta \left(\frac{\eta U}{a} \right) \cdot W - \int_{\Sigma} \alpha \left[\frac{1}{\kappa} \right] U \cdot n W \cdot n \right| \leq a C \left\| \frac{\partial U}{\partial n} \cdot n \right\|_{1/2, \Sigma} \left\| \frac{\partial W}{\partial n} \cdot n \right\|_{1/2, \Sigma}. \quad (4.6)$$

Let us summarize the result:

Theorem 4.1. *The solution of (1.2) $\{u(a), U(a)\}$, with*

$$\Sigma(a) = \{x + a\alpha(x)n(x) : x \in \Sigma\}$$

is differentiable in a in the sense that

$$v = \lim_{a \rightarrow 0} \frac{u(a) - u(0)}{a}, \quad V = \lim_{a \rightarrow 0} \frac{U(a) - U(0)}{a}$$

is solution of (1.4) where the jump $\left[\frac{1}{\kappa} \right]$ is

$$x \in \Sigma, \quad \left[\frac{1}{\kappa} \right](x) = \lim_{a \rightarrow 0} \left(\frac{1}{\kappa(x + a\alpha(x)n(x))} - \frac{1}{\kappa(x - a\alpha(x)n(x))} \right).$$

§5. Discretization and Numerical Test

5.1. Discretization

Consider a regular family of triangulations of Ω of maximum edge length h and two finite element spaces X_h and L_h to approximate and $L^2_0(\Omega)$. Although precision is increased when the triangulations approximate Σ as an internal boundary, the theory works also without this hypothesis. The numerical discretization of problem (1.4) reads

Find $(V_h, v_h) \in X_h \times L_h$ with

$$\begin{cases} \forall q \in L_h, & \int_{\Omega} (\nabla \cdot V_h) q = 0, \\ \forall W \in X_h, & \int_{\Omega} \left(\frac{1}{\kappa} V_h \cdot W + v_h \nabla \cdot W \right) = \int_{\Sigma} g W \cdot n. \end{cases} \quad (5.1)$$

Among the various admissible choices, we have selected the Raviart-Thomas element for X_h and the piecewise constant function for L_h . This couple satisfies the discrete inf-sup condition as shown in [1, Proposition 3.14]:

$$\|V - V_h\|_0 + |v - v_h|_0 \leq ch^s (\|V\|_{H^s} + \|v\|_{H^s})$$

when $(V, v) \in (H^s(\Omega))^2 \times H^s(\Omega)$, $0 < s \leq 1$.

5.2. Numerical Simulation

The numerical solution is calculated with **freefem++** (see [5]). To illustrate the theory we have solved the problem

$$-\nabla \cdot (\kappa \nabla u) = 0 \text{ in } \Omega, \quad u|_{\Gamma} = xy, \quad (5.2)$$

where Ω is the rectangle $(-5, 5) \times (-2.5, 2.5)$, κ is 6 inside an ellipse in the middle of the rectangle and 1 outside.

Then the ellipse is changed by ϵ according to

$$\{(x, y) : x = (2 + \epsilon)(\sqrt{2} + \epsilon) \cos t, \quad y = (\sqrt{2} + \epsilon) \sin t, \quad t \in (0, 2\pi)\},$$

yielding a new solution u_ϵ of (5.2). Then $u'_\epsilon = (u_\epsilon - u)/\epsilon$ is compared to the numerical solution of (5.1). The results are displayed on Fig. 2. Decreasing both ϵ and the mesh size gives convergence:

ϵ	0.1	0.05	0.025	0.0125
L^2 -error	0.96	0.71	0.56	0.46

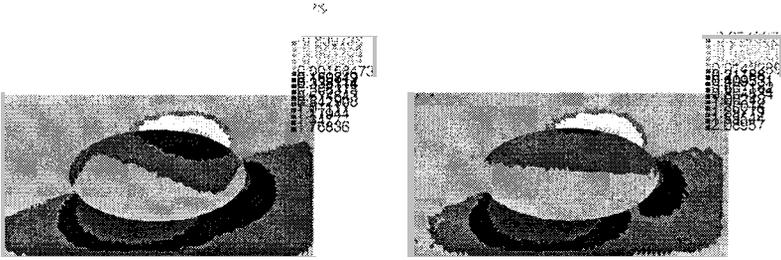


Fig. 2. **Left:** u' . **Right:** u'_ϵ when $\epsilon = 0.0125$. Observe that both solutions are clearly discontinuous across Σ .

§6. Identification of a Discontinuity

Consider an “observed” data u_d on an observation set S and the problem of finding the best κ (i.e. the best Σ) to fit this data. A least square approach leads to

$$\min_{\kappa \in K} \left\{ J(\kappa) = \int_S \|u - u_d\|^2 : -\nabla \cdot (\kappa \nabla u) = 0, \quad u|_{\Gamma} = x + y \right\}.$$

In this example the problem is driven by a non homogeneous boundary condition rather than by a right hand side f .

Obviously a normal change αa in the position of Σ induces a change in J and it is not hard to show that the derivative $J' = dJ/da$ is given by

$$J'(\kappa) = - \int_{\Sigma} \alpha \left[\frac{1}{\kappa} \right] \left(\kappa \frac{\partial u}{\partial n} \right) \left(\kappa \frac{\partial p}{\partial n} \right)$$

with p solution of

$$-\nabla \cdot (\kappa \nabla p) = 2(u - u_d) \delta_S, \quad p|_{\Gamma} = 0.$$

Assume that $\alpha = \alpha(r_1, r_2, \dots, r_m)$. A gradient method on the position of Σ with step size ρ , via the parameters $\{r_i\}_1^m$ would be to change r_i according to

$$r_i \leftarrow r_i + \rho \int_{\Sigma} \frac{\partial \alpha}{\partial r_i} \left[\frac{1}{\kappa} \right] \left(\kappa \frac{\partial u}{\partial n} \right) \left(\kappa \frac{\partial p}{\partial n} \right).$$

(a) We ran a preliminary test by taking

$$\Omega = (-5, 5) \times (-2.5, 2.5), \quad D = \{(x, y) : (x + 2)^2 + y^2 < 1\},$$

$$\Sigma(r_1, r_2) = \{(x, y) : x = (r_1 + r_2 \cos t) \cos t, \quad y = (r_1 + r_2 \cos t) \sin t\}$$

and the reference surface $\Sigma = \Sigma(\sqrt{2}, 0)$. As before $\kappa = 1$ outside Σ and 6 inside. We choose u_d to be the solution of the PDE for κ given by $\Sigma(\sqrt{2}, 0)$. Then we apply the steepest descent method with $\rho = 4$ starting from $\Sigma(0.3, 0.1)$. Fig. 3 shows the convergence curves.

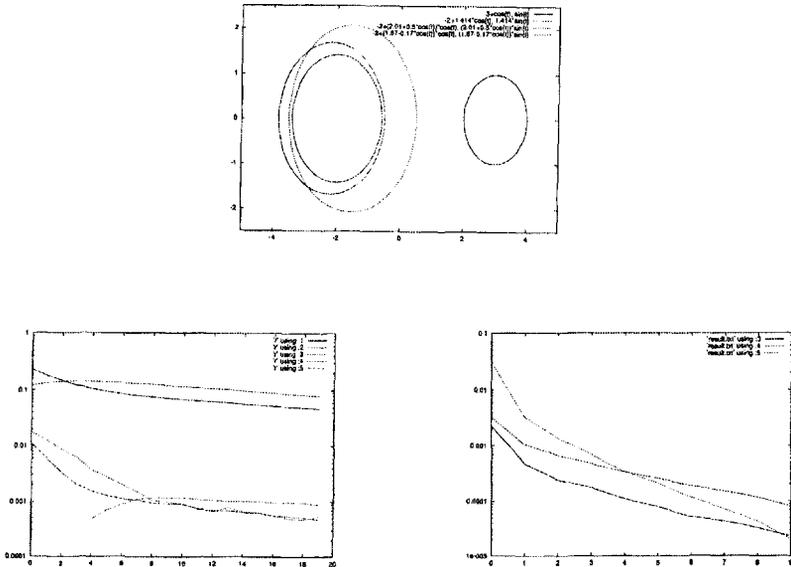


Fig. 3. **Center up:** Geometry of the second example showing the observation set (right circle), the exact and computed solution (inner circle and middle curve) and the initial shape (outer curve).

Bottom left: Convergence curves for the identification problem:

curves 1 and 2 are $r_1 - \sqrt{2}$, r_2 , curves 3 and 4 are $\frac{\partial J}{\partial r_1}$, $\frac{\partial J}{\partial r_2}$

and curve 5 is J ; all tend to zero; the x -axis is the iteration count (from 1 to 20 here). **Bottom right:** Convergence curves showing

the two gradients and the cost function for 10 iterations.

(b) In the previous configuration D intersects Σ . Now we move D to the right and Σ to the left (Fig. 2) and ran the same test with $r_1 = 0.3$ and $r_2 = 0.5$ initially. The method converges but the final shape is close to but different from $\Sigma(\sqrt{2}, 0)$. This is because the numerical approximation does not “see” the right part of Σ which is too far.

(c) In the third test where $r_1 = 0$ and the descent is only on r_2 , the exact solution is reached in 4 iterations. This indicates that the method is sound but a conjugate gradient is needed for test (b).

Acknowledgement. We would like to thank Frédéric Hecht for his help with **freefem++**.

REFERENCES

- [1] Achdou, Y., Bernardi, C., Coquel, F., Archambeau, F. & Gest, B., Discrétisation par volumes finis des équations de Darcy, Rapport EDF HI-83/00/043/A(2001).
- [2] Allaire, G., Shape optimization by the homogenization method, Springer-Verlag, New York, 2001.
- [3] Azaïez, M., Bernardi, C. & Grundmann, M., Méthodes spectrales pour les équations du milieu poreux, *East-West J. Numer. Anal.*, **2**(1994), 91–105.
- [4] Bruaset, A. M. & Nielsen, B. F., On the stability of pressure and velocity computations for heterogeneous reservoirs, *SIAM J. Appl. Math.*, **56**:4(1996), 991–1007.
- [5] Hecht, F., **freefem++**, a general purpose finite element software allowing several types of finite elements in one program, to appear.
- [6] Pironneau, O., Optimal shape design of elliptic systems, Lecture Notes in Computational Physics, Springer-Verlag, 1983.

REITERATED HOMOGENIZATION OF DEGENERATE NONLINEAR ELLIPTIC EQUATIONS

J. BYSTRÖM J. ENGSTRÖM P. WALL

*Department of Mathematics, Luleå University of Technology,
SE-971 87 Luleå, Sweden. E-mail: wall@sm.luth.se*

Abstract

The authors study homogenization of some nonlinear partial differential equations of the form $-\operatorname{div}(a(hx, h^2x, Du_h)) = f$, where a is periodic in the first two arguments and monotone in the third. In particular the case where a satisfies degenerated structure conditions is studied. It is proved that u_h converges weakly in $W_0^{1,1}(\Omega)$ to the unique solution of a limit problem as $h \rightarrow \infty$. Moreover, explicit expressions for the limit problem are obtained.

Keywords Homogenization, Reiterated, Monotone, Degenerated
2000 MR Subject Classification 35B27, 35J70, 74Q99

§1. Introduction

This paper is devoted to homogenization of partial differential operators including several periodically oscillating length scales. This type of equations appear in many fields of physics and engineering sciences where the physical phenomena occur in highly heterogeneous media. One example is heat conduction in composite materials involving two different materials which are periodically distributed. The local characteristics are then described by rapidly oscillating functions. A direct numerical treatment of such problems is often impossible due to the rapidly oscillating functions and one has to apply some type of asymptotic analysis. The branch of mathematics developed for the analysis of these types of problems is known

as homogenization. For more information concerning the homogenization theory, the reader is referred to [1, 2, 7, 10] and [12].

We will now give a short overview of previous results connected to this work and explain what our contribution is. Let us consider the class of partial differential equations of the form

$$-\operatorname{div}(a_h(x, Du_h)) = f \quad \text{on } \Omega, \quad u_h \in W_0^{1,p}(\Omega), \quad (1.1)$$

where a_h is increasingly oscillating as $h \rightarrow \infty$, Ω is an open bounded subset of \mathbf{R}^n , $1 < p < \infty$, $1/p + 1/q = 1$ and $f \in W^{-1,q}(\Omega)$. The homogenization problem for (1.1) consists of the study of the asymptotic behavior of solutions u_h as $h \rightarrow \infty$. In many important cases u_h converges weakly in $W_0^{1,p}(\Omega)$ to the solution u of the homogenized problem

$$-\operatorname{div}(b(Du)) = f \quad \text{on } \Omega, \quad u \in W_0^{1,p}(\Omega).$$

In [6] and [11] the following situation was studied: a_h is of the form $a_h(x, \xi) = a(hx, \xi)$, where a is monotone, continuous and satisfies suitable coerciveness and growth conditions in the second argument and is periodic in the first argument. A corresponding homogenization result, with the difference that a only satisfies degenerate structure conditions, was obtained in [9]. In this situation it is natural to work with weighted spaces which means that instead of (1.1) we have

$$-\operatorname{div}(a(hx, Du_h)) = f \quad \text{on } \Omega, \quad u_h \in W_0^{1,p}(\Omega, \lambda_h),$$

where (λ_h) is a sequence of periodic weights.

In the case when $a_h(x, \xi) = a(hx, h^2x, \xi)$, where a is periodic in the first two variables, one speaks about reiterated homogenization. This concept was introduced by Bensoussan, Lions and Papanicolaou in [1], where it was stated a result for linear operators. Concerning reiterated homogenization of nonlinear problems we refer to [13] and [14]. One important application of reiterated homogenization is that it has been an indispensable tool in the construction of structures with extreme effective material properties. Concerning this topic we refer to the collection of classical papers in [5], where the introduction gives a good selection of references. We remark that some of the homogenization problems above also have been studied by Γ -convergence for the corresponding variational problems and by two-scale convergence, but leave out this discussion since it is out of the scope of this work.

In this paper we study reiterated homogenization where a only satisfies degenerate structure conditions. More precisely we prove that the solutions

u_h of

$$-\operatorname{div}(a(hx, h^2x, Du_h)) = f \quad \text{on } \Omega, \quad u_h \in W_0^{1,p}(\Omega, \lambda_h),$$

converges weakly to u in $W_0^{1,1}(\Omega)$, where u is the solution of a homogenized problem

$$-\operatorname{div}(b(Du)) = f \quad \text{on } \Omega, \quad u \in W_0^{1,p}(\Omega).$$

This paper is organized in the following way: In Section 2 we fix some notation and present necessary preliminary results. Section 3 contains the homogenization result described above, which also is the main result of this paper. In Section 4 we derive a homogenization result for an auxiliary problem. A key ingredient in the proof of the main result is that the solutions of the auxiliary problem are used to define a special type of test function. Finally, in Section 5 we give some properties of the homogenized operator b .

§2. Preliminaries and Notation

Let Ω be a regular bounded open subset of \mathbf{R}^n and $|E|$ denote the Lebesgue measure of the set E in \mathbf{R}^n . Moreover let $\langle \cdot, \cdot \rangle$ denote the Euclidean scalar product on \mathbf{R}^n and χ_E the characteristic function of the set E . Let p be a real constant $1 < p < \infty$ and let q be its conjugate exponent, $1/p + 1/q = 1$. We will denote by C and C_i constants that may change from one place to another.

Furthermore, let $Y = Z = (0, 1)^n$ be the unit cube in \mathbf{R}^n . Let $\{\Omega_i \subset Y : i = 1, \dots, N\}$ be a family of disjoint open sets such that $|Y \setminus \bigcup_{i=1}^N \Omega_i| = 0$ and $|\partial\Omega_i| = 0$.

Let λ be a weight on \mathbf{R}^n , i.e. λ is measurable and

$$\lambda > 0 \text{ a.e., } \lambda \text{ and } \lambda^{-1/(p-1)} \text{ are in } L_{\text{loc}}^1(\mathbf{R}^n). \quad (2.1)$$

We denote by $L^p(\Omega, \lambda)$ the set of real functions u in $L_{\text{loc}}^1(\Omega)$ such that $u\lambda^{1/p}$ is in $L^p(\Omega)$, by $W^{1,p}(\Omega, \lambda)$ the set of the functions u in $W_{\text{loc}}^{1,1}(\Omega)$ such that $u \in L^p(\Omega, \lambda)$ and $Du \in [L^p(\Omega, \lambda)]^n$. Moreover, we denote by $W_0^{1,p}(\Omega, \lambda)$ the completion of $C_0^1(\Omega)$ with respect to the norm in $W^{1,p}(\Omega, \lambda)$, i.e.

$$\|u\|_{W^{1,p}(\Omega, \lambda)} = \left(\int_{\Omega} (|u|^p + |Du|^p) \lambda \, dx \right)^{1/p}.$$

By $C_{\text{per}}^1(Y)$ we mean the set of all Y -periodic functions in $C^1(\mathbf{R}^n)$ with mean value zero. We also define $W_{\text{per}}^{1,p}(\Omega, \lambda)$ as the set of real functions u in $W_{\text{loc}}^{1,1}(\mathbf{R}^n)$ with mean value zero such that u is Y -periodic and $u \in W^{1,p}(Y, \lambda)$.

We now define the Muckenhoupt A_p class:

Definition 2.1. Let $p > 1$, $K \geq 1$ and let λ be a weight on \mathbf{R}^n . Then λ is in the class $A_p(K)$ if

$$\left(\frac{1}{|Q|} \int_Q \lambda \, dz \right) \left(\frac{1}{|Q|} \int_Q \lambda^{-\frac{1}{p-1}} \, dz \right)^{p-1} \leq K$$

for every cube $Q \in \mathbf{R}^n$ with faces parallel to the coordinate planes. We set $A_p = \bigcup_{K \geq 1} A_p(K)$.

Let λ_i be a Y -periodic weight on \mathbf{R}^n , i.e. λ_i satisfies (2.1) and is Y -periodic. We define the weights $\bar{\lambda}_h$ and λ_h as

$$\bar{\lambda}_h(x) = \sum_{i=1}^N \chi_{\Omega_i}(x) \lambda_i(hx), \quad \lambda_h(x) = \sum_{i=1}^N \chi_{\Omega_i}(hx) \lambda_i(h^2x). \quad (2.2)$$

Then it follows that $\bar{\lambda}_h, \bar{\lambda}_h^{-1/(p-1)}, \lambda_h$ and $\lambda_h^{-1/(p-1)}$ all are in $L^1_{\text{loc}}(\mathbf{R}^n)$. Moreover, we assume that $\lambda_i, \bar{\lambda}_h$ and λ_h are in $A_p(K)$ for some K .

Let $a : \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a function such that

$$a(y, z, \xi) = \sum_{i=1}^N \chi_{\Omega_i}(y) a_i(z, \xi). \quad (2.3)$$

We assume that $a(\cdot, z, \xi)$ is Y -periodic and $a(y, \cdot, \xi)$ is Z -periodic. We also assume that a satisfies certain continuity and monotonicity conditions. To be more specific, assume that there exist constants $c_1^i, c_2^i > 0$ and constants α and β with $0 \leq \alpha \leq \min(1, p-1)$ and $\max(p, 2) \leq \beta < \infty$ such that

$$|a_i(z, \xi_1) - a_i(z, \xi_2)| \leq c_1^i \lambda_i(z) (1 + |\xi_1| + |\xi_2|)^{p-1-\alpha} |\xi_1 - \xi_2|^\alpha, \quad (2.4)$$

$$\langle a_i(z, \xi_1) - a_i(z, \xi_2), \xi_1 - \xi_2 \rangle \geq c_2^i \lambda_i(z) (1 + |\xi_1| + |\xi_2|)^{p-\beta} |\xi_1 - \xi_2|^\beta, \quad (2.5)$$

for a.e. $z \in \mathbf{R}^n$, every $\xi \in \mathbf{R}^n$. Moreover we assume that

$$a_i(z, 0) = 0 \quad (2.6)$$

for a.e. $z \in \mathbf{R}^n$.

As a direct consequence of (2.4), (2.5), and (2.6) the following inequalities hold:

$$|a_i(z, \xi)| \leq c_a^i \lambda_i(z) \left(1 + |\xi|^{p-1} \right), \quad (2.7)$$

$$\lambda_i(z) |\xi|^p \leq c_b^i (\lambda_i(z) + \langle a_i(z, \xi), \xi \rangle), \quad (2.8)$$

$$\int_Z |\xi + Dv_i^\xi|^p \lambda_i(z) \, dz \leq c_c^i (1 + |\xi|^p). \quad (2.9)$$

In [8] the following result is proved.

Lemma 2.1. *Let $p > 1$ and $K \geq 1$. Then there exist two positive constants $\delta = \delta(n, p, K)$ and $C = C(n, p, K)$ such that*

$$\left(\frac{1}{|Q|} \int_Q \lambda^{1+\delta} dy \right)^{\frac{1}{1+\delta}} \leq C \frac{1}{|Q|} \int_Q \lambda dy, \quad (2.10)$$

$$\left(\frac{1}{|Q|} \int_Q \lambda^{-(1+\delta)/(p-1)} dy \right)^{\frac{1}{1+\delta}} \leq C \frac{1}{|Q|} \int_Q \lambda^{-1/(p-1)} dy, \quad (2.11)$$

for every cube with faces parallel to the coordinate planes and every $\lambda \in A_p(K)$.

In [9] the following weighted compensated compactness result is proved.

Lemma 2.2. *Let $\nu \in A_p$, $K \geq 1$, let (λ_h) be a family in $A_p(K)$ and let Ω be an open bounded set. Let (u_h) be a family of functions satisfying*

(1) $\int_{\Omega} |Du_h|^p \lambda_h dy \leq C_1 < \infty$ for every $h \in \mathbf{N}$,

(2) there exists a function $u \in W^{1,p}(\Omega, \nu)$ such that $u_h \rightarrow u$ in $L^1(\Omega)$.

Moreover, let (a_h) be a family of vector functions in \mathbf{R}^n such that

(3) $\int_{\Omega} |a_h|^q \lambda_h^{-1/(p-1)} dy \leq C_2 < \infty$ for every $h \in \mathbf{N}$,

(4) $\operatorname{div}(a_h) = f \in L^\infty(\Omega)$ on $C_0^1(\Omega)$ for every $h \in \mathbf{N}$,

(5) there exists $a \in [L^q(\Omega, \nu^{-1/(p-1)})]^n$ such that $a_h \rightarrow a$ weakly in $[L^1(\Omega)]^n$.

Then

$$\int_{\Omega} \langle a_h, Du_h \rangle \phi dy \rightarrow \int_{\Omega} \langle a, Du \rangle \phi dy$$

for every $\phi \in C_0^\infty(\Omega)$.

In [15] the following convergence result for periodic functions is proved.

Lemma 2.3. *Let $1 \leq p \leq \infty$ and let $u_h \in L_{\text{loc}}^p(\mathbf{R}^n)$ be Y -periodic for $h \in \mathbf{N}$. Moreover, suppose that $u_h \rightarrow u$ weakly in $L^p(Y)$ (weakly* if $p = \infty$) as $h \rightarrow \infty$. Let w_h be defined by $w_h(x) = u_h(hx)$. Then as $h \rightarrow \infty$ it holds that $w_h \rightarrow \frac{1}{|Y|} \int_Y u(y) dy$ weakly in $L^p(\Omega)$ (weakly* if $p = \infty$).*

We end this section with a simple extension lemma.

Lemma 2.4. *Let λ be a Y -periodic weight on \mathbf{R}^n ; let $g : Y \rightarrow \mathbf{R}^n$ be a function such that*

$$g \in [L^q(Y, \lambda^{-1/(p-1)})]^n, \quad \int_Y \langle g, Dw \rangle dy = 0, \quad \forall w \in W_{\text{per}}^{1,p}(Y, \lambda),$$

and let \tilde{g} be the Y -periodic extension to \mathbf{R}^n of g . Then we have

$$\tilde{g} \in [L_{\text{loc}}^q(\mathbf{R}^n, \lambda^{-1/(p-1)})]^n, \quad \int_{\mathbf{R}^n} \langle \tilde{g}, Dv \rangle dy = 0, \quad \forall v \in C_0^1(\mathbf{R}^n).$$

In situations where no confusion can occur we will use the same notation for the extended function as for the original one.

§3. The Main Theorem

Let us consider the following Dirichlet problems:

$$\begin{cases} \int_{\Omega} \langle a(hx, h^2x, Du_h), D\phi \rangle dx = \int_{\Omega} f\phi dx, \quad \forall \phi \in W_0^{1,p}(\Omega, \lambda_h), \\ u_h \in W_0^{1,p}(\Omega, \lambda_h), \end{cases} \quad (3.1)$$

where $f \in L^\infty(\Omega)$. By standard results in existence theory there exist unique solutions for each h . Below we state the main result of this paper.

Theorem 3.1. *Let (u_h) be the solutions of (3.1). Then*

$$\begin{aligned} u_h &\rightarrow u \text{ weakly in } W_0^{1,1}(\Omega), \\ a(hx, h^2x, Du_h) &\rightarrow b(Du) \text{ weakly in } [L^1(\Omega)]^n, \end{aligned}$$

as $h \rightarrow \infty$, where u is the unique solution of

$$\begin{cases} \int_{\Omega} \langle b(Du), D\phi \rangle dx = \int_{\Omega} f\phi dx \text{ for every } \phi \in W_0^{1,p}(\Omega), \\ u \in W_0^{1,p}(\Omega). \end{cases} \quad (3.2)$$

The operator $b: \mathbf{R}^n \rightarrow \mathbf{R}^n$ is defined as

$$b(\xi) = \frac{1}{|Y|} \int_Y b_Y(y, \xi + Du^\xi(y)) dy,$$

where u^ξ is the unique solution of the Y -cell problem

$$\begin{cases} \int_Y \langle b_Y(y, \xi + Du^\xi(y)), D\phi \rangle dy = 0 \text{ for every } \phi \in W_{\text{per}}^{1,p}(Y), \\ u^\xi \in W_{\text{per}}^{1,p}(Y). \end{cases}$$

The operator $b_Y: Y \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ is defined as $b_Y(y, \xi) = \sum_{i=1}^N \chi_{\Omega_i}(y) b_i(\xi)$,

where

$$b_i(\xi) = \frac{1}{|Z|} \int_Z a_i(z, \xi + Dv_i^\xi(z)) dz,$$

and v_i^ξ are the unique solutions of the Z -cell problems

$$\begin{cases} \int_Z \langle a_i(z, \xi + Dv_i^\xi(z)), D\phi(z) \rangle dz = 0 \text{ for every } \phi \in W_{\text{per}}^{1,p}(Z, \lambda_i), \\ v_i^\xi \in W_{\text{per}}^{1,p}(Z, \lambda_i). \end{cases}$$

Proof. Let us first prove that $\|Du_h\|_{L^p(\Omega, \lambda_h)} \leq C$. By periodicity it follows that

$$\int_{\Omega} \lambda_h(x) dx \leq C, \quad \int_{\Omega} (\lambda_h(x))^{-1/(p-1)} dx \leq C. \quad (3.3)$$

Thus, by (2.8), (3.1), (3.3), Poincaré's and Hölder's inequalities we obtain

$$\begin{aligned} \int_{\Omega} |Du_h|^p \lambda_h(x) dx &\leq c_b \left(\int_{\Omega} \lambda_h(x) dx + \int_{\Omega} f u_h dx \right) \\ &\leq C \left(\int_{\Omega} \lambda_h(x) dx + \|f\|_{L^\infty(\Omega)} \int_{\Omega} |Du_h| dx \right) \\ &\leq C \left(1 + \left(\int_{\Omega} |Du_h|^p \lambda_h(x) dx \right)^{1/p} \right). \end{aligned}$$

Hence it is clear that

$$\|Du_h\|_{L^p(\Omega, \lambda_h)} \leq C. \quad (3.4)$$

Let η_h be defined as $\eta_h = a(hx, h^2x, Du_h)$. Then (2.7) and (3.4) implies

$$\|\eta_h\|_{[L^q(\Omega, \lambda_h^{-1/(p-1)})]^n} \leq C. \quad (3.5)$$

Take $\delta > 0$ such that (2.11) holds. Now choose σ_1 such that

$$\frac{1 + \sigma_1}{p - 1 - \sigma_1} = \frac{1 + \delta}{p - 1}.$$

Then $\sigma_1 > 0$ and $p - 1 - \sigma_1 > 0$. Let Q be a cube in \mathbf{R}^n containing Ω . Hölder's inequality and (3.4) then gives

$$\begin{aligned} \int_{\Omega} |Du_h|^{1+\sigma_1} dx &\leq \left(\int_{\Omega} |Du_h|^p \lambda_h dx \right)^{\frac{1+\sigma_1}{p}} \left(\int_{\Omega} \lambda_h^{-\frac{1+\sigma_1}{p-1-\sigma_1}} dx \right)^{\frac{p-1-\sigma_1}{p}} \\ &\leq C \left(\int_{\Omega} \lambda_h^{-\frac{1+\delta}{p-1}} dx \right)^{\frac{p-1}{p+\delta}} \leq C \left(\int_Q \lambda_h^{-\frac{1+\delta}{p-1}} dx \right)^{\frac{p-1}{p+\delta}}. \end{aligned}$$

By applying (2.11) and (3.3) in the inequality above, we obtain

$$\int_{\Omega} |Du_h|^{1+\sigma_1} dx \leq C \left(\int_Q \lambda_h^{1/(1-p)} dx \right)^{(p-1)(1+\delta)/(p+\delta)} \leq C. \quad (3.6)$$

Next, choose $\delta > 0$ such that (2.10) holds and choose σ_2 such that

$$1 + \delta = (1 + \sigma_2) \frac{q - 1}{q - 1 - \sigma_2}.$$

Then $\sigma_2 > 0$ and $q - 1 - \sigma_2 > 0$. By using (2.10) and arguing similarly as for (3.6) we obtain

$$\int_{\Omega} |\eta_h|^{1+\sigma_2} dx \leq C.$$

This means that (u_h) and (η_h) are bounded in $W_0^{1,1+\sigma_1}(\Omega)$ and $[L^{1+\sigma_2}(\Omega)]^n$ respectively. Since these spaces are reflexive, we have that

there exist subsequences, still denoted by (u_h) and (η_h) , such that

$$u_h \rightarrow u_* \text{ weakly in } W_0^{1,1+\sigma_1}(\Omega), \quad (3.7)$$

$$\eta_h \rightarrow \eta_* \text{ weakly in } [L^{1+\sigma_2}(\Omega)]^n. \quad (3.8)$$

From (3.7) and (3.8) it follows that

$$u_h \rightarrow u_* \text{ weakly in } W_0^{1,1}(\Omega), \quad (3.9)$$

$$\eta_h \rightarrow \eta_* \text{ weakly in } [L^1(\Omega)]^n. \quad (3.10)$$

From our original problem (3.1) we have

$$\int_{\Omega} \langle \eta_h, D\phi \rangle dx = \int_{\Omega} f\phi dx \text{ for every } \phi \in W_0^{1,p}(\Omega, \lambda_h). \quad (3.11)$$

By using the fact $C_0^1(\Omega) \subset W_0^{1,p}(\Omega, \lambda_h)$ and (3.10) we can pass to the limit in (3.11), thus

$$\int_{\Omega} \langle \eta_*, D\phi \rangle dx = \int_{\Omega} f\phi dx \text{ for every } \phi \in C_0^1(\Omega).$$

Density and the fact that $\eta_* \in [L^q(\Omega)]^n$ (see (3.17)) then gives

$$\int_{\Omega} \langle \eta_*, D\phi \rangle dx = \int_{\Omega} f\phi dx \text{ for every } \phi \in W_0^{1,p}(\Omega). \quad (3.12)$$

Let us now observe that by (3.9), (3.10) and (3.12) the theorem is proved if we show that

$$u_* \in W_0^{1,p}(\Omega), \quad (3.13)$$

$$\eta_* = b(Du_*) \text{ a.e. on } \Omega, \quad (3.14)$$

since the uniqueness of the solution of the homogenized problem (3.2) then implies that $u_* = u$ a.e. on Ω .

We start with the proof of (3.13). We observe that since Ω is a regular bounded open set it is sufficient to show that $Du_* \in [L^p(\Omega)]^n$. Let $\phi \in C_0(\Omega)$. Then Hölder's inequality and (3.4) gives

$$\begin{aligned} \|Du_h\phi\|_{[L^1(\Omega)]^n} &\leq \left(\int_{\Omega} |Du_h|^p \lambda_h(x) dx \right)^{\frac{1}{p}} \left(\int_{\Omega} (\lambda_h(x))^{-\frac{1}{p-1}} |\phi|^q dx \right)^{\frac{1}{q}} \\ &\leq C \left(\int_{\Omega} (\lambda_h(x))^{-\frac{1}{p-1}} |\phi|^q dx \right)^{\frac{1}{q}}. \end{aligned} \quad (3.15)$$

Applying \liminf on both sides of (3.15) and using the weak lower semicontinuity of the norm on the left hand side and periodicity on the right hand side we obtain

$$\int_{\Omega} |Du_*| |\phi| dx \leq C \|\phi\|_{L^q(\Omega)} \text{ for every } \phi \in C_0(\Omega).$$

By density and Landau's theorem we then have that

$$Du_* \in [L^p(\Omega)]^n. \quad (3.16)$$

By using (3.5) and arguments similar to those employed in the proof of (3.16) it can also be deduced that

$$\eta_* \in [L^q(\Omega)]^n. \quad (3.17)$$

It remains to prove (3.14). For this purpose let us define the test function

$$w_h^\xi(x) = (\xi, x) + \frac{1}{h} u_h^\xi(hx),$$

where u_h^ξ is defined as in the auxiliary problem (see Section 4). To be able to apply the compensated compactness result (Lemma 2.2) we have to prove certain facts about w_h^ξ and $a(hx, h^2x, Dw_h^\xi)$. Therefore, by periodicity, (4.5) and the fact that $\lambda_h(x) = \bar{\lambda}_h(hx)$ we get

$$\int_{\Omega} \left| Dw_h^\xi \right|^p \lambda_h(x) dx \leq C. \quad (3.18)$$

Moreover, by using (2.7) and (3.18) we obtain

$$\int_{\Omega} |a(hx, h^2x, Dw_h^\xi)|^q (\lambda_h(x))^{-1/(p-1)} dx \leq C.$$

By periodicity and Lemma 2.3 we have that

$$\begin{aligned} w_h^\xi(\cdot) &\rightarrow (\xi, \cdot) \text{ in } L^1(\Omega), \\ a(hx, h^2x, Dw_h^\xi) &\rightarrow \frac{1}{|Y|} \int_Y b_Y(y, \xi + Du^\xi) dy = b(\xi) \\ &\text{weakly in } [L^1(\Omega)]^n. \end{aligned}$$

Finally, due to (4.6), we can apply Lemma 2.4 on (4.1) and obtain

$$\operatorname{div} \left(a(hx, h^2x, Dw_h^\xi) \right) = 0 \text{ on } C_0^1(\Omega),$$

and we are now ready to apply the compensated compactness result. Indeed, by the monotonicity of a we have for a fixed ξ that

$$\int_{\Omega} \left\langle \eta_h - a(hx, h^2x, Dw_h^\xi(x)), Du_h(x) - Dw_h^\xi(x) \right\rangle \phi(x) dx \geq 0$$

for every $\phi \in C_0^\infty(\Omega)$, $\phi \geq 0$. By the compensated compactness lemma (Lemma 2.2) with the weight $\nu = 1$, we get in the limit

$$\int_{\Omega} \langle \eta_* - b(\xi), Du_*(x) - \xi \rangle \phi(x) dx \geq 0$$

for every $\phi \in C_0^\infty(\Omega)$, $\phi \geq 0$. Hence for our fixed $\xi \in \mathbf{R}^n$ we have that

$$\langle \eta_* - b(\xi), Du_*(x) - \xi \rangle \geq 0 \text{ for a.e. } x \in \Omega.$$

By density and the continuity of b (see Lemma 5.2), it follows that

$$\langle \eta_* - b(\xi), Du_*(x) - \xi \rangle \geq 0 \text{ for a.e. } x \in \Omega \text{ and every } \xi \in \mathbf{R}^n.$$

Since b is monotone (5.3) and continuous (5.4), we have that b is maximal monotone and hence (3.14) follows. Finally, let us observe that we have proved the theorem only up to a subsequence, but since the homogenized operator is uniquely defined and the solution of the homogenized problem is unique we can conclude that the theorem holds for the whole sequence.

§4. An Auxiliary Problem

In this section we prove a homogenization result for the auxiliary problem. This result was used in the definition of the special type of test functions defined in the proof of the main result (Theorem 3.1) of this paper.

Fix ξ and consider the following Dirichlet problems:

$$\begin{cases} \int_Y \langle a(y, hy, \xi + Du_h^\xi), D\phi \rangle dy = 0, \forall \phi \in W_{\text{per}}^{1,p}(Y, \bar{\lambda}_h), \\ u_h^\xi \in W_{\text{per}}^{1,p}(Y, \bar{\lambda}_h). \end{cases} \quad (4.1)$$

By standard results in existence theory there exist unique solutions for each h . Below we state the auxiliary result of this paper.

Theorem 4.1. *Let (u_h^ξ) be the solutions of (4.1). We then have that*

$$u_h^\xi \rightarrow u^\xi \text{ weakly in } W_{\text{per}}^{1,1}(Y),$$

$$a(y, hy, \xi + Du_h^\xi) \rightarrow b_Y(y, \xi + Du^\xi) \text{ weakly in } [L^1(Y)]^n,$$

as $h \rightarrow \infty$, where u^ξ is the unique solution of

$$\begin{cases} \int_Y \langle b_Y(y, \xi + Du^\xi), D\phi \rangle dy = 0 \text{ for every } \phi \in W_{\text{per}}^{1,p}(Y), \\ u^\xi \in W_{\text{per}}^{1,p}(Y). \end{cases} \quad (4.2)$$

The operator $b_Y : Y \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ is defined as $b_Y(y, \tau) = \sum_{i=1}^N \chi_{\Omega_i}(y) b_i(\tau)$,

where

$$b_i(\tau) = \frac{1}{|Z|} \int_Z a_i(z, \tau + Dv_i^\tau(z)) dz$$

and v_i^τ are the unique solutions of the Z -cell problems

$$\begin{cases} \int_Z \langle a_i(z, \tau + Dv_i^\tau(z)), D\phi(z) \rangle dz = 0, \forall \phi \in W_{\text{per}}^{1,p}(Z, \lambda_i), \\ v_i^\tau \in W_{\text{per}}^{1,p}(Z, \lambda_i). \end{cases} \quad (4.3)$$

Proof. By (2.8), (4.1) and (2.7) we have that

$$\begin{aligned} & \int_Y |\xi + Du_h^\xi|^p \bar{\lambda}_h(y) dy \\ & \leq C_b \left(\int_Y \bar{\lambda}_h(y) dy + C_a \int_Y \left(1 + |\xi + Du_h^\xi|^{p-1}\right) \bar{\lambda}_h(y) |\xi| dy \right). \end{aligned}$$

Moreover, if we use Young's inequality on the right hand side and rearrange the resulting inequality we obtain

$$\begin{aligned} & \left(1 - \frac{C_a C_b 2^{q-1} \eta}{q}\right) \int_Y |\xi + Du_h^\xi|^p \bar{\lambda}_h(y) dy \\ & \leq C_b \left(1 + \frac{C_a 2^{q-1} \eta}{q} + \frac{C_a |\xi|^p \eta^{-(p-1)}}{p}\right) \int_Y \bar{\lambda}_h(y) dy, \end{aligned}$$

where η is a positive real number. By choosing η small enough we get that

$$\int_Y |\xi + Du_h^\xi|^p \bar{\lambda}_h(y) dy \leq C. \quad (4.4)$$

In particular this implies

$$\int_Y |Du_h^\xi|^p \bar{\lambda}_h(y) dy \leq C. \quad (4.5)$$

Let us define $\eta_h^i = a_i(hy, \xi + Du_h^\xi)$. By using (2.7) and (4.4), it follows that

$$\int_{\Omega_i} |\eta_h^i|^q (\lambda_i(hy))^{-1/(p-1)} dy \leq C. \quad (4.6)$$

Moreover, by using (2.11), (4.5), (2.10), (4.6) and arguments similar to those employed in the proof of (3.6), it can be deduced that

$$\int_Y |Du_h^\xi|^{1+\kappa_1} dy \leq C, \quad \int_{\Omega_i} |\eta_h^i|^{1+\kappa_2} dy \leq C.$$

Thus we have that (u_h^ξ) and (η_h^i) are bounded in $W_{\text{per}}^{1,1+\kappa_1}(Y)$ and $[L^{1+\kappa_2}(\Omega_i)]^n$ respectively. Since these spaces are reflexive, there exist subsequences, still denoted by (u_h^ξ) and (η_h^i) , such that

$$u_h^\xi \rightarrow u_*^\xi \text{ weakly in } W_{\text{per}}^{1,1+\kappa_1}(Y), \quad \eta_h^i \rightarrow \eta_*^i \text{ weakly in } [L^{1+\kappa_2}(\Omega_i)]^n.$$

Hence we can conclude that

$$u_h^\xi \rightarrow u_*^\xi \text{ weakly in } W_{\text{per}}^{1,1}(Y), \quad \eta_h^i \rightarrow \eta_*^i \text{ weakly in } [L^1(\Omega_i)]^n. \quad (4.7)$$

Using similar ideas as in the proof of (3.16), it can be shown that

$$\eta_*^i \in [L^q(Y)]^n. \quad (4.8)$$

From (4.1) and (4.7) it follows that

$$\sum_{i=1}^N \int_{\Omega_i} \langle \eta_*^i, D\phi \rangle dy = 0 \text{ for every } \phi \in C_{\text{per}}^1(Y).$$

Density arguments in conjunction with (4.8) then results in

$$\sum_{i=1}^N \int_{\Omega_i} \langle \eta_*^i, D\phi \rangle dy = 0 \text{ for every } \phi \in W_{\text{per}}^{1,p}(Y).$$

Thus the theorem is proved if we show that

$$u_*^\xi \in W_{\text{per}}^{1,p}(Y), \quad (4.9)$$

$$\eta_*^i = b_i(\xi + Du_*^\xi) \text{ a.e. on } \Omega_i, \quad (4.10)$$

since the uniqueness of the solution of the homogenized problem (4.2) then implies that $u_*^\xi = u^\xi$ a.e. on Y .

Let us start with (4.9). We observe that since Y is a regular bounded open set it is sufficient to show that $Du_*^\xi \in [L^p(Y)]^n$ which is obtained using the same ideas as in the proof of (3.16).

It remains to prove (4.10). Therefore let us define the test function $w_h^{\tau,i}$ by

$$w_h^{\tau,i}(y) = (\tau, y) + \frac{1}{h} v_i^\tau(hy), \quad (4.11)$$

where $v_i^\tau \in W_{\text{per}}^{1,p}(Z, \lambda_i)$ is defined as in (4.3). To be able to apply the compensated compactness result (Lemma 2.2) we have to prove certain facts about $w_h^{\tau,i}$ and $a_i(hy, Dw_h^{\tau,i})$. Indeed,

$$\int_{\Omega_i} |Dw_h^{\tau,i}|^p \lambda_i(hy) dy \leq C \quad (4.12)$$

follows from (4.11). Moreover, we also have that

$$\int_{\Omega_i} |a_i(hy, Dw_h^{\tau,i})|^q (\lambda_i(hy))^{-1/(p-1)} dy \leq C$$

by (2.7) and (4.12). By periodicity we obtain

$$\begin{aligned} w_h^{\tau,i}(\cdot) &\rightarrow (\tau, \cdot) \text{ strongly in } L^1(\Omega_i), \\ a_i(hy, Dw_h^{\tau,i}) &\rightarrow \frac{1}{|Z|} \int_Z a_i(z, \tau + Dv_i^\tau(z)) dz = b_i(\tau) \\ &\text{weakly in } [L^1(\Omega_i)]^n. \end{aligned}$$

Application of Lemma 2.4 on (4.3) gives $\text{div}(a_i(hy, Dw_h^{\tau,i})) = 0$ on $C_0^1(\Omega_i)$.

By the monotonicity of a_i we have for a fixed τ that

$$\int_{\Omega_i} \langle \eta_*^i - a_i(hy, Dw_h^{\tau, i}(y)), \xi + Du_h^\xi(y) - Dw_h^{\tau, i}(y) \rangle \phi(y) dy \geq 0$$

for every $\phi \in C_0^\infty(\Omega_i)$, $\phi \geq 0$. By the compensated compactness lemma (Lemma 2.2) with $\nu = 1$, we then get in the limit

$$\int_{\Omega_i} \langle \eta_*^i - b_i(\tau), \xi + Du_*^\xi(y) - \tau \rangle \phi(y) dy \geq 0$$

for every $\phi \in C_0^\infty(\Omega_i)$, $\phi \geq 0$. Hence for our fixed $\tau \in \mathbf{R}^n$ we have that

$$\langle \eta_*^i - b_i(\tau), \xi + Du_*^\xi(y) - \tau \rangle \geq 0 \text{ for a.e. } y \in \Omega_i.$$

By density and the continuity of b_i (5.2), it follows that

$$\langle a_* - b_i(\tau), \xi + Du_*^\xi(y) - \tau \rangle \geq 0 \text{ for a.e. } y \in \Omega_i \text{ and every } \tau \in \mathbf{R}^n.$$

Since b_i is monotone (5.1) and continuous (5.2), we have that b_i is maximal monotone and hence (4.10) follows.

§5. Properties of the Homogenized Operators b_i and b

In this section we list some properties of the homogenized operators b_i and b . In particular these properties imply the existence and uniqueness of the solution of the homogenized problem (in the auxiliary and main problem respectively).

Lemma 5.1. *Let b_i be the homogenized operator defined in Theorem 4.1. Then*

(a) $b_i(\cdot)$ is strictly monotone. In particular, we have that

$$\langle b_i(\tau_1) - b_i(\tau_2), \tau_1 - \tau_2 \rangle \geq \tilde{c}_2 (1 + |\tau_1| + |\tau_2|)^{p-\beta} |\tau_1 - \tau_2|^\beta \tag{5.1}$$

for every $\tau_1, \tau_2 \in \mathbf{R}^n$.

(b) $b_i(\cdot)$ is continuous. In particular, we have for $\gamma = \frac{\alpha}{\beta-\alpha}$ that

$$|b_i(\tau_1) - b_i(\tau_2)| \leq \tilde{c}_1 (1 + |\tau_1| + |\tau_2|)^{p-1-\gamma} |\tau_1 - \tau_2|^\gamma \tag{5.2}$$

for every $\tau_1, \tau_2 \in \mathbf{R}^n$.

(c) $b_i(0) = 0$.

Proof. These properties follow by using the same ideas as in [3] and [4].

Lemma 5.2. *Let b be the homogenized operator defined in Theorem 3.1.*

Then

(a) $b(\cdot)$ is strictly monotone. In particular, we have that

$$\langle b(\xi_1) - b(\xi_2), \xi_1 - \xi_2 \rangle \geq \tilde{C}_2 (1 + |\xi_1| + |\xi_2|)^{p-\beta} |\xi_1 - \xi_2|^\beta \tag{5.3}$$

for every $\xi_1, \xi_2 \in \mathbf{R}^n$.

(b) $b(\cdot)$ is continuous. In particular, we have for $\delta = \frac{\gamma}{\beta-\gamma} = \frac{\alpha}{(\beta-\alpha)\beta-\alpha}$ that

$$|b(\xi_1) - b(\xi_2)| \leq \tilde{C}_1 (1 + |\xi_1| + |\xi_2|)^{p-1-\delta} |\xi_1 - \xi_2|^\delta \quad (5.4)$$

for every $\xi_1, \xi_2 \in \mathbf{R}^n$.

(c) $b(0) = 0$.

Proof. These properties follow by using (5.1) and (5.2) in the corresponding theorem given in for example [3] and [4].

REFERENCES

- [1] Bensoussan, A., Lions, J. L. & Papanicolaou, G., Asymptotic analysis for periodic structures, North Holland, Amsterdam, 1978.
- [2] Braides, A. & Defranceschi, A., Homogenization of multiple integrals, Oxford University Press, New York, 1998.
- [3] Byström, J., Correctors for some nonlinear monotone operators, Research report, No. 11, ISSN: 1400-4003, Department of Mathematics, Luleå University of Technology, 1999.
- [4] Byström, J., Correctors for some nonlinear monotone operators, *J. Nonlinear Math. Phys.*, **8**:1(2001), 8-30.
- [5] Cherkaev, A. & Kohn, R., Topics in the mathematical modelling of composite materials, Birkhäuser, Boston, 1997.
- [6] Chiado Piat, V. & Defranceschi, A., Homogenization of monotone operators, *Nonlinear Anal.*, **14**:9(1990), 717-732.
- [7] Cioranescu, D. & Donato, P., An introduction to homogenization, Oxford University Press, Oxford, 1999.
- [8] Coifman, R. R. & Fefferman, C., Weighted norm inequalities for maximal function and singular integrals, *Studia Math.*, **51**(1974), 241-250.
- [9] De Arcangelis, R. & Serra Cassano, F., On the homogenization of degenerate elliptic equations in divergence form, *J. Math. Pures Appl.*, **71**(1992), 119-138.
- [10] Defranceschi, A., An introduction to homogenization and G -convergence, Lecture notes, School on homogenization, ICTP, Trieste, 1993.
- [11] Fusco, N. & Moscariello, G., On the homogenization of quasilinear divergence structure operators, *Annali Mat. Pura Appl.*, **146**(1987), 1-13.
- [12] Jikov, V., Kozlov, S. & Oleinik, O., Homogenization of differential operators and integral functionals, Springer-Verlag, Berlin-Heidelberg-New York, 1994.
- [13] Lions, J. -L., Lukkassen, D., Persson, L. -E. & Wall, P., Reiterated homogenization of monotone operators, *C. R. Acad. Sci. Paris Series I*, **330**:8(2000), 675-680.
- [14] Lions, J. -L., Lukkassen, D., Persson, L. -E. & Wall, P., Reiterated homogenization of nonlinear monotone operators, *Chin. Ann. Math.*, **22B**:1(2001), 1-12.
- [15] Lukkassen, D. & Wall, P., On weak convergence of locally periodic functions, *J. Nonlinear. Math. Phys.*, **9**:1(2002), 42-57 2002.

ON THE CONNECTION IN FINSLER SPACE

S. S. CHERN

Nankai Institute of Mathematics, Tianjin 300071, China

Abstract

A simple derivation of the Connection in Finsler space.

Keywords Finsler space, Connection, Riemann-Finsler geometry
2000 MR Subject Classification 53C60

The study of Riemann-Finsler geometry has recently been enhanced by the publication of a substantial book [2]. In this book we made essential use of a connection introduced in 1948 [3]. The connection is a natural generalization of the Levi-Civita connection in the Riemannian case and seems to be the right analytical basis of the subject. We have given a derivation of it. According to Anastesie it coincides with the one introduced by Rund, who kindly gave an exposition of the paper in his book.

The aim of this paper is to give a short derivation of the connection. We will also show how it gives a solution of the local congruence, i.e., a complete system of local invariants which ensures that two Finsler structures differ by a change of coordinates.

§1. A Simple Equivalence Problem

Problem. Given in R^n with the coordinates x^i n Pfaffian forms ω^i , linearly independent, and in R^n with the coordinates x^{*i} also n linearly

independent Pfaffian forms ω^{*i} , $1 \leq i \leq n$. Find the conditions that there exists a coordinate transformation

$$x^{*i} = x^{*i}(x^1, \dots, x^n), \quad (1.1)$$

such that

$$\omega^{*i} = \omega^i. \quad (1.2)$$

(Our Latin subscripts and superscripts have the range $1, \dots, n$.)

The idea is to construct invariants under the transformation (1.1). We have, since the ω 's are linearly independent,

$$d\omega^i = c_{jk}^i \omega^j \wedge \omega^k, \quad (1.3)$$

where we can suppose

$$c_{jk}^i + c_{kj}^i = 0. \quad (1.4)$$

With the condition (1.4), the c_{jk}^i are completely determined. If the corresponding quantities in R^{*n} are denoted with asterisks, we have

$$c_{jk}^i = c_{jk}^{*i}. \quad (1.5)$$

Differentiating, we have

$$dc_{jk}^i = dc_{jk}^{*i} \quad (1.6)$$

so that

$$c_{jkl}^i = c_{jkl}^{*i} \quad (1.7)$$

if

$$dc_{jk}^i = c_{jkl}^{*i} \omega^l \quad (1.8)$$

and similar equations with asterisks. Continuing this process, we get a sequence of functions,

$$c_{jk}^i, c_{jkl}^i, c_{jklp}^i, \dots, \quad (1.9)$$

which are equal to the corresponding functions with asterisks. The solution of our problem is thus given by the following theorem:

Theorem 1.1. *The transformation (1.1) has the invariant functions (1.9). If one of the functions is a constant, the corresponding function with asterisk must be equal to the same constant. If some of the functions are independent and another one is a function of them, the same must be true with the functions with asterisks, by the same functional relation.*

§2. The Connection in a Riemann-Finsler Space

Let M be a manifold and TM its tangent bundle. By SM we mean the manifold of its rays, i.e., the set of non-zero tangent vectors differing by a positive factor. If $n = \dim M$, then $\dim TM = 2n$ and $\dim SM = 2n - 1$. We use the local coordinates x^i of M , then TM has the local coordinates x^i, y^i , if the vector is $y^i \frac{\partial}{\partial x^i}$, and SM has the same local coordinates, y^i being then homogeneous coordinates, up to a positive factor. In this section we will agree on the following ranges of indices:

$$1 \leq i, j, k, \dots \leq n; \quad 1 \leq \alpha, \beta, \gamma, \dots \leq n - 1. \quad (2.1)$$

A Riemann-Finsler metric on M is given by the function

$$ds = F(x^1, \dots, x^n, dx^1, \dots, dx^n), \quad (2.2)$$

where $F(x, y)$ is supposed to be smooth and positively homogeneous in the second variable, i.e.,

$$F(x, \lambda y) = \lambda F(x, y), \quad \lambda > 0. \quad (2.3)$$

We introduce the quantities

$$g_{ij} = \frac{\partial^2}{\partial y^i \partial y^j} \left(\frac{1}{2} F^2 \right), \quad (2.4)$$

which are functions on SM , and we make the regularity hypothesis that the matrix (g_{ij}) is positive definite (or more generally non-singular). The quadratic differential form $Q = g_{ij}(x, y) dx^i dx^j$ will be called the Riemann form.

The projection π pulls TM back:

$$\begin{array}{ccc} \pi^* TM & \longrightarrow & TM \\ \downarrow & & \downarrow \\ SM & \xrightarrow{\pi} & M \end{array} \quad (2.5)$$

and we will use the bundle at the left-hand side. In this bundle the g_{ij} in (2.4), being homogeneous of degree 0 in y^k and therefore functions on SM , define an inner product. SM has the distinguished one-form

$$H = \frac{\partial f}{\partial y^i} dx^i. \quad (2.6)$$

It will be called the Hilbert form.

Lemma 2.1. *Under the regularity hypothesis, the Hilbert form satisfies*

$$H \wedge (dH)^{n-1} \neq 0, \quad (2.7)$$

and hence define a contact structure on SM .

For proof, refer to [2, p. 272].

In the bundle at the left-hand side of (2.5) we take a frame field e_i and let ω^i be the coframe field dual to e_i .

A connection D is by definition the absolute differential

$$De_i = \omega_i^j e_j. \quad (2.8)$$

Then the tensor $\sum \omega^i \otimes e_i$ is independent of the choice of e_i and the invariant condition

$$D(\omega^i \otimes e_i) = 0 \quad (2.9)$$

is called the vanishing of torsion. This condition becomes, when written explicitly,

$$d\omega^i = \omega^j \wedge \omega_j^i. \quad (2.10)$$

We wish to introduce a torsionless connection in the bundle at the left column of (2.5). Analytically this is to determine the forms ω_j^i so that (2.10) are satisfied. We will make use of the local coordinates x^i, y^j described above and choose an orthonormal frame $x e_i$ such that e_n is the unit vector along the vector $y^i \frac{\partial}{\partial x^i}$. On SM , $\omega^i, \omega_n^\alpha$ form a base of the exterior algebra of differential forms.

We suppose our connection to preserve the length of e_n and the orthogonality of e_n and e_α . The connection forms therefore satisfy the conditions

$$\omega_{nn} = 0, \quad \omega_{\alpha n} + \omega_{n\alpha} = 0. \quad (2.11)$$

Here and later we use the Kronecker indices δ_{ij} to raise or lower indices. Notice that in the connection forms ω_j^i the second index is an upper index.

We complete the Hilbert form into a coframe

$$\omega^i = v_k^i dx^k, \quad (2.12)$$

with

$$\omega^n = H, \quad \text{i.e., } v_i^n = \frac{\partial F}{\partial y^i}, \quad (2.13)$$

$$y^k v_k^\alpha = 0, \quad (2.14)$$

i.e., $\langle e_n, \omega^\alpha \rangle = 0$. Let (u_i^k) be the inverse matrix of (v_i^k) , so that

$$u_i^k v_k^j = v_i^k u_k^j = \delta_i^j. \quad (2.15)$$

Then

$$u_n^k = \frac{y^k}{F} \quad (2.16)$$

and we have

$$\begin{aligned} d\omega^n &= \frac{\partial^2 F}{\partial x^i \partial y^k} dx^i \wedge dx^k + \frac{\partial^2 F}{\partial y^i \partial y^k} dy^i \wedge dx^k \\ &= \frac{\partial^2 F}{\partial x^i \partial y^k} u_p^i u_q^k \omega^p \wedge \omega^q + \frac{\partial^2 F}{\partial y^i \partial y^k} dy^i \wedge u_q^k \omega^q. \end{aligned}$$

Since $\frac{\partial F}{\partial y^i}$ is homogeneous of degree zero in y^k , $\frac{\partial^2 F}{\partial y^k \partial y^i} y^k = 0$ by Euler's theorem, and we can write

$$d\omega^n = \omega^\alpha \wedge \omega_\alpha^n, \quad (2.17)$$

where

$$\begin{aligned} \omega_\alpha^n &= -u_\alpha^k \frac{\partial^2 F}{\partial y^j \partial y^k} dy^j + \frac{1}{F} u_\alpha^j \left(\frac{\partial F}{\partial x^j} - \frac{\partial^2 F}{\partial x^j \partial y^k} y^k \right) \omega^n \\ &\quad + u_\alpha^j u_\beta^k \frac{\partial^2 F}{\partial x^j \partial y^k} \omega^\beta + \lambda_{\alpha\beta} \omega^\beta, \end{aligned} \quad (2.18)$$

where $\lambda_{\alpha\beta} = \lambda_{\beta\alpha}$ are to be determined.

On the other hand, we have

$$d\omega^\alpha = dv_k^\alpha \wedge dx^k = -v_k^\alpha du_i^k \wedge \omega^i = -v_k^\alpha du_\beta^k \wedge \omega^\beta - v_k^\alpha d\left(\frac{y^k}{F}\right) \omega^n. \quad (2.19)$$

We now study the equations (2.11). The first equation can clearly be satisfied. For the existence of ω_n^α satisfying the second equation of (2.11) and

$$d\omega^\alpha = \omega^\beta \wedge \omega_\beta^\alpha + \omega^n \wedge \omega_n^\alpha, \quad (2.20)$$

it is necessary and sufficient that $-\frac{1}{F} v_k^\alpha$ is equal to the coefficient of dy^k in the expression (2.18) for ω_n^α . This gives

$$u_\alpha^j G_{jk} = \delta_{\alpha\beta} v_k^\beta, \quad (2.21)$$

where

$$G_{jk} = F F_{jk}, \quad F_{jk} = \frac{\partial^2 F}{\partial y^i \partial y^k} \quad (2.22)$$

are functions on SM .

Notice that

$$v_i^\beta u_\alpha^j u_\beta^k G_{jk} = v_i^l u_l^k u_\alpha^j G_{jk} - v_i^n u_n^k u_\alpha^j G_{jk} = u_\alpha^j G_{ji},$$

since $u_n^k G_{jk} = \frac{1}{F} y^k (F F_{jk}) = 0$. Hence (2.21) can be rewritten

$$u_\alpha^j u_\beta^k G_{jk} = \delta_{\alpha\beta}. \quad (2.23)$$

It can also be written

$$\delta_{\alpha\beta} v_i^\alpha v_j^\beta = G_{ij}. \quad (2.24)$$

In forms the last equation becomes $\sum_\alpha \omega^\alpha{}^2 = Q - H^2$. Comparing (2.19) and (2.20), we get

$$\omega_\beta^\alpha = v_k^\alpha du_\beta^k - \delta^{\alpha\gamma} \left(u_\gamma^j u_\beta^k \frac{\partial^2 F}{\partial x^j \partial y^k} + \lambda_{\beta\gamma} \right) \omega^n + \mu_{\beta\gamma}^\alpha \omega^\gamma,$$

where $\mu_{\beta\gamma}^\alpha = \mu_{\gamma\beta}^\alpha$, but are otherwise arbitrary.

It remains to determine ω_β^α . We find

$$\begin{aligned} \delta_{\alpha\sigma} \omega_\rho^\alpha + \delta_{\alpha\rho} \omega_\sigma^\alpha &= -dG_{ij} u_\rho^i u_\sigma^j - u_\rho^i u_\sigma^j \left(\frac{\partial^2 F}{\partial x^i \partial y^j} + \frac{\partial^2 F}{\partial x^j \partial y^i} \right) \omega^n \\ &\quad - 2\lambda_{\rho\sigma} \omega^n + (\delta_{\alpha\sigma} \mu_{\rho\gamma}^\alpha + \delta_{\alpha\rho} \mu_{\sigma\gamma}^\alpha) \omega^\gamma. \end{aligned} \quad (2.25)$$

Since G_{ij} are functions on SM , we can write

$$dG_{ij} = G_{ij}^\alpha \omega_\alpha^k + G_{ijk} \omega^k. \quad (2.26)$$

We choose $\lambda_{\rho\sigma}$, $\mu_{\rho\sigma}^\alpha$ so that the following equation holds

$$\omega_{\rho\sigma} + \omega_{\sigma\rho} = H_{\rho\sigma}^\alpha \omega_\alpha^n. \quad (2.27)$$

This determines $\lambda_{\rho\sigma}$, $\mu_{\rho\sigma}^\alpha$ completely by

$$\begin{aligned} \lambda_{\rho\sigma} &= -\frac{1}{2} u_\rho^i u_\sigma^j \left(G_{ijn} + \frac{\partial^2 F}{\partial x^i \partial y^j} + \frac{\partial^2 F}{\partial x^j \partial y^i} \right), \\ \mu_{\rho\sigma}^\alpha &= \frac{1}{2} \delta^{\alpha\beta} (\xi_{\rho\beta\alpha} + \xi_{\alpha\beta\rho} - \xi_{\sigma\beta\alpha}), \\ \xi_{\rho\sigma\beta} &= G_{ij\beta} u_\rho^i u_\sigma^j. \end{aligned} \quad (2.28)$$

Thus all the ω_j^i are determined. We state the result as the theorem:

Theorem 2.1. *Given the Riemann-Finsler metric, there is a uniquely defined connection in the bundle $\pi^*TM \rightarrow SM \ni (x, y)$ characterized by the conditions:*

- (1) *It is torsionless;*
- (2) *The length of the vector y and the property of a vector \perp to y are preserved;*
- (3) *Relative to orthonormal frames the conditions*

$$\omega_{\alpha\beta} + \omega_{\beta\alpha} = H_{\alpha\beta\gamma} \omega_{\gamma n} \quad (2.27a)$$

are satisfied.

§3. Cartan Tensor

Condition (2.27a) in Theorem 2.1 in the last section means that the inner product is preserved by the parallelism defined by the connection when $\omega_{\alpha n} = 0$, i.e., when the parallelism preserves the vector y . We wish to calculate the function $H_{\alpha\beta\gamma}$ in terms of F .

By (2.25) and (2.27) we have

$$H_{\rho\sigma}^{\alpha} = -G_{ij}^{\alpha} u_{\rho}^i u_{\sigma}^j. \tag{3.1}$$

Comparing the coefficients of dy^k in (2.26), we get $G_{ij}^{\alpha} u_{\alpha}^k F_{kl} = -F F_{ijl} - F_l F_{ij}$. Here and later subscripts of F mean partial differentiation with respect to the corresponding y^i . It following that

$$G_{ij}^{\alpha} y^j u_{\alpha}^k F_{kl} = F F_{il} \text{ or } (G_{ij}^{\alpha} y^j u_{\alpha}^k - \delta_i^k F) F_{kl} = 0.$$

Since the matrix (F_{kl}) is of rank $n-1$, this holds only when $G_{ij}^{\alpha} y^j u_{\alpha}^k - \delta_i^k F = p_i y^k$. Multiplication of this equation by F_k and subsequent summation give $p_i = -F_i$. Hence

$$G_{ij}^{\alpha} y^j u_{\alpha}^k = -F_i y^k + \delta_i^k F, \tag{3.2}$$

$$G_{ij}^{\alpha} y^j = v^{\alpha} F. \tag{3.3}$$

With the help of this relation we find

$$v_l^{\rho} v_m^{\sigma} H_{\rho\sigma}^{\alpha} = -G_{ij}^{\alpha} \left(\delta_l^i - \frac{y^i}{F} F_l \right) \left(\delta_m^j - \frac{y^j}{F} F_m \right) = -G_{lm}^{\alpha} + v_l^{\alpha} F_m + v_m^{\alpha} F_l,$$

$$v_l^{\rho} v_m^{\alpha} H_{\rho\sigma}^{\alpha} u_{\alpha}^k F_{kj} = \left(\frac{1}{2} F^2 \right)_{lmj}.$$

Multiplying this by u_{β}^j , we get $v_l^{\rho} v_m^{\sigma} H_{\rho\alpha\beta} = F \left(\frac{1}{2} F^2 \right)_{lmj} u_{\beta}^j$, which gives

$$H_{\rho\sigma\alpha} = F \left(\frac{1}{2} F^2 \right)_{ijk} u_{\rho}^i u_{\sigma}^j u_{\alpha}^k, \tag{3.4}$$

$$F \left(\frac{1}{2} F^2 \right)_{ijk} = H_{\rho\sigma\alpha} v_i^{\rho} v_j^{\sigma} v_k^{\alpha}. \tag{3.5}$$

$H_{\rho\sigma\alpha}$ is usually called the Cartan tensor. For a Riemannian metric it is zero and our connection reduces to the connection of Levi-Civita.

§4. Equivalence Theorem

The following theorem is immediate.

Theorem 4.1. *Consider the bundle $\pi^*TM \rightarrow SM$ at the left-hand side of (2.5). Let*

$$P \rightarrow SM \tag{4.1}$$

be its principal bundle of orthonormal frames. Then $\dim P = \frac{1}{2}n(n+1)$ and in it are the forms ω^i , ω_{ij} , which are $\frac{1}{2}n(n+1)$ in number and are linearly independent. If the corresponding entities in M^* are denoted by asterisks, the two Riemann-Finsler structures differ by a coordinate transformation if and only if there is a coordinate transformation from P to P^* such that $\omega^{*i} = \omega^i$, $\omega_{ij}^* = \omega_{ij}$.

This reduces the equivalence problem to the problem solved in §1.

In the principal bundle P the forms ω^i , ω_i^j constitute a basis of $\wedge(T^*P)$, the exterior algebra of its cotangent bundle. By our Theorem 1.1 the local invariants of our Finsler structure are obtained through the exterior derivatives of ω^i , ω_i^j . The exterior derivatives $d\omega^i$ are given by (2.10). To find $d\omega_i^j$ we differentiate (2.10), obtaining

$$\omega^j \wedge (d\omega_j^i - \omega_j^k \wedge \omega_k^i) = 0.$$

It follows that

$$d\omega_j^i = \omega_j^k \wedge \omega_k^i + R_{jkl}^i \omega^k \wedge \omega^l + P_{jk\alpha}^i \omega^k \wedge \omega_k^\alpha, \quad (4.2)$$

where we suppose

$$R_{jkl}^i + R_{jlk}^i = 0, \quad (4.3)$$

and we have

$$R_{jkl}^i + R_{kjl}^i + R_{ljk}^i = 0, \quad (4.4)$$

$$P_{jk\alpha}^i = P_{kj\alpha}^i. \quad (4.5)$$

The R_{jkl}^i from the Riemann curvature tensor.

From the Riemann curvature one defines by contraction the Ricci curvature. The Ricci curvature is a scalar function on SM and is the most important local invariant in Finsler geometry. For details refer to [2, p. 190].

REFERENCES

- [1] Bao, D., Chern, S. S. & Shen, Z., An introduction to Riemann-Finsler geometry, Graduate Texts in Mathematics, No. 200, Springer 2000.
- [2] Chern, S. S., Chen, W. H. & Lam, K. S., Lectures on differential geometry, World Scientific, Singapore, 1999.
- [3] Chern, S. S., Local equivalence and Euclidean connections in Finsler spaces, in Chern Selected Papers, II, 95–121.

OPERATOR-SPLITTING METHODS FOR THE SIMULATION OF BINGHAM VISCO-PLASTIC FLOW

E. J. DEAN

*Department of Mathematics, University of Houston, 651 P. G. Hoffman
Hall, Houston, TX 77204-3008, USA. E-mail: dean@math.uh.edu*

R. GLOWINSKI

*Department of Mathematics, University of Houston, 651 P. G. Hoffman
Hall, Houston, TX 77204-3008, USA. E-mail: roland@math.uh.edu*

Abstract

This article discusses computational methods for the numerical simulation of unsteady Bingham visco-plastic flow. These methods are based on time-discretization by operator-splitting and take advantage of a characterization of the solutions involving some kind of Lagrange multipliers. The full discretization is achieved by combining the above operator-splitting methods with finite element approximations, the advection being treated by a wave-like equation “equivalent” formulation easier to implement than the method of characteristics or high order upwinding methods. The authors illustrate the methodology discussed in this article with the results of numerical experiments concerning the simulation of wall driven cavity Bingham flow in two dimensions.

Keywords Bingham visco-plastic flow, Operator-splitting methods,
Finite element approximations, Variational inequality
2000 MR Subject Classification 35Q30, 76d05, 35R45

§1. Introduction

From the early seventies to his untimely death in 2001, Jacques-Louis Lions was always highly interested, not to say intrigued, in the system of equations and inequalities modeling Bingham visco-plastic flow (one

of the success stories of the Variational Inequality Theory). Evidences of this interest can be found in the Chapters 6 of [1] and [2], which still are (to the best of our knowledge) the fundamental references concerning the mathematical properties of the variational inequalities modeling Bingham visco-plastic flow. These facts would have justified by themselves a Bingham flow related article dedicated to J. L. Lions. Actually, J. L. Lions was always concerned with the relevance of mathematics to applications and from that point of view we have been witnessing during these last years a surge of interest in Bingham visco-plastic fluids. It is very likely that this interest is motivated by the fact that material as diverse as fresh concrete, tortilla dough, fruits in syrup, blood in the capillaries, some muds used in drilling technologies, toothpastes, \dots , have a Bingham medium behavior.

The content of this article is as follows:

In Section 2, we shall provide the Bingham flow model, and a multiplier characterization of the solutions, very useful from a computational point of view. The system of partial differential equations and inequalities modeling Bingham flow will be time-discretized in Section 3, using an operator splitting scheme. The finite element approximation will be discussed in Section 4, and the solution of the subproblems encountered at each time step in Section 5. Finally, the results of numerical experiments will be presented in Section 6.

§2. On the Modeling of Bingham Visco-Plastic Flow

Let Ω be a bounded domain of \mathbf{R}^d ($d = 2$ or 3 in applications); we denote by Γ , the boundary of Ω . The isothermal flow of an incompressible Bingham visco-plastic medium, during the time interval $(0, T)$, is modeled by the following system of equations (clearly of the Navier-Stokes type):

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f} \quad \text{in } \Omega \times (0, T), \quad (2.1)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \times (0, T), \quad (2.2)$$

$$\boldsymbol{\sigma} = -p\mathbf{I} + \sqrt{2g} \frac{\mathbf{D}(\mathbf{u})}{|\mathbf{D}(\mathbf{u})|} + 2\mu\mathbf{D}(\mathbf{u}), \quad (2.3)$$

$$\mathbf{u}(0) = \mathbf{u}_0 \text{ (with } \nabla \cdot \mathbf{u}_0 = 0\text{);} \tag{2.4}$$

for simplicity, we shall consider only Dirichlet boundary conditions, namely,

$$\mathbf{u} = \mathbf{u}_\Gamma \text{ on } \Gamma \times (0, T), \text{ with } \int_\Gamma \mathbf{u}_\Gamma(t) \cdot \mathbf{n} d\Gamma = 0, \text{ a.e. on } (0, T). \tag{2.5}$$

In system (2.1)–(2.5):

- ρ (resp., μ and g) is the density (resp., viscosity and plasticity yield) of the Bingham medium; we have $\rho > 0$, $\mu > 0$ and $g > 0$.

- \mathbf{f} is a density of external forces.

- $\mathbf{D}(\mathbf{v}) = (\nabla \mathbf{v} + (\nabla \mathbf{v})^t)/2 = (D_{ij}(\mathbf{v}))_{1 \leq i, j \leq d}$, $\forall \mathbf{v} \in (H^1(\Omega))^d$, and $|\mathbf{D}(\mathbf{v})|$ is the Frobenius norm of tensor $\mathbf{D}(\mathbf{v})$, i.e.,

$$|\mathbf{D}(\mathbf{v})| = \left(\sum_{1 \leq i, j \leq d} |D_{ij}(\mathbf{v})|^2 \right)^{1/2}.$$

We clearly have trace $\mathbf{D}(\mathbf{v}) = 0$ if $\nabla \cdot \mathbf{v} = 0$.

We observe that if $g = 0$, system (2.1)–(2.5) reduces to the Navier-Stokes equations modeling isothermal incompressible Newtonian viscous fluid flow. Having said all that, if $g > 0$, the above model makes no sense on the set

$$Q_0 = \{ \{x, t\} | \{x, t\} \in \Omega \times (0, T), \mathbf{D}(\mathbf{u})(x, t) = \mathbf{0} \}.$$

Following Duvaut and Lions [1 and 2, Chapter 6] we eliminate the above difficulty by considering, instead of the (doubly) nonlinear equations (2.1)–(2.5), the following variational inequality model:

Find $\{\mathbf{u}(t), p(t)\} \in (H^1(\Omega))^d \times L^2(\Omega)$ such that a.e. on $(0, T)$ we have

$$\begin{cases} \rho \int_\Omega \frac{\partial \mathbf{u}}{\partial t}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx + \rho \int_\Omega (\mathbf{u}(t) \cdot \nabla) \mathbf{u}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx \\ + \mu \int_\Omega \nabla \mathbf{u}(t) : \nabla (\mathbf{v} - \mathbf{u}(t)) dx + \sqrt{2}g(j(\mathbf{v}) - j(\mathbf{u}(t))) \\ - \int_\Omega p(t) \nabla \cdot (\mathbf{v} - \mathbf{u}(t)) dx \geq \int_\Omega \mathbf{f}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx, \forall \mathbf{v} \in V_{\mathbf{u}_\Gamma(t)}, \end{cases} \tag{2.6}$$

$$\nabla \cdot \mathbf{u}(t) = 0 \text{ in } \Omega, \tag{2.7}$$

$$\mathbf{u}(0) = \mathbf{u}_0, \tag{2.8}$$

$$\mathbf{u}(t) = \mathbf{u}_\Gamma(t) \text{ on } \Gamma, \tag{2.9}$$

with, in system (2.6)–(2.9),

$$j(\mathbf{v}) = \int_{\Omega} |\mathbf{D}(\mathbf{v})| dx, \quad \forall \mathbf{v} \in (H^1(\Omega))^d, \quad (2.10)$$

$$V_{\mathbf{u}_\Gamma}(t) = \{\mathbf{v} | \mathbf{v} \in (H^1(\Omega))^d, \mathbf{v} = \mathbf{u}_\Gamma(t) \text{ on } \Gamma\}, \quad (2.11)$$

and $\mathbf{S} : \mathbf{T} = \sum_{i=1}^d \sum_{j=1}^d s_{ij} t_{ij}$, $\forall \mathbf{S} = (s_{ij})$, $\mathbf{T} = (t_{ij})$.

Let us be honest, formulation (2.6)–(2.9) is definitely an improvement compared to formulation (2.1)–(2.5), in the sense that we shall be able to derive from it computational methods “which work” (if $d = 2$, at least), however it is still partly formal. The rigorous formulation is more complicated and is thoroughly discussed in [1 and 2, Chapter 6, Section 3]; it is assumed there that $\mathbf{u}_\Gamma = \mathbf{0}$ on $\Gamma \times (0, T)$, and $\mathbf{u}_0 = \mathbf{0}$ if $d = 3$. If the above assumptions hold, it is shown in the above references that for $d = 2$, the time dependent variational inequality modeling the Bingham flow (a simple variant of problem (2.6)–(2.9)) has a unique solution, while uniqueness is still an open problem if $d = 3$ (as it is for the “ordinary” Navier-Stokes equations). Suppose that $d = 2$ and $\frac{\partial \mathbf{f}}{\partial t} = \mathbf{0}$; it is worthwhile emphasizing the fact that the uniqueness of the time dependent solution does not imply a similar property for the corresponding steady state flow problem.

For those readers who are already experts at solving the “ordinary” Navier-Stokes equations the main difficulty with model (2.6)–(2.9) is clearly the non-differentiable functional $j(\cdot)$. A simple way to overcome the above difficulty is to approximate $j(\cdot)$ by regularization, i.e., to replace it by a differentiable functional such as $j_\epsilon(\cdot)$ defined by

$$j_\epsilon(\mathbf{v}) = \int_{\Omega} \sqrt{\epsilon^2 + |\mathbf{D}(\mathbf{v})|^2} dx, \quad \forall \mathbf{v} \in (H^1(\Omega))^d. \quad (2.12)$$

Since, $\forall \mathbf{v} \in (H^1(\Omega))^d$, we have

$$|j_\epsilon(\mathbf{v}) - j(\mathbf{v})| = \epsilon^2 \int_{\Omega} \frac{dx}{\sqrt{\epsilon^2 + |\mathbf{D}(\mathbf{v})|^2} + |\mathbf{D}(\mathbf{v})|} \leq \epsilon |\Omega|, \quad (2.13)$$

where $|\Omega| = \text{meas.}(\Omega)$, $j_\epsilon(\cdot)$ is clearly an approximation of $j(\cdot)$. Concerning the differentiability of $j_\epsilon(\cdot)$ one can show that the differential $j'_\epsilon(\mathbf{v})$ of $j_\epsilon(\cdot)$ at $\mathbf{v} \in (H^1(\Omega))^d$ verifies:

$$(j'_\epsilon(\mathbf{v}), \mathbf{w}) = \int_{\Omega} \frac{\mathbf{D}(\mathbf{v}) : \mathbf{D}(\mathbf{w})}{\sqrt{\epsilon^2 + |\mathbf{D}(\mathbf{v})|^2}} dx, \quad \forall \mathbf{v} \in (H^1(\Omega))^d, \quad \forall \mathbf{w} \in (H_0^1(\Omega))^d, \quad (2.14)$$

where, in (2.14), $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $(H^{-1}(\Omega))^d$ and $(H_0^1(\Omega))^d$. Taking advantage of (2.14) it is tempting to “approximate” problem (2.6)–(2.9) (and indeed (2.1)–(2.5)) by

$$\left\{ \begin{array}{l} \rho \int_{\Omega} \frac{\partial \mathbf{u}_{\epsilon}}{\partial t}(t) \cdot (\mathbf{v} - \mathbf{u}_{\epsilon}(t)) dx + \rho \int_{\Omega} (\mathbf{u}_{\epsilon}(t) \cdot \nabla) \mathbf{u}_{\epsilon}(t) \cdot (\mathbf{v} - \mathbf{u}_{\epsilon}(t)) dx \\ + \mu \int_{\Omega} \nabla \mathbf{u}_{\epsilon}(t) : \nabla (\mathbf{v} - \mathbf{u}_{\epsilon}(t)) dx + \sqrt{2}g(j_{\epsilon}(\mathbf{v}) - j_{\epsilon}(\mathbf{u}_{\epsilon}(t))) \\ - \int_{\Omega} p_{\epsilon}(t) \nabla \cdot (\mathbf{v} - \mathbf{u}_{\epsilon}(t)) dx \geq \int_{\Omega} \mathbf{f}(t) \cdot (\mathbf{v} - \mathbf{u}_{\epsilon}(t)) dx, \quad \forall \mathbf{v} \in \mathbf{V}_{\mathbf{u}_{\Gamma}}(t), \end{array} \right. \quad (2.15)$$

$$\nabla \cdot \mathbf{u}_{\epsilon}(t) = 0 \text{ in } \Omega, \quad (2.16)$$

$$\mathbf{u}_{\epsilon}(0) = \mathbf{u}_0, \quad (2.17)$$

$$\mathbf{u}_{\epsilon}(t) = \mathbf{u}_{\Gamma}(t) \text{ on } \Gamma. \quad (2.18)$$

Replacing, in (2.15), \mathbf{v} by $\mathbf{u}_{\epsilon}(t) + \theta \mathbf{w}$ with $\theta > 0$ and $\mathbf{w} \in (H_0^1(\Omega))^d$, dividing by θ , and taking (2.14) into account, we obtain at the limit as $\theta \rightarrow 0_+$ that $\{\mathbf{u}_{\epsilon}, p_{\epsilon}\}$ is solution of the following nonlinear variational problem:

$$\left\{ \begin{array}{l} \rho \int_{\Omega} \frac{\partial \mathbf{u}_{\epsilon}}{\partial t}(t) \cdot \mathbf{w} dx + \rho \int_{\Omega} (\mathbf{u}_{\epsilon}(t) \cdot \nabla) \mathbf{u}_{\epsilon}(t) \cdot \mathbf{w} dx \\ + \mu \int_{\Omega} \nabla \mathbf{u}_{\epsilon}(t) : \nabla \mathbf{w} dx + \sqrt{2}g \int_{\Omega} \frac{\mathbf{D}(\mathbf{u}_{\epsilon}(t)) : \mathbf{D}(\mathbf{w})}{\sqrt{\epsilon^2 + |\mathbf{D}(\mathbf{u}_{\epsilon}(t))|^2}} dx \\ - \int_{\Omega} p_{\epsilon}(t) \nabla \cdot \mathbf{w} dx = \int_{\Omega} \mathbf{f}(t) \cdot \mathbf{w} dx, \quad \forall \mathbf{w} \in (H_0^1(\Omega))^d, \end{array} \right. \quad (2.19)$$

$$\nabla \cdot \mathbf{u}_{\epsilon}(t) = 0 \text{ in } \Omega, \quad (2.20)$$

$$\mathbf{u}_{\epsilon}(0) = \mathbf{u}_0, \quad (2.21)$$

$$\mathbf{u}_{\epsilon}(t) = \mathbf{u}_{\Gamma}(t) \text{ on } \Gamma. \quad (2.22)$$

Since tensor $(\epsilon^2 + |\mathbf{D}(\mathbf{u}_{\epsilon})|^2)^{-1/2} \mathbf{D}(\mathbf{u}_{\epsilon})$ is symmetric, we clearly have

$$\begin{aligned} (\epsilon^2 + |\mathbf{D}(\mathbf{u}_{\epsilon})|^2)^{-1/2} \mathbf{D}(\mathbf{u}_{\epsilon}) : \mathbf{D}(\mathbf{w}) &= (\epsilon^2 + |\mathbf{D}(\mathbf{u}_{\epsilon})|^2)^{-1/2} \mathbf{D}(\mathbf{u}_{\epsilon}) : \nabla \mathbf{w}, \\ &\quad \forall \mathbf{w} \in (H_0^1(\Omega))^d. \end{aligned} \quad (2.23)$$

Combining relations (2.19) and (2.23) implies that $\{\mathbf{u}_{\epsilon}, p_{\epsilon}\}$ verifies:

$$\rho \left[\frac{\partial \mathbf{u}_{\epsilon}}{\partial t} + (\mathbf{u}_{\epsilon} \cdot \nabla) \mathbf{u}_{\epsilon} \right] - \mu \Delta \mathbf{u}_{\epsilon} - \sqrt{2}g \nabla \cdot \frac{\mathbf{D}(\mathbf{u}_{\epsilon})}{\sqrt{\epsilon^2 + |\mathbf{D}(\mathbf{u}_{\epsilon})|^2}} + \nabla p_{\epsilon} = \mathbf{f} \text{ in } \Omega \times (0, T), \quad (2.24)$$

$$\nabla \cdot \mathbf{u}_\epsilon = 0 \text{ in } \Omega \times (0, T), \quad (2.25)$$

$$\mathbf{u}_\epsilon(0) = \mathbf{u}_0, \quad (2.26)$$

$$\mathbf{u}_\epsilon = \mathbf{u}_\Gamma \text{ on } \Gamma \times (0, T), \quad (2.27)$$

a regularized variant of problem (2.1)–(2.5) that could have been obtained directly. From a computational point of view, the situation looks good since we have replaced the variational inequality problem (2.6)–(2.9) by (2.24)–(2.27), which looks like a “not too complicated” variant of the usual Navier-Stokes equations. However, a closer inspection shows that the second derivative of $j_\epsilon(\cdot)$ at \mathbf{v} is given by

$$\left\{ \begin{array}{l} \langle j''_\epsilon(\mathbf{v})\mathbf{w}, \mathbf{z} \rangle \\ = \int_\Omega \frac{(\epsilon^2 + |\mathbf{D}(\mathbf{v})|^2)\mathbf{D}(\mathbf{w}) : \mathbf{D}(\mathbf{z}) - (\mathbf{D}(\mathbf{v}) : \mathbf{D}(\mathbf{w}))(\mathbf{D}(\mathbf{v}) : \mathbf{D}(\mathbf{z}))}{(\epsilon^2 + |\mathbf{D}(\mathbf{v})|^2)^{3/2}} dx, \\ \forall \mathbf{v} \in (H^1(\Omega))^d, \forall \{\mathbf{w}, \mathbf{z}\} \in (H_0^1(\Omega))^d \times (H_0^1(\Omega))^d, \end{array} \right. \quad (2.28)$$

which implies that close to those \mathbf{v} such that $\mathbf{D}(\mathbf{v})$ is “small” we have

$$\|j''(\mathbf{v})\| \simeq 1/\epsilon. \quad (2.29)$$

The situation is quite clear now: For \mathbf{u}_ϵ to be a good approximation of the solution \mathbf{u} of problem (2.6)–(2.9), we have to use small ϵ 's; on the other hand, relation (2.29) shows that we can expect problem (2.19)–(2.22), (2.24)–(2.27) to be badly conditioned for those situations where the rigid set

$$Q_0 = \{\{x, t\} | \{x, t\} \in \Omega \times (0, T), \mathbf{D}(\mathbf{u})(x, t) = \mathbf{0}\}$$

is large, implying that derivative based iterative methods such as Newton's, quasi-Newton's, and conjugate gradient will perform poorly. Fortunately for the practitioner, there exists an elegant way to overcome the computational difficulties associated to the non-differentiability of functional $j(\cdot)$, and make the solution of problem (2.6)–(2.9) almost as simple as that of the usual Navier-Stokes equations. This simplification is a direct consequence of Theorem 9.1 in [1 and 2, Chapter 6, Section 9]. When applied to problem (2.6)–(2.9), the Duvaut and Lions' results can be formulated as follows:

Theorem 2.1. *Let $\{\mathbf{u}, p\}$ be a solution of problem (2.6)–(2.9); there exists then a tensor-valued function $\lambda (= (\lambda_{ij})_{1 \leq i, j \leq d})$, not necessarily unique, such that*

$$\lambda \in (L^\infty(\Omega \times (0, T)))^{d \times d}, \quad \lambda = \lambda^t, \quad (2.30)$$

$$|\lambda| \leq 1 \text{ a.e. in } \Omega \times (0, T), \quad (2.31)$$

$$\lambda : \mathbf{D}(\mathbf{u}) = |\mathbf{D}(\mathbf{u})| \text{ a.e. in } \Omega \times (0, T), \quad (2.32)$$

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] - \mu \Delta \mathbf{u} - \sqrt{2}g \nabla \cdot \lambda + \nabla p = \mathbf{f} \text{ in } \Omega \times (0, T), \quad (2.33)$$

$$\nabla \cdot \mathbf{u} = 0 \text{ in } \Omega \times (0, T), \quad (2.34)$$

$$\mathbf{u}(0) = \mathbf{u}_0, \quad (2.35)$$

$$\mathbf{u} = \mathbf{u}_\Gamma \text{ on } \Gamma \times (0, T), \quad (2.36)$$

with $|\lambda| = \left(\sum_{1 \leq i, j \leq d} \lambda_{ij}^2 \right)^{1/2}$ in (2.31). Conversely, if a triple $\{\mathbf{u}, p, \lambda\}$ verifies relations (2.30)–(2.36), then $\{\mathbf{u}, p\}$ is a solution of problem (2.6)–(2.9).

Proof. (i) Relations (2.30)–(2.36) imply (2.6)–(2.9): Observe that the symmetry of λ implies that

$$\lambda : \nabla \mathbf{v} = \lambda : \mathbf{D}(\mathbf{v}), \quad \forall \mathbf{v} \in (H^1(\Omega))^d. \quad (2.37)$$

Multiplying both sides of relation (2.33) by $\mathbf{v} - \mathbf{u}(t)$, with $\mathbf{v} \in V_{\mathbf{u}_\Gamma}(t)$, integrating by parts, and taking relation (2.37) into account, we obtain

$$\left\{ \begin{aligned} & \rho \int_{\Omega} \frac{\partial \mathbf{u}}{\partial t}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx + \rho \int_{\Omega} (\mathbf{u}(t) \cdot \nabla) \mathbf{u}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx \\ & + \mu \int_{\Omega} \nabla \mathbf{u}(t) : \nabla (\mathbf{v} - \mathbf{u}(t)) dx - \int_{\Omega} p(t) \nabla \cdot (\mathbf{v} - \mathbf{u}(t)) dx \\ & + \sqrt{2}g \left(\int_{\Omega} \lambda(t) : \mathbf{D}(\mathbf{v}) dx - \int_{\Omega} \lambda(t) : \mathbf{D}(\mathbf{u}(t)) dx \right) \\ & = \int_{\Omega} \mathbf{f}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx, \quad \forall \mathbf{v} \in V_{\mathbf{u}_\Gamma}(t). \end{aligned} \right. \quad (2.38)$$

From (2.31) and (2.32), we clearly have

$$\begin{aligned} \int_{\Omega} \lambda(t) : \mathbf{D}(\mathbf{v}) dx &\leq \int_{\Omega} |\lambda(t)| |\mathbf{D}(\mathbf{v})| dx \leq \int_{\Omega} |\mathbf{D}(\mathbf{v})| dx, \quad \forall \mathbf{v} \in (H^1(\Omega))^d, \\ \int_{\Omega} \lambda(t) : \mathbf{D}(\mathbf{u}(t)) dx &= \int_{\Omega} |\mathbf{D}(\mathbf{u}(t))| dx, \end{aligned}$$

which, combined with relation (2.38), imply relation (2.6). We have thus shown that (2.30)–(2.36) implies (2.6)–(2.9).

(ii) Relations (2.6)–(2.9) imply (2.30)–(2.36): If $\mathbf{u}_\Gamma = \mathbf{0}$ on $\Gamma \times (0, T)$, the implication (2.6)–(2.9) \Rightarrow (2.30)–(2.36) is a relatively simple consequence of the Hahn-Banach theorem and of the fact that $j(\theta \mathbf{v}) = \theta j(\mathbf{v})$, $\forall \theta \geq 0$, $\forall \mathbf{v} \in (H^1(\Omega))^d$; we shall say no more sending the interested reader to [1 and 2, Chapter 6, Section 9] for the details of the proof. If $\mathbf{u}_\Gamma \neq \mathbf{0}$, the above result still holds, but is more complicated to prove.

Remark 2.1. It is shown in the above references that $\text{trace}(\boldsymbol{\lambda}) = 0$; the main reasons we did not mention this property earlier are that:

(i) Relation $\text{trace}(\boldsymbol{\lambda}) = 0$ is not necessary to prove the reciprocal implication (2.30)–(2.36) \Rightarrow (2.6)–(2.9).

(ii) It plays no role from a computational point of view.

On the other hand, what will play an important computational role is the fact that relations (2.31) and (2.32) imply

$$\boldsymbol{\lambda}(t) = P_\Lambda(\boldsymbol{\lambda}(t) + r\sqrt{2}g\mathbf{D}(\mathbf{u}(t))), \quad \forall r > 0, \text{ a.e. on } (0, T), \quad (2.39)$$

where, in (2.39), Λ is the closed convex set of $(L^2(\Omega))^{d \times d}$ (and $(L^\infty(\Omega))^{d \times d}$) defined by

$$\Lambda = \{\mathbf{q} | \mathbf{q} = (q_{ij})_{1 \leq i, j \leq d} \in (L^2(\Omega))^{d \times d}, |\mathbf{q}(x)| \leq 1 \text{ a.e. on } \Omega\}, \quad (2.40)$$

and $P_\Lambda : (L^2(\Omega))^{d \times d} \rightarrow \Lambda$ is the orthogonal-projection operator defined by

$$P_\Lambda(\mathbf{q})(x) = \begin{cases} \mathbf{q}(x) & \text{if } |\mathbf{q}(x)| \leq 1, \\ \mathbf{q}(x)/|\mathbf{q}(x)| & \text{if } |\mathbf{q}(x)| > 1, \end{cases} \quad (2.41)$$

a.e. on Ω , $\forall \mathbf{q} \in (L^2(\Omega))^{d \times d}$. We observe that operator P_Λ is symmetry preserving.

§3. Time-Discretization of Problem (2.6)–(2.9) by Operator Splitting

There are many ways to time-discretize problem (2.6)–(2.9) by operator splitting. Among the many possible schemes, we shall discuss only one, of the Marchuk-Yanenko type; this scheme reads as follows (with, as usual, $t^{n+\alpha} = (n + \alpha)\Delta t$):

$$\mathbf{u}^0 = \mathbf{u}_0, \quad (3.1)$$

then, for $n \geq 0$, \mathbf{u}^n being known, we compute $\{\mathbf{u}^{n+1/3}, p^{n+1}\}$, $\mathbf{u}^{n+2/3}$ and \mathbf{u}^{n+1} as follows:

Solve the generalized Stokes problem

$$\begin{cases} \rho \frac{\mathbf{u}^{n+1/3} - \mathbf{u}^n}{\Delta t} - \frac{\mu}{2} \Delta \mathbf{u}^{n+1/3} + \nabla p^{n+1} = \mathbf{f}^{n+1} (= \mathbf{f}(t^{n+1})) & \text{in } \Omega, \\ \nabla \cdot \mathbf{u}^{n+1/3} = 0 & \text{in } \Omega, \\ \mathbf{u}^{n+1/3} = \mathbf{u}_\Gamma^{n+1} (= \mathbf{u}_\Gamma(t^{n+1})) & \text{on } \Gamma, \end{cases} \quad (3.2)$$

then the transport problem

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}^{n+1/3} \cdot \nabla) \mathbf{u} = \mathbf{0} & \text{in } \Omega \times (t^n, t^{n+1}), \\ \mathbf{u}(t^n) = \mathbf{u}^{n+1/3}, \\ \mathbf{u} = \mathbf{u}_\Gamma^{n+1} & \text{on } \Gamma_-^{n+1} \times (t^n, t^{n+1}), \end{cases} \quad (3.3.1)$$

and set

$$\mathbf{u}^{n+2/3} = \mathbf{u}(t^{n+1}); \quad (3.3.2)$$

finally, solve the elliptic variational inequality

$$\begin{cases} \mathbf{u}^{n+1} \in (H^1(\Omega))^d, \mathbf{u}^{n+1} = \mathbf{u}_\Gamma^{n+1} & \text{on } \Gamma, \\ \rho \int_\Omega \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+2/3}}{\Delta t} \cdot (\mathbf{v} - \mathbf{u}^{n+1}) dx + \frac{\mu}{2} \int_\Omega \nabla \mathbf{u}^{n+1} : \nabla (\mathbf{v} - \mathbf{u}^{n+1}) dx \\ + g\sqrt{2}(j(\mathbf{v}) - j(\mathbf{u}^{n+1})) \geq 0, \forall \mathbf{v} \in (H^1(\Omega))^d, \mathbf{v} = \mathbf{u}_\Gamma^{n+1} & \text{on } \Gamma; \end{cases} \quad (3.4)$$

in (3.3.1), we have $\Gamma_-^{n+1} = \{x | x \in \Gamma, (\mathbf{u}_\Gamma^{n+1} \cdot \mathbf{n})(x) < 0\}$. Closely related operator splitting techniques have been used in [3] for the simulation of Bingham flow in two-dimensional square cavities.

Remark 3.1. It follows from, e.g., [4, Chapters 1 and 2] that the variational inequality problem (3.4) has a unique solution, characterized by the existence of a $d \times d$ tensor-valued function λ^{n+1} such that:

$$\rho \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+2/3}}{\Delta t} - \frac{\mu}{2} \Delta \mathbf{u}^{n+1} - g\sqrt{2} \nabla \cdot \lambda^{n+1} = \mathbf{0} \text{ in } \Omega, \quad (3.5)$$

$$\mathbf{u}^{n+1} = \mathbf{u}_\Gamma^{n+1} \text{ on } \Gamma, \quad (3.6)$$

$$\lambda^{n+1} \in (L^\infty(\Omega))^{d \times d}, \lambda^{n+1} = (\lambda^{n+1})^t, \quad (3.7)$$

$$|\lambda^{n+1}(x)| \leq 1 \text{ a.e. on } \Omega, \quad (3.8)$$

$$\lambda^{n+1}(x) : \mathbf{D}(\mathbf{u}^{n+1})(x) = |\mathbf{D}(\mathbf{u}^{n+1})(x)| \text{ a.e. on } \Omega. \quad (3.9)$$

The multiplier λ^{n+1} is not necessarily unique.

§4. On the Finite Element Approximation of Problem (2.6)–(2.9)

In this section (assuming that Ω is a bounded polygonal domain of \mathbb{R}^2) we are going to space-approximate problem (2.6)–(2.9) by a variant of the Bercovier-Pironneau finite element method discussed in, e.g., [4, Chapter 7]. The fundamental discrete spaces are thus:

$$V_h = \{\mathbf{v}_h | \mathbf{v}_h \in (C^0(\bar{\Omega}))^2, \mathbf{v}_h|_T \in (P_1)^2, \forall T \in \mathcal{T}_{h/2}\}, \quad (4.1)$$

$$V_{0h} = \{\mathbf{v}_h | \mathbf{v}_h \in V_h, \mathbf{v}_h = \mathbf{0} \text{ on } \Gamma\} (= V_h \cap (H_0^1(\Omega))^2), \quad (4.2)$$

$$P_h = \{q_h | q_h \in C^0(\bar{\Omega}), q_h|_T \in P_1, \forall T \in \mathcal{T}_h\}. \quad (4.3)$$

In (4.1), P_1 is the space of the polynomials in two variables of degree ≤ 1 . The continuous in time approximation of problem (2.6)–(2.9), associated to the above finite element spaces, is defined as follows:

For $t \in (0, T)$ find $\{\mathbf{u}_h(t), p_h(t)\} \in V_h \times P_h$ such that

$$\left\{ \begin{array}{l} \rho \int_{\Omega} \left[\frac{\partial \mathbf{u}_h}{\partial t}(t) + (\mathbf{u}_h(t) \cdot \nabla) \mathbf{u}_h(t) \right] \cdot (\mathbf{v}_h - \mathbf{u}_h(t)) dx \\ + \mu \int_{\Omega} \nabla \mathbf{u}_h(t) : \nabla (\mathbf{v}_h - \mathbf{u}_h(t)) dx - \int_{\Omega} p_h(t) \nabla \cdot (\mathbf{v}_h - \mathbf{u}_h(t)) dx \\ + g\sqrt{2}(j(\mathbf{v}_h) - j(\mathbf{u}_h(t))) \geq \int_{\Omega} \mathbf{f}_h(t) \cdot (\mathbf{v}_h - \mathbf{u}_h(t)) dx, \\ \forall \mathbf{v}_h \in V_h, \mathbf{v}_h = \mathbf{u}_{\Gamma h}(t) \text{ on } \Gamma, \end{array} \right. \quad (4.4)$$

$$\int_{\Omega} \nabla \cdot \mathbf{u}_h(t) q_h dx = 0, \quad \forall q_h \in P_h, \quad (4.5)$$

$$\mathbf{u}_h(t) = \mathbf{u}_{0h}, \quad (4.6)$$

$$\mathbf{u}_h(t) = \mathbf{u}_{\Gamma h}(t) \text{ on } \Gamma; \quad (4.7)$$

in (4.4)–(4.7):

- \mathbf{f}_h is an approximation of \mathbf{f} .
- $\mathbf{u}_{\Gamma h}$ is an approximation of \mathbf{u}_{Γ} so that

$$\left\{ \begin{array}{l} \int_{\Gamma} \mathbf{u}_{\Gamma h}(t) \cdot \mathbf{n} d\Gamma = 0, \quad \forall t \in (0, T), \\ \mathbf{u}_{\Gamma h}(t) \in \gamma V_h = \{\boldsymbol{\mu}_h | \boldsymbol{\mu}_h = \mathbf{v}_h|_{\Gamma}, \mathbf{v}_h \in V_h\}. \end{array} \right.$$

- \mathbf{u}_{0h} is an approximation of \mathbf{u}_0 so that $\mathbf{u}_{0h} \in V_h$, $\mathbf{u}_{0h} = \mathbf{u}_{\Gamma h}(0)$ on Γ .

• It is easy to compute $j(v_h)$, $\forall \mathbf{v}_h \in V_h$, since (4.1) implies that, $\forall T \in \mathcal{T}_{h/2}$, we have $\mathbf{D}(\mathbf{v}_h|_T) \in \mathbf{R}^k$ and therefore $|\mathbf{D}(\mathbf{v}_h|_T)| \in \mathbf{R}$, which implies in turn that $j(v_h) = \int_{\Omega} |\mathbf{D}(\mathbf{v}_h)| dx = \sum_{T \in \mathcal{T}_{h/2}} \text{meas.}(T) |\mathbf{D}(\mathbf{v}_h|_T)|$, $\forall \mathbf{v}_h \in V_h$.

There is thus no need for numerical integration to compute $j(v_h)$. The convergence, as $h \rightarrow 0$, of $\{\mathbf{u}_h, p_h\}_h$ to its continuous counterpart $\{\mathbf{u}, p\}$ is discussed in, e.g., [5], [6 and 7, Chapter 6].

§5. Solution of the Subproblems Encountered at Each Time Step of Scheme (3.1)–(3.4)

5.1. Solution of the Generalized Stokes Subproblems (3.2)

Combining scheme (3.1)–(3.4) with the finite element spaces described in Section 4 leads to the following approximation of the generalized Stokes problem (3.2):

Find $\{\mathbf{u}_h^{n+1/3}, p_h^{n+1}\} \in V_h \times P_h$ such that

$$\begin{cases} \rho \int_{\Omega} \frac{\mathbf{u}_h^{n+1/3} - \mathbf{u}_h^n}{\Delta t} \cdot \mathbf{v}_h dx + \frac{\mu}{2} \int_{\Omega} \nabla \mathbf{u}_h^{n+1/3} : \nabla \mathbf{v}_h dx - \int_{\Omega} p_h^{n+1} \nabla \cdot \mathbf{v}_h dx \\ = \int_{\Omega} \mathbf{f}_h^{n+1} \cdot \mathbf{v}_h dx, \quad \forall \mathbf{v}_h \in V_{0h}, \end{cases} \tag{5.1}$$

$$\int_{\Omega} \nabla \cdot \mathbf{u}_h^{n+1/3} q_h dx = 0, \quad \forall q_h \in P_h, \tag{5.2}$$

$$\mathbf{u}_h^{n+1/3} = \mathbf{u}_{\Gamma_h}^{n+1} \quad \text{on } \Gamma. \tag{5.3}$$

The approximate generalized Stokes problem (5.1)–(5.3) is clearly of the Bercovier-Pironneau type; it can be solved using the discrete analogues of the preconditioned conjugate gradient algorithms discussed in, e.g., [8]–[10].

5.2. Solution of the Transport Sub-Problems (3.3)

To solve the transport problem (3.3) we shall combine the finite element spaces described in Section 4 to the wave-like equation approach advocated in [11–13]; we obtain then the following discrete wave-like equation problem:

Find $\mathbf{u}_h(t) \in V_h$, such that, $\forall t \in (t^n, t^{n+1})$,

$$\begin{cases} \int_{\Omega} \frac{\partial^2 \mathbf{u}_h}{\partial t^2}(t) \cdot \mathbf{v}_h dx + \int_{\Omega} (\mathbf{u}_h^{n+1/3} \cdot \nabla) \mathbf{u}_h(t) \cdot (\mathbf{u}_h^{n+1/3} \cdot \nabla) \mathbf{v}_h dx \\ + \int_{\Gamma \setminus \Gamma_h^{n+1}} \mathbf{u}_h^{n+1/3} \cdot \mathbf{n} \frac{\partial \mathbf{u}_h}{\partial t}(t) \cdot \mathbf{v}_h d\Gamma = 0, \quad \forall \mathbf{v}_h \in V_{0h}^{-,n+1}, \end{cases} \tag{5.4}$$

$$\mathbf{u}_h(t^n) = \mathbf{u}_h^{n+1/3}, \tag{5.5}$$

$$\begin{cases} \frac{\partial \mathbf{u}_h}{\partial t}(t^n) \in V_{0h}^-,{}^{n+1}, \\ \int_{\Omega} \frac{\partial \mathbf{u}_h}{\partial t}(t^n) \cdot \mathbf{v}_h dx = - \int_{\Omega} (\mathbf{u}_h^{n+1/3} \cdot \nabla) \mathbf{u}_h^{n+1/3} \cdot \mathbf{v}_h dx, \forall \mathbf{v}_h \in V_{0h}^-,{}^{n+1}, \end{cases} \quad (5.6)$$

$$\mathbf{u}_h(t) = \mathbf{u}_{\Gamma_h}^{n+1} \quad \text{on } \Gamma_-^{n+1}, \quad (5.7)$$

with, in (5.4)–(5.7),

$$\Gamma_-^{n+1} = \{x | x \in \Gamma, (\mathbf{u}_h^{n+1/3} \cdot \mathbf{n})(x) < 0\},$$

$$V_{0h}^-,{}^{n+1} = \{\mathbf{v}_h | \mathbf{v}_h \in V_h, \mathbf{v}_h = \mathbf{0} \text{ on } \Gamma_-^{n+1}\}.$$

The solution of discrete wave-like equation problems such as (5.4)–(5.7) has been addressed in [11]–[13].

5.3. Solution of the Elliptic Variational Inequalities (3.4)

We approximate problem (3.4) by the following discrete elliptic variational inequality

$$\begin{cases} \mathbf{u}_h^{n+1} \in V_h, \mathbf{u}_h^{n+1} = \mathbf{u}_{\Gamma_h}^{n+1} \text{ on } \Gamma, \\ \rho \int_{\Omega} \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n+2/3}}{\Delta t} \cdot (\mathbf{v}_h - \mathbf{u}_h^{n+1}) dx + \frac{\mu}{2} \int_{\Omega} \nabla \mathbf{u}_h^{n+1} : \nabla (\mathbf{v}_h - \mathbf{u}_h^{n+1}) dx \\ + g\sqrt{2}(j(\mathbf{v}_h) - j(\mathbf{u}_h^{n+1})) \geq 0, \forall \mathbf{v}_h \in V_h, \mathbf{v}_h = \mathbf{u}_{\Gamma_h}^{n+1} \text{ on } \Gamma. \end{cases} \quad (5.8)$$

Problem (5.8) has a unique solution. To solve the above problem we are going to take advantage of its equivalence with:

$$\mathbf{u}_h^{n+1} \in V_h, \mathbf{u}_h^{n+1} = \mathbf{u}_{\Gamma_h}^{n+1} \text{ on } \Gamma, \lambda_h^{n+1} \in L_h, \lambda_h^{n+1} = (\lambda_h^{n+1})^t, \quad (5.9)$$

$$\begin{cases} \rho \int_{\Omega} \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n+2/3}}{\Delta t} \cdot \mathbf{v}_h dx + \frac{\mu}{2} \int_{\Omega} \nabla \mathbf{u}_h^{n+1} : \nabla \mathbf{v}_h dx \\ + g\sqrt{2} \int_{\Omega} \lambda_h^{n+1} : \mathbf{D}(\mathbf{v}_h) dx = 0, \forall \mathbf{v}_h \in V_{0h}, \end{cases} \quad (5.10)$$

$$|\lambda_h^{n+1}| \leq 1 \text{ a.e. in } \Omega, \lambda_h^{n+1} : \mathbf{D}(\mathbf{u}_h^{n+1}) = |\mathbf{D}(\mathbf{u}_h^{n+1})| \text{ a.e. in } \Omega, \quad (5.11)$$

where, in (5.9), space L_h is defined by

$$L_h = \{\mathbf{q}_h | \mathbf{q}_h \in (L^\infty(\Omega))^4, \mathbf{q}_h|_T \in \mathbf{R}^d, \forall T \in \mathcal{T}_{h/2}\}; \quad (5.12)$$

we have thus $\nabla \mathbf{v}_h$ and $\mathbf{D}(\mathbf{v}_h)$ belonging to L_h , $\forall \mathbf{v}_h \in V_h$. It follows from the symmetry of λ_h^{n+1} that

$$\int_{\Omega} \lambda_h^{n+1} : \mathbf{D}(\mathbf{v}_h) dx = \int_{\Omega} \lambda_h^{n+1} : \nabla \mathbf{v}_h dx, \forall \mathbf{v}_h \in V_h, \quad (5.13)$$

and from relations (5.11) that

$$\lambda_h^{n+1} = P_{\Lambda_h}(\lambda_h^{n+1} + rg\sqrt{2}\mathbf{D}(\mathbf{u}_h^{n+1})), \quad \forall r \geq 0, \quad (5.14)$$

where $\Lambda_h = \Lambda \cap L_h$, i.e.,

$$\Lambda_h = \{\mathbf{q}_h | \mathbf{q}_h \in L_h, |(\mathbf{q}_h|_T)| \leq 1, \forall T \in \mathcal{T}_{h/2}\}, \quad (5.15)$$

and where the orthogonal-projection operator from L_h onto Λ_h verifies

$$P_{\Lambda_h}(\mathbf{q}_h)|_T = \begin{cases} \mathbf{q}_h|_T & \text{if } |(\mathbf{q}_h|_T)| \leq 1, \\ \mathbf{q}_h|_T / |(\mathbf{q}_h|_T)| & \text{if } |(\mathbf{q}_h|_T)| > 1. \end{cases} \quad (5.16)$$

Denote by Λ_h^σ the (closed convex) subset of Λ_h defined by

$$\Lambda_h^\sigma = \{\mathbf{q}_h | \mathbf{q}_h \in \Lambda_h, \mathbf{q}_h = \mathbf{q}_h^t\}; \quad (5.17)$$

it is an easy exercise to show that

$$P_{\Lambda_h^\sigma}(\mathbf{q}_h) = P_{\Lambda_h}\left(\frac{\mathbf{q}_h + \mathbf{q}_h^t}{2}\right), \quad \forall \mathbf{q}_h \in L_h. \quad (5.18)$$

Combining relation (5.18) with (5.14) yields

$$\lambda_h^{n+1} = P_{\Lambda_h^\sigma}(\lambda_h^{n+1} + rg\sqrt{2}\nabla\mathbf{u}_h^{n+1}), \quad \forall r \geq 0. \quad (5.19)$$

We have thus shown that problem (5.8), (5.9)–(5.11) is equivalent to

$$\mathbf{u}_h^{n+1} \in V_h, \quad \mathbf{u}_h^{n+1} = \mathbf{u}_{\Gamma h}^{n+1} \quad \text{on } \Gamma, \quad \lambda_h^{n+1} \in L_h, \quad (5.20)$$

$$\begin{cases} \rho \int_{\Omega} \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n+2/3}}{\Delta t} \cdot \mathbf{v}_h dx + \frac{\mu}{2} \int_{\Omega} \nabla \mathbf{u}_h^{n+1} : \nabla \mathbf{v}_h dx \\ + g\sqrt{2} \int_{\Omega} \lambda_h^{n+1} : \nabla \mathbf{v}_h dx = 0, \quad \forall \mathbf{v}_h \in V_{0h}, \end{cases} \quad (5.21)$$

$$\lambda_h^{n+1} = P_{\Lambda_h^\sigma}(\lambda_h^{n+1} + rg\sqrt{2}\nabla\mathbf{u}_h^{n+1}), \quad \forall r \geq 0. \quad (5.22)$$

Following, e.g., [6], [7], and [14], we shall use the following iterative method à la Uzawa to solve problem (5.8):

$$\lambda_h^{n+1,0} \text{ is given in } \Lambda_h^\sigma; \quad (5.23)$$

then, for $k \geq 0$, assuming that $\lambda_h^{n+1,k} \in \Lambda_h^\sigma$ is known, solve

$$\begin{cases} \mathbf{u}_h^{n+1,k} \in V_h, \quad \mathbf{u}_h^{n+1,k} = \mathbf{u}_{\Gamma h}^{n+1} \quad \text{on } \Gamma, \\ \rho \int_{\Omega} \mathbf{u}_h^{n+1,k} \cdot \mathbf{v}_h dx + \frac{\mu\Delta t}{2} \int_{\Omega} \nabla \mathbf{u}_h^{n+1,k} : \nabla \mathbf{v}_h dx \\ = \rho \int_{\Omega} \mathbf{u}_h^n \cdot \mathbf{v}_h dx - g\sqrt{2}\Delta t \int_{\Omega} \lambda_h^{n+1,k} : \nabla \mathbf{v}_h dx, \quad \forall \mathbf{v}_h \in V_{0h}, \end{cases} \quad (5.24)$$

and compute

$$\lambda_h^{n+1,k+1} = P_{\Lambda_h^\sigma}(\lambda_h^{n+1,k} + rg\sqrt{2}\nabla\mathbf{u}_h^{n+1,k}). \quad (5.25)$$

Concerning the convergence of algorithm (5.23)–(5.25), we have the following

Theorem 5.1. *Suppose that*

$$0 < r < \frac{\mu}{2g^2}; \quad (5.26)$$

we have then, $\forall \lambda_h^{n+1,0} \in \Lambda_h^\sigma$,

$$\lim_{k \rightarrow +\infty} \{\mathbf{u}_h^{n+1,k}, \lambda_h^{n+1,k}\} = \{\mathbf{u}_h^{n+1}, \lambda_h^{n+1,*}\}, \quad (5.27)$$

where, in (5.27), the pair $\{\mathbf{u}_h^{n+1}, \lambda_h^{n+1,*}\}$ is a solution of problem (5.9)–(5.11), \mathbf{u}_h^{n+1} being then the unique solution of problem (5.8).

Proof. Proving the convergence of $\{\mathbf{u}_h^{n+1,k}\}_{k \geq 0}$ is fairly easy:

Suppose that $\mathbf{q}_h \in L_h$; we shall denote by $\|\mathbf{q}_h\|_0$ the $L^2(\Omega)$ -norm of \mathbf{q}_h defined by $\|\mathbf{q}_h\|_0 = \left(\int_\Omega |\mathbf{q}_h|^2 dx\right)^{1/2}$; operator $P_{\Lambda_h^\sigma}$ is a contraction for the above norm. Next, we denote by $\bar{\mathbf{u}}_h^{n+1,k}$ and $\bar{\lambda}_h^{n+1,k}$ the differences $\mathbf{u}_h^{n+1,k} - \mathbf{u}_h^{n+1}$ and $\lambda_h^{n+1,k} - \lambda_h^{n+1}$, where $\{\mathbf{u}_h^{n+1}, \lambda_h^{n+1}\} \in V_h \times \Lambda_h^\sigma$ is a solution of problem (5.9)–(5.11). By subtraction, we clearly obtain

$$\left\{ \begin{array}{l} \bar{\mathbf{u}}_h^{n+1,k} \in V_{0h}, \\ \rho \int_\Omega \bar{\mathbf{u}}_h^{n+1,k} \cdot \mathbf{v}_h dx + \mu \frac{\Delta t}{2} \int_\Omega \nabla \bar{\mathbf{u}}_h^{n+1,k} : \nabla \mathbf{v}_h dx \\ = -g\sqrt{2}\Delta t \int_\Omega \bar{\lambda}_h^{n+1,k} : \nabla \mathbf{v}_h dx, \quad \forall \mathbf{v}_h \in V_{0h}, \end{array} \right. \quad (5.28)$$

$$\|\bar{\lambda}_h^{n+1,k+1}\|_0 \leq \|\bar{\lambda}_h^{n+1,k}\|_0 + rg\sqrt{2}\|\nabla \bar{\mathbf{u}}_h^{n+1,k}\|_0, \quad \forall r \geq 0. \quad (5.29)$$

Taking $\mathbf{v}_h = \bar{\mathbf{u}}_h^{n+1,k}$ in (5.28) and combining with (5.29) we obtain

$$\left\{ \begin{array}{l} \|\bar{\lambda}_h^{n+1,k}\|_0^2 - \|\bar{\lambda}_h^{n+1,k+1}\|_0^2 \\ \geq -2rg\sqrt{2} \int_\Omega \bar{\lambda}_h^{n+1,k} : \nabla \bar{\mathbf{u}}_h^{n+1,k} dx - 2r^2g^2 \|\nabla \bar{\mathbf{u}}_h^{n+1,k}\|_0^2 \\ \geq r\mu \left(\frac{2\rho}{\mu\Delta t} \|\bar{\mathbf{u}}_h^{n+1,k}\|_{(L^2(\Omega))^2}^2 + \|\nabla \bar{\mathbf{u}}_h^{n+1,k}\|_0^2 \right) - 2r^2g^2 \|\nabla \bar{\mathbf{u}}_h^{n+1,k}\|_0^2 \\ \geq r(\mu - 2rg^2) \left(\frac{2\rho}{\mu\Delta t} \|\bar{\mathbf{u}}_h^{n+1,k}\|_{(L^2(\Omega))^2}^2 + \|\nabla \bar{\mathbf{u}}_h^{n+1,k}\|_0^2 \right). \end{array} \right. \quad (5.30)$$

Suppose that inequalities (5.26) hold; it follows then from (5.30) that the sequence $\{\|\bar{\lambda}_h^{n+1,k}\|_0\}_{k \geq 0}$ is decreasing. Since it is bounded from below by 0, it converges to some limit, implying that

$$\lim_{k \rightarrow +\infty} (\|\bar{\lambda}_h^{n+1,k}\|_0^2 - \|\bar{\lambda}_h^{n+1,k+1}\|_0^2) = 0; \quad (5.31)$$

since (5.26) implies $r(\mu - 2rg^2) > 0$, combining (5.30) with (5.31) shows that $\lim_{k \rightarrow +\infty} \bar{\mathbf{u}}_h^{n+1,k} = \mathbf{0}$, i.e., $\lim_{k \rightarrow +\infty} \mathbf{u}_h^{n+1,k} = \mathbf{u}_h^{n+1}$. To prove the convergence of $\{\lambda_h^{n+1,k}\}_{k \geq 0}$ we should proceed as in, e.g., [7, Appendix 2, Section 3].

Remark 5.1. Actually, the upper bound in (5.26) is pessimistic. Indeed, we can easily show (from relation (5.30)) that the convergence result (5.27) still holds if r verifies

$$0 < r < \left(1 + \frac{2\rho}{\mu \Delta t \beta_h^M}\right) \frac{\mu}{2g^2}, \quad (5.32)$$

where, in (5.32), β_h^M is the largest eigenvalue of the following discrete eigenvalue/eigenfunction problem:

$$\left\{ \begin{array}{l} \{\mathbf{w}_h, \beta\} \in V_{0h} \times \mathbf{R}_+, \\ \int_{\Omega} \nabla \mathbf{w}_h : \nabla \mathbf{v}_h dx = \beta \int_{\Omega} \mathbf{w}_h \cdot \mathbf{v}_h dx, \quad \forall \mathbf{v}_h \in V_{0h}, \\ \int_{\Omega} |\mathbf{w}_h|^2 dx = 1. \end{array} \right. \quad (5.33)$$

We recall that $\beta_h^M = O(h^{-2})$.

§6. Numerical Experiments

The numerical simulation of Bingham flow has not motivated as many publications as the solution of the “ordinary” Navier-Stokes equations. Besides [5], relevant publications are, e.g., [15], [16], [17], [18], [19, Chapter 6] and [3]; some of the results reported in the above references have been obtained using a stream-function formulation. The test problems considered here (and, actually, the methodology to solve them) are closely related to those in [3]. These test problems are all particular cases of the following problem:

Find $\{\mathbf{u}(t), p(t)\} \in (H^1(\Omega))^2 \times L^2(\Omega)$ such that a.e. on $(0, T)$ we have

$$\left\{ \begin{array}{l} \int_{\Omega} \frac{\partial \mathbf{u}}{\partial t}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx + \int_{\Omega} (\mathbf{u}(t) \cdot \nabla) \mathbf{u}(t) \cdot (\mathbf{v} - \mathbf{u}(t)) dx \\ + \mu \int_{\Omega} \nabla \mathbf{u}(t) : \nabla (\mathbf{v} - \mathbf{u}(t)) dx + \sqrt{2}g (j(\mathbf{v}) - j(\mathbf{u}(t))) \\ - \int_{\Omega} p(t) \nabla \cdot (\mathbf{v} - \mathbf{u}(t)) dx \geq 0, \forall \mathbf{v} \in V_{\mathbf{u}_\Gamma}, \end{array} \right. \quad (6.1)$$

$$\nabla \cdot \mathbf{u}(t) = 0 \text{ in } \Omega, \quad (6.2)$$

$$\mathbf{u}(0) = \mathbf{0}, \quad (6.3)$$

$$\mathbf{u}(t) = \mathbf{u}_\Gamma \text{ on } \Gamma, \quad (6.4)$$

with $j(\mathbf{v}) = \int_{\Omega} |\mathbf{D}(\mathbf{v})| dx$, $\forall \mathbf{v} \in (H^1(\Omega))^2$. In system (6.1)–(6.4), we have

- $\Omega = (0, 1)^2$, $\Gamma = \partial\Omega$.
- $\Gamma_N = \{x | x = \{x_1, x_2\}, x_2 = 1, 0 < x_1 < 1\}$, and

$$\mathbf{u}_\Gamma(x) = \begin{cases} \mathbf{0} & \text{if } x \in \Gamma \setminus \Gamma_N, \\ 16U \{x_1^2(1-x_1)^2, 0\} & \text{if } x \in \Gamma_N, \end{cases}$$

with $U > 0$.

- $V_{\mathbf{u}_\Gamma} = \{\mathbf{v} | \mathbf{v} \in (H^1(\Omega))^2, \mathbf{v} = \mathbf{u}_\Gamma \text{ on } \Gamma\}$.

For the time-discretization of problem (6.1)–(6.4), we have employed the Marchuk-Yanenko scheme (3.1)–(3.4). For the space discretization we have used a 128×128 uniform grid to define the finite element spaces V_h , V_{0h} and P_h (see relations (4.1)–(4.3)); from these spaces we proceeded as in Sections 4 and 5 to approximate problem (6.1)–(6.4) and compute its solutions. We have, in particular, used $r = \mu/r_g^2$ when computing $\{\mathbf{u}_h^{n+1}, \lambda_h^{n+1}\}$ by algorithm (5.23)–(5.25).

First Test Problem. It is the particular case of problem (6.1)–(6.4) corresponding to $U = 1/16$, $\mu = 1$ and $g = 0.1$; for the time discretization we have used $\Delta t = 10^{-3}$. Recalling that $\mathbf{u}(0) = \mathbf{0}$, we have shown in Fig. 6.1(a) the variation of the computed kinetic energy; it is clear from the above figure that “we” converge quickly to a steady state solution. The streamlines of the computed solution at $t = 2.39$ are shown in Fig. 6.1(b). The rigidity (black) and plastic (white) regions have been visualized in Fig. 6.1(c). The rigidity region (3-connected here) is the one where $\mathbf{D}(\mathbf{u}) = \mathbf{0}$; it is also the region where $|\lambda(x)| < 1$, as shown in Fig. 6.1(d) where the graph of $|\lambda|$ has been visualized. To conclude this presentation of the results associated to this first test problem, let us report on the following numerical

experiment: The parameters \mathbf{u}_0 , μ , g , U being as above, we solved problem (6.1)–(6.4) up to $t = 1.2$; let us denote by $\mathbf{u}(1.2)$ the velocity field at $t = 1.2$. At $t = 1.2$, we froze the motion of the upper wall implying that for $t > 1.2$ the Bingham flow is still modeled by relations (6.1), (6.2) completed by the boundary condition

$$\mathbf{u}(t) = \mathbf{0} \text{ on } \Gamma, \text{ if } t > 1.2,$$

with $\mathbf{u}(1.2)$ as initial condition at $t = 1.2$. In principle, due to the absence of body forces and to the immobility of the boundary, the medium should return to rest in finite time (see Remark 6.1, hereafter), i.e., we should have $\mathbf{u}(t) = \mathbf{0}$, $\forall t \geq t_c$, t_c being finite. Fig. 6.1(e) shows that indeed $\|\mathbf{u}(t)\|_{(L^2(\Omega))^2}$ converges to zero very quickly as $t \rightarrow +\infty$, but finite time convergence is doubtful from the above figure. Actually convergence in finite time takes place as shown in, e.g., [7, Appendix 6], [17], [18] and [19, Chapter 6]. In the above references time discretization was achieved with a fully implicit scheme à la backward Euler. It seems that for the calculation presented here, the splitting errors associated to scheme (3.1)–(3.4) prevent convergence to zero in finite time.

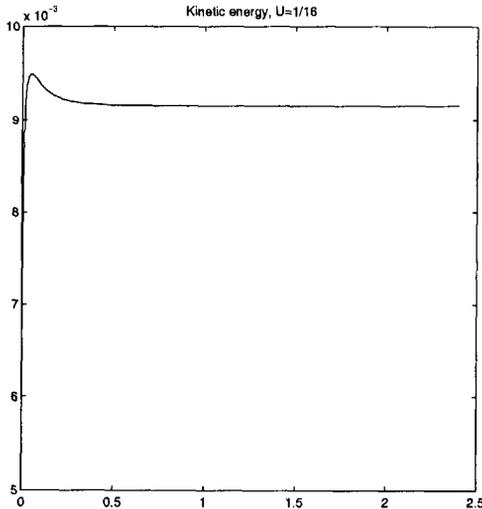


Fig. 6.1(a) Variation of the computed kinetic energy ($\mu = 1$, $g = 0.1$, $U = 1/16$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

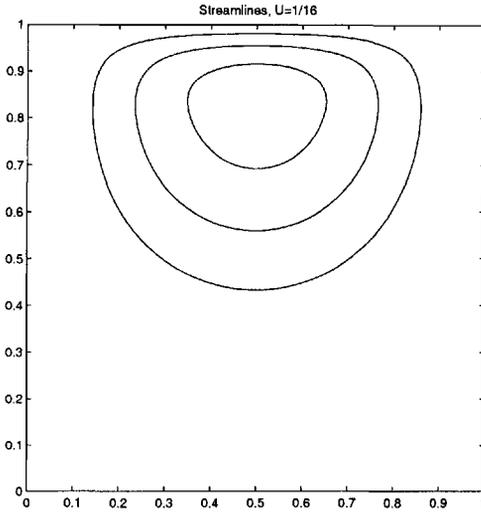


Fig. 6.1(b) Streamlines of the computed steady state velocity field ($\mu = 1$, $g = 0.1$, $U = 1/16$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

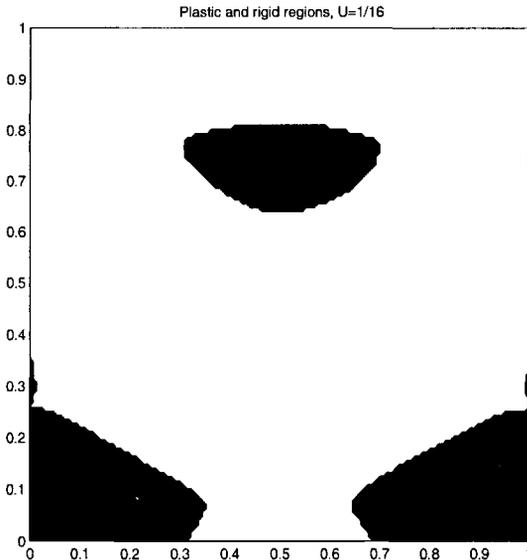


Fig. 6.1(c) Visualization of the computed plastic (white) and rigid (black) regions at steady state ($\mu = 1$, $g = 0.1$, $U = 1/16$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

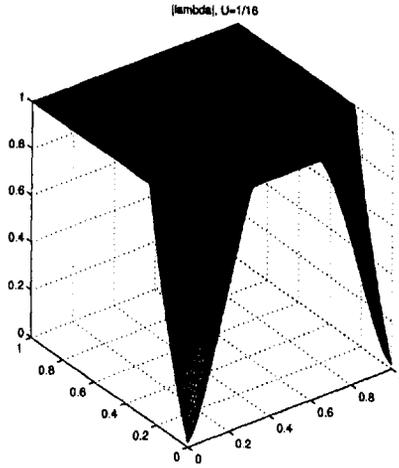


Fig. 6.1(d) Graph of $|\lambda_h|$ at steady state ($\mu = 1$, $g = 0.1$, $U = 1/16$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

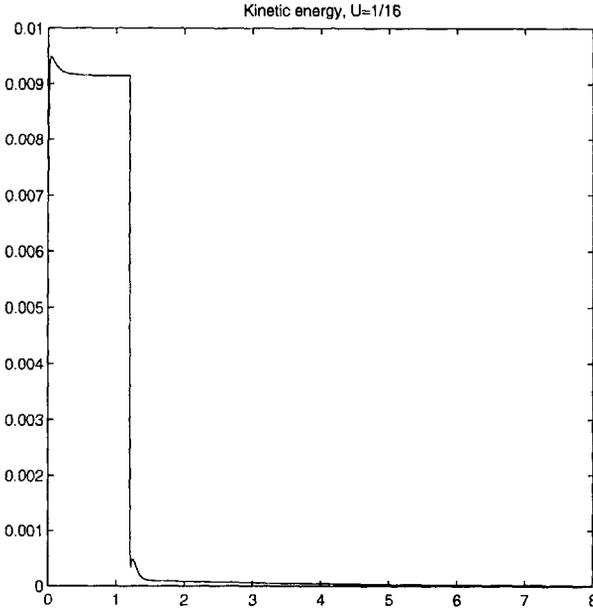


Fig. 6.1(e) Decay of the computed kinetic energy after the sliding of the upper wall has been stopped at $t = 1.2$ ($\mu = 1$, $g = 0.1$, $U = 1/16$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

Second Test Problem. This test problem is the variation of the first one obtained by taking $U = 1$ instead of $1/16$. Besides this modification, all the other physical and numerical parameters are the same. The kinetic energy variation, the streamlines, the plastic and rigid regions and the multiplier λ_h have been visualized in Fig. 6.2(a) to Fig. 6.2(d). The velocity of the upper wall being much larger the kinetic energy reaches much higher values than in the first test problem. Similarly, due to the higher level of stress, the plastic region is much larger than in the first case (compare Fig. 6.2(c) to Fig. 6.1(c)). We observe that in both cases, the viscous effects are so strong that the advection plays practically no role as shown by the symmetry of the computed results with respect to the line $x_1 = 0.5$.

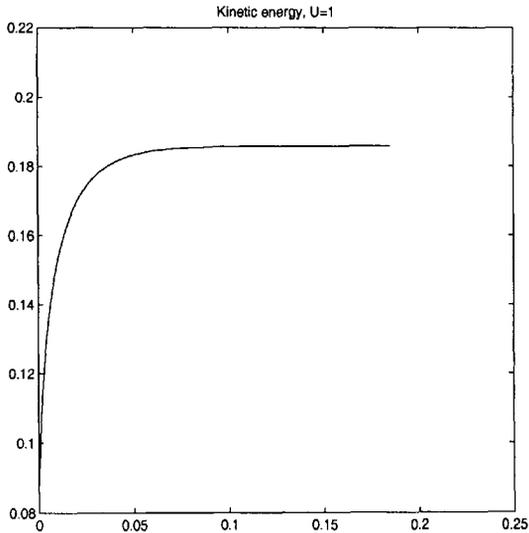


Fig. 6.2(a) Variation of the computed kinetic energy ($\mu = 1$, $g = 0.1$, $U = 1$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

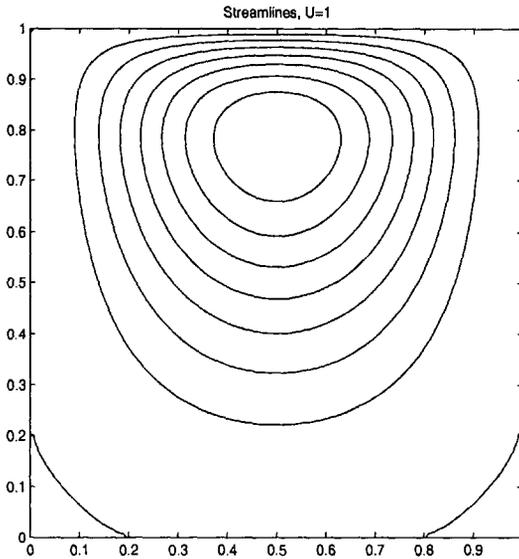


Fig. 6.2(b) Streamlines of the computed steady state velocity field ($\mu = 1$, $g = 0.1$, $U = 1$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

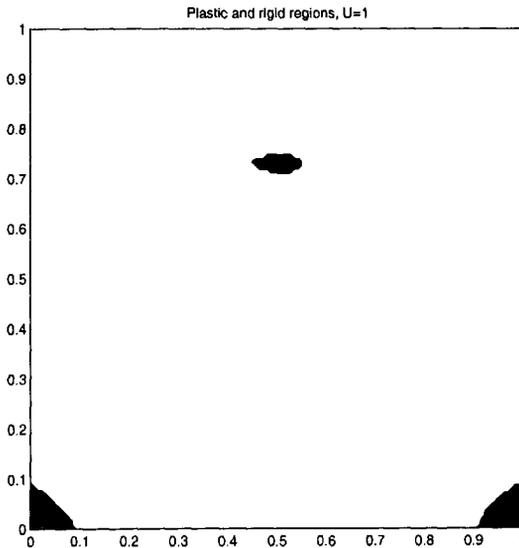


Fig. 6.2(c) Visualization of the computed plastic (white) and rigid (black) regions at steady state ($\mu = 1$, $g = 0.1$, $U = 1$, $\Delta x_1 = \Delta x_2 = 1/128$, $\Delta t = 10^{-3}$)

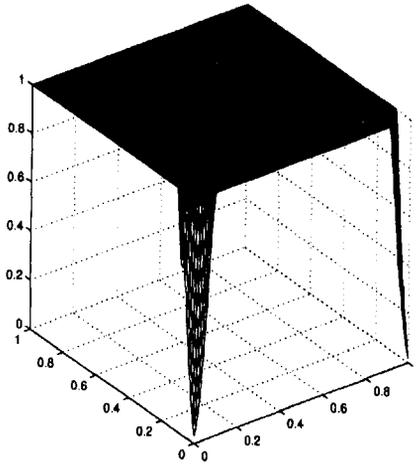


Fig. 6.2(d) Graph of $|\lambda_h|$ at steady state ($\mu = 1, g = 0.1, U = 1, \Delta x_1 = \Delta x_2 = 1/128, \Delta t = 10^{-3}$)

Remark 6.1. (On the convergence to zero in finite time). Consider problem (2.6)–(2.9) and suppose that $d = 2, \mathbf{f} = \mathbf{0}, \mathbf{u}_\Gamma = \mathbf{0}$ and $T = +\infty$. If the above assumptions hold, then $\mathbf{u}(t)$ converges to $\mathbf{0}$ in finite time as t increases, $\forall \mathbf{u}_0 \in (L^2(\Omega))^d$ such that $\nabla \cdot \mathbf{u}_0 = 0$ and $\mathbf{u}_0 \cdot \mathbf{n} = 0$ on Γ . To prove the above result, observe that $\mathbf{u}(t) \in (H_0^1(\Omega))^2$, a.e. $t \in (0, +\infty)$, and take $\mathbf{v} = \mathbf{0}$ and $\mathbf{v} = 2\mathbf{u}(t)$ in (2.6). We obtain then

$$\begin{cases} \frac{\rho}{2} \frac{d}{dt} \|\mathbf{u}(t)\|_{(L^2(\Omega))^2}^2 + \rho \int_{\Omega} (\mathbf{u}(t) \cdot \nabla) \mathbf{u}(t) \cdot \mathbf{u}(t) dx + \mu \int_{\Omega} |\nabla \mathbf{u}(t)|^2 dx \\ + g\sqrt{2}j(\mathbf{u}(t)) - \int_{\Omega} p(t) \nabla \cdot \mathbf{u}(t) dx = 0, \quad \text{a.e. } t \in (0, +\infty), \\ \mathbf{u}(t) = \mathbf{u}_0. \end{cases}$$

From now on, we shall denote $\|\cdot\|_{(L^2(\Omega))^2}$ by $\|\cdot\|_{0,\Omega}$; from $\nabla \cdot \mathbf{u}(t) = 0$ the above relation reduces to

$$\begin{cases} \frac{\rho}{2} \frac{d}{dt} \|\mathbf{u}(t)\|_{0,\Omega}^2 + \mu \int_{\Omega} |\nabla \mathbf{u}(t)|^2 dx + g\sqrt{2}j(\mathbf{u}(t)) = 0, \quad \text{a.e. } t \in (0, +\infty), \\ \mathbf{u}(t) = \mathbf{u}_0. \end{cases} \tag{6.5}$$

On the other hand, we have

$$\|\mathbf{v}\|_{0,\Omega}^2 \leq \lambda_0^{-1} \int_{\Omega} |\nabla \mathbf{v}|^2 dx, \quad \forall \mathbf{v} \in (H_0^1(\Omega))^2, \tag{6.6}$$

$$\|\mathbf{v}\|_{0,\Omega} \leq \gamma j(\mathbf{v}), \quad \forall \mathbf{v} \in (H_0^1(\Omega))^2, \tag{6.7}$$

where, in (6.6) and (6.7), $\lambda_0 (> 0)$ is the smallest eigenvalue of operator $-\Delta$ "acting" on $H_0^1(\Omega)$, and γ is a positive constant; inequality (6.7) is known as the Nirenberg-Strauss inequality and is proved in [20]. Combining relations (6.5), (6.6), and (6.7) yields

$$\begin{cases} \frac{\rho}{2} \frac{d}{dt} \|\mathbf{u}(t)\|_{0,\Omega}^2 + \mu\lambda_0 \|\mathbf{u}(t)\|_{0,\Omega}^2 + g\gamma^{-1}\sqrt{2} \|\mathbf{u}(t)\|_{0,\Omega} \leq 0, \text{ a.e. } t \in (0, +\infty), \\ \|\mathbf{u}(0)\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}. \end{cases} \quad (6.8)$$

Suppose that $\mathbf{u}(t)$ never vanishes; we have then $\|\mathbf{u}(t)\| > 0, \forall t \geq 0$ and

$$\frac{d}{dt} \|\mathbf{u}(t)\|_{0,\Omega}^2 = 2\|\mathbf{u}(t)\|_{0,\Omega} \frac{d}{dt} \|\mathbf{u}(t)\|_{0,\Omega}. \quad (6.9)$$

Combining (6.8) and (6.9) we obtain

$$\begin{cases} \rho \frac{d}{dt} \|\mathbf{u}(t)\|_{0,\Omega} + \mu\lambda_0 \|\mathbf{u}(t)\|_{0,\Omega} + g\gamma^{-1}\sqrt{2} \leq 0, \text{ a.e. } t \in (0, +\infty), \\ \|\mathbf{u}(0)\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}. \end{cases} \quad (6.10)$$

Observe now that (6.10) is equivalent to

$$\begin{cases} \frac{d}{dt} [\|\mathbf{u}(t)\|_{0,\Omega} + g\sqrt{2}(\mu\lambda_0\gamma)^{-1}] + \frac{\mu\lambda_0}{\rho} [\|\mathbf{u}(t)\|_{0,\Omega} + g\sqrt{2}(\mu\lambda_0\gamma)^{-1}] \leq 0, \\ \text{a.e. } t \in (0, +\infty), \\ \|\mathbf{u}(0)\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}. \end{cases} \quad (6.11)$$

Integrating the differential inequality in (6.11) from 0 to t we obtain

$$\|\mathbf{u}(t)\|_{0,\Omega} + g\sqrt{2}(\mu\lambda_0\gamma)^{-1} \leq e^{-\frac{\mu\lambda_0}{\rho}t} [\|\mathbf{u}_0\|_{0,\Omega} + g\sqrt{2}(\mu\lambda_0\gamma)^{-1}], \quad \forall t \geq 0. \quad (6.12)$$

Since $\lim_{t \rightarrow +\infty} e^{-\frac{\mu\lambda_0}{\rho}t} = 0$, relation (6.12) makes no sense as soon as $t > t_c$, with t_c defined by

$$t_c = \frac{\rho}{\lambda_0\mu} \ln \left(1 + \frac{\lambda_0\mu\gamma}{g\sqrt{2}} \|\mathbf{u}_0\|_{0,\Omega} \right); \quad (6.13)$$

we have then $\mathbf{u}(t) = \mathbf{0}$ if $t \geq t_c$.

The assumptions on $d, \mathbf{f}, \mathbf{u}_r$ and T staying the same, suppose now that we time-discretize problem (2.6)-(2.9) by the backward Euler scheme; we obtain then

$$\mathbf{u}^0 = \mathbf{u}_0; \quad (6.14)$$

then, for $n \geq 1$, \mathbf{u}^{n-1} being known, find $\{\mathbf{u}^n, p^n\} \in (H_0^1(\Omega))^d \times L^2(\Omega)$ such that

$$\begin{cases} \rho \int_{\Omega} \frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\Delta t} \cdot (\mathbf{v} - \mathbf{u}^n) dx + \rho \int_{\Omega} (\mathbf{u}^n \cdot \nabla) \mathbf{u}^n \cdot (\mathbf{v} - \mathbf{u}^n) dx \\ + \mu \int_{\Omega} \nabla \mathbf{u}^n : \nabla (\mathbf{v} - \mathbf{u}^n) dx + g\sqrt{2}(j(\mathbf{v}) - j(\mathbf{u}^n)) \\ - \int_{\Omega} p^n \nabla \cdot (\mathbf{v} - \mathbf{u}^n) dx \geq 0, \quad \forall \mathbf{v} \in (H_0^1(\Omega))^2, \end{cases} \quad (6.15)$$

$$\nabla \cdot \mathbf{u}^n = 0 \quad \text{on } \Omega. \quad (6.16)$$

Assuming that problem (6.14)–(6.16) has a solution, $\forall n \geq 1$ (it is not very difficult to prove that it is, indeed, the case), take $\mathbf{v} = \mathbf{0}$ and $2\mathbf{u}^n$ in (6.15), then take into account $\nabla \cdot \mathbf{u}^n = 0$, relations (6.6) and (6.7) and

$\int_{\Omega} \mathbf{u}^n \cdot \mathbf{u}^{n-1} dx \leq \|\mathbf{u}^n\|_{0,\Omega} \|\mathbf{u}^{n-1}\|_{0,\Omega}$; it follows then from (6.15) that

$$\begin{cases} \frac{\rho}{\Delta t} \|\mathbf{u}^n\|_{0,\Omega} (\|\mathbf{u}^n\|_{0,\Omega} - \|\mathbf{u}^{n-1}\|_{0,\Omega}) + \lambda_0 \mu \|\mathbf{u}^n\|_{0,\Omega}^2 + g\sqrt{2}\gamma^{-1} \|\mathbf{u}^n\|_{0,\Omega} \leq 0, \\ \forall n \geq 1, \\ \|\mathbf{u}^0\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}. \end{cases} \quad (6.17)$$

Suppose that $\mathbf{u}^n \neq \mathbf{0}$, $\forall n \geq 0$. We have then $\|\mathbf{u}^n\|_{0,\Omega} > 0$, $\forall n \geq 1$, which combined with (6.17) yields

$$\begin{cases} \frac{\rho}{\Delta t} (\|\mathbf{u}^n\|_{0,\Omega} - \|\mathbf{u}^{n-1}\|_{0,\Omega}) + \lambda_0 \mu \|\mathbf{u}^n\|_{0,\Omega} + g\sqrt{2}\gamma^{-1} \leq 0, \quad \forall n \geq 1, \\ \|\mathbf{u}^0\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}. \end{cases} \quad (6.18)$$

It follows from (6.18) that

$$\begin{cases} \|\mathbf{u}^n\|_{0,\Omega} + g\sqrt{2}(\lambda_0 \mu \gamma)^{-1} \leq \left(1 + \frac{\lambda_0 \mu}{\rho} \Delta t\right)^{-1} [\|\mathbf{u}^{n-1}\|_{0,\Omega} + g\sqrt{2}(\lambda_0 \mu \gamma)^{-1}], \\ \forall n \geq 1, \\ \|\mathbf{u}^0\|_{0,\Omega} = \|\mathbf{u}_0\|_{0,\Omega}, \end{cases}$$

which implies in turn that

$$\begin{aligned} & \|\mathbf{u}^n\|_{0,\Omega} + g\sqrt{2}(\lambda_0 \mu \gamma)^{-1} \\ & \leq \left(1 + \frac{\lambda_0 \mu}{\rho} \Delta t\right)^{-n} [\|\mathbf{u}_0\|_{0,\Omega} + g\sqrt{2}(\lambda_0 \mu \gamma)^{-1}], \quad \forall n \geq 0. \end{aligned} \quad (6.19)$$

Since $\lim_{n \rightarrow +\infty} \left(1 + \frac{\lambda_0 \mu}{\rho} \Delta t\right)^{-n} = 0$, relation (6.19) makes no sense if $n > n_c$,

with

$$n_c = \frac{\ln \left(1 + \frac{\lambda_0 \mu \gamma}{g \sqrt{2}} \|\mathbf{u}_0\|_{0,\Omega} \right)}{\ln \left(1 + \frac{\lambda_0 \mu}{\rho} \Delta t \right)}; \tag{6.20}$$

we have thus

$$\mathbf{u}^n = \mathbf{0}, \quad \forall n > n_c; \tag{6.21}$$

relation (6.21) is a discrete analogue of $\mathbf{u}(t) = \mathbf{0}, \forall t \geq t_c$. It is worth while noticing that, as expected,

$$\lim_{\Delta t \rightarrow 0_+} n_c \Delta t = \frac{\rho}{\lambda_0 \mu} \ln \left(1 + \frac{\lambda_0 \mu \gamma}{g \sqrt{2}} \|\mathbf{u}_0\|_{0,\Omega} \right) = t_c.$$

We have shown thus that the solution $\{\mathbf{u}^n\}_{n \geq 1}$ of problem (6.14)–(6.16) behaves “discretely” like the solution of problem (2.6)–(2.9). To prove (and have) the same result after space discretization it will definitely help to have

$$\int_{\Omega} (\mathbf{u}_h^n \cdot \nabla) \mathbf{u}_h^n \cdot \mathbf{u}_h^n dx = 0, \quad \forall n \geq 1. \tag{6.22}$$

This will not be the case, in general, if one employs the Hood-Taylor or Bercovier-Pironneau finite element methods to approximate problem (2.6)–(2.9). An easy way to overcome this difficulty, and recover the convergence to zero in finite discrete time, would be to replace $\int_{\Omega} (\mathbf{u}_h^n \cdot \nabla) \mathbf{u}_h^n \cdot (\mathbf{v}_h - \mathbf{u}_h^n) dx$ by

$$\int_{\Omega} \left[(\mathbf{u}_h^n \cdot \nabla) \mathbf{u}_h^n + \frac{1}{2} (\nabla \cdot \mathbf{u}_h^n) \mathbf{u}_h^n \right] \cdot (\mathbf{v}_h - \mathbf{u}_h^n) dx,$$

an idea (due to R. Temam) used by many authors.

REFERENCES

- [1] Duvaut, G. & Lions, J. L., *Les inéquations en mécanique et en physique*, Dunod, Paris, 1972.
- [2] Duvaut, G. & Lions, J. L., *Inequalities in mechanics and physics*, Springer-Verlag, Berlin, 1976.
- [3] Sanchez, F. J., *On some splitting methods for the numerical simulation of the Navier-Stokes equations*, Ph.D. dissertation, University of Houston, Department of Mathematics, 1996.
- [4] Glowinski, R., *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, NY, 1984.
- [5] Fortin, M., *Calcul numérique des écoulements des fluides Newtoniens incompressibles par la méthode des éléments finis*, Thèse d’Etat, Université Pierre et Marie Curie, Paris, France, 1972.

- [6] Glowinski, R., Lions, J. L. & Tremolières, R., Analyse numérique des inéquations variationnelles, Volumes I and II, Dunod, Paris, 1976.
- [7] Glowinski, R., Lions, J. L. & Tremolières, R., Numerical analysis of variational inequalities, North-Holland, Amsterdam, 1981.
- [8] Cahouet, J. & Chabard, J. P., Some fast 3-D solvers for the generalized Stokes problem, *Int. J. Numer. Meth. in Fluids*, **8**(1988), 269–295.
- [9] Glowinski, R., Finite element methods for the numerical simulation of incompressible viscous flow. Introduction to the control of the Navier-Stokes equations, In Anderson, C. R. and Greengard, C. (eds.) *Vortex Dynamics and Vortex Methods*, Lecture in Applied Mathematics, Vol. 28, American Mathematical Society, Providence, RI, 1991, 219–301.
- [10] Glowinski, R. & Pironneau, O., Finite element methods for Navier-Stokes equations, *Annual Review of Fluid Mechanics*, **24**(1992), 167–204.
- [11] Dean, E.J. & Glowinski, R., A wave equation approach to the numerical solution of the Navier-Stokes equations for incompressible viscous flow, *C.R. Acad. Sci. Paris*, **325** Série I, (1997), 789–797.
- [12] Dean, E. J., Glowinski, R. & Pan, T. W., A wave equation approach to the numerical simulation of incompressible viscous fluid flow modeled by the Navier-Stokes equations, In De Santo, J.A. (ed.) *Mathematical and Numerical Aspects of Wave Propagation*, SIAM, Philadelphia, PA, 1998, 65–74.
- [13] Pan, T. W. & Glowinski, R., A projection/wave-like equation method for natural convection flow in enclosures, In Bathe, K. J. (ed.) *Computational Fluid and Solid Mechanics, Vol. 2*, Elsevier, Amsterdam, 2001, 1493–1496.
- [14] He, J. W. & Glowinski, R., Steady Bingham fluid flow in cylindrical pipes: a time dependent approach to the iterative solution, *Num. Lin. Algebra with Appl.*, **7**(2000), 381–428.
- [15] Glowinski, R., Méthodes numériques pour l'écoulement stationnaire d'un fluide rigide visco-plastique incompressible, In Holt, M. (ed.) *Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics, September 15-19, 1970, University of California*, Berkeley, Lecture Notes in Physics, Vol. 8, Springer-Verlag, Berlin, 1971, 385–394.
- [16] Begis, D., Analyse Numérique de l'Écoulement d'un Fluide de Bingham, Thèse de 3^e Cycle, Université Pierre et Marie Curie, Paris, France, 1972.
- [17] Begis, D. & Glowinski, R., Application des méthodes de Lagrangien augmenté à la résolution numérique d'écoulements bi-dimensionnels de fluides visco-plastiques incompressibles, In Fortin, M. and Glowinski, R. (eds.) *Méthodes de Lagrangien Augmenté: Application à la Résolution Numérique de Problèmes aux Limites*, Dunod-Bordas, Paris, 1983, 219–240.
- [18] Begis, D. & Glowinski, R., Application to the numerical solution of the two-dimensional flow of incompressible viscoplastic fluids, In Fortin, M. and Glowinski, R., (eds.) *Augmented Lagrangian methods: Application to the numerical solution of boundary value problems*, North-Holland, Amsterdam, 1983, 233–255.
- [19] Glowinski, R. & Le Tallec, P., Augmented Lagrangians and operator splitting methods in nonlinear mechanics, SIAM, Philadelphia, PA, 1989.
- [20] Strauss, M. J., Variation of Korn's and Sobolev's Inequalities, In *Proceedings of Symposia in Pure Mathematics*, Vol. 23, American Math. Society, Providence, RI, 1973, 207–214.

ON THE UNIQUENESS OF THE WEAK SOLUTIONS OF A QUASILINEAR HYPERBOLIC SYSTEM WITH A SINGULAR SOURCE TERM*

J. P. DIAS M. FIGUEIRA

CMAF/UL, Av. Prof. Gama Pinto, 2, 1649-003

Lisboa-Portugal. E-mail: dias@lmc.fc.ul.pt

Abstract

This paper is a continuation of the authors' previous paper [1]. In this paper the authors prove, assuming additional conditions on the initial data, some results about the existence and uniqueness of the entropy weak solutions of the Cauchy problem for the singular hyperbolic system

$$\begin{cases} a_t + (au)_x + \frac{2au}{x} = 0, \\ u_t + \frac{1}{2} (a^2 + u^2)_x = 0, \end{cases} \quad x > 0, t \geq 0.$$

Keywords Cauchy problem, Weak solution, Quasilinear hyperbolic system

2000 MR Subject Classification 35L45, 35L67

§1. Introduction and Main Results

We consider the Cauchy problem for the quasilinear hyperbolic system

$$\begin{cases} a_t + (au)_x + \frac{2au}{x} = 0, \\ u_t + \frac{1}{2} (a^2 + u^2)_x = 0, \end{cases} \quad x > 0, t \geq 0, \quad (1.1)$$

*Project supported by the FCT and FEDER.

with the initial data

$$(a(x, 0), u(x, 0)) = (a_0(x), u_0(x)), \quad x > 0. \quad (1.2)$$

The system (1.1) appears in the study of the radial symmetric solutions in $\mathbf{R}^3 \times \mathbf{R}_+$ for a conservative system modelling the isentropic flow introduced by G.B. Whitham in [7, Chap.9] where a is the sound speed and u is the radial velocity. If $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ is defined by $f(a, u) = (au, \frac{1}{2}(a^2 + u^2))$, then two eigenvalues of ∇f are

$$\lambda_1 = u - a, \quad \lambda_2 = u + a \quad (1.3)$$

and so the strict hyperbolicity fails if $a = 0$, but the system is genuinely nonlinear with Riemann invariants

$$l = -u + a, \quad r = u + a \quad (1.4)$$

which satisfy the equivalent system (for classical solutions) :

$$\begin{cases} r_t + \frac{1}{2}(r^2)_x + \frac{r^2 - l^2}{2x} = 0, \\ l_t - \frac{1}{2}(l^2)_x + \frac{r^2 - l^2}{2x} = 0, \end{cases} \quad x > 0, \quad t \geq 0, \quad (1.5)$$

with initial data

$$(r(x, 0), l(x, 0)) = (r_0(x), l_0(x)), \quad x > 0, \quad (1.6)$$

with $r_0 = u_0 + a_0$, $l_0 = -u_0 + a_0$.

Following [1], if $a_0, u_0 \in H^1_0(\mathbf{R}_+)$, we will say that

$$v = (a, u) \in (L^\infty_{\text{loc}}([0, +\infty[\times [0, +\infty[))^2$$

is a weak solution for the Cauchy problem (1.1), (1.2) in $\mathbf{R}_+ \times [0, +\infty[$ if, for each pair $\varphi \in C^\infty_0(\mathbf{R}_+ \times [0, +\infty[)$, $\psi \in C^\infty_0([0, +\infty[\times [0, +\infty[)$,

$$\begin{aligned} & \int_{\mathbf{R}_+ \times [0, +\infty[} \left(a\varphi_t + au\varphi_x - \frac{2au}{x}\varphi \right) dx dt \\ & + \int_{\mathbf{R}_+ \times [0, +\infty[} \left(u\psi_t + \frac{1}{2}(a^2 + u^2)\psi_x \right) dx dt \\ & + \int_{\mathbf{R}_+} a_0(x)\varphi(x, 0) dx + \int_{\mathbf{R}_+} u_0(x)\psi(x, 0) dx = 0. \end{aligned} \quad (1.7)$$

A weak notion of null boundary condition for v (at $x = 0$) is contained in (1.7). Moreover, we will say that $v = (a, u)$ verifying (1.7) is an entropy weak solution if, for every pair of smooth functions $\eta, q : \mathbf{R}^2 \rightarrow \mathbf{R}$, η

convex (entropy/entropy flux pair) such that $\nabla\eta \cdot \nabla f = \nabla q$ in \mathbf{R}^2 , we have

$$\frac{\partial}{\partial t} \eta(v) + \frac{\partial}{\partial t} q(v) + \nabla\eta(v) \cdot \left(\frac{2au}{x}, 0 \right) \leq 0 \quad (1.8)$$

in $\mathcal{D}'(\mathbf{R}_+ \times \mathbf{R}_+)$. By applying the compensated compactness method of Tartar, Murat and DiPerna (cf. [2] and [3]) and some ideas of M. E. Schonbek in [6] we have proved in [1] the following result:

Theorem 1.1. *Assume $a_0, u_0 \in H_0^1(\mathbf{R}_+)$, $u_0(x) \geq a_0(x) \geq 0$, $x \in \mathbf{R}_+$. Then, there exists $v = (a, u) \in (L^\infty(\mathbf{R}_+ \times [0, +\infty[))^2$, with $u \geq a \geq 0$ a.e. in $\mathbf{R}_+ \times [0, +\infty[$, which is an entropy weak solution for the Cauchy problem (1.1), (1.2) in $\mathbf{R}_+ \times [0, +\infty[$. Moreover there exists a sequence $v_\varepsilon = (a_\varepsilon, u_\varepsilon) \in (C([0, +\infty[; H^3 \cap H_0^1) \cap C^1([0, +\infty[; H^1) \cap L^\infty(\mathbf{R}_+^2))^2$ such that $0 \leq a_\varepsilon \leq u_\varepsilon \leq M$, $v_\varepsilon \rightarrow v$ a.e. in $\mathbf{R}_+ \times [0, +\infty[$ and in $(L^\infty(\mathbf{R}_+^2))^2$ weak *, $v_\varepsilon(\cdot, 0) \rightarrow v(\cdot, 0)$ in $(H_0^1(\mathbf{R}_+))^2$ and v_ε is the solution of the approximate parabolic system*

$$\begin{cases} a_{\varepsilon t} + (a_\varepsilon u_\varepsilon)_x + \frac{2a_\varepsilon u_\varepsilon}{x+\varepsilon} = \varepsilon a_{\varepsilon xx}, \\ u_{\varepsilon t} + \frac{1}{2}(a_\varepsilon^2 + u_\varepsilon^2)_x = \varepsilon u_{\varepsilon xx}, \end{cases} \quad x > 0, t \geq 0, \quad (1.9)$$

with initial data

$$v_{0\varepsilon} = (a_{0\varepsilon}, u_{0\varepsilon}) \in (H^3(\mathbf{R}_+) \cap H_0^2(\mathbf{R}_+))^2, \quad u_{0\varepsilon} \geq a_{0\varepsilon} \geq 0.$$

In the framework of Theorem 2.1, we have, for

$$l_\varepsilon = -u_\varepsilon + a_\varepsilon, \quad r_\varepsilon = u_\varepsilon + a_\varepsilon,$$

$$l = -u + a, \quad r = u + a; \quad (1.10)$$

$$l_\varepsilon \leq 0, \quad 0 \leq r_\varepsilon \leq M_1, \quad r_\varepsilon^2 - l_\varepsilon^2 \geq 0, \quad l \leq 0, \quad 0 \leq r \leq M_1, \quad r^2 - l^2 \geq 0, \quad \text{a.e.},$$

$$\begin{cases} r_{\varepsilon t} + \frac{1}{2}(r_\varepsilon^2)_x + \frac{r_\varepsilon^2 - l_\varepsilon^2}{2(x+\varepsilon)} = \varepsilon r_{\varepsilon xx}, \\ l_{\varepsilon t} - \frac{1}{2}(l_\varepsilon^2)_x + \frac{r_\varepsilon^2 - l_\varepsilon^2}{2(x+\varepsilon)} = \varepsilon l_{\varepsilon xx}, \end{cases} \quad x > 0, t \geq 0, \quad (1.11)$$

with $(r_\varepsilon, l_\varepsilon) \in (C([0, +\infty[; H^3 \cap H_0^1) \cap C^1([0, +\infty[; H^1))^2$ and

$$r_\varepsilon(\cdot, 0) = r_{0\varepsilon} = u_{0\varepsilon} + a_{0\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} r_0,$$

$$l_\varepsilon(\cdot, 0) = l_{0\varepsilon} = -u_{0\varepsilon} + a_{0\varepsilon} \xrightarrow{\varepsilon \rightarrow 0} l_0 \quad \text{in } H_0^1(\mathbf{R}_+). \quad (1.12)$$

We will prove the following estimate:

Theorem 1.2. *In the framework of Theorem 1.1, assume $\frac{a_0}{x}, \frac{u_0}{x} \in L^\infty(\mathbf{R}_+)$. Then, we have, for each $t \geq 0$,*

$$\left\| \frac{r_\varepsilon(\cdot, t)}{x + \varepsilon} \right\|_\infty \leq \left\| \frac{r_{0\varepsilon}}{x + \varepsilon} \right\|_\infty \leq \left\| \frac{r_0}{x} \right\|_\infty. \tag{1.13}$$

Hence,

$$\left\| \frac{l_\varepsilon(\cdot, t)}{x + \varepsilon} \right\|_\infty \leq \left\| \frac{r_\varepsilon(\cdot, t)}{x + \varepsilon} \right\|_\infty \leq \left\| \frac{r_0}{x} \right\|_\infty$$

and, a.e. in $t \in [0, +\infty[$, $\frac{l(\cdot, t)}{x}, \frac{r(\cdot, t)}{x} \in L^\infty(\mathbf{R}_+)$ with

$$\left\| \frac{l(\cdot, t)}{x} \right\|_\infty \leq \left\| \frac{r(\cdot, t)}{x} \right\|_\infty \leq \left\| \frac{r_0}{x} \right\|_\infty.$$

Now, by an adaptation of Kruzkov’s methods (cf. [5]), we can prove the following theorem.

Theorem 1.3. *Under the hypothesis of Theorems 1.1 and 1.2 assume*

$$a_0, u_0 \in BV(\mathbf{R}_+) = \{w \in L^1(\mathbf{R}_+) \mid TV(w) < +\infty\}$$

where TV denotes the total variation in \mathbf{R}_+ . Then, we have

$$\left\| \frac{\partial r_\varepsilon(\cdot, t)}{\partial t} \right\|_1 + \left\| \frac{\partial l_\varepsilon(\cdot, t)}{\partial t} \right\|_1 \leq c e^{c_1 t}, \quad t \geq 0, \tag{1.14}$$

where $c = c\left(\|a_0\|_\infty, \|u_0\|_\infty, \left\| \frac{a_0}{x} \right\|_\infty, \left\| \frac{u_0}{x} \right\|_\infty, TV(a_0), TV(u_0)\right) > 0$ and $c_1 = c_1\left(\left\| \frac{a_0}{x} \right\|_\infty, \left\| \frac{u_0}{x} \right\|_\infty\right) > 0$ do not depend on ε .

From Theorem 1.3 it is easy to derive

Corollary 1.1. *Under the assumptions of Theorem 1.3, the weak entropy solution (a, u) of the Cauchy problem (1.1), (1.2) obtained by the vanishing viscosity method verifies $a(\cdot, t), u(\cdot, t) \in L^1(\mathbf{R}_+)$ a.e. in t and there exists $E \subset [0, +\infty[$ such that $m([0, +\infty[\setminus E) = 0$ and*

$$\lim_{\substack{t \rightarrow 0^+ \\ t \in E}} \int_{\mathbf{R}_+} (|a(x, t) - a_0(x)| + |u(x, t) - u_0(x)|) dx = 0.$$

Finally, also by adaptation of Kruzkov’s method (cf. [5]), we will prove the following theorem.

Theorem 1.4. *Let the initial conditions a_0, u_0 be in $H_0^1(\mathbf{R}_+)$ and such that $\frac{a_0}{x}, \frac{u_0}{x} \in L^\infty(\mathbf{R}_+)$, and let $(a_1, u_1), (a_2, u_2)$ be two weak entropy solu-*

tions of the Cauchy problem (1.1), (1.2) such that, for $i = 1, 2$,

$$\frac{a_i}{x}, \frac{u_i}{x} \in L_{\text{loc}}^\infty([0, +\infty[\times [0, +\infty[)$$

and there exists $E \subset [0, +\infty[$, with $m([0, +\infty[\setminus E) = 0$, such that for each $R > 0$,

$$\lim_{\substack{t \rightarrow 0^+ \\ t \in E}} \int_{0 < x < R} (|a_i(x, t) - a_0(x)| + |u_i(x, t) - u_0(x)|) dx = 0. \quad (1.15)$$

Then, $(a_1, u_1) = (a_2, u_2)$ a.e. in $\mathbf{R}_+ \times [0, +\infty[$.

§2. Proof of Theorem 1.2.

With $r = r_\varepsilon$, $0 < \varepsilon < 1$, let us consider the first equation of (1.11):

$$r_t + r r_x + \frac{r^2 - l^2}{2(x + \varepsilon)} = \varepsilon r_{xx}. \quad (2.1)$$

With $v = v_\varepsilon = \frac{r_\varepsilon^2}{(x + \varepsilon)^2}$ we obtain, multiplying (2.1) by $\frac{2r}{(x + \varepsilon)^2}$, and since $r^2 - l^2 \geq 0$,

$$v_t + 2v r_x - 8\varepsilon \frac{r_x r}{(x + \varepsilon)^3} + 6\varepsilon \frac{r^2}{(x + \varepsilon)^4} + 2\varepsilon \frac{r_x^2}{(x + \varepsilon)^2} \leq \varepsilon v_{xx}. \quad (2.2)$$

If we multiply (2.2) by v^p , $p \geq 1$, and integrate in \mathbf{R}_+ , we obtain with $\int \cdot = \int_{\mathbf{R}_+} \cdot dx$,

$$\begin{aligned} & \frac{1}{p+1} \frac{\partial}{\partial t} \int v^{p+1} + 2 \int v^{p+1} r_x - 8\varepsilon \int \frac{r_x r v^p}{(x + \varepsilon)^3} \\ & + 6\varepsilon \int \frac{r^2 v^p}{(x + \varepsilon)^4} + 2\varepsilon \int \frac{r_x^2 v^p}{(x + \varepsilon)^2} + p\varepsilon \int (v_x)^2 v^{p+1} \leq 0. \end{aligned} \quad (2.3)$$

We have

$$\begin{aligned} 2 \int v^{p+1} r_x &= 2 \int v^p r_x \left(\frac{r}{x + \varepsilon} \right)^2 = \frac{2}{3} \int v^p \frac{1}{(x + \varepsilon)^2} \frac{\partial}{\partial x} (r^3) \\ &= -\frac{2}{3} \int \frac{\partial}{\partial x} (v^p) \frac{1}{(x + \varepsilon)^2} r^3 + \frac{4}{3} \int v^p \frac{1}{(x + \varepsilon)^3} r^3 \\ &= -\frac{2}{3} \int \frac{\partial}{\partial x} (v^p) v r + \frac{4}{3} \int v^p v^{3/2} \\ &= -\frac{2}{3} p \int v^p v_x r + \frac{4}{3} \int v^{p+3/2} \end{aligned}$$

$$\begin{aligned}
&= -\frac{2}{3} \frac{p}{p+1} \int \frac{\partial}{\partial x} (v^{p+1}) r + \frac{4}{3} \int v^{p+3/2} \\
&= \frac{2}{3} \frac{p}{p+1} \int (v^{p+1}) r_x + \frac{4}{3} \int v^{p+3/2}
\end{aligned}$$

and so

$$2 \int v^{p+1} r_x = \frac{4(p+1)}{2p+3} \int v^{p+3/2}. \quad (2.4)$$

Moreover,

$$\begin{aligned}
-8 \int \frac{r_x r v^p}{(x+\varepsilon)^3} &= -4 \int \frac{\frac{\partial}{\partial x} (r^2) v^p}{(x+\varepsilon)^3} = -12 \int \frac{r^2 v^p}{(x+\varepsilon)^4} + 4 \int \frac{r^2 \frac{\partial}{\partial x} (v^p)}{(x+\varepsilon)^3} \\
&= -12 \int \frac{v^{p+1}}{(x+\varepsilon)^2} + 4p \int \frac{r^2 v^{p-1} v_x}{(x+\varepsilon)^3} \\
&= -12 \int \frac{v^{p+1}}{(x+\varepsilon)^2} + 4 \frac{p}{p+1} \int \frac{\frac{\partial}{\partial x} (v^{p+1})}{(x+\varepsilon)} \\
&= -12 \int \frac{v^{p+1}}{(x+\varepsilon)^2} + 4 \frac{p}{p+1} \int \frac{v^{p+1}}{(x+\varepsilon)^2},
\end{aligned}$$

and so

$$-8\varepsilon \int \frac{r_x r v^p}{(x+\varepsilon)^3} = \left(-12 + 4 \frac{p}{p+1}\right) \varepsilon \int \frac{v^{p+1}}{(x+\varepsilon)^2}. \quad (2.5)$$

We have also

$$v_x = \frac{2rr_x}{(x+\varepsilon)^2} - 2 \frac{r^2}{(x+\varepsilon)^3}$$

and so, with $p \geq 2$,

$$\begin{aligned}
v_x v^{p/2-1} r &= 2v^{p/2} r_x - 2v^{\frac{p+1}{2}}, \\
v^{p/2} r_x &= \frac{1}{2} v_x v^{\frac{p-2}{2}} r + v^{\frac{p+1}{2}}, \\
v^p r_x^2 &= \frac{1}{4} v_x^2 v^{p-1} (x+\varepsilon)^2 + v^{p+1} + v_x v^p (x+\varepsilon),
\end{aligned}$$

$$\begin{aligned}
2 \int \frac{r_x^2 v^p}{(x+\varepsilon)^2} &= \frac{1}{2} \int v^{p-1} v_x^2 + 2 \int \frac{v^{p+1}}{(x+\varepsilon)^2} + \frac{2}{p+1} \int \frac{\frac{\partial}{\partial x} (v^{p+1})}{(x+\varepsilon)}, \\
2\varepsilon \int \frac{r_x^2 v^p}{(x+\varepsilon)^2} &= \frac{1}{2} \varepsilon \int v^{p-1} v_x^2 + \left(2 + \frac{2}{p+1}\right) \varepsilon \int \frac{v^{p+1}}{(x+\varepsilon)^2}. \quad (2.6)
\end{aligned}$$

By (2.3), \dots , (2.6), we obtain

$$\frac{1}{p+1} \frac{\partial}{\partial t} \int v^{p+1} + \frac{4(p+1)}{2p+3} \int v^{p+3/2} \leq \frac{2}{p+1} \varepsilon \int \frac{v^{p+1}}{(x+\varepsilon)^2}. \quad (2.7)$$

But, since

$$\frac{v^{p+1}}{(x+\varepsilon)^2} = \frac{v^{\frac{p+1}{2}}}{(x+\varepsilon)} \cdot v^{\frac{p+1}{2}} \cdot \frac{1}{(x+\varepsilon)}$$

and $\frac{1}{2} + \frac{p+1}{2p+3} + \frac{1}{4p+6} = 1$, we derive by Hölder's inequality,

$$\begin{aligned} \frac{2\varepsilon}{p+1} \int \frac{v^{p+1}}{(x+\varepsilon)^2} &\leq \frac{2\varepsilon^{1/2}}{(p+1)^{1/2}} \left(\int \frac{v^{p+1}}{(x+\varepsilon)^2} \right)^{\frac{1}{2}} \left(\int v^{p+3/2} \right)^{\frac{p+1}{2p+3}} \\ &\quad \frac{\varepsilon^{1/2}}{(p+1)^{1/2}} \left(\int \frac{1}{(x+\varepsilon)^{4p+6}} \right)^{\frac{1}{4p+6}} \end{aligned}$$

and so

$$\frac{2\varepsilon}{p+1} \int \frac{v^{p+1}}{(x+\varepsilon)^2} \leq 2 \left(\int v^{p+3/2} \right)^{\frac{2p+2}{2p+3}} \frac{1}{p+1} \frac{1}{(4p+5)^{2p+3}} \varepsilon^{2\left(\frac{1}{2} - \frac{4p+5}{4p+6}\right)}.$$

We derive, by the inequality $b^{1/q} c^{1/q'} \leq \frac{1}{q} b + \frac{1}{q'} c$, with $q = \frac{2p+3}{2p+2}$, $q' = 2p+3$,

$$\begin{aligned} &\frac{2\varepsilon}{p+1} \int \frac{v^{p+1}}{(x+\varepsilon)^2} \\ &\leq 2 \left[\frac{2p+2}{2p+3} \int v^{p+3/2} + \frac{1}{2p+3} \frac{1}{(p+1)^{2p+3}} \frac{1}{4p+5} \varepsilon^{-2(p+1)} \right]. \end{aligned} \quad (2.8)$$

If we fix $n \geq 1$, let p be such that

$$\frac{\varepsilon^{-2(p+1)}}{(p+1)^{2p+3}} \leq \varepsilon^{n(p+1)} \left(\text{that is, } -\log \varepsilon \leq \frac{2p+3}{(p+1)(n+2)} \log(p+1) \right).$$

We deduce, from (2.8),

$$\frac{2\varepsilon}{p+1} \int \frac{v^{p+1}}{(x+\varepsilon)^2} \leq \frac{4(p+1)}{2p+3} \int v^{p+3/2} + \frac{2}{(4p+5)(2p+3)} \varepsilon^{n(p+1)}. \quad (2.9)$$

From (2.7) and (2.9) we derive

$$\frac{\partial}{\partial t} \int v^{p+1} \leq \frac{2(p+1)}{(4p+5)(2p+3)} \varepsilon^{n(p+1)}. \quad (2.10)$$

Hence, for $t \geq 0$, we obtain

$$\|v_\varepsilon(\cdot, t)\|_{p+1}^{p+1} \leq \|v_{0\varepsilon}\|_{p+1}^{p+1} + \frac{2(p+1)t}{(4p+5)(2p+3)} \varepsilon^{n(p+1)}$$

and so, by the inequality $(b^{p+1} + a^{p+1})^{1/(p+1)} \leq b + c$, we deduce

$$\|v_\varepsilon(\cdot, t)\|_{p+1} \leq \|v_{0\varepsilon}\|_{p+1} + \left[\frac{2(p+1)t}{(4p+5)(2p+3)} \right]^{\frac{1}{p+1}} \varepsilon^n. \tag{2.11}$$

Letting $p \rightarrow +\infty$ we obtain from (2.11), $\|v_\varepsilon(\cdot, t)\|_\infty \leq \|v_{0\varepsilon}\|_\infty + \varepsilon^n$. Now we let $n \rightarrow +\infty$ and we derive

$$\left\| \frac{r_\varepsilon(\cdot, t)}{x + \varepsilon} \right\|_\infty^2 = \|v_\varepsilon(\cdot, t)\|_\infty \leq \|v_{0\varepsilon}\|_\infty = \left\| \frac{r_{0\varepsilon}}{x + \varepsilon} \right\|_\infty^2 \leq \left\| \frac{r_0}{x} \right\|_\infty^2.$$

Finally we let $\varepsilon \rightarrow 0$ and we deduce a.e. on $t \in [0, +\infty[$, $\frac{r(\cdot, t)}{x} \in L^\infty(\mathbf{R}_+)$ and

$$\left\| \frac{r(\cdot, t)}{x} \right\|_\infty \leq \left\| \frac{r_0}{x} \right\|_\infty$$

and so

$$\left\| \frac{l(\cdot, t)}{x} \right\|_\infty \leq \left\| \frac{r(\cdot, t)}{x} \right\|_\infty \leq \left\| \frac{r_0}{x} \right\|_\infty,$$

and this achieves the proof of Theorem 1.2.

§3. Proofs of Theorems 1.3 and 1.4

3.1. Sketch of the proof of Theorem 1.3

We follow the lines of the proofs of Theorems 2.3 and 3.1 of Chap.II in [4], applying the method of Kruzkov for the case of scalar conservation laws (cf. [5]). We take the t derivative in both equations of the approximate system (1.11) in r_ε and l_ε ; we multiply the first equation by $\text{sgn}(r_{\varepsilon t})$, the second equation by $\text{sgn}(l_{\varepsilon t})$ and both equations by $\psi_R(x) = \chi(x/R)$, $R > 0$, where sgn denotes the usual sign function and χ is the cut function introduced in (2.25) of Chap.II in [4]. We integrate in \mathbf{R}_+ and we add the two equations. If we point out the estimate

$$\begin{aligned} & \left| \int_{\mathbf{R}_+} \frac{2(r_\varepsilon r_{\varepsilon t} - l_\varepsilon l_{\varepsilon t})}{2(x + \varepsilon)} [\text{sgn}(r_{\varepsilon t}) + \text{sgn}(l_{\varepsilon t})] \psi_R(x) dx \right| \\ & \leq c_1 \int_{\mathbf{R}_+} (|r_{\varepsilon t}| + |l_{\varepsilon t}|) \psi_R(x) dx, \end{aligned}$$

by Theorem 1.2, we easily deduce, with the help of Lemma 3.1 in Chap.II of [4],

$$\int_{\mathbf{R}_+} (|r_{\varepsilon t}(x, t)| + |l_{\varepsilon t}(x, t)|) \psi_R(x) dx$$

$$\leq c + \left(\frac{c_0}{R} + c_1\right) \int_0^t \int_{\mathbf{R}_+} (|r_{\varepsilon t}(x, \tau)| + |l_{\varepsilon t}(x, \tau)|) \psi_R(x) dx d\tau,$$

where

$$c_0 = c_0(\|a_0\|_\infty, \|u_0\|_\infty),$$

$$c = c\left(\|a_0\|_\infty, \|u_0\|_\infty, \left\|\frac{a_0}{x}\right\|_\infty, \left\|\frac{u_0}{x}\right\|_\infty, TV(a_0), TV(u_0)\right),$$

$$c_1 = c_1\left(\left\|\frac{a_0}{x}\right\|_\infty, \left\|\frac{u_0}{x}\right\|_\infty\right)$$

are positive constants not depending on ε . The result follows if we apply Gronwall's inequality and then let $R \rightarrow \infty$.

Before proving Theorem 1.4, we need the following lemma that can be proved like the Lemma 4.2 in Chap.II of [4] (cf. also the inequality (3.12) in [5]).

Lemma 3.1. *Let $(a_i, u_i) \in (L_{loc}^\infty([0, +\infty[\times]0, +\infty]^2), i = 1, 2$, be two solutions in $\mathcal{D}'(\mathbf{R}_+ \times \mathbf{R}_+)$ of the system (1.1) verifying the entropy condition (1.8) and let $r_i = u_i + a_i, l_i = -u_i + a_i, i = 1, 2$. Then, we have, in $\mathcal{D}'(\mathbf{R}_+ \times \mathbf{R}_+)$,*

$$\frac{\partial}{\partial t} |r_1 - r_2| + \frac{\partial}{\partial x} \left[\operatorname{sgn}(r_1 - r_2) \frac{1}{2} (r_1^2 - r_2^2) \right]$$

$$+ \operatorname{sgn}(r_1 - r_2) \frac{1}{2x} [(r_1^2 - l_1^2) - (r_2^2 - l_2^2)] \leq 0,$$

$$\frac{\partial}{\partial t} |l_1 - l_2| - \frac{\partial}{\partial x} \left[\operatorname{sgn}(l_1 - l_2) \frac{1}{2} (l_1^2 - l_2^2) \right]$$

$$+ \operatorname{sgn}(l_1 - l_2) \frac{1}{2x} [(r_1^2 - l_1^2) - (r_2^2 - l_2^2)] \leq 0.$$

3.2. Sketch of the proof of Theorem 1.4

We follow the lines of the proof of Theorem 4.1 in Chap.II of [4], applying the method of Kruzkov for the case of scalar conservation laws (cf. [5]). With $T > 0, R > 0$ and $W = L^\infty([0, R + MT + 1] \times [0, T + 1])$, let us put

$$M = M(T, R) = \max_{1 \leq i \leq 2} (\|r_i\|_W, \|l_i\|_W),$$

where $r_i = u_i + a_i, l_i = -u_i + a_i, i = 1, 2$. For $\delta, \varepsilon, \theta > 0, \varepsilon < \delta <$

$\min(1, T/3)$, $\theta < \min(1, R/2)$ and $x > 0$, we set

$$\varphi(x, t) = (1 - h(x/\theta))(Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T))(1 - Y_\theta(x - R - M(T - t)))$$

where $Y_\varepsilon(t) = \int_{-\infty}^t \zeta_\varepsilon(s) ds$, $\zeta_\varepsilon \in \mathcal{D}(\mathbf{R})$ is a positive cut-off function with support in $[-\varepsilon, \varepsilon]$, $h \in C^\infty([0, +\infty[)$, $0 \leq h \leq 1$, $h(x) = 0$ if $x \geq 2$ and $h(x) = 1$ if $x \leq 1$. With this choice of test function we deduce from the inequalities in Lemma 3.1:

$$\begin{aligned} & \int_0^{+\infty} \int_0^{+\infty} (|r_1 - r_2| + |l_1 - l_2|) (1 - h(x/\theta)) (\zeta_\varepsilon(t - \delta) - \zeta_\varepsilon(t - T)) \\ & \cdot Y_\theta(x - R - M(T - t)) dx dt \\ & - \int_0^{+\infty} \int_0^{+\infty} (1 - h(x/\theta)) (Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)) \left[M(|r_1 - r_2| + |l_1 - l_2|) \right. \\ & \left. + \operatorname{sgn}(r_1 - r_2) \frac{1}{2} (r_1^2 - r_2^2) - \operatorname{sgn}(l_1 - l_2) \frac{1}{2} (l_1^2 - l_2^2) \right] \\ & \cdot \zeta_\theta(x - R - M(T - t)) dx dt \\ & - \int_0^{+\infty} \int_0^{+\infty} \frac{1}{\theta} h'(x/\theta) (Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)) (1 - Y_\theta(x - R - M(T - t))) \\ & \cdot \left[\operatorname{sgn}(r_1 - r_2) \frac{1}{2} (r_1^2 - r_2^2) - \operatorname{sgn}(l_1 - l_2) \frac{1}{2} (l_1^2 - l_2^2) \right] dx dt \\ & - \int_0^{+\infty} \int_0^{+\infty} (1 - h(x/\theta)) (Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)) \\ & (1 - Y_\theta(x - R - M(T - t))) \\ & \cdot [\operatorname{sgn}(r_1 - r_2) + \operatorname{sgn}(l_1 - l_2)] \frac{1}{2x} [(r_1^2 - l_1^2) - (r_2^2 - l_2^2)] dx dt \geq 0. \end{aligned}$$

Now, we point out that in the third integral in the previous inequality we can put $r_1^2 - r_2^2 = (r_1 - r_2) \frac{(r_1 + r_2)}{x} x$, $l_1^2 - l_2^2 = (l_1 - l_2) \frac{(l_1 + l_2)}{x} x$. Hence, by the assumptions, this integral can be estimated in modulus by

$$c \int_0^{T+1} \int_\theta^{2\theta} \frac{x}{\theta} dx dt \leq c(T+1) \frac{3}{2} \theta \xrightarrow{\theta \rightarrow 0^+} 0.$$

Moreover, regarding the last integral in the same inequality, we observe that

$$\begin{aligned} \left| \frac{1}{x} [(r_1^2 - l_1^2) - (r_2^2 - l_2^2)] \right| & \leq \frac{|r_1| + |r_2|}{x} |r_1 - r_2| + \frac{|l_1| + |l_2|}{x} |l_1 - l_2| \\ & \leq c(|r_1 - r_2| + |l_1 - l_2|). \end{aligned}$$

We can now continue as in the proof of Theorem 4.1 in Chap.II of [4] to deduce, for almost T and δ ,

$$\begin{aligned} & \int_{|x| \leq R} (|r_1(x, T) - r_2(x, T)| + |l_1(x, T) - l_2(x, T)|) dx \\ & \leq \int_{|x| \leq R+M(T-\delta)} (|r_1(x, \delta) - r_2(x, \delta)| + |l_1(x, \delta) - l_2(x, \delta)|) dx \\ & \quad + c \int_0^T \int_{|x| \leq R+M(T-\tau)} (|r_1(x, \tau) - r_2(x, \tau)| + |l_1(x, \tau) - l_2(x, \tau)|) dx \end{aligned}$$

with $c = c(T, R) > 0$, increasing function of T and R . Hence, by the assumption (1.15), we derive, a.e. in $T > 0$, since the two solutions have the same initial data,

$$\begin{aligned} & \int_{|x| \leq R} (|r_1(x, T) - r_2(x, T)| + |l_1(x, T) - l_2(x, T)|) dx \\ & \leq c \int_0^T \int_{|x| \leq R+M(T-\tau)} (|r_1(x, \tau) - r_2(x, \tau)| + |l_1(x, \tau) - l_2(x, \tau)|) dx \end{aligned}$$

and this implies

$$\int_{|x| \leq R} (|r_1(x, T) - r_2(x, T)| + |l_1(x, T) - l_2(x, T)|) dx = 0$$

for all $R > 0$ and a.e. in $T > 0$, which achieves the proof of Theorem 1.4.

REFERENCES

- [1] Dias, J. P. & Figueira, M., On the radial weak solutions of a conservative system modeling the isentropic flow, *Rendic. di Mat.*, **21**(2001), 245–258.
- [2] DiPerna, R. J., Convergence of approximate solutions to conservation laws, *Arch. Rat. Mech. Anal.*, **82**(1983), 27–70.
- [3] DiPerna, R. J., Convergence of the viscosity method for isentropic gas dynamics, *Commun. Math. Phys.*, **91**(1983), 1–30.
- [4] Godlewski, E. & Raviart, P. A., Hyperbolic systems of conservation laws, *Mathématiques et Applications*, SMAI, Ellipses, 1991.
- [5] Kruzkov, S. N., First order quasilinear equations in several independent variables, *Math. USSR Sbornik*, **10**(1970), 217–243.
- [6] Schonbek, M. E., Existence of solutions to singular conservation laws, *SIAM J. Math. Anal.*, **15**(1984), 1125–1139.
- [7] Whitham, G. B., *Linear and nonlinear waves*, Wiley Interscience, New York, 1974.

This page intentionally left blank

ON THE CLASSIFICATION OF INITIAL DATA FOR NONLINEAR WAVE EQUATIONS*

GU CHAOHAO (C. H. GU)

*Institute of Mathematics, Fudan University,
Shanghai 200433, China. E-mail: guch@fudan.ac.cn*

Abstract

The purpose of the present paper is to call for attention to the following question: Which of the initial data (non-small) admit global smooth solutions to the Cauchy problem for nonlinear wave equations. A few cases and examples are sketched, showing that the general answer of this question may be quite complicated.

Keywords Cauchy problem, Initial data, Global smooth solution
2000 MR Subject Classification 35Q51, 35Q58

§1. Introduction

For hyperbolic equations the most fundamental problem is the Cauchy problem. In the linear case, the well-posedness of the Cauchy problem has been established successfully long time ago [2, 6]. In the nonlinear case the existence and uniqueness of the local solution is also well-known [2, 12]. One of main problems to be settled is the existence of global smooth solutions. There are a series of important works in this direction. In particular, for a big class of nonlinear wave equations, the existence of global solutions to

*Project supported by the Chinese Special Funds for Major State Basic Research Project "Nonlinear Science".

the Cauchy problem with sufficient small initial data is quite clear [7 – 14]. However, the existence of global solutions to the simplest wave equation

$$u_{tt} - u_{xx} = f(u) \quad (1.1)$$

is still to be elucidated if the initial data

$$u(0, x) = \phi_0(x), \quad u_t(0, x) = \phi_1(x) \quad (1.2)$$

are non-small. For simplifying the statement we suppose that each function appeared is of C^∞ . Define $\Sigma[f]$ as the set of initial data for which the Cauchy problem (1.1) and (1.2) admits a global smooth solution on the entire t - x plane. The main problem is to identify $\Sigma[f]$. The purpose of the present paper is to call for attention to this problem. Some examples will be sketched so as to explain some approaches to the construction of $\Sigma[f]$. It is expected that these examples will give some insight to the further study.

§2. Hyperbolic Liouville Equation

The hyperbolic Liouville equation is

$$u_{tt} - u_{xx} = -e^u. \quad (2.1)$$

The equation has a general solution of the form [1]

$$u = \ln \frac{-2\psi'(\xi)\phi'(\eta)}{[\psi(\xi) + \phi(\eta)]^2}, \quad (2.2)$$

where $\xi = \frac{t+x}{2}$, $\eta = \frac{t-x}{2}$. It is seen that

(i) The expression (2.2) is meaningful if and only if

$$\psi'(\xi)\phi'(\eta) < 0, \quad \psi(\xi) + \phi(\eta) \neq 0. \quad (2.3)$$

(ii) u is unchanged if ψ and ϕ are replaced by

$$\psi_1 = a\psi + b, \quad \phi_1 = a\phi - b, \quad (a, b = \text{const.}) \quad (2.4)$$

(iii) u is a global solution if ψ, ϕ are defined on $(-\infty, \infty)$ and (2.3) holds for every $\xi \in (-\infty, \infty)$, $\eta \in (-\infty, \infty)$.

(iv) The expression (2.2) exhausts all C^∞ solutions not only in local sense but also and in global sense if (iii) holds.

(i), (ii) and (iii) are obvious, (iv) follows from the fact that the unique solution to the Goursat problem with the boundary condition

$$u|_{\xi=0} = \alpha(\eta), \quad u|_{\eta=0} = \beta(\xi) \quad (\alpha(0) = \beta(0)) \quad (2.5)$$

can be expressed in the form (2.2). Thus the initial data

$$u(0, x) = \phi_0(x), \quad u_t(0, x) = \phi_1(x) \tag{2.6}$$

admit a global solution on the t - x plane iff there exist ϕ, ψ defined on $(-\infty, \infty)$ and satisfy (iii) such that

$$\phi_0(x) = \ln \frac{-2\psi'(\frac{x}{2})\phi'(-\frac{x}{2})}{[\psi(\frac{x}{2}) + \phi(-\frac{x}{2})]^2}, \tag{2.7}$$

$$\phi_1(x) = \frac{1}{2} \frac{\psi''(\frac{x}{2})}{\psi'(\frac{x}{2})} + \frac{1}{2} \frac{\phi''(-\frac{x}{2})}{\phi'(-\frac{x}{2})} - \frac{\psi'(\frac{x}{2}) + \phi'(-\frac{x}{2})}{\psi(\frac{x}{2}) + \phi(-\frac{x}{2})}. \tag{2.8}$$

Consequently, the set of all C^∞ initial data with global solutions, i.e. $\Sigma[-e^u]$, has been obtained explicitly. However, it is quite difficult to check whether a given pair $(\phi_0(x), \phi_1(x))$ can be initial data of a global smooth solution. The problem is reduced to seeing whether the solution (ϕ, ψ) of (2.7) and (2.8) is a global one and satisfies (iii) or not.

Remark: In [1], there is a table of special nonlinear PDEs with explicit solutions. Some of them can be treated in a similar way.

§3. Some Cases for $\Sigma[f]=C^\infty \times C^\infty$

There is a big class of nonlinear wave equations, for which any given initial data ϕ_0, ϕ_1 will admit a global smooth solution. Among these nonlinear wave equations we mention

(i) Equations with $f(u)$ of weak nonlinearity

Let $h(t, x)$ be the solution of wave equation

$$\frac{\partial^2 h}{\partial t^2} - \frac{\partial^2 h}{\partial x^2} = 0$$

satisfying the initial data (ϕ_0, ϕ_1) . The equation (1.1) with the same initial data is equivalent to the integral equation

$$u(x, t) = h(t, x) + \frac{1}{2} \int_{\Delta_{(t,x)}} f(u(\tau, \xi)) d\tau d\xi, \tag{3.1}$$

where $\Delta_{(t,x)}$ is the characteristic triangle bounded by the x -axis and two characteristic lines passing through (t, x) . The function $f(u)$ is called to be of weak nonlinearity if

$$|f(u)| < |u|^{1+r}, \text{ for } |u| > A \text{ (} A \text{ is a positive constant, } r < 1). \tag{3.2}$$

It is well-known that the solution $u(x, t)$ exists globally. Example: sine-Gordon equation

$$u_{tt} - u_{xx} = \sin u.$$

(ii) Equations with strong nonlinear dissipation

Suppose that $f(u) = -F'(u)$ and $F(u) > ku^2$ (k is a positive constant). Let u be a solution on the characteristic triangle $\Delta_{(a,0)}$ with the vertices $(a, 0)$, $(0, -a)$ and $(0, a)$. Then we have the energy inequality

$$\int_{l_\tau} (u_t^2 + u_x^2 + F(u))dx \leq \int_{l_0} (\phi_1^2 + \phi_{0x}^2 + F(\phi_0))dx \quad (0 \leq \tau \leq a), \quad (3.3)$$

where l_τ is the interval $t = \tau$, $-a + \tau \leq x \leq a - \tau$. It follows that u is bounded in $\Delta_{(a,0)}$ and hence the solution $u(t, x)$ exists on $\Delta_{(a,0)}$ for any $a > 0$. Example: $f(u) = -\sinh u$, i.e. the minus sinh-Gordon equation

$$u_{tt} - u_{xx} = -\sinh u. \quad (3.4)$$

§4. Dual Relation

If we know $\Sigma[f]$, then $\Sigma[-f]$ can be constructed in the following way.

Let $(\phi_0, \phi_1) \in \Sigma[f]$. By solving the Cauchy problem for (1.1) with the initial data (ϕ_0, ϕ_1) we obtain the solution $u(t, x)$ which is defined on the whole x - t plane. Let

$$\psi_0 = u(t, 0), \quad \psi_1 = u_x(t, 0). \quad (4.1)$$

By interchanging t and x , it is obvious that the Cauchy problem for

$$v_{tt} - v_{xx} = -f(v) \quad (4.2)$$

with the initial data $v(0, x) = \psi_0(x)$, $v_t(0, x) = \psi_1(x)$ on the line $t = 0$ admits the global solution

$$v(t, x) = u(x, t). \quad (4.3)$$

The transformation $(\phi_0, \phi_1) \rightarrow (\psi_0, \psi_1)$ is denoted by L_f . Thus $\Sigma[-f]$ can be obtained from $\Sigma[f]$ via the transformation

$$L_f\{\phi_0, \phi_1\} = \{\psi_0, \psi_1\}. \quad (4.4)$$

The domain and range of L_f are $\Sigma[f]$ and $\Sigma[-f]$ respectively.

It is obvious that if $\Sigma[f]$ is nonempty, then $\Sigma[-f]$ is nonempty too and $L_{-f} = (L_f)^{-1}$. For example, let $f = \sinh u$. Then $\Sigma[\sinh u] = L_{(-\sinh u)} : \Sigma[\sinh(-u)] = L_{(-\sinh u)} : C^\infty \times C^\infty$.

Even in this case, the structure of $\Sigma[\sinh u]$ is not very clear and it is difficult to answer whether a given $\{\phi_0, \phi_1\}$ belongs to $\Sigma[\sinh u]$ or not. However, $\Sigma[\sinh u]$ is quite big and there exists a 1-1 mapping from $C^\infty \times C^\infty$ to $\Sigma[\sinh u]$. In $\Sigma[\sinh u]$, $|\phi_0|$ may be infinitely large, since in $\Sigma[-\sinh u]$, ϕ_0 can be arbitrary and $\psi_0(0) = \phi_0(0)$ if $(\psi_0, \psi_1) \in L_{(-\sinh u)}(\phi_0, \phi_1)$.

§5. Case of Hyperbolic Systems

The problem is more difficult even in the case of 1 + 1 dimension. However, we can mention the following facts. Besides those, in [9,10,12] there are a series of results on global classical solutions.

(i) For the Euler equations for isentropic gas, the only global smooth solutions are the trivial ones (i.e. $u=\text{const}$, $\rho=\text{const}$.) [3.10].

(ii) For the harmonic maps from R^{1+1} to any complete Riemann manifold (so called wave maps) any smooth initial data admits a global solution [4].

(iii) For the harmonic maps from R^{1+1} to the Lorentz surface $S^{1,1}$, the classification of initial data according to the global existence has been discussed in [5]. The situation is quite complicated.

In short, the problem of determining initial data such that the Cauchy problem admits global smooth solution is of significance and worthy to be studied further.

REFERENCES

- [1] Ames, W. F., Nonlinear partial differential equations in engineering, Mathematics in Science and Engineering 18, Academic Press, New York, 1965.
- [2] Courant, R. & Hilbert, D., Method on mathematical physics, Vol. II, International Publishers, New York, London, 1962.
- [3] Gu, C. H., Shock-free global solutions for quasilinear hyperbolic systems, Collections of Mathematical Papers of Fudan University, 1960, 36-39 (in Chinese).
- [4] Gu, C. H., On the Cauchy problem for harmonic map defined on two dimensional Minkowski space, *Comm. Pure Appl. Math.*, **33**(1980), 727-737.
- [5] Gu, C. H., On the harmonic maps from $R^{1,1}$ to $S^{1,1}$, *J. Reine Angw. Math.*, **346** (1984), 101-107.
- [6] Hörmander, L., Linear partial differential operators, Springer-Verlag, Berlin, 1963.
- [7] John, F., Solutions of quasi-linear wave equation with small initial data, The third phase, Colloque international Problems hyperboliques non lineaires, Bordeaux, 1988, 155-184.
- [8] Klainerman, S. & Ponce, G., Global, small amplitude solutions to nonlinear evolution equations, *Comm. Pure Appl. Math.*, **36**(1983), 133-141.
- [9] Li, T. T., Gobar regularity and formation of singularities of solutions to first order quasi-linear hyperbolic system, *Proc. Royal Soc. Edinbaegh*, **87A**(1981), 255-261.

- [10] Li, T. T., Global classical solutions for quasilinear hyperbolic systems, *Research in Applied Mathematics* 32, John Wiley and Sons/Masson, 1994.
- [11] Li, T. T. & Chen, Y. M., Global classical solutions for nonlinear evolution equations, *Pitman Monographs and Surveys in Pure and Applied Mathematics* 45, Longman Scientific and Technical, 1992.
- [12] Lions, J. L., *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod/Gauthier-Villars, 1969.
- [13] Li, T. T. & Qin, T. H., Global smooth solutions for a class of quasilinear hyperbolic systems with dissipative term, *Chin. Ann. of Math.*, **6B**:2(1985), 199–210.
- [14] Zhou, Y., Blow up of solutions to the Cauchy problem for nonlinear wave equations, *Chin. Ann. of Math.*, **22B**:3(2001), 275–280.

REGULARITY RESULTS FOR LINEAR ELLIPTIC PROBLEMS RELATED TO THE PRIMITIVE EQUATIONS

HU CHANGBING

*The Institute for Scientific Computing & Applied Mathematics,
Indiana University, Rawles Hall, Bloomington, IN 47405, USA*

R. TEMAM

*The Institute for Scientific Computing & Applied Mathematics,
Indiana University, Rawles Hall, Bloomington, IN 47405, USA
Laboratoire d'Analyse Numérique, Université Paris-Sud, Bâtiment
425, 91405 Orsay, France. E-mail: temam@indiana.edu*

M. ZIANE

*Department of Mathematics, University of
Southern California, Los Angeles, CA 90089, USA*

Abstract

The authors study the regularity of solutions of the GFD-Stokes problem and of some second order linear elliptic partial differential equations related to the Primitive Equations of the ocean. The present work generalizes the regularity results in [18] by taking into consideration the non-homogeneous boundary conditions and the dependence of solutions on the thickness ε of the domain occupied by the ocean and its varying bottom topography. These regularity results are important tools in the study of the PEs (see e.g. [6]), and they seem also to possess their own interest.

Keywords Primitive equations of the ocean, Oceanography,
GFD Stokes problem, Non-smooth domains
2000 MR Subject Classification 35J65, 35B65, 35Q30

§1. Introduction

We establish regularity results for the GFD-Stokes system and some second order elliptic partial differential equations related to the primitive equations of the ocean in thin domains with varying bottom topography. These equations constitute the core part of the Primitive Equations (PEs for brevity) for the atmosphere and the ocean. Motivated by the smallness of the aspect ratio of the domain occupied by the ocean, we studied the global existence of solutions for the PEs in thin domains, in [6]. As a preliminary step, we study, in this article, the H^2 -regularity of solutions of the Stokes-type problem and some second order elliptic partial differential equations related to the primitive equations. Compared to the work [18, 19], an important aspect of our results in this article is that we determine the dependence on the thickness parameter ε and the bottom topography for the constants of the H^2 regularity. We also derive as in [16] some Sobolev inequalities with the exact dependence on ε of the constants appearing in the aforementioned inequalities. These results play a crucial role in establishing the global existence of strong solutions of the PEs in thin domains (see [6]).

We present an outline of this article. In the rest of this section we describe the problem and present the main results. Then, in Section 2, we study the H^2 regularity of two nonhomogeneous elliptic boundary value problems. In Section 3 we use the results of Section 2 and give the proof of the main Theorem. Finally in Section 4, we derive various Sobolev type inequalities together with some applications of the results of Section 3.

Notations. The domain occupied by the ocean is of the form

$$M_\varepsilon = \{(x_1, x_2, x_3) \in \mathbb{R}^3, (x_1, x_2) \in \Gamma_i, -\varepsilon h(x_1, x_2) < x_3 < 0\},$$

and its boundary $\partial M_\varepsilon = \Gamma_i \cup \Gamma_b \cup \Gamma_l$, where Γ_i is the interface between the ocean and the atmosphere, assumed to be a bounded smooth open subset of \mathbb{R}^2 , Γ_b is the bottom boundary of the ocean, and Γ_l is its lateral boundary. Throughout this article, we will assume that h is independent of ε , where ε is the small parameter representing the thickness of the domain. In the final results we assume that h is a positive constant but in several parts of the article we will make the following assumptions concerning h : there exist positive constants $\bar{h}, \underline{h}, h_1$ such that

$$h \in C^2(\bar{\Gamma}_i), \quad 0 < \underline{h} \leq h \leq \bar{h}, \quad \text{and } \|h\|_{C^2(\bar{\Gamma}_i)} \leq h_1. \quad (1.1)$$

We are concerned with the regularity of solutions of the GFD-Stokes problem, namely,

$$\begin{cases} -\Delta v - \frac{\partial^2 v}{\partial x_3^2} + \text{grad } p = f_1 & \text{in } M_\varepsilon, \\ \int_{-\varepsilon h}^0 \text{div } v = 0 & \text{on } \Gamma_i, \\ \frac{\partial v}{\partial x_3} + \alpha_v v = g_v & \text{on } \Gamma_i, v = 0, \quad \text{on } \Gamma_b \cup \Gamma_l, \end{cases} \quad (1.2)$$

where $v = v(x_1, x_2, x_3) \in \mathbb{R}^2$, and $p = p(x_1, x_2) \in \mathbb{R}$ are the unknown functions, f_1 is the external volume force, $\alpha_v > 0$, g_v are given. We are also interested in the regularity of solutions of the following elliptic problem related to the equation of the temperature T or the equation of salinity:

$$\begin{cases} -\Delta T - \frac{\partial^2 T}{\partial x_3^2} = f_2 & \text{in } M_\varepsilon, \\ \frac{\partial T}{\partial x_3} + \alpha_T T = g_T & \text{on } \Gamma_i, \\ \frac{\partial T}{\partial \mathbf{n}} = 0 & \text{on } \Gamma_b \cup \Gamma_l, \end{cases} \quad (1.3)$$

where $T = T(x_1, x_2, x_3) \in \mathbb{R}$ is the unknown function, f_2 is the heating source inside the ocean, $\alpha_T > 0$, g_T is given, \mathbf{n} is the unit outward normal to the boundary. Throughout this article, we use $\nabla, \Delta, \text{div}$ to denote the two dimensional gradient, Laplacian and divergence operators on the horizontal plane, and use ∇_3, Δ_3 and div_3 for the corresponding 3D differential operators. The spaces $H^s(M_\varepsilon), H_0^s(M_\varepsilon), s \geq 0$, are the usual Sobolev spaces constructed on $L^2(M_\varepsilon)$, and

$$\mathbb{L}^2(M_\varepsilon) = (L^2(M_\varepsilon))^2, \quad \mathbb{H}^s(M_\varepsilon) = (H^s(M_\varepsilon))^2.$$

Furthermore, we define the space (see [8–10])

$$\mathcal{V}_1 = \left\{ v \in C^\infty(M_\varepsilon) : v \text{ is zero near } \Gamma_l \cup \Gamma_b, \text{div} \int_{-\varepsilon h}^0 v \, dz = 0 \right\}, \quad (1.4)$$

and H_1 is defined to be the closure of \mathcal{V}_1 in $\mathbb{L}^2(M_\varepsilon)$, $H_2 = L^2(M_\varepsilon)$, and V_1 is the closure of \mathcal{V}_1 in $\mathbb{H}^1(M_\varepsilon)$, $V_2 = H^1(M_\varepsilon)$. The norms and inner products for the spaces H and H_k ($k = 1, 2$) are the L^2 ones, all denoted

by $(\cdot, \cdot)_\varepsilon$ and $|\cdot|_\varepsilon$. Throughout the paper, c_0 will stand for a numerical constant that may vary from line to line.

The main result of this paper is the following

Theorem 1.1. *The hypotheses are those above, and we assume that M_ε is convex and that h is a positive constant. Let $(v, p) \in \mathbb{H}^1(M_\varepsilon) \times L^2(\Gamma_i)$ (resp. $T \in H^1(M_\varepsilon)$) be a weak solution of (1.2) (resp. (1.3)). Then*

$$(v, p) \in \mathbb{H}^2(M_\varepsilon) \times H^1(M_\varepsilon), \quad T \in H^2(M_\varepsilon). \tag{1.5}$$

Moreover the following inequalities hold:

$$|v|_{\mathbb{H}^2(M_\varepsilon)}^2 + \varepsilon |p|_{H^1(\Gamma_i)}^2 \leq c_0 [|f_1|_\varepsilon^2 + |g_v|_{L^2(\Gamma_i)}^2 + \varepsilon |\nabla g_v|_{L^2(\Gamma_i)}^2], \tag{1.6}$$

$$|T|_{H^2(M_\varepsilon)}^2 \leq c_0 (|f_2|_\varepsilon^2 + |\nabla g_T|_{L^2(\Gamma_i)}^2 + |g_T|_{L^2(\Gamma_i)}^2). \tag{1.7}$$

Remark 1.1. We will study in a separate article the case where M_ε is not convex and h satisfies (1.1).

§2. Preliminary Results

A preliminary step in the proof of Theorem 1.1 is the study of the H^2 -regularity of the solution of an auxiliary elliptic boundary value problem, which is obtained by setting $p = 0$ in (1.2), and deleting the second equation (compare also to (1.3)).

Lemma 2.1. *Assume that M_ε is convex, and $h \in C^2(\bar{\Gamma}_i)$. For $f \in L^2(M_\varepsilon)$ and $g \in H_0^1(\Gamma_i)$, there exists a unique $\Psi \in H^2(M_\varepsilon)$ solution of*

$$\begin{cases} -\Delta_3 \Psi = f & \text{in } M_\varepsilon, \\ \frac{\partial \Psi}{\partial x_3} + \alpha \Psi = g & \text{on } \Gamma_i, \\ \Psi = 0 & \text{on } \Gamma_b \cup \Gamma_l. \end{cases} \tag{2.1}$$

Furthermore, there exists a constant $c(h, \alpha)$ depending only on α and h (and Γ_i), such that

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \Psi}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha) [|f|_\varepsilon^2 + |g|_i^2 + |\nabla g|_i^2].$$

For a function in $L^2(\Gamma_i)$, its norm is denoted by $|\cdot|_i$.

The proof of Lemma 2.1 is given below. We first construct a function Ψ^* satisfying the boundary conditions in (2.1) and with the exact dependence on ε of the L^2 -norm of the second order derivatives (see Lemma 2.2 below).

Then $e^{\alpha x_3}(\Psi - \Psi^*) = \widehat{\Phi}$ satisfies the homogeneous Neumann condition on Γ_i and the homogeneous Dirichlet boundary condition on $\Gamma_l \cup \Gamma_b$. By a reflection argument, we extend the force f to $x_3 > 0$ to be an odd function. We then consider a homogeneous Dirichlet problem on the convex domain

$$\widehat{M} = \{(x_1, x_2, x_3) \in \mathbb{R}^3; (x_1, x_2) \in \Gamma_i, -\varepsilon h(x_1, x_2) < x_3 < \varepsilon h(x_1, x_2)\},$$

the solution of which \widehat{W} coincides with $\widehat{\Phi}$ on M_ε . Finally we use the classical H^2 regularity results in convex domains (see [4]) to obtain that $\widehat{W} \in H^2(\widehat{M}_\varepsilon)$ and thus $\widehat{\Phi}$ and Ψ are in $H^2(M_\varepsilon)$ together with the bounds on the L^2 -norm of their second derivatives. We start with the following lifting lemma.

Lemma 2.2. *Let $h \in C^2(\overline{\Gamma}_i)$ and $g \in H_0^1(\Gamma_i)$. There exists $\Psi^* \in H^2(M_\varepsilon)$, such that*

$$\frac{\partial \Psi^*}{\partial x_3} + \alpha \Psi^* = g \quad \text{on } \Gamma_i, \quad \Psi^* = 0 \quad \text{on } \Gamma_l \cup \Gamma_b.$$

Furthermore, there exists a constant $c(h)$ depending only on h , such that for $0 < \varepsilon \leq 1$,

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \Psi^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq (c(h)\varepsilon^2 + 1)(|g|_i^2 + |\nabla g|_i^2). \quad (2.2)$$

Proof. We first construct a function $\widetilde{\Psi}$ as a solution of the heat equation with $-x_3$ corresponding to time:

$$\begin{cases} \frac{\partial \widetilde{\Psi}}{\partial x_3} = -\Delta \widetilde{\Psi} & \text{in } \Gamma_i \times (-\infty, 0), \\ \widetilde{\Psi} = 0 & \text{on } \partial\Gamma_i \times (-\infty, 0), \\ \widetilde{\Psi}(x_1, x_2, 0) = g(x_1, x_2) & \text{on } \Gamma_i. \end{cases} \quad (2.3)$$

We set for $(x_1, x_2, x_3) \in \Gamma_i \times (-\infty, 0)$,

$$\Psi^*(x_1, x_2, x_3) = e^{-\alpha x_3} \int_{-\varepsilon h(x_1, x_2)}^{x_3} \widetilde{\Psi}(x_1, x_2, z) dz.$$

Note that $\Psi^*(x_1, x_2, x_3) = 0$ when $(x_1, x_2, x_3) \in \Gamma_l \cup \Gamma_b$. Furthermore

$$\frac{\partial \Psi^*}{\partial x_3} + \alpha \Psi^* = e^{-\alpha x_3} \widetilde{\Psi}(x_1, x_2, x_3), \quad (2.4)$$

and therefore, if $x_3 = 0$, we have $\frac{\partial \Psi^*}{\partial x_3} + \alpha \Psi^* = g$; that is, Ψ^* satisfies the boundary conditions in (2.1). Now we recall the classical energy estimates

for $\tilde{\Psi}$ solution of the heat equation with $-x_3$ corresponding to time.

We have

$$\frac{1}{2}|\tilde{\Psi}|_i^2(x_3) + \int_{x_3}^0 |\nabla \tilde{\Psi}|_i^2 = \frac{1}{2}|g|_i^2, \quad (2.5)$$

$$\frac{1}{2}|\nabla \tilde{\Psi}|_i^2(x_3) + \int_{x_3}^0 \left| \frac{\partial \tilde{\Psi}}{\partial x_3} \right|_i^2(z) dz = \frac{1}{2}|\nabla g|_i^2, \quad (2.6)$$

$$-\frac{x_3}{2} \left| \frac{\partial \tilde{\Psi}}{\partial x_3} \right|_i^2(x_3) - \int_{x_3}^0 z \left| \frac{\partial \nabla \tilde{\Psi}}{\partial x_3} \right|_i^2(z) dz = \frac{1}{2} \int_{x_3}^0 \left| \frac{\partial \tilde{\Psi}}{\partial x_3} \right|_i^2(z) dz, \quad (2.7)$$

$$\frac{x_3^2}{2} \left| \frac{\partial \nabla \tilde{\Psi}}{\partial x_3} \right|_i^2(x_3) + \int_{x_3}^0 z^2 \left| \frac{\partial^2 \tilde{\Psi}}{\partial x_3^2} \right|_i^2(z) dz = - \int_{x_3}^0 z \left| \frac{\partial \nabla \tilde{\Psi}}{\partial x_3} \right|_i^2(z) dz. \quad (2.8)$$

Hence (2.5), (2.6), and (2.3) yield

$$|\tilde{\Psi}|_i^2(-\varepsilon \bar{h}) \leq |g|_i^2, \quad \frac{1}{2}|\nabla \tilde{\Psi}|_i^2(-\varepsilon \bar{h}) + \frac{1}{2}|\Delta \tilde{\Psi}|_i^2 + \frac{1}{2} \left| \frac{\partial \tilde{\Psi}}{\partial x_3} \right|_i^2 \leq \frac{1}{2}|\nabla g|_i^2, \quad (2.9)$$

and by integration of (2.5) and (2.6) with respect to x_3 from $-\varepsilon \bar{h}$ to 0, we obtain

$$|\tilde{\Psi}|_i^2 + |\nabla \tilde{\Psi}|_i^2 \leq \varepsilon \bar{h} (|g|_i^2 + |\nabla g|_i^2). \quad (2.10)$$

Now, taking $x_3 = -\varepsilon \bar{h}$ in (2.7) and (2.8), we can write, using (2.6),

$$\frac{\varepsilon \bar{h}}{2} \left| \frac{\partial \tilde{\Psi}}{\partial x_3} \right|_i^2(-\varepsilon \bar{h}) - \int_{-\varepsilon \bar{h}}^0 z \left| \frac{\partial \nabla \tilde{\Psi}}{\partial x_3} \right|_i^2(z) dz \leq \frac{1}{4}|\nabla g|_i^2, \quad (2.11)$$

$$\frac{\varepsilon^2 \bar{h}^2}{2} \left| \frac{\partial \nabla \tilde{\Psi}}{\partial x_3} \right|_i^2(-\varepsilon \bar{h}) + \int_{-\varepsilon \bar{h}}^0 z^2 \left| \frac{\partial^2 \tilde{\Psi}}{\partial x_3^2} \right|_i^2(z) dz \leq \frac{1}{4}|\nabla g|_i^2, \quad (2.12)$$

and, in particular,

$$\varepsilon^2 \bar{h}^2 \int_{-\varepsilon \bar{h}}^{-\varepsilon \bar{h}} \left| \frac{\partial^2 \tilde{\Psi}}{\partial x_3^2} \right|_i^2(z) dz \leq \frac{1}{4}|\nabla g|_i^2. \quad (2.13)$$

Therefore, using

$$|\varphi(x_1, x_2, -\varepsilon h(x_1, x_2))| \leq |\varphi(x_1, x_2, -\varepsilon \bar{h})| + \left| \int_{-\varepsilon \bar{h}}^{-\varepsilon h} \frac{\partial \varphi}{\partial x_3}(x_1, x_2, z) dz \right|,$$

we obtain

$$\int_{\Gamma_i} |\tilde{\Psi}(x_1, x_2, -\varepsilon h(x_1, x_2))|^2 dx_1 dx_2 \leq c_0 [1 + \varepsilon \bar{h}] (|g|_i^2 + |\nabla g|_i^2),$$

$$\int_{\Gamma_i} |\nabla \tilde{\Psi}(x_1, x_2, -\varepsilon h(x_1, x_2))|^2 dx_1 dx_2 \leq c_0 \left(1 + \frac{\bar{h}}{h}\right) (|g|_i^2 + |\nabla g|_i^2),$$

$$\int_{\Gamma_1} \left| \frac{\partial \tilde{\Psi}}{\partial x_3}(x_1, x_2, -\varepsilon h(x_1, x_2)) \right|^2 dx_1 dx_2 \leq c_0 \left(\frac{\bar{h} + \underline{h}}{\varepsilon \bar{h} \underline{h}} \right) (|g|_i^2 + |\nabla g|_i^2). \quad (2.14)$$

We now derive estimates on Ψ^* . Since

$$\Psi^* = e^{-\alpha x_3} \int_{-\varepsilon h(x_1, x_2)}^{x_3} \tilde{\Psi}(x_1, x_2, z) dz,$$

we have, by the Cauchy-Schwarz inequality and (2.10),

$$|\Psi^*|_\varepsilon^2 \leq \varepsilon^2 e^{2\bar{h}\alpha} \bar{h}^2 |\tilde{\Psi}|_\varepsilon^2 \leq \varepsilon^3 e^{2\bar{h}\alpha} \bar{h}^3 (|g|_i^2 + |\nabla g|_i^2).$$

Furthermore,

$$\begin{aligned} e^{\alpha x_3} \frac{\partial \Psi^*}{\partial x_k} &= \int_{-\varepsilon h(x_1, x_2)}^{x_3} \frac{\partial \tilde{\Psi}}{\partial x_k}(x_1, x_2, z) dz \\ &\quad + \varepsilon \frac{\partial h}{\partial x_k} \tilde{\Psi}(x_1, x_2, -\varepsilon h(x_1, x_2)), \quad k = 1, 2, \end{aligned}$$

and for $k, j = 1, 2$,

$$\begin{aligned} e^{\alpha x_3} \frac{\partial^2 \Psi^*}{\partial x_k \partial x_j} &= \int_{-\varepsilon h(x_1, x_2)}^{x_3} \frac{\partial^2 \tilde{\Psi}}{\partial x_k \partial x_j}(x_1, x_2, z) dz \\ &\quad + \varepsilon \frac{\partial h}{\partial x_j} \frac{\partial \tilde{\Psi}}{\partial x_k}(x_1, x_2, -\varepsilon h(x_1, x_2)) \\ &\quad + \varepsilon \frac{\partial^2 h}{\partial x_k \partial x_j} \tilde{\Psi}(x_1, x_2, -\varepsilon h(x_1, x_2)) \\ &\quad - \varepsilon^2 \frac{\partial h}{\partial x_k} \frac{\partial h}{\partial x_j} \frac{\partial \tilde{\Psi}}{\partial x_3}(x_1, x_2, -\varepsilon h(x_1, x_2)). \end{aligned}$$

Therefore using (2.14), we have

$$\begin{aligned} \left| \frac{\partial \Psi^*}{\partial x_k} \right|_\varepsilon^2 &\leq \varepsilon^2 e^{2\bar{h}\alpha} \bar{h}^2 (\varepsilon \bar{h} + h_1(1/2 + \bar{h} + \sqrt{\varepsilon \bar{h}})) (|g|_i^2 + |\nabla g|_i^2) \\ &\leq c(h) \varepsilon^2 (|g|_i^2 + |\nabla g|_i^2), \end{aligned} \quad (2.15)$$

$$\begin{aligned} \left| \frac{\partial^2 \Psi^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 &\leq c_0 \varepsilon^2 \left(\bar{h}^2 + h_1 \bar{h} \left(1 + \frac{\bar{h}}{\underline{h}} \right) + \bar{h} h_2 (1 + \sqrt{\varepsilon \bar{h}}) + \bar{h} h_1^2 \frac{\bar{h} + \underline{h}}{\bar{h} \underline{h}} \right) \\ &\quad (|g|_i^2 + |\nabla g|_i^2) \\ &\leq c(h) \varepsilon^2 (|g|_i^2 + |\nabla g|_i^2). \end{aligned} \quad (2.16)$$

Furthermore, since $\frac{\partial \Psi^*}{\partial x_3} = -\alpha \Psi^* + e^{-\alpha x_3} \tilde{\Psi}$, we have

$$\begin{aligned} \left| \frac{\partial \Psi^*}{\partial x_3} \right|_\varepsilon^2 &\leq 2\alpha |\Psi^*|_\varepsilon^2 + 2e^{2\alpha h} |\tilde{\Psi}|_\varepsilon^2 \leq c(h)\varepsilon^2 (|g|_i^2 + |\nabla g|_i^2), \\ \left| \nabla \frac{\partial \Psi^*}{\partial x_3} \right|_\varepsilon^2 &\leq 2\alpha |\nabla \Psi^*|_\varepsilon^2 + 2|\nabla \tilde{\Psi}|_\varepsilon^2 \leq c(h)\varepsilon^2 (|g|_i^2 + |\nabla g|_i^2). \end{aligned} \quad (2.17)$$

Finally, since

$$\frac{\partial^2 \Psi^*}{\partial x_3^2} = -\alpha \frac{\partial \Psi^*}{\partial x_3} - \alpha e^{-\alpha x_3} \tilde{\Psi} + e^{-\alpha x_3} \frac{\partial \tilde{\Psi}}{\partial x_3}, \quad (2.18)$$

we have

$$\left| \frac{\partial^2 \Psi^*}{\partial x_3^2} \right|_\varepsilon^2 \leq (c(h)\varepsilon^2 + 1)(|g|_i^2 + |\nabla g|_i^2). \quad (2.19)$$

Proof of Lemma 2.1. Back to the proof of Lemma 2.1, we construct Ψ by setting $\Psi = \Phi + \Psi^*$, where Ψ^* is the H^2 -function constructed in Lemma 2.2 and Φ is the unique solution of

$$\begin{cases} -\Delta_3 \Phi = \bar{f} & \text{in } M_\varepsilon, \\ \frac{\partial \Phi}{\partial x_3} + \alpha \Phi = 0 & \text{on } \Gamma_i, \\ \Phi = 0 & \text{on } \Gamma_b \cup \Gamma_l, \end{cases} \quad (2.20)$$

where $\bar{f} = f - \Delta_3 \Psi^* \in L^2(M_\varepsilon)$, and $|\bar{f}|_\varepsilon \leq |f|_\varepsilon + c(h)(|g|_i + |\nabla g|_i)$. Note that

$$\frac{1}{2} |\nabla_3 \Phi|_\varepsilon^2 + \alpha^2 |\Phi|_{L^2(\Gamma_i \times \{0\})} \leq \frac{\varepsilon^2}{2} |\bar{f}|^2. \quad (2.21)$$

In order to prove that $\Phi \in H^2(M_\varepsilon)$, we define

$$\hat{\Phi}(x_1, x_2, x_3) = e^{\alpha x_3} \Phi(x_1, x_2, x_3). \quad (2.22)$$

Note that

$$\frac{\partial \hat{\Phi}}{\partial x_3} = e^{\alpha x_3} \left(\frac{\partial \Phi}{\partial x_3} + \alpha \Phi \right) \quad \text{and} \quad \frac{\partial^2 \hat{\Phi}}{\partial x_3^2} = e^{\alpha x_3} \left(\frac{\partial^2 \Phi}{\partial x_3^2} + 2\alpha \frac{\partial \Phi}{\partial x_3} + \alpha^2 \Phi \right). \quad (2.23)$$

The function $\hat{\Phi}$ satisfies the boundary conditions

$$\frac{\partial \hat{\Phi}}{\partial x_3} = 0 \quad \text{on } \Gamma_i, \quad \hat{\Phi} = 0 \quad \text{on } \Gamma_l \cup \Gamma_b. \quad (2.24)$$

Moreover

$$\Delta_3 \hat{\Phi} = e^{\alpha x_3} \Delta_3 \Phi + e^{\alpha x_3} \left(2\alpha \frac{\partial \Phi}{\partial x_3} + \alpha^2 \Phi \right). \quad (2.25)$$

Hence

$$-\Delta_3 \widehat{\Phi} = e^{\alpha x_3} \bar{f} - e^{\alpha x_3} \left(2\alpha \frac{\partial \Phi}{\partial x_3} + \alpha^2 \Phi \right) = \hat{f}. \quad (2.26)$$

It is easy to see that $\hat{f} \in L^2(M_\varepsilon)$, and

$$|\hat{f}|_\varepsilon \leq |f|_\varepsilon + c(h, \alpha)(|g|_i + |\nabla g|_i).$$

Therefore, we write the equation satisfied by $\widehat{\Phi}$ in the following form:

$$\begin{cases} -\Delta_3 \widehat{\Phi} = \hat{f} \in L^2(M_\varepsilon), \\ \frac{\partial \widehat{\Phi}}{\partial x_3} = 0 \quad \text{on } \Gamma_i, \\ \widehat{\Phi} = 0 \quad \text{on } \Gamma_b \cup \Gamma_l. \end{cases} \quad (2.27)$$

Let

$$\widehat{F}(x_1, x_2, x_3) = \begin{cases} \hat{f}(x_1, x_2, x_3), & -\varepsilon h(x_1, x_2) < x_3 < 0, \\ \hat{f}(x_1, x_2, -x_3), & 0 < x_3 < \varepsilon h(x_1, x_2). \end{cases}$$

Let

$$\widehat{M}_\varepsilon = \{(x_1, x_2, x_3) \in \mathbb{R}^3, (x_1, x_2) \in \Gamma_i, -\varepsilon h(x_1, x_2) < x_3 < \varepsilon h(x_1, x_2)\},$$

and consider the following Laplace problem:

$$\begin{cases} -\Delta_3 \widehat{W} = \widehat{F} \in L^2(\widehat{M}_\varepsilon), \\ \widehat{W} = 0 \quad \text{on } \partial \widehat{M}_\varepsilon. \end{cases}$$

The convexity of M_ε implies that \widehat{M}_ε is convex. Moreover since $\widehat{F} \in L^2(\widehat{M}_\varepsilon)$, we obtain, thanks to [4], $\widehat{W} \in H^2(\widehat{M}_\varepsilon)$, and

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \widehat{W}}{\partial x_k \partial x_j} \right|_{L^2(\widehat{M}_\varepsilon)}^2 \leq |\widehat{F}|_{L^2(\widehat{M}_\varepsilon)}^2 \leq 2|\hat{f}|_\varepsilon^2 \leq c(h, \alpha)[|f|_\varepsilon^2 + |g|_i^2 + |\nabla g|_i^2].$$

Since \widehat{F} is even in x_3 , the solution \widehat{W} is also even in x_3 . Therefore, $\frac{\partial \widehat{W}}{\partial x_3} = 0$ at $x_3 = 0$. By the uniqueness of solutions of (2.27), we obtain $\widehat{W}|_{M_\varepsilon} = \widehat{\Phi}$, and therefore

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \widehat{\Phi}}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha)[|f|_\varepsilon^2 + |g|_i^2 + |\nabla g|_i^2]. \quad (2.28)$$

Finally, since $\Psi = \Phi + \Psi^*$, we obtain, thanks to Lemma 2.2 and (2.28),

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \Psi}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha)[|f|_\varepsilon^2 + |g|_i^2 + |\nabla g|_i^2].$$

Remark 2.1. As a corollary of Lemma 2.1, by an interpolation argument (see [7]), we have for $g = 0$: if $f \in H^{-1/2-\delta}(M_\varepsilon)$ with $0 < \delta < \frac{1}{2}$, then $\widehat{\Psi} \in H^{3/2-\delta}(M_\varepsilon)$, and if $f \in H^{-1/2+\delta}(M_\varepsilon)$ with $0 < \delta < \frac{1}{2}$, then $\widehat{\Psi} \in H^{2-\delta}(M_\varepsilon)$.

We will also need, in the proof of the main result, the following regularity result:

The next lemma establishes the H^2 regularity of the temperature.

Lemma 2.3. *Assume that M_ε is convex, and that h is constant. For $f \in L^2(M_\varepsilon)$ and $g \in H^1_0(\Gamma_i)$, there exists a unique $T \in H^2(M_\varepsilon)$ solution of*

$$\begin{cases} -\Delta_3 T = f_2 & \text{in } M_\varepsilon, \\ \frac{\partial T}{\partial x_3} + \alpha_T T = g_T & \text{on } \Gamma_i, \\ \frac{\partial T}{\partial \mathbf{n}} = 0 & \text{on } \Gamma_b \cup \Gamma_l. \end{cases} \tag{2.29}$$

Furthermore, there exists a constant $c(h, \alpha_T)$ depending only on α_T and h , such that for all $\varepsilon > 0$,

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 T}{\partial x_j \partial x_k} \right|_\varepsilon^2 \leq c(h, \alpha_T) [|f_2|_\varepsilon^2 + |g_T|_i^2 + |\nabla g_T|_i^2].$$

Proof. We start with the case $g_T = 0$. Thus let \widetilde{T} be the unique solution of

$$\begin{cases} -\Delta_3 \widetilde{T} = f_2 & \text{in } M_\varepsilon, \\ \frac{\partial \widetilde{T}}{\partial x_3} + \alpha_T \widetilde{T} = 0 & \text{on } \Gamma_i, \\ \frac{\partial \widetilde{T}}{\partial \mathbf{n}} = 0 & \text{on } \Gamma_b \cup \Gamma_l. \end{cases} \tag{2.30}$$

We note that

$$|\nabla \widetilde{T}|_\varepsilon^2 + \left| \frac{\partial \widetilde{T}}{\partial x_3} \right|_\varepsilon^2 + \alpha_T |\widetilde{T}|_i^2 \leq |f_2|_\varepsilon |T|_\varepsilon, \tag{2.31}$$

but since (see Appendix)

$$|\widetilde{T}|_\varepsilon \leq \sqrt{2\varepsilon \bar{h}} |\widetilde{T}|_i + \varepsilon \bar{h} \left| \frac{\partial \widetilde{T}}{\partial x_3} \right|_\varepsilon, \tag{2.32}$$

we obtain

$$|\nabla \widetilde{T}|_\varepsilon^2 + \frac{1}{2} \left| \frac{\partial \widetilde{T}}{\partial x_3} \right|_\varepsilon^2 + \frac{\alpha_T}{2} |\widetilde{T}|_i^2 \leq \frac{\varepsilon \bar{h}}{\alpha_T} |f_2|_\varepsilon^2 + \varepsilon^2 \bar{h}^2 |f_2|_\varepsilon^2. \tag{2.33}$$

Therefore

$$|\tilde{T}|_\varepsilon^2 \leq 4\varepsilon^2 \bar{h}^2 \left[\frac{1}{\alpha_T} + \varepsilon \bar{h} \right]^2 |f_2|_\varepsilon^2. \quad (2.34)$$

For the sake of simplicity, we will assume now that the function h is constant and let $\eta(x_3)$ and T^* be defined by

$$\eta(x_3) = \exp \left[\frac{\alpha_T}{\varepsilon h} \left(\varepsilon h x_3 + \frac{1}{2} x_3^2 \right) \right] \quad \text{and} \quad T^* = \eta \tilde{T}. \quad (2.35)$$

Since h is constant, it is easy to check that

$$\begin{aligned} -\Delta_3 T^* &= 2 \frac{\eta'}{\eta} \frac{\partial \tilde{T}}{\partial x_3} + \frac{\eta''}{\eta} \tilde{T} + f_2, \\ \frac{\partial T^*}{\partial x_3} &= 0 \quad \text{on } \Gamma_b \cup \Gamma_i \quad \text{and} \quad \frac{\partial T^*}{\partial \mathbf{n}} = 0 \quad \text{on } \Gamma_l. \end{aligned} \quad (2.36)$$

Noting that $\eta'(x_3) = \frac{\alpha_T}{\varepsilon h} (\varepsilon h + x_3) \eta$ and $\eta''(x_3) = \frac{\alpha_T}{\varepsilon h} \eta + \frac{\alpha_T^2}{\varepsilon^2 h^2} (\varepsilon h + x_3)^2 \eta$, we can write

$$\left\{ \begin{array}{l} -\Delta_3 T^* = 2 \frac{\alpha_T}{\varepsilon h} (\varepsilon h + x_3) \frac{\partial \tilde{T}}{\partial x_3} + \frac{\alpha_T^2}{\varepsilon^2 h^2} (\varepsilon h + x_3)^2 \tilde{T} + \frac{\alpha_T}{\varepsilon h} \tilde{T} + f_2 = \tilde{f}_2, \\ \frac{\partial T^*}{\partial x_3} = 0 \quad \text{on } \Gamma_b \cup \Gamma_i, \\ \frac{\partial T^*}{\partial \mathbf{n}} = 0 \quad \text{on } \Gamma_l. \end{array} \right. \quad (2.37)$$

Since, by assumption, the domain M_ε is convex, we can apply the classical results of the H^2 -regularity in convex domains (see (3,1,2,2) in [4]) and obtain

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 T^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq |\Delta T^*|_\varepsilon^2 = |\tilde{f}_2|_\varepsilon^2.$$

Now, using (1.34), there exists a constant $c(h, \alpha_T)$ independent of ε , such that for $0 < \varepsilon \leq 1$, we have $|\tilde{f}_2|_\varepsilon^2 \leq c(h, \alpha_T) |f_2|_\varepsilon^2$ and

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 T^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha_T) |f_2|_\varepsilon^2.$$

Finally, noting that $\exp(-\alpha_T \varepsilon h) \leq \eta(x_3) \leq 1$, we have

$$|T^*|_\varepsilon^2 \leq \exp(2\alpha_T h) |\tilde{T}|_\varepsilon^2 \leq c(h, \alpha_T) \varepsilon^2 |f_2|_\varepsilon^2$$

and we obtain easily

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \tilde{T}}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha_T) |\tilde{f}_2|_\varepsilon^2.$$

Now we treat the general case with $g_T \neq 0$. We use a lemma similar to Lemma 2.2.

Lemma 2.4. *Assume that h is constant and $g_T \in H^1(\Gamma_i)$. There exists $\bar{T} \in H^2(M_\varepsilon)$, such that*

$$\frac{\partial \bar{T}}{\partial x_3} + \alpha_T \bar{T} = g_T \quad \text{on } \Gamma_i, \quad \frac{\partial \bar{T}}{\partial \mathbf{n}} = 0 \quad \text{on } \Gamma_l \cup \Gamma_b.$$

Furthermore, there exists a constant $c(h)$ depending only on h , such that for $0 < \varepsilon \leq 1$,

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 \bar{T}}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq \frac{(c(h)\varepsilon^2 + 1)}{\alpha^2} (|g|_i^2 + |\nabla g|_i^2). \tag{2.38}$$

Proof. We proceed as in the proof of Lemma 2.2, and construct a function $\bar{\Psi}$ as a solution of the heat equation with $-x_3$ corresponding to time:

$$\begin{cases} \frac{\partial \bar{\Psi}}{\partial x_3} = -\Delta \bar{\Psi} & \text{in } \Gamma_i \times (-\infty, 0), \\ \frac{\partial \bar{\Psi}}{\partial \mathbf{n}} = 0 & \text{on } \partial \Gamma_i \times (-\infty, 0), \\ \bar{\Psi}(x_1, x_2, 0) = g_T(x_1, x_2) & \text{on } \Gamma_i. \end{cases} \tag{2.39}$$

We then set for $(x_1, x_2, x_3) \in \Gamma_i \times (-\infty, 0)$,

$$\bar{T}(x_1, x_2, x_3) = e^{-\alpha x_3} \int_0^{x_3} \bar{\Psi}(x_1, x_2, z) dz - \left(x_3 - \frac{1}{\alpha}\right) e^{\alpha \varepsilon h} \bar{\Psi}(x_1, x_2, -\varepsilon h). \tag{2.40}$$

We can easily check that $\frac{\partial \bar{T}}{\partial x_3} + \alpha_T \bar{T} = g_T$ for $x_3 = 0$, $\frac{\partial \bar{T}}{\partial x_3} = 0$ for $x_3 = -\varepsilon h$.

Following the same lines of the proof of Lemma 2.2, we have

$$|\bar{\Psi}|_i^2(-\varepsilon h) \leq |g_T|_i^2, \quad \frac{1}{2} |\nabla \bar{\Psi}|_i^2(-\varepsilon h) + \frac{1}{2} |\Delta \bar{\Psi}|_\varepsilon^2 + \frac{1}{2} \left| \frac{\partial \bar{\Psi}}{\partial x_3} \right|_\varepsilon^2 \leq \frac{1}{2} |\nabla g_T|_i^2, \tag{2.41}$$

$$|\bar{\Psi}|_\varepsilon^2 + |\nabla \bar{\Psi}|_\varepsilon^2 \leq \varepsilon h (|g_T|_i^2 + |\nabla g_T|_i^2). \tag{2.42}$$

From this we can easily check that $\bar{T} \in H^2(M_\varepsilon)$ and that (2.38) holds.

§3. Proof of the Main Result

In this section, we prove the main result of this paper as stated in Theorem 1.1. We have already established in Lemma 2.4 the H^2 regularity of the temperature T and the estimates (the dependence on ε) of the L^2 -norms of its second derivatives. Thus we need only to study the regularity of the velocity, governed by the problem (1.2), i.e.,

$$\begin{cases} -(\Delta v + \frac{\partial^2 v}{\partial x_3^2}) + \nabla p = f_1 & \text{in } M_\varepsilon, \\ \operatorname{div} \int_{-\varepsilon h}^0 v \, dz = 0 & \text{in } \Gamma_i, \\ v = 0 & \text{on } \Gamma_l \cup \Gamma_b, \quad \frac{\partial v}{\partial x_3} + \alpha_v v = g_v & \text{on } \Gamma_i. \end{cases} \quad (3.1)$$

This result has been proven in [18] where $\varepsilon = 1$ and $g_v = 0$. As indicated in the Introduction we study here the general case where $g_v \neq 0$ and we carefully investigate the dependence on ε of the constants.

Our approach to obtain the H^2 regularity is the same as in [18] and is based on the following observation: assume that the solution v of (3.1) satisfies $\frac{\partial v}{\partial x_3} \Big|_{\Gamma_i} \in \mathbb{L}^2(\Gamma_i)$ and $\frac{\partial v}{\partial x_3} \Big|_{\Gamma_b} \in \mathbb{L}^2(\Gamma_b)$, then integrating (3.1) in x_3 over $(-\varepsilon h, 0)$ yields a standard Stokes in 2-D with homogeneous boundary condition on Γ_i . By the classical regularity theory of the 2D-Stokes problem in smooth domains (see for instance [15] and [3]), p belongs to $H^1(\Gamma_i)$. Then, by moving the pressure term to the right hand side, the problem (3.1) reduces to an elliptic problem of the type studied in Lemma 2.1, and the H^2 regularity of v follows. The estimates on the L^2 norms of the second derivatives are then obtained using the trace theorem and the estimates in Lemma 2.1. Therefore, we start with proving that

$$\frac{\partial v}{\partial x_3} \Big|_{\Gamma_i} \in \mathbb{L}^2(\Gamma_i), \quad \frac{\partial v}{\partial x_3} \Big|_{\Gamma_b} \in \mathbb{L}^2(\Gamma_b).$$

Lemma 3.1. *Assume that $h \in C^2(\bar{\Gamma}_i)$ and M_ε is convex. Let (v, p) be the weak solution of (3.1), then*

$$\frac{\partial v}{\partial x_3} \Big|_{\Gamma_i} \in \mathbb{L}^2(\Gamma_i), \quad \frac{\partial v}{\partial x_3} \Big|_{\Gamma_b} \in \mathbb{L}^2(\Gamma_b). \quad (3.2)$$

Proof. By integration by parts, we have

$$|\nabla_3 v|_\varepsilon^2 + \alpha_v |v|_i^2 = (f, v)_\varepsilon + (g_v, v)_i, \quad (3.3)$$

and therefore the existence and uniqueness of the weak solution (v, p) follows from the Lax-Milgram theorem and De Rham's theorem. Hence $\nabla p \in H^{-1}(M_\varepsilon)$ and thus $\nabla p \in H^{-1}(\Gamma_i)$ since p is independent of x_3 .

Let v_i be the unique solution of

$$\begin{cases} \Delta v_i = \nabla p & \text{in } \Gamma_i, \\ v_i = 0 & \text{on } \partial\Gamma_i. \end{cases} \tag{3.4}$$

Here v_i satisfies a 2D Laplace equation on Γ_i . Hence $v_i \in H_0^1(\Gamma_i)$. Let $\tilde{v} = v - v_i$, then \tilde{v} satisfies

$$\begin{cases} \Delta_3 \tilde{v} = f_1, \\ \tilde{v} = 0 & \text{on } \Gamma_l, \\ \tilde{v} = -v_i & \text{on } \Gamma_b, \\ \frac{\partial \tilde{v}}{\partial x_3} + \alpha_v \tilde{v} = g_v - \alpha v_i & \text{on } \Gamma_i. \end{cases} \tag{3.5}$$

Thanks to Lemma 3.2, with $\psi_i = v_i$ and $\gamma = -\delta$ for some $0 < \delta < \frac{1}{2}$, we have $\tilde{v} \in H^{3/2-\delta}(M_\varepsilon)$.

$$g_i = -\frac{1}{\varepsilon h} \int_{-\varepsilon h}^0 \operatorname{div} \tilde{v} dx_3 \in H^{1/2-\delta}(\Gamma_i). \tag{3.6}$$

Therefore, since $\operatorname{div} v_i = g_i$, with this new information on v_i , we rewrite the equation for v_i in the form of a 2D Stokes problem:

$$\begin{cases} -\Delta v_i + \nabla p = 0 & \text{in } \Gamma_i, \\ \operatorname{div} v_i = g_i \in H^{\frac{1}{2}-\delta}(\Gamma_i) \end{cases} \tag{3.7}$$

and thanks the classical regularity on the non-homogeneous Stokes problem on the Γ_i (see [15]), we have $v_i \in H^{\frac{3}{2}-\delta}(\Gamma_i) \cap H_0^1(\Gamma_i) = H_0^{\frac{3}{2}-\delta}(\Gamma_i)$. With this new information on the regularity of v_i , we go back to the problem (3.5) and using Lemma 3.2 with $\psi_i = v_i$ and $\gamma = \frac{1}{2} - \delta$, we conclude that $\tilde{v} \in H^{2-\delta}(M_\varepsilon)$ and thus

$$\frac{\partial \tilde{v}}{\partial x_3} \Big|_{\Gamma_i} \in H^{1/2-\delta}(\Gamma_i), \quad \frac{\partial \tilde{v}}{\partial x_3} \Big|_{\Gamma_b} \in H^{1/2-\delta}(\Gamma_b). \tag{3.8}$$

Lemma 3.2. *Assume that M_ε is convex, $h \in C^2(\bar{\Gamma}_i)$. For $f \in L^2(M_\varepsilon)$ and $g \in H_0^1(\Gamma_i)$, and $\psi_i \in H_0^{1+\gamma}(\Gamma_b)$, $-\frac{1}{2} < \gamma < \frac{1}{2}$, $\gamma \neq 0$, there exists a unique $\Psi_i \in H^{3/2+\gamma}(M_\varepsilon)$ solution of*

$$\begin{cases} -\Delta_3 \Psi_i = f & \text{in } M_\varepsilon, \\ \frac{\partial \Psi_i}{\partial x_3} + \alpha_v \Psi_i = g - \alpha_v \psi_i & \text{on } \Gamma_i, \\ \Psi_i = -\psi_i & \text{on } \Gamma_b, \\ \Psi_i = 0 & \text{on } \Gamma_l. \end{cases} \quad (3.9)$$

Proof. Thanks to Lemma 2.1, the problem is reduced to the case $f = 0$ and $g = 0$, by replacing Ψ_i with $\Psi_i - \Psi$, where Ψ is the function constructed in Lemma 2.1. Thus, without loss of generality, we may assume that $f = 0$ and $g = 0$.

Let Q_ε be the cylinder $Q_\varepsilon = \Gamma_i \times (-\varepsilon, 0)$, and let v_p be the unique solution of

$$\begin{cases} \Delta_3 v_p = 0, & \text{in } Q_\varepsilon, \\ v_p = 0 & \text{on } \partial\Gamma_i \times (-\varepsilon, 0), \\ v_p = -\psi_i & \text{on } \Gamma_i \times \{-\varepsilon\}, \\ v_p = \varepsilon h \alpha_v \psi_i & \text{on } \Gamma_i \times \{0\}. \end{cases} \quad (3.10)$$

We will show that $v_p \in H^{3/2+\gamma}(Q_\varepsilon)$ for all $-\frac{1}{2} < \gamma < \frac{1}{2}$, $\gamma \neq 0$. To this end, let \widehat{Q}_ε be any C^2 -domain containing Q_ε such that $\Gamma_i \times \{-\varepsilon, 0\} \subset \partial\widehat{Q}_\varepsilon$. Since ψ_i (resp. $h\alpha_v\psi_i$) is in $H_0^{1+\gamma}(\Gamma_i \times \{-\varepsilon\})$ (resp. $H_0^{1+\gamma}(\Gamma_i \times \{0\})$), we can define a function $V_i \in H^1(\partial\widehat{Q}_\varepsilon)$ by setting $V_i = -\psi_i$ on $\Gamma_i \times \{-\varepsilon\}$, $V_i = \varepsilon h \alpha_v \psi_i$ on $\Gamma_i \times \{0\}$, and $V_i = 0$ on $\partial\widehat{Q}_\varepsilon - \Gamma_i \times \{-\varepsilon, 0\}$. Now let V_p be the unique solution of $\Delta_3 V_p = 0$ in \widehat{Q}_ε and $V_p = V_i$ on $\partial\widehat{Q}_\varepsilon$. Since $\partial\widehat{Q}_\varepsilon$ is of class C^2 , the classical regularity results (see [7]) yield $V_p \in H^{3/2+\gamma}(\widehat{Q}_\varepsilon)$ for $-\frac{1}{2} < \gamma < \frac{1}{2}$, $\gamma \neq 0$. Now let \widetilde{V}_i be the trace of V_p on $\partial\Gamma_i \times (-\varepsilon, 0)$. It is easy to see that $\widetilde{V}_i \in H_0^{1+\gamma}(\partial\Gamma_i \times (-\varepsilon, 0))$. Let $\widetilde{V}_p = V_p - v_p$, we have

$$\begin{cases} \Delta_3 \widetilde{V}_p = 0 & \text{in } Q_\varepsilon, \\ \widetilde{V}_p = 0 & \text{on } \Gamma_i \times \{-\varepsilon, 0\}, \\ \widetilde{V}_p = \widetilde{V}_i & \text{on } \partial\Gamma_i \times (-\varepsilon, 0). \end{cases} \quad (3.11)$$

Using a reflection argument around $x_3 = 0$ (resp. $x_3 = -\varepsilon$) by extending \widetilde{V}_i in a ‘‘symmetrically’’ odd function defined on $\partial\Gamma_i \times (-\varepsilon, \varepsilon)$ (resp. $\partial\Gamma_i \times (-2\varepsilon, 0)$), and using the classical local regularity theory (see [7]), we conclude that $\widetilde{V}_p \in H^{3/2+\gamma}(Q_\varepsilon)$ for $-\frac{1}{2} < \gamma < \frac{1}{2}$, $\gamma \neq 0$. Hence since $V_p \in H^{3/2+\gamma}(Q_\varepsilon)$, we have $v_p = V_p - \widetilde{V}_p \in H^{3/2+\gamma}(Q_\varepsilon)$.

Now let

$$\tilde{v}_p(x_1, x_2, x_3) = -\frac{x_3}{\varepsilon h(x_1, x_3)} v_p\left(x_1, x_2, \frac{x_3}{h(x_1, x_2)}\right) \quad \text{for } (x_1, x_2, x_3) \in M_\varepsilon. \tag{3.12}$$

It is obvious that $\tilde{v}_p \in H^{3/2+\gamma}(M_\varepsilon)$,

$$\tilde{v}_p(x_1, x_2, -\varepsilon h(x_1, x_2)) = -\psi_i(x_1, x_2) \quad \text{and} \quad \frac{\partial \tilde{v}_p}{\partial x_3} + \alpha_v \tilde{v}_p = -\alpha_v v_i \quad \text{on } \Gamma_i.$$

Therefore setting $\tilde{V} = \tilde{v} - \tilde{v}_p$, we have

$$\begin{cases} \Delta_3 \tilde{V} = f_1 - \Delta_3 \tilde{v}_p \in H^{-1/2+\gamma}(M_\varepsilon), \\ \tilde{V} = 0 \quad \text{on } \Gamma_l \cup \Gamma_b, \\ \frac{\partial \tilde{V}}{\partial x_3} + \alpha_v \tilde{V} = 0 \quad \text{on } \Gamma_i. \end{cases} \tag{3.13}$$

Hence, thanks to Lemma 2.1 and Remark 2.1, we see that \tilde{V} and thus \tilde{v} are in $H^{3/2+\gamma}(M_\varepsilon)$ for $-\frac{1}{2} < \gamma < \frac{1}{2}$, $\gamma \neq 0$.

Proof of Theorem 1.1. The proof is divided into two steps. In Step 1, we prove the H^2 regularity of solutions, i.e., $v \in H^2(M_\varepsilon)$ and $p \in H^1(\Gamma_i)$. Then, in Step 2, we establish the Cattabriga-Solonnikov type inequality on the solutions, i.e., prove the bounds on the L^2 -norms of the second derivative of v and the H^1 -norm on the pressure, in particular we establish their (non) dependence on ε .

Step 1. The H^2 -Regularity of Solutions

Let

$$\bar{v} = \int_{-\varepsilon h}^0 v \, dz,$$

we have

$$\frac{\partial^2 \bar{v}(x_1, x_2, x_3)}{\partial x_i^2} = \int_{-\varepsilon h}^0 \frac{\partial^2 v(x_1, x_2, z)}{\partial x_i^2} \, dz + \varepsilon \frac{\partial h}{\partial x_i} \frac{\partial v(x_1, x_2, -\varepsilon h(x_1, x_2))}{\partial x_i}, \tag{3.14}$$

$i = 1, 2.$

Integrating the first equation in (3.1) with respect to x_3 we obtain the 2D Stokes problem:

$$\begin{cases} -\Delta \bar{v} + \varepsilon h \nabla p = \bar{f} \quad \text{in } \Gamma_i, \\ \operatorname{div} \bar{v} = 0 \quad \text{in } \Gamma_i, \quad v = 0 \quad \text{on } \partial \Gamma_i, \end{cases} \tag{3.15}$$

where

$$\tilde{f} = \int_{-\varepsilon h}^0 f dz + \frac{\partial v}{\partial x_3} \Big|_{x_3=0} - \frac{\partial v}{\partial x_3} \Big|_{x_3=-\varepsilon h} + \varepsilon \sum_{i=1}^2 \frac{\partial h}{\partial x_i} \frac{\partial v(x_1, x_2, -\varepsilon h(x_1, x_2))}{\partial x_i}. \tag{3.16}$$

We have, thanks to Lemma 3.1, $\tilde{f} \in L^2(\Gamma_i)$. Therefore from the classical regularity theory of the 2D Stokes problem, we conclude that $\nabla p \in L^2(\Gamma_i)$. We return to the problem (3.1), and move the gradient of the pressure to the right hand side and obtain, thanks to Lemma 2.1, $v \in H^2(M_\varepsilon)$ and

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 v}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha) [|f|_\varepsilon^2 + |g|_i^2 + |\nabla g|_i^2] + c(h, \alpha) \varepsilon |\nabla p|_\varepsilon^2.$$

Furthermore,

$$|\tilde{f}|_i^2 \leq \varepsilon \bar{h} |f_1|_\varepsilon^2. \tag{3.17}$$

Step 2. The Cattabriga-Solomnikov Type Inequality

First we homogenize the boundary condition. Let $v_l = (\Psi_1, \Psi_2)$ where Ψ_1 and Ψ_2 are constructed by using Lemma 2.1, i.e.,

$$\begin{cases} -\Delta_3 \Psi_k = f_{1,k} & \text{in } M_\varepsilon, \quad k = 1, 2, \\ \frac{\partial \Psi_k}{\partial x_3} + \alpha \Psi_k = g_{v,k} & \text{on } \Gamma_i, \quad k = 1, 2, \\ \Psi_k = 0 & \text{on } \Gamma_b \cup \Gamma_l, \quad k = 1, 2, \end{cases}$$

where $f_1 = (f_{1,1}, f_{1,2})$ $g_v = (g_{v,1}, g_{v,2})$. Thanks to Lemma 2.1, we have

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 v_l}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha_v) (|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2). \tag{3.18}$$

Let $v^* = v - v_l$, we have

$$\begin{cases} -\left(\Delta v^* + \frac{\partial^2 v^*}{\partial x_3^2}\right) + \nabla p = 0 & \text{in } M_\varepsilon, \\ \operatorname{div} \int_{-\varepsilon h}^0 v^* dz = g^* & \text{in } \Gamma_i, \\ v^* = 0 & \text{on } \Gamma_l \cup \Gamma_b, \\ \frac{\partial v^*}{\partial x_3} + \alpha_v v^* = 0 & \text{on } \Gamma_i, \end{cases} \tag{3.19}$$

where $g^* = -\operatorname{div} \int_{-\varepsilon h}^0 v_l dx_3$. Note that inequality (3.18) implies

$$\|g^*\|_{H^1(\Gamma_i)}^2 \leq c(h, \alpha_v) \varepsilon [|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2]. \tag{3.20}$$

Define $V^* = \frac{1}{h} \int_{-\varepsilon h}^0 v^* dx_3$, which satisfies the 2D-Stokes problem

$$\begin{cases} -\Delta V^* + \nabla(\varepsilon p) = F^* & \text{in } \Gamma_i, \\ \operatorname{div} V^* = G^* \\ V^* = 0 & \text{on } \partial\Gamma_i, \end{cases} \quad (3.21)$$

where

$$\begin{aligned} F^* &= \sum_{i=1}^2 \left[\frac{\partial}{\partial x_i} \left(\frac{1}{h^2} \frac{\partial h}{\partial x_i} \right) \int_{-\varepsilon h}^0 v^* dx_3 + \frac{2}{h^2} \frac{\partial h}{\partial x_i} \int_{-\varepsilon h}^0 \frac{\partial v^*}{\partial x_i} dx_3 \right. \\ &\quad \left. - \frac{\varepsilon}{h} \frac{\partial h}{\partial x_i} \frac{\partial v^*}{\partial x_i}(x_1, x_2, -\varepsilon h(x_1, x_2)) \right] + \frac{1}{h} \frac{\partial v^*}{\partial x_3} \Big|_{x_3=0} - \frac{1}{h} \frac{\partial v^*}{\partial x_3} \Big|_{x_3=-\varepsilon h}, \\ G^* &= \frac{1}{h} \operatorname{div} \int_{-\varepsilon h}^0 v_l dx_3 + \nabla \left(\frac{1}{h} \right) \cdot \int_{-\varepsilon h}^0 v^* dx_3. \end{aligned} \quad (3.22)$$

Clearly

$$\begin{aligned} |F^*|_{L^2(\Gamma_i)}^2 &\leq c(h)\varepsilon \left[|v^*|_\varepsilon^2 + \left| \frac{\partial v^*}{\partial x_i} \right|_\varepsilon^2 + \left| \frac{\partial v^*}{\partial x_i} \right|_{L^2(\Gamma_b)}^2 \right] \\ &\quad + c(h) \left[\left| \frac{\partial v^*}{\partial x_3} \right|_{L^2(\Gamma_i)}^2 + \left| \frac{\partial v^*}{\partial x_3} \right|_{L^2(\Gamma_b)}^2 \right]. \end{aligned} \quad (3.23)$$

Now, since $v^* = 0$ on Γ_b , we have $\frac{\partial v^*}{\partial x_i} = \varepsilon \frac{\partial h}{\partial x_i} \frac{\partial v^*}{\partial x_3}$ on γ_b and by the Poincaré inequality, we also have

$$\left| \frac{\partial v^*}{\partial x_3} \right|_{L^2(\Gamma_i)}^2 \leq 2\alpha_v^2 \varepsilon \left| \frac{\partial v^*}{\partial x_3} \right|_\varepsilon^2.$$

Now since

$$\begin{aligned} \left| \frac{\partial v^*}{\partial x_3} \right|_{L^2(\Gamma_b)}^2 &\leq \left| \frac{\partial v^*}{\partial x_3} \right|_{L^2(\Gamma_i)}^2 + 2 \left| \frac{\partial v^*}{\partial x_3} \right|_\varepsilon \left| \frac{\partial^2 v^*}{\partial x_3^2} \right|_\varepsilon \\ &\leq 2\alpha_v^2 \varepsilon \left| \frac{\partial v^*}{\partial x_3} \right|_\varepsilon^2 + \theta \varepsilon \left| \frac{\partial^2 v^*}{\partial x_3^2} \right|_\varepsilon^2 + \frac{c_0}{\theta} \varepsilon \left| \frac{\partial v^*}{\partial x_3} \right|_\varepsilon^2, \end{aligned} \quad (3.24)$$

where θ is a small positive constant independent of ε , we have

$$\left| \frac{\partial v^*}{\partial x_i} \right|_{L^2(\Gamma_b)}^2 \leq c(h, \theta) \left| \frac{\partial v^*}{\partial x_3} \right|_\varepsilon^2 + c(h)\varepsilon^2 \left| \frac{\partial^2 v^*}{\partial x_3^2} \right|_\varepsilon^2. \quad (3.25)$$

Thus

$$|F^*|_{L^2(\Gamma_i)}^2 \leq c(h)\varepsilon \left[|v^*|_\varepsilon^2 + \left| \frac{\partial v^*}{\partial x_i} \right|_\varepsilon^2 \right] + c(h)\theta\varepsilon \left| \frac{\partial^2 v^*}{\partial x_3^2} \right|_\varepsilon^2. \quad (3.26)$$

We estimate the H^1 -norm of v^* , using $v^* = v - v_l$ and the H^1 -estimates of

v and v_l . We obtain easily

$$|F^*|_{L^2(\Gamma_i)}^2 \leq c(h)\varepsilon[|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2] + c(h)\theta\varepsilon\left|\frac{\partial^2 v^*}{\partial x_3^2}\right|_\varepsilon^2. \quad (3.27)$$

Similarly, we have

$$|G^*|_{H^1(\Gamma_i)}^2 \leq c(h)\varepsilon[|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2]. \quad (3.28)$$

Now using the Cattabriga-Solonnikov inequality for the 2D Stokes problem (3.21), we see that there exists a constant c_0 independent of ε such that

$$|V^*|_{H^2(\Gamma_i)}^2 + \varepsilon^2|\nabla p|_{L^2(\Gamma_i)}^2 \leq c_0[|F^*|_{L^2(\Gamma_i)}^2 + |G^*|_{H^1(\Gamma_i)}^2]. \quad (3.29)$$

Therefore

$$\varepsilon^2|\nabla p|_{L^2(\Gamma_i)}^2 \leq c(h, \theta)\varepsilon[|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2] + c(h)\theta\varepsilon\left|\frac{\partial^2 v^*}{\partial x_3^2}\right|_\varepsilon^2. \quad (3.30)$$

Since $\Delta_3 v^* = \nabla p$, in M_ε , $v^* = 0$ on $\Gamma_b \cup \Gamma_l$ and $\frac{\partial v^*}{\partial x_3} + \alpha_v v^* = 0$ on Γ_i , we have thanks to Lemma 2.1

$$\begin{aligned} \sum_{k,j=1}^3 \left| \frac{\partial^2 v^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 &\leq c(h, \alpha)\varepsilon|\nabla p|_{L^2(\Gamma_i)}^2 \\ &\leq c(h, \alpha)[|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2] + c(h, \alpha_v)\theta\left|\frac{\partial^2 v^*}{\partial x_3^2}\right|_\varepsilon^2, \end{aligned} \quad (3.31)$$

and therefore for θ small enough, so that $c(h, \alpha_v)\theta \leq \frac{1}{2}$, we conclude that

$$\sum_{k,j=1}^3 \left| \frac{\partial^2 v^*}{\partial x_k \partial x_j} \right|_\varepsilon^2 \leq c(h, \alpha)\varepsilon|\nabla p|_{L^2(\Gamma_i)}^2 \leq c(h, \alpha)[|f_1|_\varepsilon^2 + |g_v|_i^2 + |\nabla g_v|_i^2]. \quad (3.32)$$

The proof of the main result is now complete.

§4. Appendix

We present in this appendix some Sobolev type inequalities satisfied by solutions of (1.2) and (1.3).

Lemma 4.1. *For v satisfying the boundary condition in (1.2), we have*

$$|v|_\varepsilon \leq 2\varepsilon \left| \frac{\partial v}{\partial x_3} \right|_\varepsilon, \quad v \in H^1(M_\varepsilon), \quad (4.1)$$

$$\left| \frac{\partial v}{\partial x_3} \right|_\varepsilon^2 + 2\alpha_v \int_{\Gamma_i} v^2 dx_1 dx_2 \leq 2\varepsilon^2 \left| \frac{\partial^2 v}{\partial x_3^2} \right|_\varepsilon^2 + 4\varepsilon|g_v|^2, \quad v \in H^2(M_\varepsilon). \quad (4.2)$$

Proof. Inequality (4.1) is the classical Poincaré inequality for v , we omit its proof (note that $v = 0$ on Γ_b but not necessarily on Γ_i). To prove (4.2) we establish this inequality for v smooth, and the result follows by density in the general case; when v is smooth the proof consists in integrating by parts $\int_{M_\varepsilon} v \frac{\partial^2 v}{\partial x_3^2} dx$, and using the Cauchy-Schwarz inequality and the boundary condition in (1.2).

Lemma 4.2. *For any $T \in H^1(M_\varepsilon)$, we have*

$$|T|_\varepsilon^2 \leq 2\bar{h}\varepsilon |T|_{L^2(\Gamma_i \times \{0\})}^2 + 4\bar{h}^2\varepsilon^2 \left| \frac{\partial T}{\partial x_3} \right|_\varepsilon^2, \tag{4.3}$$

$$|T|_{L^2(\Gamma_i \times \{0\})}^2 \leq \frac{2}{\varepsilon \underline{h}} |T|_\varepsilon^2 + \frac{2\bar{h}^2\varepsilon}{\underline{h}} \left| \frac{\partial T}{\partial x_3} \right|_\varepsilon^2. \tag{4.4}$$

The proof of (4.3) and (4.4) is similar to the proof of the Poincaré inequality.

Lemma 4.3 (Agmon’s inequality). *For $v \in \mathbb{H}^2(M_\varepsilon)$ satisfying the boundary condition in (1.2),*

$$|v|_{L^\infty} \leq c_0 [\varepsilon^{\frac{1}{2}} |\Delta_3 v|_\varepsilon + |g_v|_i + |\nabla g_v|_i]. \tag{4.5}$$

Proof. The proof is an easy extension of the following anisotropic inequality established in [16]:

$$|\bar{v}|_{L^\infty(\Gamma_i \times (-\varepsilon, 0))} \leq |c_0 \bar{v}|_\varepsilon^{\frac{1}{4}} \left(\left| \frac{\partial^2 \bar{v}}{\partial x_3'^2} \right|_\varepsilon + \frac{1}{\varepsilon} \left| \frac{\partial \bar{v}}{\partial x_3'} \right|_\varepsilon + \frac{1}{\varepsilon^2} |\bar{v}| \right)^{\frac{1}{4}} \cdot \prod_{i=1}^2 \left(\left| \frac{\partial^2 \bar{v}}{\partial x_i'^2} \right|_\varepsilon + \left| \frac{\partial \bar{v}}{\partial x_i'} \right|_\varepsilon + |\bar{v}|_\varepsilon \right)^{\frac{1}{4}},$$

where \bar{v} is the function corresponding to v via the change of variables to flatten the boundary, and the inequality

$$|\bar{v}|_{L^2(\Gamma_i \times (-\varepsilon, 0))} \leq \varepsilon \left| \frac{\partial \bar{v}}{\partial x_3'} \right|_\varepsilon \leq c_0 \varepsilon^2 \left[\left| \frac{\partial^2 \bar{v}}{\partial x_3'^2} \right|_\varepsilon + \varepsilon^{-\frac{1}{2}} |g_v|_{L^2(\Gamma_i)} \right].$$

We skip the details.

We recall the following version of Ladyzhenskaya’s inequality established in [16]:

$$|u|_{L^6(\Omega)} \leq c_0 \prod_{i=1}^3 \left(\frac{1}{b_i - a_i} |u|_{L^2(\Omega)} + \left| \frac{\partial u}{\partial x_i} \right|_{L^2(\Omega)} \right)^{\frac{1}{3}}, \quad \forall u \in H^1(\Omega), \tag{4.6}$$

where $\Omega = \prod_{i=1}^3 (a_i, b_i)$, and c_0 is a numerical constant. As a corollary of (4.6), we prove the following

Lemma 4.4. *There exists a constant c_0 independent of ε such that*

$$|v|_{L^6(M_\varepsilon)} \leq c_0 \|v\|_\varepsilon, \quad \forall v \in V_1, \quad (4.7)$$

$$|\nabla v|_{L^6(M_\varepsilon)} \leq c_0 \bar{h}^2 |\Delta_3 v|_\varepsilon, \quad \forall v \in D(A_1). \quad (4.8)$$

Proof. Inequality (4.7) is an easy consequences of (4.6) and the fact that v satisfies the Poincaré inequality $|v|_\varepsilon \leq \varepsilon \|v\|_\varepsilon$. The second inequality follows from the fact that $v = 0$ on Γ_b , which implies that $\nabla v = \varepsilon \nabla h \frac{\partial v}{\partial x_3}$. We skip the details.

Now we derive some inequalities concerning a scalar function T which satisfies the following boundary condition on ∂M_ε :

$$\frac{\partial T}{\partial n} = 0 \quad \text{on} \quad \Gamma_t \cup \Gamma_b, \quad \frac{\partial T}{\partial x_3} + \alpha_T T = g_T \quad \text{on} \quad \Gamma_i. \quad (4.9)$$

Lemma 4.5. *For $T \in H^2(M_\varepsilon)$ satisfying the boundary conditions (4.9), we have*

$$\sum_{i=1}^3 \left| \frac{\partial T}{\partial x_i} \right|_\varepsilon^2 + \frac{\alpha_T}{2} \int_{\Gamma_i} T^2 dx_1 dx_2 \leq |T|_\varepsilon |\Delta_3 T|_\varepsilon + \frac{1}{2\alpha_T} |g_T|_{L^2(\Gamma_i)}^2. \quad (4.10)$$

Acknowledgement. This work was partially supported by the National Science Foundation under the grant NSF - DMS 0074334, and by the Research Fund of Indiana University.

REFERENCES

- [1] Agmon, S., Douglis, A. & Nirenberg, L., Estimates near the boundary for solutions of the elliptic partial differential equations satisfying general boundary conditions, *Comm. Pure Appl. Math.*, **12**(1959), 623–723.
- [2] do Carmo, M., Riemman geometry, Birkhäuser, 1992.
- [3] Constantin, P. & Foias, C., Navier-Stokes equations, Univ. of Chicago Press, Chicago, 1988.
- [4] Grisvard, P., Elliptic problems in non-smooth domains, Monographs and studies in Mathematics, vol. 24, Pitman, London, 1985.
- [5] Grisvard, P. & Iooss, G., Problèmes aux limites unilatéraux dans des domaines non réguliers, *Publications des Séminaires de Mathématiques, Université de Rennes, exposé*, **9**(1975), 1–26.
- [6] Hu, C., Temam, R. & Ziane, M., The primitive equations for large scale ocean under small depth hypothesis, *Discrete and Continuous Dynamical Systems (DCDS-A)* (to appear).

- [7] Lions, J. L. & Magenes, E., Nonhomogeneous boundary value problems and applications, Vol 1, Springer-Verlag, New York, 1972.
- [8] Lions, J. L., Temam, R. & Wang, S., New formulation of the primitive equations of the atmosphere and applications, *Nonlinearity*, **5**(1992a), 237–288.
- [9] Lions, J. L., Temam, R. & Wang, S., On the equations of the large scale ocean, *Nonlinearity*, **5** (1992b), 1007–1053.
- [10] Lions, J. L., Temam, R. & Wang, S., Models of the coupled atmosphere and ocean (CAO I), *Computation Mechanics Advances*, **1**(1993a), 5–54.
- [11] Lions, J. L., Temam, R. & Wang, S., Numerical analysis of the coupled atmosphere and ocean models (CAO II), *Computation Mechanics Advances*, **1**(1993a), 55–120.
- [12] Lions, J. L., Temam, R. & Wang, S., Mathematical study of the coupled models of atmosphere and ocean (CAO III), *J. Math. Pures Appl.*, **74** (1995), 105–163.
- [13] Lions, J. L., Temam, R. & Wang, S., A simple global model for the general circulation of atmosphere, *Comm. Pure Appl. Math.*, **50**(1997), 707–752.
- [14] Lions, J. L., Temam, R. & Wang, S., Geostrophic asymptotic of the primitive equations of the atmosphere, *Topol. Methods Nonlinear Anal.*, **24**(1994), 253–287.
- [15] Temam, R., Navier-Stokes equations, theory and numerical analysis, AMS-Chelsea Series, AMS, Providence, 2001.
- [16] Temam, R. & Ziane, M., Navier-Stokes equations in three dimensional thin domains with various boundary conditions, *Advances in Differential Equations*, **1**(1996), 499–546.
- [17] Temam, R. & Ziane, M., Navier-Stokes equations in thin spherical shells, *Optimization methods in partial differential equations*, (1996), 281–314, *Contemporary Mathematics*, **209**.
- [18] Ziane, M., Regularity results for Stokes type systems, *Applicable Analysis*, **58** (1995), 263–292.
- [19] Ziane, M., Regularity results for stationary primitive equations of atmosphere and the ocean, *Nonlinear Analysis*, **28**(1997), 289–313.

LOCAL EXACT BOUNDARY CONTROLLABILITY FOR A CLASS OF QUASILINEAR HYPERBOLIC SYSTEMS*

LI TATSIEN

*Department of Mathematics, Fudan University,
Shanghai 200433, China. E-mail: dqli@fudan.edu.cn*

RAO BOPENG

*Institut de Recherche Mathématique Avancée, Université Louis Pasteur
de Strasbourg, 7 rue René-Descartes, 67084 Strasbourg, France*

Abstract

For a class of mixed initial-boundary value problem for general quasilinear hyperbolic systems, this paper establishes the local exact boundary controllability with boundary controls only acting on one end. As an application, the authors show the local exact boundary controllability for a kind of nonlinear vibrating string problem.

Keywords Exact boundary controllability, Quasilinear hyperbolic system, Nonlinear vibrating string equation

2000 MR Subject Classification 35L50, 35L20, 93B05, 93C20

§1. Introduction

Let us consider the following first order quasilinear hyperbolic system

$$\frac{\partial u}{\partial t} + A(u) \frac{\partial u}{\partial x} = F(u), \quad (1.1)$$

where $u = (u_1, \dots, u_n)^T$ is an unknown vector function of (t, x) , $A(u) = (a_{ij}(u))$ is an $n \times n$ matrix with suitably smooth elements $a_{ij}(u)(i, j =$

*Project supported by the Special Funds for Major State Basic Research Projects of China.

$1, \dots, n)$ and $F : R^n \rightarrow R^n$ is a vector function of u with suitably smooth components $f_i(u)$ ($i = 1, \dots, n$) such that

$$F(0) = 0. \quad (1.2)$$

By the definition of hyperbolicity, on the domain under consideration, the matrix $A(u)$ has n real eigenvalues $\lambda_i(u)$ ($i = 1, \dots, n$) and a complete set of left eigenvectors $l_i(u) = (l_{i1}(u), \dots, l_{in}(u))$ ($i = 1, \dots, n$):

$$l_i(u)A(u) = \lambda_i(u)l_i(u). \quad (1.3)$$

We have

$$\det |l_{ij}(u)| \neq 0. \quad (1.4)$$

Moreover, we assume that on the domain under consideration, the eigenvalues satisfy the following conditions:

$$\lambda_r(u) < 0 < \lambda_s(u) \quad (r = 1, \dots, m; s = m + 1, \dots, n). \quad (1.5)$$

Let

$$v_i = l_i(u)u \quad (i = 1, \dots, n). \quad (1.6)$$

We set the boundary conditions as follows:

$$x = 0 : \quad v_s = G_s(t, v_1, \dots, v_m) + H_s(t) \quad (s = m + 1, \dots, n), \quad (1.7)$$

$$x = 1 : \quad v_r = G_r(t, v_{m+1}, \dots, v_n) + H_r(t) \quad (r = 1, \dots, m), \quad (1.8)$$

where, without loss of generality, we assume that

$$G_i(t, 0, \dots, 0) \equiv 0 \quad (i = 1, \dots, n). \quad (1.9)$$

There is a number of publications on the exact controllability for linear hyperbolic systems (see [11] and the references therein). Especially, the exact boundary controllability for first order linear hyperbolic systems has been established by the characteristic method. J.-L. Lions introduced his Hilbert Uniqueness Method (HUM)(see [9,10]) which gives a more general and systematic framework for the study of the exact boundary controllability and the stabilisation for wave equations. Combining the HUM and Schauder's fixed point theorem, the first work on the exact controllability for semilinear wave equations was given by Zuazua [12, 13]. Later, using a global inversion theorem, Lasiecka and Triggiani [3] gave an abstract result on the global exact controllability for semilinear wave equations.

However, the exact controllability for the quasilinear hyperbolic systems remains quite open. To our knowledge, the first work in this direction was done by Cirinà [2] (see also [1]). Under linear boundary controls, he proved the local null controllability essentially for quasilinear hyperbolic systems of diagonal form. In [5] and [6] the exact boundary controllability for reducible quasilinear hyperbolic systems was established by a characteristic method. Recently, these results were generalized to the case of general quasilinear hyperbolic systems. The following result was proved in [8].

Proposition 1.1. *Assume that $l_{ij}(u)$, $\lambda_i(u)$, $f_i(u)$ and $G_i(t, \cdot)$ ($i, j = 1, \dots, n$) are all C^1 functions with respect to their arguments. Assume furthermore that (1.2), (1.4)–(1.5) and (1.9) hold. Let*

$$T_0 > \max_{i=1, \dots, n} \frac{1}{|\lambda_i(0)|}. \quad (1.10)$$

Then, for any given initial data $\phi \in C^1[0, 1]$ and final data $\psi \in C^1[0, 1]$ with small C^1 norm, there exist boundary controls $H_i(t) \in C^1[0, T_0]$ ($i = 1, \dots, n$) with small C^1 norm, such that the mixed initial-boundary value problem for system (1.1) with the initial condition

$$t = 0: \quad u = \phi(x) \quad (1.11)$$

and the boundary conditions (1.7)–(1.8) admits a unique C^1 solution $u = u(t, x)$ with small C^1 norm on the domain

$$R(T_0) = \{(t, x) \mid 0 \leq t \leq T_0, \quad 0 \leq x \leq 1\},$$

which satisfies the final condition

$$t = T_0: \quad u = \psi(x). \quad (1.12)$$

As mentioned in [8], the exact controllability time T_0 given in Proposition 1.1 is optimal, however, the boundary controls are not unique.

We will prove in §2 that for a class of mixed initial-boundary value problem, the number of boundary controls can be diminished, provided that the exact controllability time is doubled. In §3 this result will be applied to show the local exact boundary controllability for a class of nonlinear vibrating string problem.

§2. Main Results

We now suppose that

$$n = 2m. \quad (2.1)$$

We suppose furthermore that the boundary condition (1.7) (resp. (1.8)) can be equivalently rewritten as

$$x = 0: \quad v_r = \overline{G}_r(t, v_{m+1}, \dots, v_n) + \overline{H}_r(t) \quad (r = 1, \dots, m) \quad (2.2)$$

$$[\text{resp. } x = 1: \quad v_s = \overline{G}_s(t, v_1, \dots, v_m) + \overline{H}_s(t) \quad (s = m + 1, \dots, n)],$$

where

$$\overline{G}_r(t, 0, \dots, 0) \equiv 0 \quad (r = 1, \dots, m) \quad (2.3)$$

$$[\text{resp. } \overline{G}_s(t, 0, \dots, 0) \equiv 0 \quad (s = m + 1, \dots, n)],$$

consequently,

$$\text{small } C^1 \text{ norm of } H_s \iff \text{small } C^1 \text{ norm of } \overline{H}_r \quad (2.4)$$

$$[\text{resp. } \text{small } C^1 \text{ norm of } H_r \iff \text{small } C^1 \text{ norm of } \overline{H}_s],$$

where $r = 1, \dots, m; s = m + 1, \dots, n$.

Theorem 2.1. *Under the assumptions of Proposition 1.1, we suppose furthermore that conditions (2.1)–(2.4) hold and $\overline{G}_r(t, \cdot)$ ($r = 1, \dots, m$) (resp. $\overline{G}_s(t, \cdot)$ ($s = m + 1, \dots, n$)) are C^1 functions with respect to their arguments. Let*

$$T > 2 \max_{i=1, \dots, n} \frac{1}{|\lambda_i(0)|}. \quad (2.5)$$

Suppose finally that $H_s(t)$ ($s = m + 1, \dots, n$) (resp. $H_r(t)$ ($r = 1, \dots, m$)) are given $C^1[0, T]$ functions with small C^1 norm. Then, for any given initial data $\phi \in C^1[0, 1]$ and final data $\psi \in C^1[0, 1]$ with small C^1 norm, such that the conditions of C^1 compatibility are satisfied at points $(0, 0)$ and $(T, 0)$ (resp. $(0, 1)$ and $(T, 1)$) respectively, there exist boundary controls $H_r(t) \in C^1[0, T]$ ($r = 1, \dots, m$) (resp. $H_s(t) \in C^1[0, T]$ ($s = m + 1, \dots, n$)) with small C^1 norm, such that the mixed initial boundary value problem (1.1), (1.11) and (1.7)–(1.8) admits a unique C^1 solution $u = u(t, x)$ with small C^1 norm on the domain

$$R(T) = \{(t, x) \mid 0 \leq t \leq T, \quad 0 \leq x \leq 1\},$$

which satisfies the final condition

$$t = T : \quad u = \psi(x). \tag{2.6}$$

In order to prove Theorem 2.1, it suffices to establish the following

Lemma 2.1. *Under the assumptions of Theorem 2.1, for any given initial data $\phi \in C^1[0, 1]$ and final data $\psi \in C^1[0, 1]$ with small C^1 norm, such that the conditions of C^1 compatibility are satisfied at points $(0, 0)$ and $(T, 0)$ (resp. $(0, 1)$ and $(T, 1)$) respectively, the quasilinear hyperbolic system (1.1) with the boundary condition (1.7) (resp. (1.8)) admits a C^1 solution $u = u(t, x)$ with small C^1 norm on the domain $R(T)$, which satisfies (1.11) and (2.6).*

In fact, let $u = u(t, x)$ be a C^1 solution on the domain $R(T)$ given by Lemma 2.1. Taking the boundary controls as

$$H_r(t) = (v_r - G_r(t, v_{m+1}, \dots, v_n)|_{x=1} \quad (r = 1, \dots, m) \tag{2.7}$$

$$[\text{resp. } H_s(t) = (v_s - G_s(t, v_1, \dots, v_m)|_{x=0} \quad (s = m + 1, \dots, n)],$$

the C^1 norm of which is small, we obtain the exact boundary controllability desired by Theorem 2.1.

We now prove Lemma 2.1. For fixing the idea, in what follows we consider only the case that the boundary controls are given at the end $x = 1$.

Noting (2.5), there exists an $\epsilon_0 > 0$ such that

$$T > 2 \max_{|u| \leq \epsilon_0, i=1, \dots, n} \frac{1}{|\lambda_i(u)|}. \tag{2.8}$$

Let

$$T_1 = \max_{|u| \leq \epsilon_0, i=1, \dots, n} \frac{1}{|\lambda_i(u)|}. \tag{2.9}$$

We first consider the forward mixed initial-boundary value problem for system (1.1) with the initial data

$$t = 0 : \quad u = \phi(x), \quad 0 \leq x \leq 1 \tag{2.10}$$

and the boundary conditions (1.7) and

$$x = 1 : \quad v_r = \bar{f}_r(t) \quad (r = 1, \dots, m), \tag{2.11}$$

where $\bar{f}_r(t)$ are any given functions of t with small $C^1[0, T_1]$ norm. We assume that the conditions of C^1 compatibility are satisfied at point $(0, 1)$.

By [7], there exists a unique semi-global C^1 solution $u = u^{(1)}(t, x)$ with small C^1 norm on the domain

$$\{(t, x) \mid 0 \leq t \leq T_1, \quad 0 \leq x \leq 1\}. \quad (2.12)$$

Thus we can uniquely determine the corresponding value of $u = u^{(1)}(t, x)$ on $x = 0$ as

$$x = 0: \quad u = a(t), \quad 0 \leq t \leq T_1 \quad (2.13)$$

and the $C^1[0, T_1]$ norm of $a(t)$ is suitably small.

Similarly, we consider the backward mixed initial-boundary value problem for system (1.1) with the initial condition

$$t = T: \quad u = \psi(x), \quad 0 \leq x \leq 1 \quad (2.14)$$

and the boundary conditions (2.2) and

$$x = 1: \quad v_s = \bar{g}_s(t) \quad (s = m + 1, \dots, n), \quad (2.15)$$

where $\bar{g}_s(t)$ ($s = m + 1, \dots, n$) are any given functions of t with small $C^1[T - T_1, T]$ norm. We assume that the conditions of C^1 compatibility are satisfied at point $(T, 1)$. Once again by [7], there exists a unique semi-global C^1 solution $u = u^{(2)}(t, x)$ with small C^1 norm on the domain

$$\{(t, x) \mid T - T_1 \leq t \leq T, \quad 0 \leq x \leq 1\}. \quad (2.16)$$

Thus we can uniquely determine the corresponding value of $u = u^{(2)}(t, x)$ on $x = 0$ as

$$x = 0: \quad u = b(t), \quad T - T_1 \leq t \leq T \quad (2.17)$$

and the $C^1[T - T_1, T]$ norm of $b(t)$ is suitably small. Noting that both $a(t)$ and $b(t)$ satisfy the boundary condition (1.7), we can find a $C^1[0, T]$ function $c(t)$ with small C^1 norm, such that

$$c(t) = \begin{cases} a(t), & 0 \leq t \leq T_1, \\ b(t), & T - T_1 \leq t \leq T \end{cases} \quad (2.18)$$

and $c(t)$ satisfies the boundary condition (1.7) on the whole interval $[0, T]$.

Now we change the order of the variables t and x , then system (1.1) is rewritten in the following form

$$\frac{\partial u}{\partial x} + A^{-1}(u) \frac{\partial u}{\partial t} = \tilde{F}(u) := A^{-1}(u)F(u). \quad (2.19)$$

We notice that

$$\tilde{F}(0) = 0. \tag{2.20}$$

Noting (1.5), the eigenvalues of the inverse matrix $A^{-1}(u)$ satisfy

$$\frac{1}{\lambda_r(u)} < 0 < \frac{1}{\lambda_s(u)} \quad (r = 1, \dots, m; \quad s = m + 1, \dots, n). \tag{2.21}$$

Moreover, since the matrices $A(u)$ and $A^{-1}(u)$ have the same left eigenvectors, we can still define the variables $v_i (i = 1, \dots, n)$ by the same formula (1.6).

We now consider the mixed initial-boundary value problem for system (2.19) with the initial condition

$$x = 0 : \quad u = c(t), \quad 0 \leq t \leq T \tag{2.22}$$

and the boundary conditions

$$t = 0 : \quad v_s = \Phi_s(t) \quad (s = m + 1, \dots, n), \tag{2.23}$$

$$t = T : \quad v_r = \Psi_r(t) \quad (r = 1, \dots, m), \tag{2.24}$$

where

$$\Phi_i(x) = l_i(\phi(x))\phi(x) \quad (i = 1, \dots, n), \tag{2.25}$$

$$\Psi_i(x) = l_i(\psi(x))\psi(x) \quad (i = 1, \dots, n), \tag{2.26}$$

the C^1 norm of which is small. It is easy to see that the mixed initial-boundary value problem (2.19) and (2.22)–(2.24) satisfies the conditions of C^1 compatibility at points $(0, 0)$ and $(T, 0)$ respectively. Therefore, by [7] there exists a unique semi-global C^1 solution $u = u(t, x)$ with small C^1 norm on the domain

$$R(T) = \{(t, x) \mid 0 \leq t \leq T, \quad 0 \leq x \leq 1\}.$$

In order to finish the proof of Lemma 2.1, it is only necessary to check that

$$t = 0 : \quad u = \phi(x), \quad 0 \leq x \leq 1, \tag{2.27}$$

$$t = T : \quad u = \psi(x), \quad 0 \leq x \leq 1. \tag{2.28}$$

In fact, the C^1 solutions $u = u(t, x)$ and $u = u^{(1)}(t, x)$ satisfy the system (2.19) (namely, (1.1)), the initial condition

$$x = 0 : \quad u = a(t), \quad 0 \leq t \leq T_1 \tag{2.29}$$

and the boundary condition

$$t = 0 : \quad v_s = \Phi_s(t) \quad (s = m + 1, \dots, n). \quad (2.30)$$

Because of the finiteness of the speed of wave propagation and the choice of T_1 given by (2.9), the mixed-initial boundary problem (2.19) and (2.29)–(2.30) has a unique C^1 solution on the domain

$$\{(t, x) \mid 0 \leq t \leq T_1(1 - x), \quad 0 \leq x \leq 1\} \quad (2.31)$$

(see [4]). Then it follows that

$$u(t, x) \equiv u^{(1)}(t, x) \quad (2.32)$$

on this domain. In particular, we obtain (2.27). We can get (2.28) in a similar way.

Thus $u = u(t, x)$ is the desired C^1 solution. The proof of Lemma 2.1 is complete.

Remark 2.1. The exact controllability time given by Theorem 2.1 is optimal.

Remark 2.2. The boundary controls in Theorem 2.1 are not unique.

§3. Application to a Class of Nonlinear Vibrating String Problem

In this section, we will use the previous results to show the local exact boundary controllability for the following nonlinear vibrating string equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left(K \left(\frac{\partial u}{\partial x} \right) \right) = F \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial t} \right), \quad (3.1)$$

where $K = K(v)$ is a given C^2 function of v , such that

$$K'(v) > 0, \quad (3.2)$$

and $F = F(v, w)$ is a C^1 function of v and w , satisfying

$$F(0, 0) = 0. \quad (3.3)$$

We consider the exact boundary controllability only with one control applied at one end of the string. The boundary condition at the end $x = 0$ is of Dirichlet type:

$$u = h(t), \quad (3.4)$$

where $h(t)$ is a given C^2 function of t ; while the boundary condition at the end $x = 1$ is one of the following types:

$$u = \bar{h}(t), \tag{3.5.1}$$

$$u_x = \bar{h}(t), \tag{3.5.2}$$

$$u_x + \alpha u = \bar{h}(t), \tag{3.5.3}$$

$$u_x + \alpha u_t = \bar{h}(t), \tag{3.5.4}$$

where α is a positive constant and $\bar{h}(t)$, as boundary control, is a C^2 function (in case (3.5.1)) or a C^1 function (in cases (3.5.2)–(3.5.4)).

We want to find a time $T > 0$ and suitable boundary control $\bar{h}(t)$ with small C^1 norm $\|\bar{h}'\|_{C^1[0,T]}$ in case (3.5.1) or $\|\bar{h}\|_{C^1[0,T]}$ in cases (3.5.2)–(3.5.4), such that for any given initial data $(\phi(x), \psi(x))$ and final data $(\Phi(x), \Psi(x))$ with small C^1 norms

$$\|\phi'\|_{C^1[0,1]}, \quad \|\psi\|_{C^1[0,1]}, \quad \|\Phi'\|_{C^1[0,1]}, \quad \|\Psi\|_{C^1[0,1]},$$

satisfying the conditions of C^2 compatibility at points $(0,0)$ and $(T,0)$ respectively, the C^2 solution $u = u(t, x)$ to the mixed initial-boundary value problem for equation (3.1) with the initial condition

$$t = 0: \quad u = \phi(x), \quad u_t = \psi(x) \tag{3.6}$$

and the boundary conditions (3.4)–(3.5) satisfies the final condition

$$t = T: \quad u = \Phi(x), \quad u_t = \Psi(x). \tag{3.7}$$

We will prove

Theorem 3.1. *Let*

$$T > \frac{2}{\sqrt{K'(0)}}. \tag{3.8}$$

Then for any given initial data $\phi(x) \in C^2[0, 1], \psi(x) \in C^1[0, 1]$ and final data $\Phi(x) \in C^2[0, 1], \Psi(x) \in C^1[0, 1]$ with small C^1 norms

$$\|\phi'\|_{C^1[0,1]}, \quad \|\psi\|_{C^1[0,1]} \quad \text{and} \quad \|\Phi'\|_{C^1[0,1]}, \quad \|\Psi\|_{C^1[0,1]}$$

and any given function $h(t) \in C^2[0, T]$ with small C^1 norm $\|h'\|_{C^1[0,T]}$, satisfying the following conditions of C^2 compatibility at points $(0,0)$ and $(T,0)$ respectively:

$$\begin{cases} h(0) = \phi(0), & h'(0) = \psi(0), \\ h''(0) = K'(\phi'(0))\phi''(0) + F(\phi'(0), \psi(0)) \end{cases} \tag{3.9}$$

and

$$\begin{cases} h(T) = \Phi(0), & h'(T) = \Psi(0), \\ h''(T) = K'(\Phi'(0))\Phi''(0) + F(\Phi'(0), \Psi(0)), \end{cases} \quad (3.10)$$

there exists a boundary control $\bar{h}(t) \in C^2[0, T]$ with small C^1 norm $\|\bar{h}'\|_{C^1[0, T]}$ in case (3.5.1) or $\bar{h}(t) \in C^1[0, T]$ with small C^1 norm $\|\bar{h}\|_{C^1[0, T]}$ in cases (3.5.2)–(3.5.4), such that the mixed initial-boundary value problem for equation (3.1) with the initial condition (3.6), the boundary condition (3.4) at the end $x = 0$ and one of the boundary conditions (3.5) at the end $x = 1$ admits a unique C^2 solution $u = u(t, x)$ on the domain

$$R(T) = \{(t, x) \mid 0 \leq t \leq T, \quad 0 \leq x \leq 1\},$$

which satisfies the final condition (3.7).

In order to prove Theorem 3.1, setting

$$v = \frac{\partial u}{\partial x}, \quad w = \frac{\partial u}{\partial t}, \quad (3.11)$$

equation (1.1) is reduced to the following first order quasilinear system

$$\begin{cases} \frac{\partial v}{\partial t} - \frac{\partial w}{\partial x} = 0, \\ \frac{\partial w}{\partial t} - \frac{\partial K(v)}{\partial x} = F(v, w). \end{cases} \quad (3.12)$$

The system is strictly hyperbolic with two distinct real eigenvalues $\pm \lambda$, where

$$\lambda = \sqrt{K'(v)} > 0. \quad (3.13)$$

We introduce the Riemann invariants r and s as follows:

$$2r = G(v) + w, \quad 2s = -G(v) + w, \quad (3.14)$$

where

$$G(v) = \int_0^v \sqrt{K'(v)} dv \quad (3.15)$$

with

$$G(0) = 0, \quad G'(v) = \sqrt{K'(v)} > 0. \quad (3.16)$$

Let H be the inverse function of G . It follows from (3.14) that

$$w = r + s, \quad v = H(r - s). \quad (3.17)$$

Using the Riemann invariants r and s , (3.12) can be rewritten into the

diagonal form

$$\begin{cases} \frac{\partial r}{\partial t} - \lambda \frac{\partial r}{\partial x} = f(r, s), \\ \frac{\partial s}{\partial t} + \lambda \frac{\partial s}{\partial x} = f(r, s), \end{cases} \quad (3.18)$$

where

$$f(r, s) = \frac{1}{2}F(H(r - s), r + s) \quad (3.19)$$

with

$$f(0, 0) = 0. \quad (3.20)$$

Correspondingly, the initial and final conditions (3.6)–(3.7) yield

$$t = 0 : \quad \begin{cases} r = r_0(x) := \frac{1}{2}(G(\phi'(x)) + \psi(x)), \\ s = s_0(x) := \frac{1}{2}(-G(\phi'(x)) + \psi(x)), \end{cases} \quad (3.21)$$

$$t = T : \quad \begin{cases} r = r_1(x) := \frac{1}{2}(G(\Phi'(x)) + \Psi(x)), \\ s = s_1(x) := \frac{1}{2}(-G(\Phi'(x)) + \Psi(x)), \end{cases} \quad (3.22)$$

while the boundary condition (3.4) implies that

$$x = 0 : \quad r + s = h'(t). \quad (3.23)$$

Moreover it follows from the last two equalities in (3.9)–(3.10) that the conditions of C^1 compatibility at point $(0, 0)$:

$$\begin{cases} r_0(0) + s_0(0) = h'(0), \\ h''(0) = \lambda(r_0(0), s_0(0))(r'_0(0) - s'_0(0)) + 2f(r_0(0), s_0(0)) \end{cases} \quad (3.24)$$

and at point $(0, T)$:

$$\begin{cases} r_1(0) + s_1(0) = h'(T), \\ h''(T) = \lambda(r_1(0), s_1(0))(r'_1(0) - s'_1(0)) + 2f(r_1(0), s_1(0)) \end{cases} \quad (3.25)$$

are satisfied. Then it is easy to check that we can apply Lemma 2.1 to get the following

Lemma 3.1. *Assume that $h(t) \in C^2[0, T]$ with small C^1 norm $\|h'\|_{C^1[0, T]}$, where T is defined by (3.8). Then for any given initial data $r_0, s_0 \in C^1[0, 1]$ and final data $r_1, s_1 \in C^1[0, 1]$ with small C^1 norm, satisfying the conditions of C^1 compatibility (3.24)–(3.25), the quasilinear hyperbolic system (3.18) associated with the boundary condition (3.23) admits a C^1 solution $(r, s) = (r(t, x), s(t, x))$ with small C^1 norm on the domain*

$$R(T) = \{(t, x) \mid 0 \leq t \leq T, \quad 0 \leq x \leq 1\},$$

which satisfies the initial and final conditions (3.21)–(3.22).

Proof of Theorem 3.1. Applying Lemma 3.1, we can find a C^1 solution $(v, w) = (v(t, x), w(t, x))$ to system (3.12) on the domain $R(T)$, which satisfies the initial and final conditions

$$t = 0 : \quad v = \phi'(x), \quad w = \psi(x), \quad (3.26)$$

$$t = T : \quad v = \Phi'(x), \quad w = \Psi(x) \quad (3.27)$$

and the boundary condition

$$x = 0 : \quad w = h'(t). \quad (3.28)$$

Let

$$u(t, x) = h(t) + \int_0^x v(t, y) dy. \quad (3.29)$$

It is easy to see that $u = u(t, x)$ is a C^2 function on the domain $R(T)$, such that

$$u(t, 0) = h(t) \quad (3.30)$$

and

$$\frac{\partial}{\partial x} u(t, x) = v(t, x). \quad (3.31)$$

On the other hand, noting the first equation in (3.12) and using (3.28), we have

$$\frac{\partial}{\partial t} u(t, x) = h'(t) + \int_0^x \frac{\partial}{\partial x} w(t, y) dy = w(t, x). \quad (3.32)$$

Therefore, it follows from the second equation in (3.12) that the function u defined by (3.29) satisfies the string equation (3.1) and the boundary condition (3.4). Moreover, noting the first equality in (3.9)–(3.10), we check easily that u satisfies the initial and final conditions (3.6)–(3.7):

$$t = 0 : \quad u = h(0) + \int_0^x v(0, y) dy = h(0) + \int_0^x \phi'(y) dy = \phi(x), \quad (3.33)$$

$$t = T : \quad u = h(T) + \int_0^x v(T, y) dy = h(T) + \int_0^x \Phi'(y) dy = \Phi(x). \quad (3.34)$$

We now define the control $h(t)$ at the end $x = 1$ by one of the following expressions:

$$\bar{h}(t) =: u, \quad (3.35.1)$$

$$\bar{h}(t) =: u_x, \quad (3.35.2)$$

$$\bar{h}(t) =: u_x + \alpha u, \quad (3.35.3)$$

$$\bar{h}(t) =: u_x + \alpha u_t. \quad (3.35.4)$$

The proof of Theorem 3.1 is thus finished.

REFERENCES

- [1] Cirinà, M., Nonlinear hyperbolic problems with solutions on preassigned sets, *Michigan Math. J.*, **17**(1970), 193–209.
- [2] Cirinà, M., Boundary controllability of nonlinear hyperbolic systems, *SIAM J. Control Optim.*, **7**(1969), 198–212.
- [3] Lasiecka, I. & Triggiani, R., Exact controllability of semilinear abstract systems with applications to waves and plates boundary control problems, *Appl. Math. Optim.*, **23**(1991), 109–154.
- [4] Li Tatsien & Yu Wenci, Boundary value problems for quasilinear hyperbolic systems, Duck University Mathematics Series V, 1985.
- [5] Li Tatsien & Zhang Bingyu, Global exact controllability of a class of quasilinear hyperbolic systems, *J. Math. Anal. Appl.*, **225**(1998), 289–311.
- [6] Li Tatsien, Rao Bopeng & Jin Yi, Semi-global C^1 solution and exact boundary controllability for reducible quasilinear hyperbolic systems, *Math. Model. Numer. Anal.*, **34**(2000), 399–408.
- [7] Li Tatsien & Jin Yi, Semi-global C^1 solution to the mixed initial-boundary value problem for quasilinear hyperbolic systems, *Chin. Ann. of Math.*, **22B**:3(2001), 325–336.
- [8] Li Tatsien, Rao Bopeng & Jin Yi, Solution C^1 semi-globale et contrôlabilité exacte frontière de systèmes hyperboliques quasi linéaires, *C. R. Acad. Sci. Paris*, **333**:Série I(2001), 219–224.
- [9] Lions, J. -L., Exact controllability, stabilization and perturbations for distributed systems, *SIAM Rev.*, **30**(1988), 1–68.
- [10] Lions, J.-L., Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués, Vol 1, Masson, 1988.
- [11] Russell, D. L., Controllability and stabilizability theory for linear partial differential equations: Recent progress and open questions, *SIAM Rev.*, **20**(1978), 639–739.
- [12] Zuazua, E., Exact controllability for the semilinear wave equation, *J. Math. Pures et Appl.*, **69**(1990), 1–31.
- [13] Zuazua, E., Exact controllability for the semilinear wave equation in one space dimension, *Ann. Inst. H. Poincaré, Analyse Non Linéaire*, **10**(1993), 109–129.

This page intentionally left blank

ON NONLINEAR DIFFERENTIAL GALOIS THEORY

B. MALGRANGE

*Université de Grenoble I, Institut Fourier, Laboratoire de
Mathématiques, UMR 5582 CNRS-UJF, B.P. 74, 38402 ST MARTIN
D'HÈRES Cedex, France. E-mail: malgrang@fourier.ujf-grenoble.fr*

Abstract

Let X denote a complex analytic manifold, and let $\text{Aut}(X)$ denote the space of invertible maps of a germ (X, a) to a germ (X, b) ; this space is obviously a groupoid; roughly speaking, a “Lie groupoid” is a subgroupoid of $\text{Aut}(X)$ defined by a system of partial differential equations. To a foliation with singularities on X one attaches such a groupoid, e.g. the smallest one whose Lie algebra contains the vector fields tangent to the foliation. It is called “the Galois groupoid of the foliation”. Some examples are considered, for instance foliations of codimension one, and foliations defined by linear differential equations; in this last case one recuperates the usual differential Galois group.

Keywords Differential Galois group, Complex analytic manifold, Lie groupoid

2000 MR Subject Classification 22E30

This is an account of a work in course of progress. The aim is the following: define and study, for non linear differential equations, an object which generalizes the differential Galois group of linear equations. In [9], I give such a definition, and I prove the required result in the linear case. Here, I recall it shortly, and I insist on further examples and on open problems.

I should mention that another definition of a differential Galois group was proposed several years ago, by Umemura [15]; at the moment, I do not know the exact relations between both theories.

§1. General Definitions

I give the definitions in the complex analytic case; a similar theory, somewhat simpler, could be developed in the algebraic case (but, up to now, nothing is written in this context).

Let X denote a (smooth) complex analytic manifold, of dimension n ; let $\text{Aut} X$ be the space of germs of invertible maps $(X, a) \rightarrow (X, b)$ [e.g. the source is a , and the target b ; $a, b \in X$]; let also J_k , resp. J_k^* be the space of jets of order k , resp. invertible jets of order k , of maps from X to X . I provide these spaces by the sheaf (on X^2) of functions which are analytic on the variables of X^2 , and polynomial in the derivatives: precisely, in local coordinates, let x_1, \dots, x_n be the coordinates at the source, and y_1, \dots, y_n the coordinates at the target; with the standard notations, we have

$$\begin{aligned} \mathcal{O}_{J_k} &= \mathcal{O}_{X^2}[y_j^\alpha], \quad 1 \leq j \leq n, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n, \\ &|\alpha| = \alpha_1 + \dots + \alpha_n \in \{1, 2, \dots, k\}, \\ \mathcal{O}_{J_k^*} &= \mathcal{O}_{X^2}[y_j^\alpha, \delta^{-1}], \quad \delta = \det(y_j^i) \text{ the Jacobian matrix.} \end{aligned}$$

The main objects in consideration here are what I call “ D -groupoids” or “Lie groupoids” on X . Roughly speaking, there are the subgroupoids of $\text{Aut} X$ which are defined by a system of partial differential equations. More precisely, putting $\mathcal{O}_{J^*} = \cup \mathcal{O}_{J_k^*}$, we consider a sheaf of ideals $\mathfrak{J} \subset \mathcal{O}_{J^*}$ which has the following properties:

(i) $\mathfrak{J}_k = \mathcal{O}_{J_k^*} \cap \mathfrak{J}$ is coherent and reduced (i.e., equal to its radical), for every $k \geq 0$; we say for short that \mathfrak{J} is pseudocoherent, and reduced.

(ii) \mathfrak{J} is differential, e.g. stable by derivations.

[The derivations are defined in local coordinates by the standard formula

$$D_i f = \frac{\partial f}{\partial x_i} + \sum_{j, \alpha} \frac{\partial f}{\partial y_j^\alpha} y_j^{\alpha + \varepsilon_i}, \quad \varepsilon_i = (0, \dots, 1, \dots, 0);$$

it is easy to verify that the stability by derivations is independent of the coordinates chosen.]

The two preceding properties define, generally speaking, a system of partial differential equations in $\text{Aut} X$. One has to add a third condition to have a Lie groupoid, which I explain now.

The set J_k^* is obviously a groupoid for the composition of jets (e.g. the composition $\psi\varphi$ exists if target $\varphi =$ source ψ , and all elements are invert-

ible); this property can be translated into a property of $\mathcal{O}_{J_k^*}$; now, take \mathfrak{J}_k a coherent sheaf of ideals of $\mathcal{O}_{J_k^*}$, and call Y_k the ringed space $(X^2, \mathcal{O}_{J_k^*}/\mathfrak{J}_k)$. This has a sense to say that “ Y_k is a subgroupoid of J_k^* ”; we have to take this sentence in the sense of ideals, or in the “scheme sense”, and not only in the set-theoretical sense (this can be expressed by properties of \mathfrak{J}_k which I leave the reader to guess, or to look, f.i. in [7]).

So, now, the third condition to impose to \mathfrak{J} is the following

(iii-1) For every k , Y_k contains the identity (e.g. the ideal of the identity contains \mathfrak{J}_k), and Y_k is stable by the inverse.

(iii-2) For every $U \subset X$, open relatively compact (I abbreviate $U \subset\subset X$), there exists Z , closed analytic subset of codimension ≥ 1 of U such that, on $(U - Z)^2$, Y_k is a subgroupoid of J_k^* for k large.

[Note that, in the algebraic case, the “ U ” would be unnecessary, and also probably the restriction “ k large”.]

One can see that this definition has reasonably good properties, and covers the cases of interest, for the following reasons:

(a) Any increasing sequence $\mathfrak{J}^\ell \subset \mathfrak{J}^{\ell+1} \subset \dots$ of ideals defining a D -groupoid is locally stationary.

(b) Let $\mathfrak{J}_k \subset \mathcal{O}_{J_k^*}$ be a coherent sheaf of ideals, not necessarily reduced defining a subgroupoid of J_k^* outside of some $Z \subset X$ of codimension ≥ 1 ; let \mathfrak{J}' be the pseudocoherent and differential ideal generated by \mathfrak{J}_k ; then the radical $\mathfrak{J}'^{\text{rad}}$ of \mathfrak{J}' defines a Lie groupoid; and, on every $U \subset\subset X$, one has $\mathfrak{J}'^{\text{rad}} = \mathfrak{J}'$ outside a subset of codimension ≥ 1 .

This situation occurs practically when \mathfrak{J}_k is the ideal expressing the condition that the transformation stabilizes a “differential structure of order k , meromorphic on X with poles on Z ” (whichever be the reasonable meaning given to this expression); standard examples are given by the groupoids preserving a 2-form (for instance, symplectic structures), one form up to invertible factors (f.i. contact structures), etc. In the literature, these objects are usually called “Lie pseudogroups”, and considered in the nonsingular case; but here, I need to accept singularities.

(c) From (a) and (b), one deduces that any intersection of D -groupoids is locally finite, and is a D -groupoid (if the groupoids are defined by the sheaf of ideals $\mathfrak{J}^\alpha \subset \mathcal{O}_{J^*}$, their intersection is the smallest pseudocoherent reduced differential sheaf of ideals which contains the \mathfrak{J}^α).

For these properties, see [9].

§2. *D*-Envelope and Galois Groupoid

Call “*D*-Lie-algebra” any system of linear partial differential equations on Θ , the sheaf of vector fields of X , such that its solutions are stable by Lie bracket. A *D*-groupoid has a *D*-Lie algebra, which is simply the linearized differential system along the identity.

But the converse is not true. This is similar to the fact that a Lie subalgebra of $\mathrm{Gl}(n, \mathbb{C})$ is not necessarily the Lie algebra of an algebraic subgroup of $\mathrm{Gl}(n, \mathbb{C})$. Given a *D*-Lie algebra \mathcal{L} , we can therefore define its *D*-envelope, as the smallest *D*-groupoid whose Lie algebra contains \mathcal{L} ; this is meaningful, according to the property 1-c.

An especially interesting case is the case of the foliations with singularities. This is defined, f.i. by a coherent subsheaf N of Ω^1 (the sheaf of 1-forms on X) of locally constant rank outside of $Z \subset X$ of codimension ≥ 1 and verifying outside of Z the Frobenius condition $dN \subset \Omega^1 \wedge N$. We can suppose, by increasing a little bit N , that any local section which is in N outside of Z is actually in N . Then N defines a *D*-Lie algebra \mathcal{F} , the vectors fields orthogonal to N (in fact, \mathcal{F} is defined by equations of order 0).

Definition *The Galois groupoid $G(\mathcal{F})$ of \mathcal{F} is its *D*-envelope.*

Before giving examples, a few explanations are necessary. First, I will use the following facts:

- (i) The solutions of a *D*-groupoid Y make a subgroupoid, in the usual sense, of $\mathrm{Aut}X$.
- (ii) These solutions determine Y . Therefore, I will often identify both objects implicitly.

The first result is essentially obvious. The second one is a general fact of differential algebra (see [10], or [13] for the algebraic case).

Now consider, outside Z , the groupoid of automorphisms of X which preserve \mathcal{F} (or N); it is easy to prove that the Zariski closure of the corresponding *D*-groupoid is a *D*-groupoid on X ; denote it by $\mathrm{Aut}\mathcal{F}$. It is obvious that its *D*-Lie algebra contains \mathcal{F} ; therefore, $\mathrm{Aut}\mathcal{F}$ contains $\mathrm{Gal}\mathcal{F}$.

Call “admissible” (w.r.t. \mathcal{F}) any *D*-groupoid Y contained in $\mathrm{Aut}\mathcal{F}$ and containing $\mathrm{Gal}\mathcal{F}$. If we consider Y near a pair $(a, b) \in (X - Z)^2$, no

condition for Y occurs along the leaves, and Y is given by equations on the variables transversal to the foliation (f.i., in the case of $\text{Aut}\mathcal{F}$, there are no such equations).

More precisely, suppose \mathcal{F} of codimension d outside Z , and call “transversal” a locally closed submanifold T of X of dimension d which, outside of a set of codimension ≥ 1 , does not meet Z and is transverse to \mathcal{F} . Call \mathcal{T} the disjoint union of all the transversals; it is easy to see that, on \mathcal{T} , an admissible groupoid Y defines a groupoid: the “transversal groupoid” defined by Y . The philosophy is the following one: from a geometrical point of view, the interesting object is the transversal groupoid (which will correspond, f.i. to some transversal structure). But, to verify the admissibility, we have to go back to X , and to verify the possibility of extension along Z (actually, meromorphic extension in a suitable sense will be sufficient, since, then, one has just to take Zariski closures of the corresponding varieties $Y_k \hookrightarrow J_k^*$).

§3. Examples

The first example is treated in [9]; the second one will be developed elsewhere.

(i) The Linear Case

Let C be a complex nonsingular connected curve (= manifold of dimension one), and $X \xrightarrow{\pi} C$ a vector bundle over C . Let S be a discrete subset of C , and $Z = \pi^{-1}S$; we suppose given a connection ∇ on X , meromorphic on S (in local coordinates, this is simply a linear differential system $Y' = aY$, with a meromorphic S) with poles on S ; note that we could equally consider flat meromorphic connections in the sense of [4] on vector bundles over a nonsingular analytic manifold of any dimension; the result would be similar.

On the total space X of the bundle, ∇ defines a foliation \mathcal{F} of dimension one, with singularities on Z . The Galois groupoid of \mathcal{F} can be described in the following way:

(a) Outside Z , the linear structures of $X \rightarrow C$ gives a “transversal linear structure”, and a corresponding subgroupoid of $\text{Aut}\mathcal{F}$; one proves that this groupoid extends to X (in a unique way) into an admissible groupoid $\text{Lin}\mathcal{F}$.

(b) Admissible subgroupoids of $\text{Lin}\mathcal{F}$ can be described in the following way: choose a base point $a \in C - S$, and put $X_a = \pi^{-1}a$. Let G be an algebraic subgroup of $\text{Gl}(X_a)$, containing the monodromy of ∇ at a .

Now, let $\text{Iso}X$ be the groupoid consisting of the family of linear isomorphisms $X_b \xrightarrow{\sim} X_c$, for $b, c \in C$, and let \tilde{G} the subgroupoid of $\text{Iso}X|X - Z$ generated by G and the isomorphisms of monodromy (= the parallel transport along any path in $X - Z$). Then, \tilde{G} defines a subgroupoid of $\text{Lin}\mathcal{F}|X - Z$. This subgroupoid extends to an admissible groupoid if and only if \tilde{G} extends to an analytic subvariety of $\text{Iso}X$.

(c) Call “admissible” such a G ; then there is a smallest admissible G ; one proves that it is the differential Galois group of ∇ , in the “tannakian” sense (for the definition, see [5]).

Roughly speaking, this means that the differential Galois group of ∇ “depends only” on the corresponding foliation and “does not depend” on the further structures of X and ∇ .

(ii) Codimension One

The D -Lie algebras in dimension one, outside the singularities, have been determined by Lie; locally, their spaces of solutions are of dimension 0, 1, 2, 3, ∞ ; in suitable coordinates, these solutions can be written in the following way:

Dimension 0: 0

Dimension 1: $\lambda \frac{d}{dx}$, $\lambda \in \mathbb{C}$ (structure of translation)

Dimension 2: $\lambda \frac{d}{dx} + \mu x \frac{d}{dx}$ (linear affine structure)

Dimension 3: $\lambda \frac{d}{dx} + \mu x \frac{d}{dx} + \nu x^2 \frac{d}{dx}$ (projective structure)

Dimension ∞ : all vector fields; no equation.

A study of these algebras, and the corresponding D -groupoids, near a singularity, will be published by G. Casale. I just mention the following facts: in dimension 0, one can have many groupoids (f.i. actions on X of finite groups); in dimension 1, the correspondence groupoids λ algebras is neither injective nor surjective.

In dimension 2 or 3, one has just one groupoid, e.g. the groupoid of automorphisms of the corresponding Lie algebra (in case of a D -Lie algebra of dimension one, the groupoid of its automorphisms has dimension 2).

Finally, the case “ ∞ ” is trivial: the corresponding groupoid is $\text{Aut}X$.

Now, the foliation, with singularities of codimension one on a manifold

X , can be classified according to the dimension of the transversal Galois groupoid; if X is connected, it is easy to verify that this dimension is independent of the chosen transversal. Here, I will only look at the local situation, e.g. at a foliation in the neighborhood of $0 \in \mathbb{C}^n$; it can be defined by a holomorphic 1-form ω , with $\omega \wedge d\omega = 0$ (we can even suppose that ω has only singularities in codimension 2).

Even in this local case, I have no complete results when $\text{Gal}\mathcal{F}$ has transversal dimension 0 or 1.

(a) Dimension 0. This will be the case if ω has a first integral f , e.g. if there exists a meromorphic f such that $\omega \wedge df = 0$; then, $\text{Gal}\mathcal{F}$ is contained in the groupoid which fixes f . I have no necessary and sufficient condition; the question is related to the quotient by an equivalence relation.

(b) Dimension 1. This will be the case if ω has an integrating factor, e.g. if there exist g meromorphic such that $d(g\omega) = 0$; in that case, $\text{Gal}\mathcal{F}$ is contained in the groupoid which fixes $g\omega$; to prove this, it is sufficient to show that its D -Lie algebra “contains \mathcal{F} ”, e.g. contains the vector fields tangent to the foliation; but, if ξ is such a vector field, one has with standard notations $L_\xi(g\omega) = i_\xi d(g\omega) + d\langle \xi, g\omega \rangle = 0$.

A classical theorem of Lie says the following: if ξ is a meromorphic vector field not in \mathcal{F} , but preserving the foliation, e.g. $L_\xi\omega \wedge \omega = 0$, then $\langle \xi, \omega \rangle^{-1}$ is an integrating factor; in fact, write $\omega = \langle \xi, \omega \rangle \pi$; one has also $L_\xi\pi \wedge \pi = 0$; as $d\langle \xi, \pi \rangle = 0$, this can be written as $(i_\xi d\pi) \wedge \pi = 0$; but, one has $d\pi \wedge \pi = 0$, therefore $0 = i_\xi(d\pi \wedge \pi) = i_\xi d\pi \wedge \pi + d\pi \cdot \langle \xi, \pi \rangle$, and $d\pi = 0$.

Conversely, if g is an integrating factor, choose a ξ such that $\langle \xi, \omega \rangle = g^{-1}$ (ξ is determined mod \mathcal{F}); the same calculation shows that ξ preserves the foliation.

Transversally, this can be interpreted in the following way: ξ defines a vector field $\bar{\xi}$ on the transversals, which is fixed by the holonomy; then $\bar{\xi}^{-1}$ is a 1-form $\bar{\pi}$ on the transversals, and $\pi = \langle \xi, \omega \rangle^{-1}\omega$ is just the inverse image of $\bar{\pi}$ on X .

I will return to this example later (see Remark (iii)).

One can ask for a necessary and sufficient condition for \mathcal{F} to have an admissible groupoid of transversal dimension one; it seems to me likely that such a condition is: there exists a finite ramified covering $\tilde{X} \xrightarrow{\pi} X$ such that $\pi^*\omega$ admits a meromorphic integrating factor.

Of course, if $\text{Gal}\mathcal{F}$ has transversal dimension one, such an admissible groupoid exists. But I do not know if it is always the case when $\text{Gal}\mathcal{F}$ has transversal dimension 0.

(c) Dimensions 2 and 3. The answer is related with Godbillon-Vey sequence. Starting with ω meromorphic near 0, $\omega \wedge d\omega = 0$, one can construct recursively $\omega_1, \dots, \omega_n, \dots$ meromorphic, such that

$$\begin{aligned} d\omega &= \omega \wedge \omega_1, \\ d\omega_1 &= \omega \wedge \omega_2, \\ &\vdots \\ d\omega_n &= \omega \wedge \omega_{n+1} + \sum_1^n \binom{n}{k} \omega_k \wedge \omega_{n-k+1}. \end{aligned}$$

The following trick is due to J. Martinet: put $\Omega = dt + \sum \frac{t^n}{n!} \omega_n$ ($\omega_0 = \omega$); then, one has $d\Omega = \Omega \wedge \frac{\partial \Omega}{\partial t}$.

Given a germ of codimension one foliation, the Godbillon-Vey sequence is not unique: one can replace ω_0 by $g\omega_0$; and, at each step of the recurrence, one can add to ω_n a multiple of ω_0 . We say that “ $G - V$ sequence stops at order i ”, $i = 1, 2, 3$, if we can choose the ω_i ’s in such a way that $\omega_n = 0$ for $n \geq i$. To stop at order 1, it is necessary and sufficient that ω_0 has an integrating factor. To stop at order 2 (resp. 3), we must arrive at ω_1 , with $d\omega_1 = 0$ (resp. ω_2 , with $d\omega_2 = \omega_1 \wedge \omega_2$). In the nonsingular case and C^∞ context the following fact is well-known: the existence of a transversal affine (resp. projective) structure equivalent to the possibility of stopping $G - V$ sequence at order 2 (resp. 3) (cf. [6]).

The same result is true in the present context (this fact was suggested to me by [2]; see also [14], for an interpretation of “case 2”, in an algebraic context, in terms of liouvillian solutions).

Let me give a few words of explanations; the details will be published elsewhere. I will look at the “case 3”; the other one is similar, and simpler.

Let X be a neighborhood of $0 \in \mathbb{C}^n$ on which $\omega, \omega_1, \omega_2$ are meromorphic, and giving a $G - V$ sequence stopping at order 3. On $X \times P_1$, the form $\Omega = dt + \omega_0 + \omega_1 t + \omega_2 \frac{t^2}{2}$ defines a foliation transverse to the fibers $\{x\} \times \mathbb{P}_1$; outside of the singularities, this defines a transverse structure of type 3, which, on $\{x\} \times \mathbb{P}_1$, is simply the standard projective structure; [this is

well known; in fact the equation $\Omega = 0$ comes from the integrable system $dy_1 = \frac{\omega_1}{2}y_1 + \frac{\omega_2}{2}y_2$, $dy_2 = -\omega_0y_1 - \frac{\omega_1}{2}y_2$ by taking $t = \frac{y_2}{y_1}$. One sees that the corresponding groupoid extends to $X \times \mathbb{P}_1$ into an admissible one. Taking the restriction to $t = 0$ gives an admissible groupoid of transversal dimension 3.

Now, the result is: conversely any admissible groupoid of dimension 3 of the foliation defined by ω can be obtained in this way, in particular, there exists a $G - V$ -sequence $\omega_0, \omega_1, \omega_2$ stopping at order 3, with $\omega_0 = f\omega$.

(d) Dimension ∞ . The corresponding groupoid is $\text{Aut}\mathcal{F}$; and there is nothing else to say from the point of view considered here. This case, which is of course the general case, should be studied by other methods (recurrence, attractors, etc), familiars in the theory of dynamical systems.

(iii) Remark: Lie Symmetries and Galois Symmetries

The result of Lie, mentionned in (ii), (b) is sometimes stated in a slightly confusing way, as f.i. "if one has a one parameter group of symmetries of the equation, one can solve it"; this could induce a confusion between symmetries of the data ("Lie symmetries"), and Galois symmetries, e.g. the Galois groupoid.

It seems to me that the precise relations should be stated in the following way. We give on X a foliation with singularities \mathcal{F} ; first of all, call "symmetries of \mathcal{F} " the global automorphisms of \mathcal{F} which preserve \mathcal{F} , e.g. which are solutions of $\text{Aut}\mathcal{F}$. In fact here, we are not interested in them, but in the corresponding D -Lie algebra; call $\text{meraut}\mathcal{F}$ the sheaf of its meromorphic solutions; call similarly $\text{mer}\mathcal{F}$ the germs of meromorphic vector fields tangent to \mathcal{F} ; incidentally, note that $\text{meraut}\mathcal{F}/\text{mer}\mathcal{F}$ is a sheaf of Lie algebras, (which can be interpreted as vector fields on the transversals).

Now, if we have a subset $\Gamma \subset \Gamma(X, \text{meraut}\mathcal{F}/\text{mer}\mathcal{F})$, we get an admissible groupoid G by taking the solutions of $\text{Aut}\mathcal{F}$ which fix Γ . Therefore, the bigger is Γ , the smaller is G . But this is only one procedure among others to have admissible groupoids as small as possible.

Of course, if one has a Lie group acting on X , and preserving \mathcal{F} , its infinitesimal transformations give sections of $\text{meraut}\mathcal{F}$ and a fortiori sections of $\text{meraut}\mathcal{F}/\text{mer}\mathcal{F}$.

§4. Further Examples and Problems

(i) Can one extend the results of §3 (ii) (c) to some cases of higher codimension?

(ii) Let ω be a closed form (f.i. meromorphic) of any degree on X . Then, the condition $i_\xi \omega = 0$ (ξ , vector field) defines a foliation with singularities \mathcal{F} , because one has $L_\xi \omega = di_\xi \omega + i_\xi d\omega = di_\xi \omega = 0$. An admissible groupoid is therefore obtained by taking the φ , solutions of $\text{Aut } \mathcal{F}$ which verify $\varphi^* \omega = \omega$.

Find “interesting” examples where some further reduction, on no further reduction can be obtained.

The most interesting case is probably the case where ω is a 2-form; then according to a classical theorem of Darboux, one has, outside of the singularities, a transversal symplectic structure.

For instance, consider the differential equation $y'' = f(x, y)$; writing this equation as the Pfaff system $dy - y dx, dz - f dx$, put $\omega = (dy - z dx) \wedge (dz - f dx)$; one has $d\omega = 0$, because f does not contain y' ; and the foliation associated to ω is just the foliation corresponding to the equation.

In particular, it is a classical problem to prove that no further reduction occur for the “Painlevé 1” equation $y'' = y^2 + x$; in fact, Painlevé claimed this result with insufficient proof, see [12] (I owe this reference to J. P. Ramis; see also [15]).

The same problem could be considered for the other Painlevé equations, for which Okamoto has constructed transverse symplectic structures [11].

(iii) More generally, Hitchin [8] has given a transverse symplectic structure on the Schlesinger equations for isomonodromic deformations; this result has been extended by Boalch [3] to the irregular case.

Do these structures extend meromorphically to the singularities, e.g. do they give admissible groupoids? If this is the case, which seems to me likely, are there further reductions, or not?

(iv) One can consider other problems of D -envelopes that the problems arising from foliations. For instance

(a) Find the smallest D -groupoid whose given automorphisms of X are solutions. F.i., if f is a germ of automorphism of $(\mathbb{C}, 0)$, and if $f'(0)$ is not a root of unit, f can be embedded in a germ of non-trivial D -groupoid iff f

is linearizable (see the forthcoming paper by G. Casale mentioned above). The case of roots of unit is more complicated; see loc. cit.

(b) Given a vector field ξ , find its D -envelope, e.g. the smallest D -groupoid whose D -Lie algebra contains ξ as a solution. The case of symplectic vector fields on a symplectic manifold is of special interest: the D envelope is contained obviously in the D -groupoid of symplectic transformations fixing ξ ; but further reductions could occur.

It would be, f.i. quite interesting to look at the following case: ξ is the germ at 0, in \mathbb{C}^4 of the symplectic gradient of a Morse function (=case of 2 coupled oscillators).

In the same order of ideas, I mention the following result, which I owe to J. P. Ramis: Let X be a symplectic manifold of dimension $2n$, f a holomorphic function on X , and ξ its symplectic gradient. Suppose that ξ is Liouville-integrable, e.g. suppose that there exist $f = f_1, f_2, \dots, f_n$, with df_1, \dots, df_n generically independent, such that the Poisson brackets $\{f_i, f_j\}$ vanish. Then the D -Lie algebra of the D -envelope of ξ is abelian.

Actually, near a point a where the df_i are independent, we can complete f_1, \dots, f_p with p_1, \dots, p_n to have a system of coordinates such that the symplectic form is $\sum df_i \wedge dp_i$. Then, one proves easily that a local symplectic vector field at a fixing f_1, \dots, f_n is the symplectic gradient of a function $\varphi(f_1, \dots, f_n)$.

This fact is related to the methods used by Ziglin and Moralès-Ramis to prove the non-integrability of some hamiltonian systems. On this subject, a lot of very nice work has been made recently; the reader could consult the survey [1] by M. Audin.

REFERENCES

- [1] Audin, M., Intégrabilité et non-intégrabilité de systèmes hamiltoniens, d'après S. Ziglin, J. Moralès-Ruiz, J.-P. Ramis, Séminaire Bourbaki, **884**(2000/2001).
- [2] Belliard, M., Lioussse, I. & Loray, F., The generic rational differential equation $dw/dz = P(z, w)/Q(z, w)$ carries no interesting transverse structure, *Ergodic Theory and Dyn. Syst.*, **21**(2001), 1599–1607.
- [3] Boalch, P., Symplectic manifolds and isomonodromic deformations, to appear in *Advances in Mathematics*.
- [4] Deligne, P., Equations différentielles à points singuliers réguliers, *Lecture Notes in Math.*, **163**, Springer-Verlag, 1970.
- [5] Deligne, P., Catégories tannakiennes, *The Grothendieck Festschrift*, 2, *Progress in Math.*, *Birkhäuser*, **87**(1990), 111–195.

- [6] Godbillon, C., Feuilletages, Études géométriques, Birkhäuser, 1991.
- [7] Grothendieck, A., Techniques de construction et théorèmes d'existence en géométrie algébrique, III, Séminaire Bourbaki, **212**, (1960/61).
- [8] Hitchin, N. J., Frobenius manifolds (notes by D. Calderbank), in: Gauge theory and symplectic geometry, J. Hurtubise and F. Laborde editors, vol. 488 of NATO ASI series C: Maths & Physics, Kluwer, 1995.
- [9] Malgrange, B., Le groupoïde de Galois d'un feuilletage, to be published in l'Enseignement Mathématique.
- [10] Malgrange, B., La variété caractéristique d'un système différentiel analytique, *Ann. Inst. Fourier (Grenoble)* **50-2**(2000), 491–518.
- [11] Okamoto, K., Polynomial hamiltonians associated with Painlevé equations, I, Proceeding Japan Academy, **56-A-6**(1980), 264–268; II, Proceeding Japan Academy, **56-A-8** (1980), 367–371.
- [12] Painlevé, P., Démonstration de l'irréductibilité absolue de l'équation $y'' = 6y^2 + x$, **135**(1902), 642–647.
- [13] Ritt, J., Differential algebra, Coll. publications **33**, Amer. Math. Soc. Dover, 1966.
- [14] Singer, M., Liouvillian first integral of differential equations, *Trans. A. M. S.*, **333** (1992), 673–688.
- [15] Umemura, H., Differential Galois theory of infinite dimension, *Nagoya Math. J.*, **144**(1996), 59–134.

ITERATIVE ALGORITHMS FOR DATA ASSIMILATION PROBLEMS*

G. MARCHUK V. SHUTYAEV

*Institute of Numerical Mathematics, Russian Academy of
Sciences, ul. Gubkina 8, Moscow, 119991 GSP-1, Russia*

Abstract

Iterative algorithms for solving the data assimilation problems are considered, based on the main and adjoint equations. Spectral properties of the control operators of the problem are studied, the iterative algorithms are justified.

Keywords Data assimilation, Optimal control, Iterative algorithms
2000 MR Subject Classification 65K10

§0. Introduction

The investigation of global changes has increased the interest to the observation data assimilation and data processing problems, which are applied to the modeling, retrospective analysis, and forecasting various physical and geophysical processes. From the mathematical standpoint, these problems may be formulated as the optimal control problems. Starting with the studies of Bellman and Pontryagin, these problems attract the attention of many researchers. New essential ideas were contributed to the optimization theory and methods by French mathematical school. In this connection, we must mention the works by J.-L.Lions and his disciples, which became fundamental, dedicated to investigation of problems on insensitive optimal control, nonlinear sentinels for distributed systems. The general approach

*Project supported by the Russian Foundation for Basic Research (grant 00-01-00611).

(Hilbert Uniqueness Method) developed by J.-L.Lions makes it possible to prove the existence of insensitive controls in linear and nonlinear systems.

In this study, we consider numerical algorithms for the data assimilation problems based on the iterative algorithms using the main and adjoint equations. Properties of the control operators studied are used to justify the iterative algorithms.

§1. Statement of the Problem

Consider mathematical model of a physical process that is described by the evolution problem

$$\begin{cases} \frac{d\varphi}{dt} + A(t)\varphi = f, & t \in (0, T), \\ \varphi|_{t=0} = u, \end{cases} \quad (1.1)$$

where $\varphi = \varphi(t)$ is an unknown function, $A(t)$ is an operator (generally, nonlinear) acting for each t in the Hilbert space X with the definition domain $D(A) \subset X$, $u \in X$, and $f = f(t)$ is a prescribed function.

Introduce the functional

$$J(\varphi) = \frac{1}{2} \int_0^T (C(\varphi - \hat{\varphi}), \varphi - \hat{\varphi})_X dt + \frac{\alpha}{2} (\varphi|_{t=0} - \hat{\varphi}^\circ, \varphi|_{t=0} - \hat{\varphi}^\circ)_X, \quad (1.2)$$

where $\alpha = \text{const.} \geq 0$, C is a linear operator, and $(\cdot, \cdot)_X$ is an inner product in X . The function $\hat{\varphi} = \hat{\varphi}(t)$, as a rule, is determined by a priori observation data, $\hat{\varphi}^\circ \in X$. We assume hereinafter that all spaces and functions under consideration are real.

Consider problem (1.1) with an unknown function $u \in X$ in the initial condition. The data assimilation problem can be formulated as follows: find φ and u such that they satisfy (1.1) and, on the set of solutions to equation (1.1), functional (1.2) takes the minimum value. Write this problem as

$$\begin{cases} \frac{d\varphi}{dt} + A(t)\varphi = f, & t \in (0, T), \\ \varphi(0) = u, \\ J(\varphi) = \inf_{\tilde{u} \in H} J(\tilde{\varphi}), \end{cases} \quad (1.3)$$

where $\tilde{\varphi}$ is a solution of (1.1) when $\tilde{\varphi}(0) = \tilde{u}$.

Problems in the form (1.3) were analyzed by Pontryagin [2], Lions [3] (see also [5–14], etc.). To solve (1.3) a number of approaches may be used (see e.g. [14]). We will consider iterative algorithms for solving (1.3), assuming for simplicity that $A(t)$ is a linear operator.

The necessary optimality condition [3] reduces problem (1.3) to the system for finding the functions φ , φ^* , u :

$$\frac{d\varphi}{dt} + A(t)\varphi = f, \quad t \in (0, T); \quad \varphi(0) = u, \quad (1.4)$$

$$-\frac{d\varphi^*}{dt} + A(t)^*\varphi^* = C(\hat{\varphi} - \varphi), \quad t \in (0, T); \quad \varphi^*(T) = 0, \quad (1.5)$$

$$\alpha(u - \hat{\varphi}^0) - \varphi^*(0) = 0, \quad (1.6)$$

where $A(t)^*$ is the operator adjoint to $A(t)$.

§2. Control Operator and Its Properties

Let $Y = L_2(0, T; X)$ be a space of abstract functions $u(t)$ with values in X , with the inner product and the norm

$$(u, v) = \int_0^T (u, v)_X dt, \quad \|u\| = \left(\int_0^T \|u\|_X^2 dt \right)^{1/2}, \quad u, v \in Y.$$

In the forthcoming, we suppose that the original model satisfies the following conditions:

(i) the solution to the problem

$$\begin{cases} \frac{d\psi}{dt} + A(t)\psi = f, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (2.1)$$

meets the inequality

$$\|\psi\|_Y \leq c_1 (\|f\|_Y + \|v\|_X), \quad c_1 = \text{const.} > 0; \quad (2.2)$$

(ii) the solution of the adjoint problem

$$\begin{cases} -\frac{d\psi^*}{dt} + A^*(t)\psi^* = p, & t \in (0, T), \\ \psi^*|_{t=T} = 0, \end{cases} \quad (2.3)$$

satisfies

$$\|\psi^*\|_Y + \|\psi^*|_{t=0}\|_X \leq c_1^* \|p\|_Y, \quad c_1^* = \text{const.} > 0. \quad (2.4)$$

Remark 2.1. The solutions of the problems (2.1) and (2.3) are supposed to exist such that $\psi, \psi^* \in Y$, treated in a classical or a weak sense. The conditions (i), (ii) are satisfied if, for example, the operator $A(t)$ is positive definite:

$$(A(t)w, w)_Y \geq \gamma \|w\|_Y^2, \quad \gamma = \text{const.} > 0, \quad \forall w \in Y.$$

Indeed, from (2.1) we get

$$\left(\frac{d\psi}{dt}, \psi \right)_X + (A(t)\psi, \psi)_X = (f, \psi)_X,$$

whence

$$\frac{1}{2} \int_0^T \frac{d}{dt} (\psi, \psi)_X dt + \int_0^T (A(t)\psi, \psi)_X dt = \int_0^T (f, \psi)_X dt,$$

and by virtue of positive definiteness of $A(t)$,

$$\frac{1}{2} \|\psi|_{t=T}\|_X^2 + \gamma \|\psi\|_Y^2 \leq (f, \psi)_Y + \frac{1}{2} \|\psi|_{t=0}\|_X^2 \leq \|f\|_Y \|\psi\|_Y + \frac{1}{2} \|v\|_X^2,$$

or

$$\gamma \|\psi\|_Y^2 \leq \frac{1}{2\gamma} \|f\|_Y^2 + \frac{\gamma}{2} \|\psi\|_Y^2 + \frac{1}{2} \|v\|_X^2.$$

The last inequality gives (2.2) with $c_1 = \max(\gamma^{-1}, \gamma^{-1/2})$. Similarly, the inequality (2.4) is obtained. In the finite-dimensional case, when $X = \mathbf{R}^n$, $n \in \mathbf{N}$, the inequalities (2.2), (2.4) are valid without positive definiteness requirement if the $n \times n$ -matrix $A(t)$ is regular enough (for instance, having the elements continuous in t).

Let us introduce the operator $L : X \rightarrow X$ defined through the successive solutions of the following problems:

$$\begin{cases} \frac{d\psi}{dt} + A(t)\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (2.5)$$

$$\begin{cases} -\frac{d\psi^*}{dt} + A^*(t)\psi^* = -C\psi, & t \in (0, T), \\ \psi^*|_{t=T} = 0, \end{cases} \quad (2.6)$$

$$Lv = \alpha v - \psi^*(0). \quad (2.7)$$

We define also $F \in X$ as the successive solutions of the following problems:

$$\begin{cases} \frac{d\phi}{dt} + A(t)\phi = f, & t \in (0, T), \\ \phi|_{t=0} = 0, \end{cases} \quad (2.8)$$

$$\begin{cases} -\frac{d\phi^*}{dt} + A^*(t)\phi^* = C(\hat{\varphi} - \phi), & t \in (0, T), \\ \phi^*|_{t=T} = 0, \end{cases} \quad (2.9)$$

$$F = \alpha \hat{\varphi}^\circ + \phi^*(0), \quad (2.10)$$

where $f, \hat{\varphi} \in Y$, $\hat{\varphi}^\circ \in X$ are introduced in (1.4)–(1.6). We suppose that $C : Y \rightarrow Y$ is a linear bounded self-adjoint positive semi-definite operator.

Then, the system (1.4)–(1.6) is reduced to the equation for the control

u :

$$Lu = F, \quad (2.11)$$

and the operator $L : X \rightarrow X$ is called the control operator [13].

Under the hypotheses (i), (ii) the following statement is valid.

Lemma 2.1. *The operator L acts in X with domain of definition $D(L) = X$, it is bounded, self-adjoint, and positive semi-definite. If $\alpha > 0$, the operator L is positive definite.*

Proof. Let $v \in X$, and ψ be the solution to (2.5). By (2.2), $\|\psi\|_Y \leq c_1 \|v\|_X$. For the solution ψ^* of (2.6) the inequality (2.4) holds:

$$\|\psi^*\|_Y + \|\psi^*\big|_{t=0}\|_X \leq c_1^* \|C\psi\|_Y.$$

Hence, from (2.7),

$$\begin{aligned} \|Lv\|_X &= \|\alpha v - \psi^*\big|_{t=0}\|_X \leq \alpha \|v\|_X + \|\psi^*\big|_{t=0}\|_X \\ &\leq \alpha \|v\|_X + c_1^* \|C\psi\|_Y \leq \alpha \|v\|_X + c_1^* c_1 \|C\| \|v\|_X, \end{aligned}$$

and, therefore, L is bounded. Further, we have for $v, w \in X$,

$$\begin{aligned} (Lv, w)_X &= (\alpha v - \psi^*\big|_{t=0}, w)_X = \alpha (v, w)_X - (\psi^*\big|_{t=0}, w)_X \\ &= \alpha (v, w)_X + (C\psi, \psi_1)_Y = \alpha (v, w)_X + (\psi, C\psi_1)_Y = (v, Lw)_X, \end{aligned}$$

where ψ_1 is the solution to (2.5) with $v = w$. Hence, L is self-adjoint, and

$$(Lv, v)_X = \alpha (v, v)_X + (C\psi, \psi)_Y \geq 0,$$

that is, L is positive semi-definite. Moreover, L is positive definite if $\alpha > 0$.

Corollary 2.1. *The following estimate is valid:*

$$(Lv, v)_X \geq \mu_{\min}(v, v)_X, \quad \forall v \in X, \quad (2.12)$$

where μ_{\min} is the lower spectrum bound of the operator L , and $\mu_{\min} \geq \alpha$.

The following solvability result holds.

Lemma 2.2. *Under the hypotheses (i), (ii), for $\alpha > 0$, the control equation (2.11) has a unique solution $u \in X$, and*

$$\|u\|_X \leq \frac{\alpha}{\mu_{\min}} \|\hat{\varphi}^\circ\|_X + \frac{c_1^*}{\mu_{\min}} \|C\hat{\varphi}\|_Y + \frac{c_1 c_1^*}{\mu_{\min}} \|C\| \|f\|_Y. \quad (2.13)$$

Proof. If $\alpha > 0$, from Corollary 2.1, there exist a unique solution u of the control equation (2.11), and

$$\|u\|_X \leq \frac{1}{\mu_{\min}} \|F\|_X. \quad (2.14)$$

The solution ϕ^* of (2.9) satisfies the inequality (2.4), and

$$\|F\|_X = \alpha \|\hat{\varphi}^\circ\|_X + \|\phi^*(0)\|_X \leq \alpha \|\hat{\varphi}^\circ\|_X + c_1^* \|C(\hat{\varphi} - \phi)\|_Y,$$

where ϕ is the solution to (2.8). Due to (2.2), $\|\phi\|_Y \leq c_1 \|f\|_Y$, then

$$\|F\|_X \leq \alpha \|\hat{\varphi}^\circ\|_X + c_1^* \|C\hat{\varphi}\|_Y + c_1^* c_1 \|C\| \|f\|_Y. \quad (2.15)$$

From (2.14)–(2.15) we obtain (2.13). This ends the proof.

Remark 2.2. For $\alpha = 0$, the last lemma holds true also if $\mu_{\min} > 0$. It is true, for instance, in the case that $X = \mathbf{R}^n$, $n \in \mathbf{N}$, $C = E$ (the identity operator). The weight coefficient α is usually called a regularization parameter [4].

§3. Spectrum Bounds of the Control Operator

In the general case, from Corollary 2.1, for the lower spectrum bound of the operator L we have $\mu_{\min} \geq \alpha$. Moreover, from the proof of Lemma 2.1, we get

$$(Lv, v)_X = \alpha(v, v)_X + (C\psi, \psi)_Y, \quad v \in X,$$

where ψ is the solution of (2.5). Hence, due to (2.2),

$$(Lv, v)_X \leq \alpha(v, v)_X + \|C\| \|\psi\|_Y^2 \leq \alpha \|v\|_X^2 + c_1 \|C\| \|v\|_X^2,$$

and for the upper spectrum bound μ_{\max} of the operator L we get

$$\mu_{\max} \leq \alpha + c_1 \|C\|. \quad (3.1)$$

In the case that $C = E$ (the identity operator), sharper estimates may be derived. The following result is valid.

Theorem 3.1. *The spectrum $\sigma(L)$ of the operator L defined by (2.5)–(2.7) for $C = E$ satisfies the estimates*

$$m \leq \sigma(L) \leq M, \quad (3.2)$$

where

$$m = \alpha + \int_0^T e^{-\int_0^t \lambda_{\max}(\tau) d\tau} dt, \quad M = \alpha + \int_0^T e^{-\int_0^t \lambda_{\min}(\tau) d\tau} dt,$$

and λ_{\min} , λ_{\max} are the lower and the upper spectrum bounds of the operator $A + A^*$, respectively.

Proof. For the operator L defined by (2.5)–(2.7) for $C = E$ the following representation is valid:

$$(Lu, u) = \alpha(u, u) + \int_0^T (\varphi(t), \varphi(t)) dt, \quad u \in X, \quad (3.3)$$

where $\varphi(t)$ is the solution to (2.5) for $v = u$. From (2.5),

$$\frac{d}{dt} \|\varphi\|^2 + ((A + A^*)\varphi, \varphi) = 0,$$

then

$$-\lambda_{\max}(t) \|\varphi\|^2 \leq \frac{d}{dt} \|\varphi\|^2 \leq -\lambda_{\min}(t) \|\varphi\|^2,$$

where λ_{\max} and λ_{\min} are the lower and the upper spectrum bounds of the operator $A + A^*$, respectively. Therefore, the function $F(t) = \ln \|\varphi\|^2$ meets the inequality

$$-\lambda_{\max}(t) \leq \frac{dF}{dt} \leq \lambda_{\min}(t).$$

By integrating this inequality with respect to t from 0 to t , we get

$$-\int_0^t \lambda_{\max}(\tau) d\tau \leq F(t) - F(0) \leq -\int_0^t \lambda_{\min}(\tau) d\tau,$$

or

$$-\int_0^t \lambda_{\max}(\tau) d\tau \leq \ln \frac{\|\varphi(t)\|^2}{\|u\|^2} \leq -\int_0^t \lambda_{\min}(\tau) d\tau.$$

Hence

$$e^{-\int_0^t \lambda_{\max}(\tau) d\tau} \leq \frac{\|\varphi(t)\|^2}{\|u\|^2} \leq e^{-\int_0^t \lambda_{\min}(\tau) d\tau}.$$

Integrating the last inequality with respect to t from 0 to T and taking into account (3.3), we obtain

$$\int_0^T e^{-\int_0^t \lambda_{\max}(\tau) d\tau} dt \leq \frac{(\bar{L}u, u)}{(u, u)} \leq \int_0^T e^{-\int_0^t \lambda_{\min}(\tau) d\tau} dt,$$

where \bar{L} is the operator L for $\alpha = 0$. Thus, the spectrum bounds of the operator L are defined by (3.2). This ends the proof.

If $A(t) = A : X \rightarrow X$ is a linear closed operator independent of time and being unbounded self-adjoint positive definite operator in X with the compact inverse, then the eigenvalues μ_k of the operator \bar{L} are defined by the formula [13]

$$\mu_k = \frac{1 - e^{-2\lambda_k T}}{2\lambda_k},$$

where λ_k are the eigenvalues of the operator A . Then in (3.2) $\lambda_{\min} = 2\lambda_1$, $\lambda_{\max} = \infty$, and m, M are given in the explicit form

$$m = \alpha, \quad M = \alpha + \frac{1 - e^{-2\lambda_1 T}}{2\lambda_1}, \quad (3.4)$$

where λ_1 is the least eigenvalue of the operator A . By this is meant that the estimates (3.2) are exact.

§4. Iterative Algorithms

To solve (1.4)–(1.6) we consider a class of iterative algorithms:

$$\frac{d\varphi^k}{dt} + A(t)\varphi^k = f, \quad t \in (0, T); \quad \varphi^k(0) = u^k, \quad (4.1)$$

$$-\frac{d\varphi^{*k}}{dt} + A^*(t)\varphi^{*k} = C(\hat{\varphi} - \varphi^k), \quad t \in (0, T); \quad \varphi^{*k}(T) = 0, \quad (4.2)$$

$$u^{k+1} = u^k - \alpha_{k+1}B_k(\alpha(u^k - \hat{\varphi}^\circ)\varphi^{*k}|_{t=0}) + \beta_{k+1}C_k(u^k - u^{k-1}), \quad (4.3)$$

where $B_k, C_k : H \rightarrow H$ are some operators, and $\alpha_{k+1}, \beta_{k+1}$ the iterative parameters. Let m and M be the spectral bounds of the control operator L defined by (3.2). We introduce the following notations:

$$\tau_{opt} = 2(M + m)^{-1}, \quad \theta = (M + m)(M - m)^{-1}, \quad (4.4)$$

$$\tau_k = 2(M + m - (M - m)\cos\omega_k\pi)^{-1}, \quad k = 1, 2, \dots, s, \quad (4.5)$$

$$\alpha_{k+1} = \begin{cases} 2(M + m)^{-1}, & k = 0, \\ 4(M - m)^{-1} \frac{T_k(\theta)}{T_{k+1}(\theta)}, & k > 0, \end{cases} \quad (4.6)$$

$$\beta_{k+1} = \begin{cases} 0, & k = 0, \\ \frac{T_{k-1}(\theta)}{T_{k+1}(\theta)}, & k > 0, \end{cases} \quad (4.6)$$

$$e_k = \begin{cases} 0, & k = 0, \\ p_k \|\xi^k\|_H^2 / \|\xi^{k-1}\|_H^2, & k > 0, \end{cases} \quad (4.7)$$

$$p_{k+1} = \alpha + (\eta^k, \eta^k) / \|\xi^k\|_H^2 - e_k, \quad k = 0, 1, \dots, \quad (4.8)$$

where $\omega_k = (2i - 1)/2s$, T_k is the k -th degree Chebyshev polynomial of the first kind,

$$\xi^k = \alpha u^k - \varphi^{*k}(0),$$

and η^k is the solution of the problem

$$\frac{d\eta^k}{dt} + A\eta^k = 0, \quad t \in (0, T); \quad \eta^k(0) = \xi^k.$$

Theorem 4.1. (i) *If $\alpha_{k+1} = \tau$, $B_k = E$, $\beta_{k+1} = 0$, then the condition $0 < \tau < 2/(M + m)$ is a sufficient condition for the convergence of the*

iterative process (4.1)–(4.3). For $\tau = \tau_{opt}$ defined by (4.4) the following convergence rate estimates are valid:

$$\|\varphi - \varphi^k\|_W \leq c_1 q_k, \quad \|\varphi^* - \varphi^{*k}\|_W \leq c_2 q_k, \quad \|u - u^k\|_H \leq c_3 q_k, \quad (4.9)$$

where $q_k = 1/\theta^k$, θ is given by (4.4), and the constants c_1, c_2, c_3, c_4 do not depend on the number of iterations and on the functions $\varphi, \varphi^k, \varphi^*, \varphi^{*k}, u, u^k, k > 0$.

(ii) If $B_k = E$, $\beta_{k+1} = 0$, and $\alpha_{k+1} = \tau_k$, where the parameters τ_k are defined by (4.5) and repeated cyclically with the period s , then the error in the iterative process (4.1)–(4.3) is suppressed after each cycle of the length s . After $k = ls$ iterations the error estimates (4.9) are valid with $q_k = (T_s(\theta))^{-l}$.

(iii) If $B_k = C_k = E$ and $\alpha_{k+1}, \beta_{k+1}$ are defined by (4.6), then the error in the algorithm (4.1)–(4.3) is suppressed for each $k \geq 1$, and the estimates (4.9) hold for $q_k = (T_k(\theta))^{-1}$.

(iv) If $B_k = C_k = E$ and $\alpha_{k+1} = 1/p_{k+1}$, $\beta_{k+1} = e_k/p_{k+1}$, where e_k, p_{k+1} are defined by (4.7), (4.8), then the iterative process (4.1)–(4.3) is convergent, and the convergence rate estimates (4.9) are valid with $q_k = (T_k(\theta))^{-1}$.

Proof. It is not difficult to show that the iterative process (4.1)–(4.3) is equivalent to the following iterative algorithm

$$u^{k+1} = u^k - \alpha_{k+1} B_k (L u^k - F) + \beta_{k+1} C_k (u^k - u^{k-1}) \quad (4.10)$$

for solving the control equation (2.11) with the right-hand side F defined by (2.8)–(2.10).

The bounds m and M of the spectrum of the control operator L are given by (3.2). Thus, for $\alpha > 0$ for solving the equation $Lu = F$ we may use the well-known iterative algorithms with optimal choice of parameters. The theory of these methods is well developed [15]. Taking into account the explicit form of the bounds for m and M and applying for the equation $Lu = F$ the simple iterative method, the Chebyshev acceleration methods (s -cyclic and two-step ones), and the conjugate gradient method in the form (4.10), we arrive at the conclusions of the Theorem, using the well-known convergence results [15] for these methods.

Remark 4.1. In case $\alpha_k = 1/\alpha$, $B_k = E$, $\beta_k = 0$, the iterative algorithm (4.1)–(4.3) coincides with the Krylov-Chernousko method [16].

The numerical analysis of the above-formulated algorithms has been done in [17] for the data assimilation problem with a linear parabolic state equation. The numerical experiments are in good agreement with theoretical results on the convergence of the iterative algorithms.

REFERENCES

- [1] Bellman, R., *Dynamic programming*, Princeton Univ. Press, New Jersey, 1957.
- [2] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V. & Mischenko, E. F., *The mathematical theory of optimal processes*, John Wiley, New York, 1962.
- [3] Lions, J. L., *Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles*, Dunod, Paris, 1968.
- [4] Tikhonov, A. N., On the solution of ill-posed problems and the regularization method, *Dokl. Akad. Nauk SSSR*, **151**:3(1963), 501–504.
- [5] Le Dimet, F. X. & Talagrand, O., Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus*, **38A**(1986), 97–110.
- [6] Navon, I. M., A review of variational and optimization methods in meteorology [A], In: *Variational methods in geosciences* [C], (Ed. Sasaki Y.K.) – New York, Elsevier, 1986, 29–34.
- [7] Glowinski, R. & Lions, J. L., Exact and approximate controllability for distributed parameter systems, *Acta Numerica*, **1**(1994), 269–378.
- [8] Lions, J. L., On controllability of distributed system, *Proc. Natl. Acad. Sci. USA*, **94**(1997), 4828–4835.
- [9] Agoshkov, V. I. & Marchuk, G. I., On solvability and numerical solution of data assimilation problems, *Russ. J. Numer. Anal. Math. Modelling*, **8**(1993), 1–16.
- [10] Marchuk, G. I. & Zalesny, V. B., A numerical technique for geophysical data assimilation problem using Pontryagin's principle and splitting-up method, *Russ. J. Numer. Anal. Math. Modelling*, **8**(1993), 311–326.
- [11] Marchuk, G. I. & Shutyaev, V. B., Iteration methods for solving a data assimilation problem, *Russ. J. Numer. Anal. Math. Modelling*, **9**(1994), 265–279.
- [12] Marchuk, G. I., Agoshkov, V. I. & Shutyaev, V. P., *Adjoint equations and perturbation algorithms in nonlinear problems*, New York, CRC Press Inc., 1996.
- [13] Shutyaev, V. P., Some properties of the control operator in the problem of data assimilation and iterative algorithms, *Russ. J. Numer. Anal. Math. Modelling*, **10**:4(1995), 357–371.
- [14] Marchuk, G., Shutyaev, V. & Zalesny, V., Approaches to the solution of data assimilation problems [A], In: *Optimal control and partial differential equations*, Amsterdam: IOS Press, 2001, 489–497.
- [15] Marchuk, G. I. & Lebedev, V. I., *Numerical methods in the theory of neutron transport*, Harwood Academic Publishers, New York, 1986.
- [16] Krylov, I. A. & Chernousko, F. L., On a successive approximation method for solving optimal control problems, *Zh. Vychisl. Mat. Mat. Fiz.*, **2**:6(1962), 1132–1139.
- [17] Parmuzin, E. I. & Shutyaev, V. P., Numerical analysis of iterative methods for solving evolution data assimilation problems, *Russ. J. Numer. Anal. Math. Modelling*, **14**(1999), 275–289.

QUADRILATERAL MESH

MING PINGBING

*A1 South 4 street, Zhong Guan Cun, Institute of Computational
Mathematics, Chinese Academy of Sciences, Po Box
2719, Beijing 100080, China. E-mail: mpb@lsec.cc.ac.cn*

SHI ZHONGCI

*A1 South 4 street, Zhong Guan Cun, Institute of Computational
Mathematics, Chinese Academy of Sciences, Po Box
2719, Beijing 100080, China. E-mail: shi@lsec.cc.ac.cn*

Abstract

Several quadrilateral shape regular mesh conditions commonly used in the finite element method are proven to be equivalent. Their influence on the finite element interpolation error and the consistency error committed by nonconforming finite elements are investigated. The effect of the Bi-Section Condition and its extended version $(1+\alpha)$ -Section Condition on the degenerate mesh conditions is also checked. The necessity of the Bi-Section Condition in finite elements is underpinned by means of counterexamples.

Keywords Quadrilateral mesh, 4-node isoparametric element,
Nonconforming quadrilateral element

2000 MR Subject Classification 65N30, 65D05, 41A10, 41A27,
41A25, 41A63

§1. Introduction

Quadrilateral mesh is widely used in the finite element method due to its simplicity and flexibility. However, numerical accuracy cannot be achieved over an arbitrary mesh, so one has to impose certain mesh conditions. There exist several mesh conditions in the literature, which can be classified into two groups. One is the shape regular mesh condition and the other is

the degenerate condition. Roughly speaking, a shape regular condition requires that the element cannot be too narrow on the one hand ((C-R)₁ hereinafter) and the interior angle of each vertex is neither too small nor too close to π on the other hand ((C-R)₂ hereinafter). The first condition of this type belongs to Ciarlet-Raviart (C-R) [27, 28]. Another two are attributed to Girault-Raviart (G-R) [33] and Arunakirinathar-Reddy (A-R) [11], all these three conditions aim for the optimal interpolation error for the isoparametric element, whereas a similar condition of Z. Zhang (Z) [68] appeared in the study of the Wilson nonconforming element. It may be interesting to ask whether such conditions are equivalent. One of the main results of this paper is a strict proof for the equivalence of C-R, G-R, A-R and Z (see Theorem 3.1 below).

Meanwhile, there are several degenerate mesh conditions, which violate either (C-R)₁ or (C-R)₂, and sometimes even both of them. Such degenerate meshes are particularly effective for the finite element approximation of some physical problems with singularities (see [5]). These conditions are scattered in the literature, most of them are *mutatis mutandis*. We only consider two degenerate conditions, namely the Jamet condition (J) [34] and Acosta-Durans [1] Regular Decomposition Property (RDP). We will clarify their connection to the shape regular mesh condition. In particular, we show by means of a counterexample that RDP is necessary for obtaining the optimal interpolation error in the H^1 -norm for the 4-node isoparametric element, thereby we solve the open problem proposed in [1].

On the other hand, the Bi-Section Condition or its extended version $(1 + \alpha)$ -Section Condition are two mesh conditions which quantify the deviation of an arbitrary quadrilateral from a parallelogram. These two mesh conditions are used to estimate the consistency error of the Wilson nonconforming element [54] as well as the interpolation and consistency error of the nonconforming quadrilateral rotated Q_1 element (NRQ₁) (see [40,42,46]), and the interpolation error of the lowest-order Raviart-Thomas element (RT_[0]) (see [61,43]). Süli [58] proved that the Jamet degenerate mesh condition plus the Bi-Section Condition actually imply the shape regular conditions when the mesh diameter approaches zero. We show that the requirement of the Bi-Section Condition in Süli's result can be replaced by the even weaker $(1 + \alpha)$ -Section Condition.

The necessity of the Bi-Section Condition for the optimal consistency error of the Wilson element is illustrated by Z. Shi [54] with a counterexample. Likewise, P. Ming [40] showed that this is also true for the NRQ_1^p (see §6 for the definition). We will propose a series of counterexamples to demonstrate that the Bi-Section Condition is also necessary for the optimal interpolation error of NRQ_1 and $\text{RT}_{[0]}$. All these facts substantiate the necessity of the Bi-Section Condition in the finite element analysis.

We mainly focus on the 4-node isoparametric element, two low-order quadrilateral nonconforming elements, i.e., the Wilson and NRQ_1 elements, and the quadrilateral $\text{RT}_{[0]}$ element. Moreover, only 2-D mesh is taken into account (see [67], for some other isoparametric elements and [31,53] for 3-D). We only consider the finite element approximation for the coercive elliptic problem, the non-coercive problem is more involved and will be addressed elsewhere (see also [1,10] for related references.)

The remaining part of this paper is organized as follows. In §2, we state all shape regular mesh conditions mentioned above. Their equivalence is proven in §3. In §4, several degenerate mesh conditions are reviewed and their connections to the shape regular mesh condition are elucidated with the aid of the Bi-Section Condition and $(1 + \alpha)$ -Section Condition. The influence of different mesh conditions on the interpolation error and the finite element error for the 4-node isoparametric element, the nonconforming Wilson element and NRQ_1 element, and the $\text{RT}_{[0]}$ element are included in §5, §6 and §7, respectively. Some conclusions are drawn and three open problems are proposed in the last section.

Throughout this paper, the generic constant C is independent of the element geometry unless otherwise stated.

§2. Shape Regular Mesh Condition

Before introducing mesh conditions, we fix some notations. For any convex polygon Ω with Lipschitz boundary, $H^k(\Omega)$ is defined as the standard Sobolev space [2] equipped with the norm $\|\cdot\|_k$, and semi-norm $|\cdot|_{k,\Omega}$, $H_0^k(\Omega)$ is the corresponding homogeneous space. Ω will be dropped if no confusion can occur. $\bar{f}_\Omega = \frac{1}{|\Omega|} \int_\Omega u \, dx$ is defined as the integral average of u on Ω . For any vectors $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, $\mathbf{x} \otimes \mathbf{y}$ is a 2×2 matrix with

elements $(\mathbf{x} \otimes \mathbf{y})_{ij} = x_i y_j$. For any matrix $A, \|A\|$ denotes its Euclidean norm.

2.1. Geometric Facts of Quadrilateral Mesh

Let \mathcal{T}_h be a partition of $\bar{\Omega}$ by convex quadrilaterals with the mesh size $h := \max_{K \in \mathcal{T}_h} h_K$. We define h_K and \underline{h}_K as the longest and shortest edges of K , respectively. ρ_K is defined as the diameter of the largest circle inscribed in K . As in Fig.1, we denote the four vertices of K by P_i with the coordinates \mathbf{x}_i . Let their edges be $P_i P_{i+1}$ and $|P_i P_{i+1}|$ their corresponding lengths. The subtriangle of K with vertices P_{i-1}, P_i and P_{i+1} is denoted by \mathcal{T}_i (i modulo 4), i.e., $\mathcal{T}_i := \langle P_{i-1}, P_i, P_{i+1} \rangle$. Similar to K , h_i and ρ_i are defined as the longest edges of \mathcal{T}_i and the diameter for the largest circle inscribed in \mathcal{T}_i , respectively. Denote the interior angle of the vertex P_i by θ_i , and $\mu_K := \max_{1 \leq i \leq 4} |\cos \theta_i|$. Moreover, d_K is denoted as the distance between midpoints of two diagonals of K .

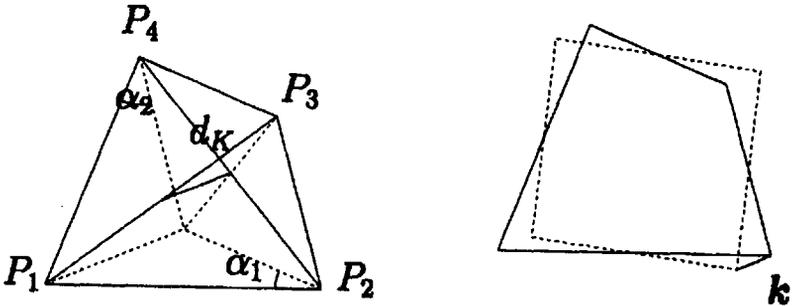


Fig.1

We define by \mathcal{P}_k the space of polynomials of degree no more than k , and by \mathcal{Q}_k the space of degree no more than k in each variable.

Let $\hat{K} = [-1, 1]^2$ be the reference square having the vertex \hat{P}_i with the coordinates $\hat{\mathbf{x}}_i (1 \leq i \leq 4)$, then there exists a unique mapping $\mathcal{F}_K(\xi, \eta) \in \mathcal{Q}_1(\hat{K})$ such that $\mathcal{F}_K(\hat{\mathbf{x}}_i) = \mathbf{x}_i, 1 \leq i \leq 4$. The Jacobian of $\mathcal{F}_K(\xi, \eta)$ is denoted by $D\mathcal{F}_K(\xi, \eta)$ which can be split as $D\mathcal{F}_K(\xi, \eta) = D\mathcal{F}_K(0, 0) + \mathbf{k} \otimes \mathbf{l}$, where $\mathbf{k} := (\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4)^T / 4$ and $\mathbf{l} := (\eta, \xi)^T$. The determinant of $D\mathcal{F}_K(\xi, \eta)$ is $J_K(\xi, \eta) = \det D\mathcal{F}_K(\xi, \eta) = J_0 + J_1 \xi + J_2 \eta$. It can be shown

that

$$\max_{(\xi, \eta) \in \hat{K}} J_K(\xi, \eta) = \max_{1 \leq i \leq 4} |\mathcal{T}_i|/2, \quad \min_{(\xi, \eta) \in \hat{K}} J_K(\xi, \eta) = \min_{1 \leq i \leq 4} |\mathcal{T}_i|/2$$

and

$$J_0 = J_K(0, 0) = |K|/4, \quad J_K(\xi, \eta) > 0.$$

2.2. Shape Regular Mesh Conditions

We mainly concern the following shape regular mesh conditions.

(1) Ciarlet-Raviart (C-R) (see [27,28]).

\mathcal{T}_h is regular [26, p. 247], if there exist two constants $\sigma_0 \geq 1$ and $0 < \mu < 1$ such that

$$h_K/\underline{h}_K \leq \sigma_0, \quad 0 < \mu_K \leq \mu < 1. \tag{2.1}$$

(2) Girault-Raviart (G-R) (see [17,33]).

Define $\bar{\rho}_K := 2 \min_{1 \leq i \leq 4} \rho_i$. \mathcal{T}_h is regular if there exists $\sigma > 0$ such that

$$\max_{K \in \mathcal{T}_h} h_K/\bar{\rho}_K \leq \sigma. \tag{2.2}$$

(3) Arunakirinathar-Reddy (A-R) (see [11]).

\mathcal{T}_h is regular if there exist four constants C, C_0, C_1 and $\gamma > -1$ such that

$$\|D\mathcal{F}_K(0, 0)\| \leq Ch_K, \quad \|D\mathcal{F}_K^{-1}(0, 0)\| \leq Ch_K^{-1}, \quad C_0 h_K^2 \leq J_K(0, 0) \leq C_1 h_K^2. \tag{2.3}$$

$$(D\mathcal{F}_K^{-1}(0, 0)\mathbf{k}, \mathbf{l}) \geq \gamma. \tag{2.4}$$

(4) Z. Zhang (Z) (see [68]).

Define $\tilde{\rho}_K := \min_{1 \leq i \leq 4} \rho_i$. \mathcal{T}_h is regular if there exists $\sigma > 0$ such that

$$\max_{K \in \mathcal{T}_h} h_K/\tilde{\rho}_K \leq \sigma. \tag{2.5}$$

Remark 2.1. A simple manipulation yields that C_1 in (2.3) can equal $1/8$, which is even sharp.

§3. Equivalence of Shape Regular Conditions

In this section, we will prove that all the aforementioned shape regular conditions are equivalent. Before the presentation, we state two lemmas for later use.

Lemma 3.1. [70] *For a family of triangular finite elements, if the ratio between the longest edge of the triangle and the radius of the biggest circle inscribed into the triangle is uniformly bounded, then there exists a constant θ_0 such that all interior angle θ_K of triangles satisfies*

$$\theta_K \geq \theta_0 > 0. \quad (3.1)$$

Lemma 3.2. *For any triangle K , the diameter of the biggest circle inscribed into K is less than the shortest edge of K .*

Theorem 3.1. *All the forgoing mentioned shape regular mesh conditions (C-R, G-R, A-R and Z) are equivalent.*

Proof. We only need to prove that all shape regular mesh conditions are equivalent to C-R. The equivalence between G-R and Z is obvious. So what we need to prove is the equivalence of C-R, G-R and A-R.

Firstly we prove that G-R implies C-R.

Given the regular condition G-R, in view of Lemma 3.2, we have

$$h_K/\underline{h}_K \leq 2h_K/\bar{\rho}_K \leq 2\sigma. \quad (3.2)$$

which gives (C-R)₁ with $\sigma_0 = 2\sigma$.

Notice that $h_i/rho_i \leq 2h_K/\bar{\rho}_K \leq 2\sigma$, invoking Lemma 3.1, we see that each interior angle of K is bounded below by θ_0 , which in turn implies an upper bound for each angle with $\pi - 2\theta_0$, so $\theta_0 \leq \theta_i \leq \pi - 2\theta_0$, thus we come to (C-R)₂ with $\mu = \max(\cos \theta_0, |\cos 2\theta_0|)$. Therefore, C-R follows from G-R.

We are in a position to prove that C-R implies G-R. For any triangle \mathcal{T}_i , we have

$$\rho_i = \frac{2|\mathcal{T}_i|}{|P_{i-1}P_i| + |P_iP_{i+1}| + |P_{i+1}P_{i-1}|} \geq \frac{\underline{h}_K \sin \theta_i}{3h_i} \geq \frac{(1 - \mu^2)^{1/2}}{3\sigma_0^2} h_i,$$

where $|\mathcal{T}_i|$ is the area of the triangle \mathcal{T}_i and $\sin \theta_i \geq (1 - \mu^2)^{1/2}$ since $|\cos \theta_i| \leq \mu$. The above inequality immediately leads to

$$h_i/\rho_i \leq 3\sigma_0^2/(1 - \mu^2)^{1/2}. \quad (3.3)$$

Let $\bar{\rho}_K = 2\rho_{i_0}$. Using (3.3), we get

$$h_K/\bar{\rho}_K \leq (h_K/h_{i_0})(h_{i_0}/2\rho_{i_0}) \leq \sigma_0 \frac{3\sigma_0^2/2}{(1 - \mu^2)^{1/2}} = : \sigma,$$

that is just the G-R condition.

We remain to prove that C-R and A-R are equivalent. Firstly we show that C-R implies A-R. The first part of A-R is easily deduced from C-R, we omit details for simplicity. Further, a simple manipulation shows $(D\mathcal{F}_K(0,0)^{-1}\mathbf{k}, \mathbf{l}) = (J_K - J_0)/J_0$, then

$$J_K/J_0 \geq \frac{\min_{(\xi,\eta) \in \hat{K}} J_K(\xi,\eta)}{|K|/4} = 2 \min_{1 \leq i \leq 4} |\mathcal{T}_i|/|K|. \quad (3.4)$$

It is seen that

$$|\mathcal{T}_i| \geq 1/2h_K^2(1 - \mu^2)^{1/2} \quad \forall 1 \leq i \leq 4, \quad |K| \leq h_K^2.$$

Inserting the above two inequalities into (3.4) leads to

$$J_K/J_0 \geq (1 - \mu^2)^{1/2}/\sigma_0^2.$$

Let $\gamma := (1 - \mu^2)^{1/2}/\sigma_0^2 - 1 > -1$, then we obtain the A-R condition.

To deduce C-R from A-R, without loss of generality, let \mathcal{T}_1 include the shortest edge \underline{h}_K . The second part of A-R implies

$$2|\mathcal{T}_1|/|K| \geq \min_{(\xi,\eta) \in \hat{K}} J_K(\xi,\eta)/J_0 \geq 1 + \gamma. \quad (3.5)$$

Using (3.2) and noticing that $|\mathcal{T}_1| \leq 1/2h_K\underline{h}_K$, we obtain (C-R)₁ with $\sigma_0 = (4C_0(1 + \gamma))^{-1}$. Since the area of any triangle \mathcal{T}_i can be expressed as $|\mathcal{T}_i| \leq 1/2h_K^2 \sin \theta_i$, so repeating the above procedure using $|\mathcal{T}_i| \leq 1/2h_K^2 \sin \theta_i$ instead of $|\mathcal{T}_i| \leq 1/2h_K\underline{h}_K$, we get $\sin \theta_i \geq 4C_0(1 + \gamma)$, which implies (C-R)₂ with $\mu = (1 - 16C_0^2(1 + \gamma)^2)^{1/2}$. We complete the proof.

§4. Degenerate Mesh Condition

In what follows, we discuss some degenerate mesh conditions. As a preparation, we introduce the $(1 + \alpha)$ -Section Condition.

Definition 4.1. $(1 + \alpha)$ -Section Condition ($0 \leq \alpha \leq 1$)

$$d_K = \mathcal{O}(h_K^{1+\alpha}),$$

uniformly for all elements K as $h \rightarrow 0$.

If $d_K = 0$, K degenerates into a parallelogram. In case $\alpha > 0$, we recover the Condition A in [54]. In case α equals to 1, we obtain Condition B, or the Bi-Section Condition [54].

Angle condition is another kind of degenerate mesh condition, which is introduced in [50] and used to measure the deviation of a quadrilateral from

a parallelogram. Define σ_K as

$$\sigma_K := \max(|\pi - \alpha_1|, |\pi - \alpha_2|).$$

Here α_1 is the angle between the outward normal of two opposite sides of K and α_2 is the angle between the outward normal of other two sides. We call a mesh satisfying the Angle Condition if $\sigma_K = \mathcal{O}(h_K)$, i.e., if σ_K/h_K is uniformly bounded for all elements. It is seen that $0 \leq \sigma_K < \pi$, and $\sigma_K = 0$ iff K is a parallelogram.

Assuming that C-R holds, H. S. Chow et al (see [24, Theorem 3.2]) proved that the Angle Condition and the Bi-Section Condition are equivalent.

Remark 4.1. Observe that the h^2 -parallelogram mesh condition in [32] is actually equivalent to the Bi-Section Condition.

Motivated by the Bi-Section Condition, we define a kind of mesh which will be shown to be quite useful for the convergence analysis of finite elements.

Definition 4.2. We call \mathcal{T}_h an asymptotically regular parallelogram mesh if it satisfies C-R as well as the Bi-Section Condition.

Notice that any polygon can be meshed by asymptotically regular parallelograms with a mesh size tending to zero. Indeed, if we begin with any mesh of convex quadrilaterals, and refine it by dividing each quadrilateral into four by connecting two midpoints of opposite edges. As in Fig.2, the resulting mesh is an asymptotically regular parallelogram mesh.

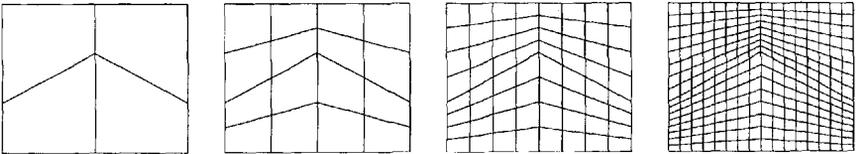


Fig.2

There is also another kind of method for generating such an asymptotically regular parallelogram mesh. For example, the C^2 -grid in [8] which results from a mapping of uniform grids is actually an asymptotically regular parallelogram mesh. This approach has already appeared in [71] (see also [25, Remark 2.3]).

As to the generation of a shape regular mesh satisfying the $(1+\alpha)$ -Section Condition, we refer to [23] for the work of Whiteman's school.

Degenerate Mesh Conditions

(1) Jamet condition (J) (see [34,36]).

\mathcal{T}_h is regular if there exists a constant $\sigma > 0$ such that

$$h_K/\rho_K \leq \sigma.$$

(2) Acosta-Duran Regular Decomposition Property (RDP) (see [1]).

\mathcal{T}_h is regular with constant $N \in \mathbb{R}$ and $0 < \psi < \pi$, or shortly $\text{RDP}(N, \psi)$, if we can divide K into two triangles along one of its diagonals, which will always be called d_1 , the other is d_2 in such a way that $|d_2|/|d_1| \leq N$ and both triangles satisfy the maximum angle condition, i.e., each interior angle of these two triangles is bounded from above by ψ .

(3) Süli condition (S) (see [58]).

\mathcal{T}_h is regular if it satisfies the J condition and the Bi-Section Condition simultaneously.

Remark 4.2. Notice that Süli’s condition was firstly appeared in the convergence proof of a kind of cell vertex finite volume method for hyperbolic problems and has recently got renewed interest in the mixed finite volume method [25].

As addressed in [36], C-R implies the J condition with $\sigma = (1 - \mu^2)^{1/2}/\sigma_0$. On the contrary, the J condition allows for the degeneration of a quadrilateral K into a triangle since the ratio h/\underline{h} can be arbitrarily large and the largest angle of K can equal π , i.e., the J condition may violate either $(\text{C-R})_1$ or $(\text{C-R})_2$. This is shown in Fig.3 below.

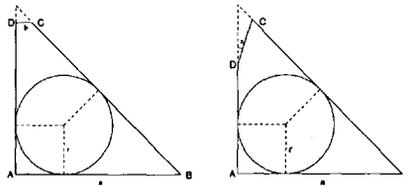


Fig.3

Both elements in Fig.3 satisfy the J condition since $h_K/\rho_K \leq \sqrt{2} + 1$. However, as to the left element, the ratio $h_K/\underline{h}_K = (a^2 + (a - b)^2)^{1/2}/b \leq \sqrt{2}a/b$ blows up as b tends to zero, which obviously violates $(\text{C-R})_1$. As to the right element, the interior angle $\angle ADC = \pi - x$, which approaches π as x tends to zero, thereby it violates $(\text{C-R})_2$. However, not the whole $(\text{C-R})_2$

is violated since the J condition excludes the interior angle from becoming too small. This fact is hidden in [58, Lemma 1] which is stated as follows.

Lemma 4.1. *If \mathcal{T}_h satisfies Süli's condition, then for sufficiently small h , \mathcal{T}_h is shape regular in the sense of C-R.*

Proceeding along the same line of the above lemma, we obtain

Corollary 4.1. *If \mathcal{T}_h satisfies the J condition as well as the $(1 + \alpha)$ -Section Condition, then for sufficiently small h , \mathcal{T}_h is shape regular in the sense of C-R.*

Notice that C-R does not imply the S condition. Indeed, considering the trapezoid mesh as that in Fig. 6, it is seen that the S condition is stronger than C-R. In fact, it is a strong shape regular mesh condition instead of a degenerate one. Invoking [1, Remark 2.7], $\text{RDP}(N, \psi)$ is weaker than both the J condition and C-R, therefore it is weaker than the S condition. Moreover, $\text{RDP}(N, \psi)$ together with the Bi-Section Condition does not imply C-R. It is due to the fact that a rectangular element which satisfies both $\text{RDP}(1, \pi/2)$ and the Bi-Section Condition simultaneously may still have its anisotropic aspect ratio arbitrarily large.

Remark 4.3. Schmidt [51] replaced the Bi-Section Condition in Lemma 4.1 by the arbitrary smallness of the ratio d_K/h_K .

Remark 4.4. Zlámal [71] proposed the k -strongly regular mesh condition for investigating the superconvergence of isoparametric elements. Following the analysis of Theorem 3.1, we find that Zlámal's 1-strongly regular condition is actually equivalent to the S condition. Notice that the 1-strongly regular condition has already appeared in [29] but in a slightly different form.

Remark 4.5. Braess [18, p.99, Remark 2] proposed a new shape regular mesh condition which consists of $(\text{C-R})_1$, the J condition and the maximum interior angle condition. Obviously, it is equivalent to C-R. However, due to its superfluous complexity this new condition is not advisable.

§5. 4-Node Isoparametric Element

Denoting by \mathcal{Q} the standard Lagrangian interpolant for the 4-node isoparametric element, we look for a geometric condition under which the

estimate

$$\|u - Qu\|_{0,K} + h_K|u - Qu|_{1,K} \leq Ch_K^2|u|_{2,K}$$

holds uniformly for $K \in \mathcal{T}_h$.

In view of [1, Theorem 4.7], the optimal interpolation error estimate with respect to the L^2 norm holds with a constant independent of the geometry of K , so we only consider

$$|u - Qu|_{1,K} \leq Ch_K|u|_{2,K}. \quad (5.1)$$

Ciarlet and Raviart [27] proved (5.1) under C-R. However, C-R prohibits the quadrilateral from either reducing to a triangle or becoming too flat. Jamet [35, 36] derived (5.1) under the J condition which allows a quadrilateral degenerating to a triangle, but not too flat. Ženíšek and Vanmaele [66] also proved (5.1). They required that the two longest sides of the element be opposite and almost parallel, but the constant C in (5.1) depends on an angle, which somehow is the minimum angle of the element K . Apel [4] derived the following estimate

$$|u - Qu|_{1,K} \leq C \left(h_1 \left\| \frac{\partial}{\partial x_1} \nabla u \right\|_{0,K} + h_2 \left\| \frac{\partial}{\partial x_2} \nabla u \right\|_{0,K} \right). \quad (5.2)$$

Here h_1 and h_2 are the element sizes in the direction of x_1 and x_2 , respectively. Acosta and Durán [1] derived (5.1) under $\text{RDP}(N, \psi)$, which seems the weakest mesh condition up to now under which (5.1.) is valid (see [1, Remark 2.4–Remark 2.7]). One may ask whether $\text{RDP}(N, \psi)$ is also necessary. Acosta and Durán put it as an open problem in [1]. The following example shows that this condition is indeed necessary.

Counterexample

Consider an element K like Fig.4 which does not satisfy $\text{RDP}(N, \psi)$. If we decompose it by the diagonal AC , then the triangles $\langle A, B, C \rangle$ and $\langle A, C, D \rangle$ indeed satisfy the maximal angle condition since all interior angles in these two triangles are bounded from above by $\pi/2$. However, $|BD|/|AC| = 1/a$ which cannot be bounded by any constant as a tends to zero. If we decompose it by the diagonal BD , a simple computation leads to $\sin \angle DAB = 2a/(1 + a^2)$, so the angle $\angle DAB$ approaches π as a tends to zero, thus it also violates $\text{RDP}(N, \psi)$.

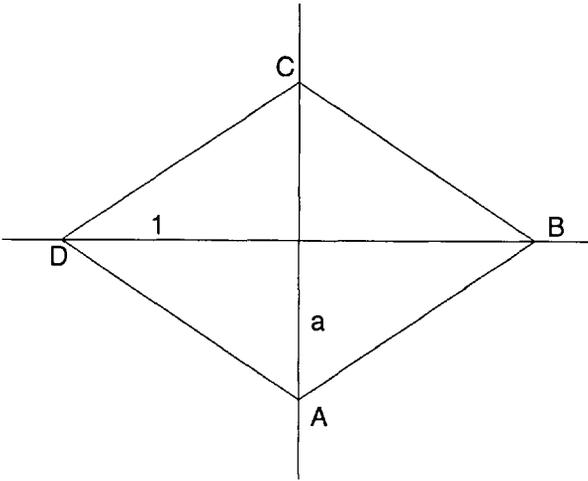


Fig.4

Let $u(x, y) = x^2$, then $|u|_{2,K}^2 = 8a$. A direct manipulation shows that $\left\| \frac{\partial(u - Qu)}{\partial y} \right\|_{0,K}^2 = (3a)^{-1}$, so we have

$$|u - Qu|_{1,K}^2 / |u|_{2,K}^2 \geq \left\| \frac{\partial(u - Qu)}{\partial y} \right\|_{0,K}^2 / |u|_{2,K}^2 = (1/24)1/a^2.$$

Since the diameter of K is 2, (5.1) does not hold with a constant independent of a .

To sum up, we have the following interpolation result for the 4-node isoparametric element.

Theorem 5.1. *For any $u \in H^2(\Omega)$, if RDP(N, ψ) holds, then there exists a constant $C = C(N, \psi)$ such that*

$$|u - Qu|_{1,K} \leq Ch_K |u|_{2,K}. \tag{5.3}$$

Moreover, RDP(N, ψ) is also necessary for the validity of (5.3).

By Céa Lemma, the error bound of the 4-node isoparametric element solution is bounded by its interpolation error, i.e.,

$$\|u - u_h\|_1 \leq C \inf_{v \in V_h} \|u - v\|_1. \tag{5.4}$$

Here

$$V_h := \{v \in H_0^1 \mid v|_K \in Q_1(K), \quad \forall K \in \mathcal{T}_h\}.$$

However, the experience with triangle elements indicates that the interpolation error actually says nothing about the approximation error of the finite element. When triangles with uncontrolled maximal angles are taken into account, examples in [12] show that the approximation error grows to infinity as the interpolation error with respect to the H^1 -norm grows to infinity. However, in [6], another example shows that the finite element solution converges while the interpolation error with respect to the H^1 -norm also grows to infinity. So it is equally interesting to ask such question for the quadrilateral element approximation.

§6. Nonconforming Quadrilateral Element

As to the nonconforming quadrilateral element approximation, the situation is less satisfactory. The main nonconforming quadrilateral elements are the Wilson element and nonconforming rotated Q_1 element. The former is well-known in the engineering community and has a long history (see [65]). The latter is the simplest nonconforming quadrilateral element, and was proposed for solving the incompressible flow problem in [50] and widely used in the software FEATFLOW [62]. It has also been applied to the crystalline microstructure problem [37], the Chappman-Ferraro problem [38] and the Reissner-Mindlin plate bending problem [41, 45].

By Strang Lemma [57], it is common to split the finite element error into two parts, i.e., the interpolation error and the consistency error, thus the impact of the mesh conditions on both errors has to be checked.

For any v belonging to a nonconforming finite element space, we define the discrete H^1 -norm as

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

The interpolation error of numerous nonconforming elements can be checked case by case. As to the consistency error, it is common to estimate the following piecewise integral functional

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \psi \cdot \mathbf{n} v \, ds.$$

Usually the nonconforming element has some sort of continuities, so to bound the above piecewise integral functional is boiled down to the estimate

of the following integral functional, i.e.,

$$\|u - \mathcal{J}(u)\|_{0,\mathcal{E}} \quad \forall \mathcal{E} \subset \partial K.$$

Here $\mathcal{J}(u)$ usually takes two forms as

$$\mathcal{J}(u) = \mathcal{J}_{\mathcal{E}}^a(u) := \int_{\mathcal{E}} u \, ds \quad \text{or} \quad \mathcal{J}(u) = \mathcal{J}_K(u) := \int_K u \, dx, \quad u \in H^1(K)$$

and $\mathcal{J}(u) = \mathcal{J}_{\mathcal{E}}^p(u) := u(\mathcal{M})$ in case of $u \in H^1(K) \cap C^0(\overline{K})$, where \mathcal{M} is the middle point of the edge \mathcal{E} . The following lemma bounds the above functional.

Lemma 6.1. *For $u \in H^1(K)$ or $u \in H^1(K) \cap C^0(\overline{K})$, and for any $\mathcal{E} \subset \partial K$, we have*

$$\|u - \mathcal{J}(u)\|_{0,\mathcal{E}} \leq C(|\mathcal{E}|/|K|)^{1/2} h_K \|\nabla u\|_{0,K}. \quad (6.1)$$

Proof. In case of $u \in H^1(K)$, since $\|u - \int_{\mathcal{E}} u \, ds\|_{0,\mathcal{E}} = \inf_{C \in \mathbb{R}} \|u - C\|_{0,\mathcal{E}}$, we have

$$\|u - \mathcal{J}_{\mathcal{E}}^a(u)\|_{0,\mathcal{E}} \leq \|u - \mathcal{J}_K(u)\|_{0,\mathcal{E}}.$$

By a sharp trace inequality in [63, Lemma 3.2], the right hand side of the above inequality is bounded by

$$(2|\mathcal{E}|/|K|)^{1/2} (\|u - \mathcal{J}_K(u)\|_{0,K} + h_K \|\nabla u\|_{0,K}).$$

Since K is convex, by the Poincaré Inequality [48] we have the first term as

$$\|u - \mathcal{J}_K(u)\|_{0,K} \leq h_K/\pi \|\nabla u\|_{0,K}. \quad (6.2)$$

A combination of the above three inequalities yields (6.1).

When $u \in H^1(K) \cap C^0(\overline{K})$, proceeding along the same line of the above procedure and employing the scaling trick instead of the Poincaré Inequality on the last step complete the proof.

To be more specific, we further bound the expression $C(|\mathcal{E}|/|K|)^{1/2} h_K$ appeared on the right hand side of (6.1) as $M\sigma^{1/2} h_K^{1/2}$ provided that the J condition holds, where M is independent of h .

The following simple example shows that the dependence on σ is essential.

Counterexample

Consider an element K centered at the origin with the lengths $2h_x$ and $2h_y$ in the x and y directions, respectively. Without loss of generality, we

assume that $h_x < h_y$, and define $\sigma := (h_x^2 + h_y^2)^{1/2} / \rho_K$. Let $u(x, y) = x^2 + y^2$ and \mathcal{E} be one vertical edge. A simple computation leads to

$$\|u - \mathcal{J}_{\mathcal{E}}^{\alpha}(u)\|_{0,\mathcal{E}}^2 = 8h_y^5/45, \quad \|\nabla u\|_{0,K}^2 = 16h_x h_y (h_x^2 + h_y^2)/3.$$

A combination of these two identities leads to

$$\|u - \mathcal{J}_{\mathcal{E}}^{\alpha}(u)\|_{0,\mathcal{E}} / \|\nabla u\|_{0,K} \geq \sigma^{1/2} (h_x^2 + h_y^2)^{1/4} / 2\sqrt{30}.$$

On the other hand, (6.1) gives the upper bound as

$$\|u - \mathcal{J}_{\mathcal{E}}^{\alpha}(u)\|_{0,\mathcal{E}} / \|\nabla u\|_{0,K} \leq C\sigma^{1/2} (h_x^2 + h_y^2)^{1/4}.$$

The above two inequalities illustrate the sharpness of the element geometry dependence of the constant in the right hand side of (6.1). Moreover, this example also works for the other two cases when $\mathcal{J}(u)$ is $\mathcal{J}_K(u)$ or $\mathcal{J}_{\mathcal{E}}^p(u)$.

6.1. Wilson Element

The Wilson nonconforming finite element space [65] is

$$\{v \in L^2(\Omega) \mid v \circ \mathcal{F}_K \in \mathcal{P}(\hat{K}) \quad \forall K \in \mathcal{T}_h\},$$

where $\mathcal{Q}_1(\hat{K}) \subset \mathcal{P}(\hat{K}) \subset \mathcal{P}_2(\hat{K})$. We write $\mathcal{P}(\hat{K}) = \mathcal{Q}_1(\hat{K}) + \mathcal{B}(\hat{K})$, where $\mathcal{B}(\hat{K})$ contains the nonconforming part:

$$\mathcal{B}(\hat{K}) = \text{Span}(\xi^2 - 1, \eta^2 - 1).$$

There are also another two types of Wilson-like elements (see [68] for a review).

Let Π_h denote the interpolant for the Wilson element, we sum up the interpolation results in the following theorem.

Theorem 6.1. *For any $u \in H^2(\Omega)$, if \mathcal{T}_h is C -R shape regular, then there exists a constant $C = C(\sigma_0, \mu)$ such that*

$$\|u - \Pi_h u\|_0 + h\|u - \Pi_h u\|_{1,h} \leq Ch^2 \|u\|_2. \tag{6.3}$$

Moreover, if $u \in H^3(\Omega)$, \mathcal{T}_h is C -R shape regular and the $(1 + \alpha)$ -Section Condition holds, then there exists a constant $C = C(\sigma_0, \mu)$ such that

$$\|u - \Pi_h u\|_0 + h\|u - \Pi_h u\|_{1,h} \leq Ch^{2+\alpha} \|u\|_3. \tag{6.4}$$

Proof. (6.3) has already been included in [39, Theorem 1]. Notice that $\mathcal{P}_2(\hat{K}) \subset \mathcal{P}(\hat{K})$, the standard interpolation argument yields (6.4).

One may ask if the shape regular condition can be relaxed in the above theorem. There is no such result for a general quadrilateral mesh, however,

the following result indicates that it is not hopeless at least for a rectangular mesh.

Lemma 6.2. *If Ω is covered by a uniform rectangular mesh with $h_x(h_1)$ and $h_y(h_2)$ in the x and y directions, respectively, then*

$$\|u - \Pi_h u\|_0 \leq Ch^2 \left(\left\| \frac{\partial^2 u}{\partial x^2} \right\|_0 + \left\| \frac{\partial^2 u}{\partial y^2} \right\|_0 \right), \tag{6.5}$$

$$\left\| \frac{\partial(u - \Pi_h u)}{\partial x_i} \right\|_0 \leq C \sum_{j=1}^2 h_j \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_0, \quad i = 1, 2. \tag{6.6}$$

Moreover, if $u \in H^3$, then

$$\|u - \Pi_h u\|_0 \leq Ch^3 |u|_3, \tag{6.7}$$

$$\left\| \frac{\partial(u - \Pi_h u)}{\partial x_i} \right\|_0 \leq C \sum_{j,k=1}^2 h_j h_k \left\| \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} \right\|_0, \quad i = 1, 2. \tag{6.8}$$

To estimate the consistency error, we follow that in [54]. The consistency functional can be decomposed into

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_{\partial K} \psi \cdot n v \, ds &= \sum_{K \in \mathcal{T}_h, \mathcal{E} \subset \partial K} \int_{\mathcal{E}} (\psi \cdot n - \int_K \psi \cdot n) (v - \int_K v) \, ds \\ &+ \sum_{K \in \mathcal{T}_h, \mathcal{E} \subset \partial K} \int_{\mathcal{E}} (\psi \cdot n - \int_K \psi \cdot n) \int_K v \, ds \\ &+ \sum_{K \in \mathcal{T}_h, \mathcal{E} \subset \partial K} \int_K \psi \cdot n \int_{\mathcal{E}} v \, ds. \end{aligned} \tag{6.9}$$

Lemma 6.1 bounds the first two terms on the right hand side of (6.9), thus the J condition is needed. Moreover, in [54, Theorem 2], a counterexample is presented to show that the Bi-Section Condition is necessary for estimating the third term on the right hand side of (6.9). Summing up and using Lemma 4.1, we see that the asymptotically regular parallelogram mesh condition is both sufficient and necessary for obtaining the optimal consistency error, at least for the above decomposition. Moreover, the optimal interpolation error requires that \mathcal{T}_h is shape regular which may be weakened as indicated by Lemma 6.1. However, it is still unknown whether or not the asymptotically regular parallelogram mesh is really necessary for the optimal error estimate of the Wilson-type elements.

6.2. Nonconforming Rotated Q_1 Element

Two types of the quadrilateral rotated Q_1 finite element spaces can be defined as follows. Let

$$\bar{Q}_1 := \{ q \circ \mathcal{F}_K^{-1} \mid q \in \text{Span}\langle 1, x, y, x^2 - y^2 \rangle \}.$$

Denote $\mathcal{J}_\mathcal{E}^{a/p}$ for $\mathcal{J}_\mathcal{E}^a$ as well as $\mathcal{J}_\mathcal{E}^p$. The finite element spaces are defined as $V_h^{a/p} := \{ v \in L^2(\Omega) \mid v|_K \in \bar{Q}_1(K), v \text{ is continuous regarding } \mathcal{J}_\mathcal{E}^{a/p}(\cdot) \}$, and the corresponding homogeneous spaces as

$$V_{0,h}^{a/p} := \{ v \in V_h^{a/p} \mid \mathcal{J}_\mathcal{E}^{a/p}(v) = 0, \text{ if } \mathcal{E} \subset \partial\Omega \}.$$

A global interpolation operator π_h is realized by the forgoing local interpolation operator $\mathcal{J}_\mathcal{E}^{a/p}$, i.e., $\pi_h|_K = \mathcal{J}_\mathcal{E}^{a/p} \quad \forall \mathcal{E} \subset \partial K$. We have the following interpolation result for π_h .

Theorem 6.2. [50, Lemma 1] *For any $u \in H^2 \cap H_0^1$, if \mathcal{T}_h is C-R shape regular, then*

$$\|u - \pi_h u\|_0 + h \|u - \pi_h u\|_{1,h} \leq Ch(h + \sigma_h) \|u\|_2. \quad (6.10)$$

Here C depends on σ_0 and μ .

As a direct consequence of the above result, we have

Corollary 6.1. *For any $u \in H^2 \cap H_0^1$, if the $(1 + \alpha)$ -Section Condition holds, then*

$$\|u - \pi_h u\|_0 \leq C_1 h^{1+\alpha} \|u\|_2. \quad (6.11)$$

Here the constant C_1 is independent of the geometry of $K \in \mathcal{T}_h$. If \mathcal{T}_h is an asymptotically regular parallelogram mesh, then

$$\|u - \pi_h u\|_{1,h} \leq C_2 h \|u\|_2. \quad (6.12)$$

Here C_2 depends on σ_0 and μ .

By Theorem 5.1, RDP(N, ψ) is sufficient and necessary for obtaining the optimal interpolation error for the 4-node isoparametric element. One may ask whether this is the same for the NR Q_1 element. The following lemma and a counterexample give a negative answer.

Lemma 6.3. *If Ω is covered by a uniform rectangular mesh with $h_x(h_1)$ and $h_y(h_2)$ in the x and y directions, respectively. Then for $u \in H^2 \cap H_0^1$, we have*

$$\|u - \pi_h u\|_0 \leq C \sum_{i,j=1}^2 h_i h_j \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_0, \quad (6.13)$$

$$|u - \pi_h u|_{1,h} \leq Ch(1 + \sigma)|u|_2. \quad (6.14)$$

Here σ is the anisotropic ratio which is defined by $\sigma := \max(h_x/h_y, h_y/h_x)$.

Proof. (6.14) is included in [42, Theorem 2.1], from which we get (6.13).

We adopt a counterexample in [42] to show that the anisotropic ratio σ appeared in the right hand side of (6.14) is essential.

Counterexample

Consider an element K centered at the origin with the lengths $2h_x$ and $2h_y$ in the x and y directions, respectively. Without loss of generality, we assume that $h_y < h_x$, and define the anisotropy ratio as $\sigma := h_x/h_y$. Let $u(x, y) = 2x^2$, then $|u|_{2,K}^2 = 64h_x h_y$. As to this u , we can verify that $|u - \pi_h u|_{1,K}^2 = 16h_x^3 h_y(1 + \sigma^2)/3$, therefore

$$|u - \pi_h u|_{1,K}/|u|_{2,K} = ((1 + \sigma^2)/3)^{1/2} h_x/2.$$

So (6.14) cannot hold with a constant independent of the anisotropic ratio σ .

Notice that the asymptotically regular parallelogram mesh condition is sufficient for the optimal interpolation error for the NRQ₁ element, the following counterexample shows that the Bi-Section Condition is also necessary.

Counterexample

Consider the element K as that in Fig. 5. Let $u(x, y) := 1/(1 + x)$. It is seen that

$$|u|_{2,K}^2 \leq 4|K| = 8(2 - a)h_K^2. \quad (6.15)$$

A simple manipulation shows that

$$\|u - \pi_h u\|_{0,K}^2 = \int_K h_K^2 (1 - a(1 + \eta)/2)(1 + \xi)(ah_K \xi \eta/2 + \mathcal{O}(h_K^2))^2 / f(\xi, \eta) d\hat{x},$$

where $f(\xi, \eta)$ is defined as

$$f(\xi, \eta) := (1 + h_K)^2 (1 + (1 - a)h_K)^2 (1 + (2 - a)h_K)^2 \\ (1 + (1 - a(1 + \eta)/2)(1 + \xi)h_K)^2.$$

Notice that $0 \leq a < 1$, so $|f(\xi, \eta)| \leq 36$, thus a combination of the above two identities yields

$$\|u - \pi_h u\|_{0,K}^2 \geq a^2 h_K^4 / 648 + \mathcal{O}(h_K^6), \quad (6.16)$$

which together with (6.15) leads to

$$\|u - \pi_h u\|_{0,K}/|u|_{2,K} \geq \sqrt{2}ah_K/18 = \sqrt{2}d_K/18. \quad (6.17)$$

Proceeding along the same line, we obtain

$$\frac{|u - \pi_h u|_{1,K}}{|u|_{2,K}} \geq \frac{\left\| \frac{\partial(u - \pi_h u)}{\partial x} \right\|_{0,K}}{|u|_{2,K}} \geq (\sqrt{3}/36)d_K/h_K. \quad (6.18)$$

The above two inequalities clearly show the necessity of the Bi-Section Condition for the optimal interpolation error bound of the NRQ₁ element with respect to both L²-norm and H¹-norm.

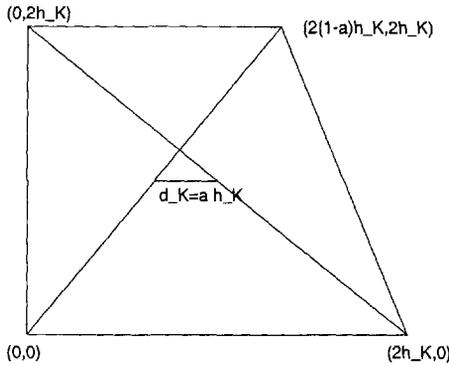


Fig.5

Degradation of the interpolation error will occur on real quadrilateral meshes, particularly, on the following trapezoid meshes. Nevertheless, as shown by Corollary 6.1, if the quadrilateral mesh is an asymptotic parallelogram (see Fig.2), such degradation will not occur.

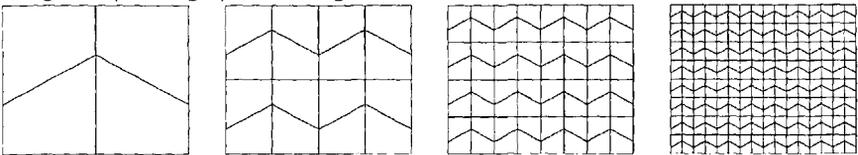


Fig.6

It remains to consider the consistency error of NRQ₁. As to NRQ₁^a, the consistency functional can be decomposed into

$$\sum_{K \in \mathcal{T}_h, \mathcal{E} \subset \partial K} \int_{\mathcal{E}} (\psi \cdot \mathbf{n} - \int_{\mathcal{E}} \psi \cdot \mathbf{n}) (v - \int_{\mathcal{E}} v) ds.$$

By Lemma 6.1, the J condition is needed for the optimal consistency error estimate. As to NRQ₁^p, besides the J condition the Bi-Section Condition

is also needed. The latter is even necessary in some sense for NRQ_1^p as shown in [42, Theorem 3.2]. However, for the rectangular NRQ_1 , a different argument yields the optimal consistency error bound which is independent of the J condition (see [44] for more details). We do not know whether a similar argument works for the quadrilateral NRQ_1 .

The above discussion indicates that there is a convergence degradation of the NRQ_1 element over a degenerate mesh. Indeed, such degradation was observed by the numerical results in [50, 62], both the interpolation and consistency error commit such degradation. There are many works on modifications of this element to accommodate the degenerate mesh (see [7, 21, 22, 44]).

To sum up, neither the Wilson element nor the NRQ_1 element can be used for a degenerate mesh. Is there another kind of lower-order nonconforming quadrilateral element which can be used over a fully degenerate mesh while retaining its excellent stability property simultaneously? Such an element would be a grail for the finite element circus.

§7. RT Element of Lowest-Order

Up to now, there are only few results available to explain the mesh dependence of the interpolation error for mixed elements, like $\text{R-T}_{[k]}$, $\text{BDFM}_{[k]}$ and $\text{BDM}_{[k]}$ et al [20].¹

Wang and Mathew [64] gave the optimal interpolation error for all these mixed finite elements over an arbitrary quadrilateral, unfortunately, the mesh dependence is not clearly stated therein.

Raviart and Thomas [49] derived the optimal interpolation error for the R-T element over a shape regular parallelogram. Proceeding along the same line of [1, Lemma 4.1, Lemma 4.2], one can easily get the following interpolation and stability estimates for the $\text{RT}_{[0]}$ element (see [1, Remark 4.1] and [8]) for the case when \mathcal{T}_h is a rectangular mesh.

Lemma 7.1. *If Ω is covered by a uniform rectangular mesh with the diameter h_1 and h_2 in the x and y directions, respectively, then*

¹see [64, 56] for the definitions of $\text{RT}_{[k]}$, $\text{BDDM}_{[k]}$ and $\text{BDFM}_{[k]}$ over an arbitrary quadrilateral mesh.

$$\|(\mathbf{u} - RT\mathbf{u})_i\|_0 \leq C \sum_{j=1}^2 h_j \left\| \frac{\partial \mathbf{u}_i}{\partial x_j} \right\|_0, \quad i = 1, 2 \quad (7.1)$$

with a constant C independent of the ratio between h_1 and h_2 . Moreover, we have the following stability estimate

$$\left\| \frac{\partial (RT\mathbf{u})_i}{\partial x_i} \right\|_0 \leq \left\| \frac{\partial \mathbf{u}_i}{\partial x_i} \right\|_0, \quad i = 1, 2. \quad (7.2)$$

Remark 7.1. As to a simple proof for (7.2), see [8, Lemma 5.7].

Similar to [43, Theorem 3.2], we have the following result for the quadrilateral $RT_{[0]}$.

Theorem 7.1. *If \mathcal{T}_h is an asymptotically regular parallelogram mesh, then for $\mathbf{u} \in \mathbf{H}^1(\text{div})$, there holds*

$$\|\mathbf{u} - RT\mathbf{u}\|_0 \leq Ch\|\mathbf{u}\|_1, \quad (7.3)$$

$$\|\text{div}(\mathbf{u} - RT\mathbf{u})\|_0 \leq Ch\|\mathbf{u}\|_{H^1(\text{div})}. \quad (7.4)$$

Here the constant C depends on σ_0 and μ .

Remark 7.2. Similar results can be found in [25, Lemma 3.2]. However, the Bi-Section Condition is missing therein, which is actually necessary for obtaining the optimal interpolation error. This will be illustrated by the example below.

Similar to [43, Theorem 3.1], we have the refined interpolation results for $RT_{[0]}$ as follows.

Theorem 7.2. *If \mathcal{T}_h satisfies C-R as well as the $(1 + \alpha)$ -Section Condition, then for any $\mathbf{u} \in \mathbf{H}^\beta(\text{div})$ with $\beta \in (0, 1]$, there holds*

$$\|\mathbf{u} - RT\mathbf{u}\|_0 \leq C(h^\beta|\mathbf{u}|_{\mathbf{H}^\beta} + h^\alpha\|\mathbf{u}\|_0 + h\|\text{div}\mathbf{u}\|_0), \quad (7.5)$$

$$\|\text{div}(\mathbf{u} - RT\mathbf{u})\|_0 \leq C(h^\beta|\text{div}\mathbf{u}|_{\mathbf{H}^\beta} + h^\alpha\|\text{div}\mathbf{u}\|_0). \quad (7.6)$$

If $\beta > 1/2$, the last term in (7.5) can be dropped. Moreover, if $\mathbf{u} \in \mathbf{H}^1(\Omega)$ and the Bi-Section Condition holds, we have

$$\|\mathbf{u} - RT\mathbf{u}\|_0 \leq Ch\|\mathbf{u}\|_1, \quad (7.7)$$

$$\|\text{div}(\mathbf{u} - RT\mathbf{u})\|_0 \leq Ch\|\text{div}\mathbf{u}\|_1. \quad (7.8)$$

Here the constant C depends on σ_0 and μ .²

²The definitions of \mathbf{H}^β and $\mathbf{H}^\beta(\text{div})$ with $\beta \in (0, 1]$ can be found in [3].

Comparing Theorem 7.1 and Theorem 7.2 with Lemma 7.1, one may ask whether the mesh conditions in Theorem 7.1 and Theorem 7.2 can be relaxed. The following example shows the necessity of the $(1 + \alpha)$ -Section Condition or the Bi-Section Condition for the optimality of the interpolation error. It seems that the argument based on the Piola transform as that in [43] and [25] is less hopeful for further relaxing the mesh conditions.

Counterexample

Consider the element as in Fig. 5. Let $\mathbf{u} = (1, 0)$, a simple manipulation yields

$$\|\mathbf{u} - RT\mathbf{u}\|_{0,K}^2 = 2(ah_K/2)^2 \int_{-1}^1 \frac{\eta^2}{1 - a/2 - a/2\eta} d\eta \geq a^2 h_K^2/3. \quad (7.9)$$

It is seen that

$$\|\mathbf{u}\|_{1,K}^2 = 2(2 - a)h_K^2.$$

A combination of the above two inequalities leads to

$$\|\mathbf{u} - RT\mathbf{u}\|_{0,K}/\|\mathbf{u}\|_{1,K} \geq (1/2\sqrt{3})a = (1/2\sqrt{3})d_K/h_K. \quad (7.10)$$

The above example clearly shows the necessity of the $(1 + \alpha)$ -Section Condition for obtaining the optimal interpolation error bound (7.5).

Let $\mathbf{u} = (x, 0)$; proceeding along the same line of the above procedure, we obtain

$$\|\operatorname{div}(\mathbf{u} - RT\mathbf{u})\|_{0,K}/\|\operatorname{div}\mathbf{u}\|_{1,K} \geq (1/\sqrt{6})d_K/h_K, \quad (7.11)$$

which shows the necessity of the $(1 + \alpha)$ -Section Condition for the optimal interpolation error bounds in (7.6). Naturally, (7.10) and (7.11) also demonstrate the necessity of the Bi-Section Condition for the optimal interpolation error bounds in (7.7) and (7.8), respectively.

§8. Conclusions and Open Problems

In this paper, some commonly used shape regular mesh conditions are proven to be equivalent and their connections to some degenerate mesh conditions are also clarified.

We have checked the influence of mesh conditions on the interpolation error for the 4-node isoparametric element, quadrilateral nonconforming element and $RT_{[0]}$ element. The asymptotically regular parallelogram mesh is found to be indispensable for the successful application of either Wilson,

NRQ_1 or $\text{RT}_{[0]}$ element, otherwise, the degradation of the convergence order will occur which is not widely appreciated, and was casually observed in numerical experiments (see [69, §8.7] and [9]).

Before closing this paper, we propose three open problems.

(1) Is $\text{RDP}(N, \psi)$ also necessary for the optimal finite element approximation error of the 4-node isoparametric element for the 2-order elliptic problem?

(2) What is the necessary and sufficient mesh condition for the convergence as well as the optimal error bounds of the Wilson and NRQ_1 element for the 2-order elliptic problem?

(3) What is the necessary and sufficient mesh condition for obtaining the optimal interpolation error bounds of $\text{RT}_{[k]}$, $\text{BDDM}_{[k]}$ and $\text{BDFM}_{[k]}$?

Acknowledgement. This work was finished when the first author was visiting the Faculty of Mathematics at the Ruhr-University Bochum on a grant by the Alexander von Humboldt Foundation. He would like to thank AvH for the support.

The authors would like to thank Professor F. Kikuchi for providing them with [34] and some suggestive discussion.

REFERENCES

- [1] Acosta, G. & Durán, R., Error estimate for \mathcal{Q}_1 isoparametric elements satisfying a weak angle condition, *SIAM J. Numer. Anal.*, **38**(2001), 1073–1088.
- [2] Adams, R. A., Sobolev spaces, Academic Press, 1975.
- [3] Amrouche, C., Bernardi, C. & Dauge, M., Vector potentials in three dimensional non-smooth domains, *Math. Meth. in Appl. Sciences*, **21**(1998), 823–864.
- [4] Apel, T., Anisotropic interpolation error estimates for isoparametric quadrilateral finite elements, *Computing*, **60**(1998), 157–174.
- [5] Apel, T., Anisotropic finite elements: local estimates and applications, Series: “Advances in Numerical Mathematics”, Teubner, Stuggart, 1999.
- [6] Apel, T. & Dobrowolski, M., Anisotropic interpolation with application to the finite element method, *Computing*, **47**(1992), 277–293.
- [7] Apel, T., Nicaise, S. & Schöberl, J., Crouzeix-Raviart type finite elements on anisotropic meshes, *Numer. Math.*, **89**(2001), 193–223.
- [8] Arbogast, T., Wheeler, M. F. & Yotov, I., Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite difference, *SIAM J. Numer. Anal.*, **34**(1997), 828–852.
- [9] Arnold, D. N., Boffi, D. & Falk, R. S., Approximation by quadrilateral finite elements, Preprint, 2000.

- [10] Arnold, D. N., Boffi, D., Falk, R. S. & Gastaldi, L., Finite element approximation on quadrilateral meshes, Preprint, 2001.
- [11] Arunakirinathar, K. & Reddy, B. D., Some geometrical results and estimates for quadrilateral finite elements, *Comput. Methods Appl. Mech. Engrg.*, **122**(1995), 307–314.
- [12] Babuska, I. & Aziz, K., On the angle condition in the finite element method, *SIAM J. Numer. Anal.*, **13**(1976), 214–226.
- [13] Bathe, K. J., Finite element procedure, Prentice-Hall, 1996.
- [14] Becker, R., An adaptive finite element method for the incompressible Navier-Stokes equations on time-dependent domains, Ph. D dissertation, Ruprecht-Karls-Universität, Heidelberg, 1995.
- [15] Becker, R. & Rannachar, R., Finite element solution of the incompressible Navier-Stokes equations on anisotropically refined meshes, In Fast solvers for flow problems, Vol. 49 of Notes on Numerical Fluid Mechanics, 52–62. Wiesbaden, 1995, Vieweg.
- [16] Berger, A., Scott, R. & Strang, G., Approximate boundary conditions in the finite element method, *Symposia Mathematica*, **10**(1972), 295–313.
- [17] Bernardi, C. & Girault, V., A local regularization operator for triangular and quadrilateral finite elements, *SIAM J. Numer. Anal.*, **36**(1998), 1893–1916.
- [18] Braess, D., Finite elements: theory, fast solvers, and applications in solid mechanics, Cambridge University Press, 2001.
- [19] Bramble, J. H. & Hilbert, S. R., Bounds for a class of linear functionals with application to Hermite interpolation, *Numer. Math.*, **16**(1972), 362–369.
- [20] Brezzi, F. & Fortin, M., Mixed and hybrid finite element methods, Springer-Verlag, 1991.
- [21] Cai, Z., Douglas, J. & Ye, X., A stable nonconforming quadrilateral finite element method for the Stokes and Navier-Stokes equations, *Calcolo.*, **36**(1999), 215–232.
- [22] Cai, Z., Douglas, J., Santos, J. E., Sheen, D. & Ye, X., Nonconforming quadrilateral finite elements: a correction, *Calcolo.*, **37**(2000), 253–254.
- [23] Cheng, Y. P., Recovery of gradient piecewise linear finite element approximation on non-uniform partitions, *Numer. Meth. Part. Diff. Equations*, **8**(1998), 169–192.
- [24] Chou, S. H. & He, A. N., On the regularity and uniformness conditions on quadrilateral meshes, preprint, 2001.
- [25] Chou, S. H., Kwak, D. & Kim, K., A general framework for the constructing and analyzing mixed finite volume methods on quadrilateral grids: the overlapping covolume case, *SIAM J. Numer. Anal.*, **39**(2001), 1170–1196.
- [26] Ciarlet, P. G., The finite element method for elliptic problems [M], North Holland, Amsterdam, 1978.
- [27] Ciarlet, P. G. & Raviart, P.-A., Interpolation theory over curved elements with applications to finite element methods, *Comput. Methods Appl. Mech. Engrg.*, **1**(1972), 217–249.
- [28] Ciarlet, P. G. & Raviart, P.-A., General Lagrange and Hermite interpolation in \mathbb{R}^n with application to finite elements, *Arch. Rational Mech. Anal.*, **46**(1972), 177–199.
- [29] Ciarlet, P. G. & Raviart, P.-A., The combined effect of curved boundary and numerical integrations in isoparametric finite element methods, The mathematics foundation of the finite element methods with applications to partial differential equations, A. K. Aziz and I. Babuská eds., Academic Press, New York and London, 1972, 409–475.

- [30] Clément, P., Approximation by finite element functions using local regularization, *RAIRO, Anal. Numér.*, **9**(1975), 75–84.
- [31] Durán, R., Error estimates for 3-d finite elements, *Math. Comp.*, **68**(1999), 187–199.
- [32] Ewing, R., Liu, M. M. & Wang, J. P., Superconvergence of mixed finite element approximations over quadrilaterals, *SIAM J. Numer. Anal.*, **36**(1999), 772–787.
- [33] Girault, V. & Raviart, P.-A., Finite element methods for the Navier-Stokes equations: Theory and algorithms, Springer-Verlag, Berlin, 1986.
- [34] Jamet, P., Estimations de erreur d'interpolation dans un domaine variable et application aux éléments finis quadrilatéraux dégénérés, in méthodes numériques en mathématiques appliquées, 55–100. Presses de l'Université de Montréal.
- [35] Jamet, P., Estimations d'erreure pour des éléments finis droits presque dégénérés, *RAIRO, Anal. Numér.*, **10**(1976), 46–61.
- [36] Jamet, P., Estimation of the interpolation error for quadrilateral finite elements which can degenerate into triangles, *SIAM J. Numer. Anal.*, **14**(1977), 925–930.
- [37] Klouček, P., Li, B. & Luskin, M., Analysis of a nonconforming finite elements for crystalline microstructures, *Math. Comp.*, **65**(1996), 1111–1135.
- [38] Klouček, P. & Toffoletto, F. R., The three dimensional nonconforming finite element solution of the Chapman-Ferraro problem, *J. Comp. Phy.*, **150**(1999), 549–560.
- [39] Lesaint, P. & Zlamal, M., Convergence of the nonconforming Wilson element for arbitrary quadrilateral meshes, *Numer. Math.*, **36**(1980), 33–52.
- [40] Ming Pingbing, Nonconforming finite elements vs locking problems, Doctorate dissertation, Institute of Computational Mathematics, Chinese Academy of Sciences, 2000.
- [41] Ming Pingbing & Shi Zhongci, Nonconforming rotated \mathcal{Q}_1 element for Reissner-Mindlin plate, *Math. Model Meth. Appl. Sci.*, **8**(2001), 1311–1342.
- [42] Ming Pingbing & Shi Zhongci, Mathematical analysis for nonconforming rotated \mathcal{Q}_1 element: (I) Convergence study, Preprint, 2000.
- [43] Ming Pingbing & Shi Zhongci, Optimal L^2 error estimate for stabilized MITC4 element, Preprint, 2001.
- [44] Ming Pingbing & Shi Zhongci, Nonconforming rotated \mathcal{Q}_1 element and its variants, Preprint, 2001.
- [45] Ming Pingbing & Shi Zhongci, Quadrilateral nonconforming rotated \mathcal{Q}_1 element for Reissner-Mindlin plate, Preprint, 2001.
- [46] Ming Pingbing & Shi Zhongci, Convergence analysis for quadrilateral rotated \mathcal{Q}_1 element [A], in Advances in computation: Theory and practice, Vol. 7: Scientific computing and applications, Peter Minev and Yanping Lin eds., Nova Science Publishers, Inc., 2001, 115–124.
- [47] Oganessian, L. A. & Rukhovets, L. A., Variational difference methods for the solution of elliptic problem, Izd. Akad. Nauk Armyanskoi SSR, Jerevan, 1979 (in Russian).
- [48] Payne, L. E. & Weinberger, H. F., An optimal Poincaré inequality for convex domains, *Arch. Rational Mech. Anal.*, **5**(1960), 286–292.
- [49] Raviart, P. A. & Thomas, J. M., A mixed finite element method for second order elliptic problems, in Mathematics aspects of the finite element method [M], I. Galigani, E. Magenes, eds., Lecture Notes in Math. 606, Springer-Verlag, New York, 1977, 292–315.

- [50] Rannachar, R. & Turek, S., Simple nonconforming quadrilateral stokes element, *Numer. Meth. Part. Diff. Equations.*, **8**(1992), 97–111.
- [51] Schmidt, T., Box schemes on quadrilateral meshes, *Computing.*, **51**(1993), 271–292.
- [52] Scott, R. & Zhang Shangyou, Finite element interpolation at the nonsmooth functions satisfying boundary conditions, *Math. Comp.*, **54**(1990), 483–493.
- [53] Al. Shenk, N., Uniform error estimates for certain narrow Lagrange finite elements, *Math. Comp.*, **63**(1994), 105–119.
- [54] Shi, Z., A convergence condition for the quadrilateral Wilson element, *Numer. Math.*, **44**(1984), 349–361.
- [55] Shi, Z., A remark on the optimal order of convergence of Wilson nonconforming element, *Math. Numer. Sinica*, **8**(1986), 159–163.
- [56] Stenberg, R. & Suri, M., An hp error analysis of MITC plate elements, *SIAM J. Numer. Anal.*, **34**(1997), 544–568.
- [57] Strang, G. & Fix, G. J., An analysis of the finite element method, Prentice Hall, Englewood Cliffs, NJ, 1973.
- [58] Süli, E., The accuracy of cell vertex finite volume methods on quadrilateral meshes, *Math. Comp.*, **59**(1992), 359–382.
- [59] Syngé, J. L., The hypercircle in mathematics physics, Cambridge University Press, New York, 1957.
- [60] Taylor, R. L., Beresford, P. J. & Wilson, E. L., A nonconforming element for stress analysis, *Int. J. Numer. Meth. Eng.*, **10**(1976), 2111–2119.
- [61] Thomas, J. M., Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes, Thèse d'Etat, Université Pierre et Marie Curie, Paris, 1977.
- [62] Turek, S., Efficient solvers for incompressible flow problems: an algorithm and computational approach, Lecture Notes in Computational Sciences and Engineering, Vol. 6. Springer-verlag, Berlin, Heidelberg, 1999.
- [63] Verfürth, R., Error estimate for some quasi-interpolation operator, *RAIRO, M²AN.*, **33**(1999), 695–713.
- [64] Wang, J. & Mathew, T., Mixed finite element method over quadrilaterals, in Proceedings of 3rd International Conference on Advances in Numerical Methods and Applications, J. T. Dimov, B. Sendov and P. Vassilevski eds., World Scientific, River Edge, NJ, 1994, 203–214.
- [65] Wilson, E. L., Taylor, R. L., Doherty, W. & Ghaboussi, J., Incompatible displacement models, in: S. J. Fenves et al. (Eds), Numerical and Computer Methods in Structural Mechanics, Academic Press, New York, 1973, 43–57.
- [66] Ženíšek, A. & Vanmaele, M., The interpolation theorem for narrow quadrilateral isoparametric finite elements, *Numer. Math.*, **72**(1995), 123–141.
- [67] Zhang, J. & Kikuchi, F., Interpolation error estimates of a modified 8-node serendipity finite element, *Numer. Math.*, **85**(2000), 503–524.
- [68] Zhang, Z., Analysis of some quadrilateral nonconforming elements for incompressible elasticity, *SIAM J. Numer. Anal.*, **34**(1997), 640–663.
- [69] Zienkiewicz, O. C. & Taylor, R. L., The finite element method, Volume 1: Basic formulation and linear problems, 4-ed, McGraw-Hill, 1989.
- [70] Zlámal, M., On the finite element method, *Numer. Math.*, **12**(1968), 394–409.
- [71] Zlámal, M., Superconvergence and reduced integration in the finite element method, *Math. Comp.*, **32**(1978), 663–685.

ON THE HYPERBOLIC OBSTACLE PROBLEM OF FIRST ORDER*

J. F. RODRIGUES

*CMAF/Universidade de Lisboa, Av. Prof. Gama Pinto, 2,
1649-003 Lisboa, Portugal. E-mail: rodrigue@lmc.fc.ul.pt*

Abstract

This paper presents new results for strong solutions and their coincidence sets of the obstacle problem for linear hyperbolic operators of first order. An inequality similar to the Lewy-Stampacchia ones for elliptic and parabolic problems is shown. Under nondegeneracy conditions the stability of the coincidence set is shown with respect to the variation of the data and with respect to approximation by semilinear hyperbolic problems. These results are applied to the asymptotic stability of the evolution problem with respect to the stationary coercive problem with obstacle.

Keywords Hyperbolic obstacle problem, Linear hyperbolic operator,
Strong solution

2000 MR Subject Classification 35L85, 35R35

§1. Introduction

The classical obstacle problem can be formulated as the problem of finding the equilibrium position of an elastic membrane constrained to lie above an obstacle. Although, in the words of J. L. Lions, this “simple, beautiful and deep” problem is naturally associated with partial differential equations of elliptic type, it arises in many other frameworks and in different kinds of free boundary problems (see [3] or [10], and their references) and it is related to variational inequalities (see [6,7]).

*Partially supported by the Project FCT-POCTI/34471/MAT/2000.

Variational inequalities of first order hyperbolic type were introduced in 1973 by Bensoussan and Lions [2] for the study of deterministic cases in problems of optimal stopping time, in which their solutions can be interpreted as optimal cost functions. More recently, motivated by physical problems in petroleum engineering, some unilateral problems for scalar conservation laws have been considered by L. Lévi in [4] (see also [5]), where the existence and uniqueness of weak entropy solutions are proven for quasi-linear hyperbolic operators.

In this work we are concerned with the problem of finding a function u defined in an open smooth domain $Q \subset \mathbb{R}^N$, such that, for a given function ψ (the obstacle)

$$u \geq \psi \quad \text{a.e. in } Q \quad (1.1)$$

and, in the a priori unknown region where the solution u does not coincide with the obstacle ψ ,

$$Hu = f \quad \text{a.e. in } \{u > \psi\} \quad (1.2)$$

with a Dirichlet boundary condition in a known part of the boundary of the domain

$$u = h \quad \text{a.e. on } \Sigma_- \subset \partial Q \quad (1.3)$$

for given functions f and h . Here H is a linear first order operator, whose principal parts determines the subset Σ_- , where the boundary condition can be imposed.

The first systematic study of the obstacle problem (1.1),(1.2),(1.3) was done by Mignot and Puel in 1976 [8] in the framework of strong and weak solutions of variational inequalities of first order. Their approach, which will be followed here, is based on the general linear theory of boundary value problems for first order partial differential equations of Bardos [1]. For linear operators the boundary condition (1.3) and the definition of Σ_- are well-known and their functional spaces are recalled in Section 2, where we introduce and show the continuous dependence of strong solutions for the hyperbolic obstacle problem (1.1),(1.2),(1.3).

Always in the framework of strong solutions, in Section 3, we extend the Lewy-Stampacchia inequalities to linear first order operators by considering the approximation by solutions of semilinear hyperbolic problems. In

Section 4, we show the stability of the set where the solution coincides with the obstacle, under a natural nondegeneracy condition on the obstacle and the nonhomogeneous term.

In Section 5 we extend to this case an estimate on the variation of the coincidence sets associated with the dependence on the data, including the variation of the (nondegenerating) obstacles. These results extend to first order obstacle problems the previous theory for second order linear operators (see [10], for instance) and can be applied to show the stability of the solution and coincidence set to the evolution first order obstacle problem with respect to the respective stationary one, as time goes to infinity. This is done in the final section.

§2. The Obstacle Problem of First Order

Let $Q \subset \mathbb{R}^N$ be an open domain and H be the linear first order operator defined on \bar{Q} by

$$Hu = \mathbf{b} \cdot \nabla u + b_0 u = Bu + b_0, u \tag{2.1}$$

where $b_0 = b_0(x) \in L^\infty(Q)$ and the vector field $B = \mathbf{b} \cdot \nabla = \sum_{i=1}^N b_i(x) \frac{\partial}{\partial x_i}$ has coefficients $b_i \in C^1(\bar{Q}) \cap W^{1,\infty}(Q)$ and the boundary $\Sigma = \partial Q$ is C^1 piecewise in the sense of [1], i.e., admits a decomposition where

$$\Sigma_+ = \{x \in \partial Q: \mathbf{b} \cdot \mathbf{n} > 0\} \quad \text{and} \quad \Sigma_- = \{x \in \partial Q: \mathbf{b} \cdot \mathbf{n} < 0\}$$

have a finite number of C^1 piecewise subboundaries of dimension $N - 2$. Here $\mathbf{n} = \mathbf{n}(x)$ is the outer normal vector at $x \in \Sigma = \partial Q$, defined almost everywhere. We define $\ell(x) = |\mathbf{b}(x) \cdot \mathbf{n}(x)|$ on Σ and we introduce the Hilbert spaces associated with the vector field B :

$$\begin{aligned} L_B^2(Q) &= \{v \in L^2(Q): Bv \in L^2(Q)\} \quad \text{and} \\ \tilde{L}_B^2(Q) &= \{v \in L_B^2(Q): v|_\Sigma \in L_\ell^2(\Sigma)\}, \end{aligned} \tag{2.2}$$

where $L_\ell^2(\Sigma) = \{v: \Sigma \rightarrow \mathbb{R} \mid \int_\Sigma v^2 \ell \, d\Sigma < \infty\}$. We recall from [1] that the graph norm in $\tilde{L}_B^2(Q)$ is equivalent to

$$\|v\|_{\tilde{L}_B^2}^2 = \|Bv\|_{L^2(Q)}^2 + \|v\|_{L_\ell^2(\Sigma_-)}^2, \tag{2.3}$$

that $C^1(\bar{Q})$ is dense in $\tilde{L}_B^2(Q)$ and the following integration by parts formula

holds in this space

$$\int_Q (\mathbf{b} \cdot \nabla u) v \, dx + \int_Q u \nabla \cdot (v \mathbf{b}) \, dx = \int_{\Sigma_+ \cup \Sigma_-} (\mathbf{b} \cdot \mathbf{n}) u v \, d\Sigma. \tag{2.4}$$

This framework allowed Bardos [1] to show the existence and uniqueness of the solution w in $\tilde{L}^2_B(Q)$ of the linear first order problem

$$Hw = f \text{ in } Q, \quad w = h \text{ on } \Sigma_- \tag{2.5}$$

under the coercivity assumption

$$b_0(x) - \frac{1}{2}(\nabla \cdot \mathbf{b})(x) \geq \beta > 0, \quad \forall x \in Q, \tag{2.6}$$

for any given data

$$f \in L^2(Q) \quad \text{and} \quad h \in L^2_t(\Sigma_-). \tag{2.7}$$

Consider now an obstacle $\psi = \psi(x)$ such that

$$\psi \in \tilde{L}^2_B(Q), \quad \psi \leq h \text{ on } \Sigma_-, \tag{2.8}$$

and introduce the non-empty convex subsets

$$K_\psi = \{v \in L^2(Q) : v \geq \psi \text{ a.e. in } Q\} \quad \text{and} \quad \tilde{K}_\psi = K_\psi \cap \tilde{L}^2_B(Q). \tag{2.9}$$

Following Mignot and Puel [8], we consider the strong formulation of the obstacle problem for the first order operator (2.1)

$$u \in \tilde{K}_\psi, \quad u|_{\Sigma_-} = h: \quad \int_Q (Hu - f)(v - u) \, dx \geq 0, \quad \forall v \in \tilde{K}_\psi. \tag{2.10}$$

Under the assumptions (2.7) and (2.8) the existence and uniqueness of a strong solution to the hyperbolic variational inequality (2.10) are shown in [8]. Actually the conditions on ψ in (2.8) can be taken in a weaker sense, since only the regularity $\tilde{\psi} = \sup(\psi, w) \in \tilde{L}^2_B(Q)$ and $\tilde{\psi} \leq h$ on Σ_- are necessary. Indeed, it was shown that if w and u are the solutions of (2.5) and (2.10), respectively, then $u \geq w$ a.e. in Q . So also $u \geq \tilde{\psi}$ and to solve the variational inequality (2.10) in \tilde{K}_ψ and in $\tilde{K}_{\tilde{\psi}}$ are equiv a lent problems, and (2.8) is, therefore, a natural assumption to obtain strong solutions.

We also recall the property of $\tilde{L}^2_B(Q)$ as a Dirichlet space, i.e., v^+, v^- and $|v| \in \tilde{L}^2_B(Q)$ if $v \in \tilde{L}^2_B(Q)$, which can be proved as in the Sobolev space $H^1(\Omega)$. Similarly, one has, for instance, $Bv^+ = Bv$ in $\{v > 0\}$ and $Bv^+ = 0$ in $\{v \leq 0\}$, provided $v \in \tilde{L}^2_B(Q)$, in the almost everywhere sense. We can also show for $v \in \tilde{L}^2_B(Q)$ that

$$Bv = 0 \quad \text{a.e. in } \{x \in Q : v(x) = 0\}. \tag{2.11}$$

Here we shall use the standard notations

$$u \vee v = \sup(u, v), \quad v^+ = v \vee 0 \quad \text{and} \quad v^- = (-v)^+.$$

We do not restrict the generality in taking

$$\psi = 0, \tag{2.12}$$

since we can reformulate the obstacle problem (2.10) into an equivalent one for the translated functions

$$\tilde{u} = u - \psi, \quad \tilde{h} = h - \psi|_{\Sigma_-} \quad \text{and} \quad \tilde{f} = f - H\psi. \tag{2.13}$$

Indeed, (2.10) is easily seen to be equivalent to

$$\tilde{u} \in \tilde{K}_0, \quad \tilde{u}|_{\Sigma_-} = \tilde{h}: \quad \int_Q (H\tilde{u} - \tilde{f})(v - \tilde{u}) \, dx \geq 0, \quad \forall v \in K_0. \tag{2.14}$$

We have the following continuous dependence estimate for strong solutions.

Proposition 2.1. *Let u_i denote the solution of (2.10) corresponding to the data f_i, h_i and ψ_i under the assumptions (2.6), (2.7), (2.8) for $i = 1, 2$, respectively. Then*

$$\|u_1 - u_2\|_{L^2(Q)} \leq C (\|f_1 - f_2\|_{L^2(Q)} + \|h_1 - h_2\|_{L^2_\ell(\Sigma_-)} + \|\psi_1 - \psi_2\|_{\sim}), \tag{2.15}$$

where $\|\cdot\|_{\sim}$ denotes the norm (2.3) and $C > 0$ is a constant independent of the data.

Proof. Using (2.13) we may assume $\psi_1 = \psi_2 = 0$, for $i = 1, 2$. We may take $v = \tilde{u}_2$ in the inequality for \tilde{u}_1 and $v = \tilde{u}_1$ in the one for \tilde{u}_2 . Setting $w = \tilde{u}_1 - \tilde{u}_2 = u_1 - u_2 - (\psi_1 - \psi_2)$ and denoting $\bar{f} = \tilde{f}_1 - \tilde{f}_2 = f_1 - f_2 - H(\psi_1 - \psi_2)$ and $\bar{h} = h_1 - h_2 - (\psi_1 - \psi_2)|_{\Sigma_-}$ we obtain

$$\int_Q w H w \, dx \leq \int_Q \bar{f} w \, dx \leq \frac{\beta}{2} \int_Q w^2 \, dx + \frac{1}{2\beta} \int_Q \bar{f}^2 \, dx. \tag{2.16}$$

On the other hand, from the coercivity condition (2.6), we find

$$\begin{aligned} \int_Q w H w \, dx &= \int_Q \left(b_0 - \frac{1}{2} \nabla \cdot \mathbf{b} \right) w^2 \, dx + \frac{1}{2} \int_{\Sigma_+ \cup \Sigma_-} (\mathbf{b} \cdot \mathbf{n}) w^2 \, d\Sigma \\ &\geq \beta \int_Q w^2 \, dx - \frac{1}{2} \int_{\Sigma_-} w^2 \ell \, d\Sigma, \end{aligned} \tag{2.17}$$

which combined with (2.16) yields

$$\beta \int_Q w^2 \, dx \leq \frac{1}{\beta} \int_Q \bar{f}^2 \, dx + \int_{\Sigma_-} \bar{h}^2 \ell \, d\Sigma.$$

This implies the conclusion (2.15) by the definitions and the equivalence of the norm $\|\cdot\|_{\sim}$ in $\tilde{L}^2_B(Q)$.

§3. An Inequality for Strong Solutions

The strong solution of the obstacle problem, solving the first order variational inequality, is also a solution to the nonlinear complementary problem

$$u \geq \psi, \quad Hu - f \geq 0 \quad \text{and} \quad (Hu - f)(u - \psi) = 0 \quad \text{a.e. in } Q. \quad (3.1)$$

Indeed, it suffices to take $v = \psi$ and $v = 2u - \psi$ in (2.10) to conclude the third condition from the first two. The second one, which follows from (2.10) with $v = u + w$ for arbitrary $w \in L^2(Q)$, $w \geq 0$, provides a lower bound for Hu . The aim of this section is to show an upper bound, extending to H the well-known Lewy–Stampacchia inequalities obtained first for second order obstacle problem of elliptic type (see [10], for references).

Theorem 3.1. *Under the assumptions (2.6), (2.7), (2.8) the strong solution u of the first order obstacle problem (2.10) satisfies the inequalities*

$$f \leq Hu \leq f \vee H\psi \quad \text{a.e. in } Q. \quad (3.2)$$

The proof of this result follows easily by recalling that $f \vee H\psi = f + (H\psi - f)^+$ and the fact that u can be approximated in $L^2(Q)$ by the solution $u_\varepsilon \in \tilde{L}_B^2(Q)$ of the semilinear first order equation

$$Hu_\varepsilon + \xi \vartheta_\varepsilon(u_\varepsilon - \psi) = f + \xi \quad \text{in } Q, \quad u_\varepsilon = h \quad \text{on } \Sigma_-. \quad (3.3)$$

Here we consider for each $\varepsilon > 0$, the nondecreasing Lipschitz function $\vartheta_\varepsilon: \mathbb{R} \rightarrow [0, 1]$ defined by

$$\vartheta_\varepsilon(t) = 0, \quad t \leq 0, \quad \vartheta_\varepsilon(t) = t/\varepsilon, \quad 0 < t \leq \varepsilon \quad \text{and} \quad \vartheta_\varepsilon(t) = 1, \quad t > \varepsilon, \quad (3.4)$$

and the nonnegative function $\xi \in L^2(Q)$ given by

$$\xi = (H\psi - f)^+. \quad (3.5)$$

We can prove the very precise approximation result.

Theorem 3.2. *If u and u_ε denote the solutions of (2.10) and (3.3), respectively, under the previous assumptions we have*

$$u_\varepsilon \in \tilde{K}_\psi, \quad (3.6)$$

$$u_\varepsilon \geq u_{\hat{\varepsilon}} \quad \text{in } Q \quad \text{if } \varepsilon > \hat{\varepsilon} > 0, \quad (3.7)$$

$$\|u_\varepsilon - u\|_{L^2(Q)}^2 \leq \frac{\varepsilon}{\beta} \|(H\psi - f)^+\|_{L^1(Q)} \quad \text{as } \varepsilon \rightarrow 0. \quad (3.8)$$

Proof. Since ϑ_ε is monotone and H is coercive, the existence and uniqueness of $u_\varepsilon \in \tilde{L}_B^2(Q)$ follows by the results of Bardos [1].

To prove (3.6), we must show that $u_\varepsilon \geq \psi$ in Q . Take $z = (\psi - u_\varepsilon)^+ \in \tilde{L}^2_B(Q)$ and note that by (2.8) we have $z|_{\Sigma_-} = 0$. Since $\vartheta_\varepsilon(u_\varepsilon - \psi) = 0$ whenever $\psi > u_\varepsilon$ we obtain $Hu_\varepsilon = f + \xi \geq H\psi$ if $\psi > u_\varepsilon$ and

$$\int_Q z Hz dx = \int_Q (H\psi - Hu_\varepsilon)(\psi - u_\varepsilon)^+ dx \leq 0,$$

since $z Hz = (\psi - u_\varepsilon)^+ H(\psi - u_\varepsilon)$ a.e. in Q . Hence, using (2.17)

$$0 \geq \int_Q z Hz dx \geq \beta \int_Q z^2 dx$$

we conclude $z = 0$ a.e. in Q and (3.6) follows.

A similar argument applies to $z = (u_\varepsilon - u_\varepsilon)^+$, by using

$$z Hz = z H(u_\varepsilon - u_\varepsilon) = z \xi [\vartheta_\varepsilon(u_\varepsilon - \psi) - \vartheta_\varepsilon(u_\varepsilon - \psi)] \leq 0 \quad \text{in } Q,$$

since if $u_\varepsilon > u_\varepsilon$, then $\vartheta_\varepsilon(u_\varepsilon - \psi) \geq \vartheta_\varepsilon(u_\varepsilon - \psi) \geq \vartheta_\varepsilon(u_\varepsilon - \psi)$.

Finally, remarking that for any $v \in K_\psi$ we have

$$[1 - \vartheta_\varepsilon(u_\varepsilon - \psi)](v - u_\varepsilon) \geq [1 - \vartheta_\varepsilon(u_\varepsilon - \psi)](\psi - u_\varepsilon) \geq -\varepsilon,$$

we first obtain

$$\int_Q (Hu_\varepsilon - f)(v - u_\varepsilon) dx = \int_Q \xi [1 - \vartheta_\varepsilon(u_\varepsilon - \psi)](v - u_\varepsilon) dx \geq -\varepsilon \int_Q \xi dx. \tag{3.9}$$

Setting $v = u$ in (3.9) and $v = u_\varepsilon$ in (2.10), we conclude (3.8) with the help of (2.17) for $w = u_\varepsilon - u$

$$\beta \int_Q w^2 dx \leq \int_Q w Hw dx = \int_Q (u_\varepsilon - u) H(u_\varepsilon - u) dx \leq \varepsilon \int_Q \xi dx.$$

Remark 3.1. The proof of Theorem 3.2 actually also shows the existence of the solution u to (2.10), since $0 \leq \vartheta_\varepsilon \leq 1$ implies that the approximating solution u_ε of (3.3) are bounded in $W = \{v \in \tilde{L}^2_B(Q) : v|_{\Sigma_-} = h\}$, uniformly in $\varepsilon > 0$. Hence the lower semi-continuity of

$$w \mapsto \int_Q w Hw dx \quad \text{in } W$$

allows to pass to the limit $\varepsilon \rightarrow 0$ in (3.9) by showing that if $u_\varepsilon \rightharpoonup u$ in $\tilde{L}^2_B(Q)$ then u solves (2.10). By uniqueness, which is a consequence of the coercivity in W , the whole sequence converges.

Remark 3.2. In [8] the existence of a strong solution was obtained with a different approximation (see [6]) by considering the penalized problem for

$\varepsilon > 0$,

$$Hw_\varepsilon - \frac{1}{\varepsilon}(\psi - w_\varepsilon)^+ = f \quad \text{in } Q,$$

with the same boundary condition $w_\varepsilon = h$ on Σ . While this is a more natural way to penalize the constraint $u \geq \psi$, this method does not allow to conclude the second inequality in (3.2).

§4. Stability of the Coincidence Set

A main feature in the obstacle problem is the presence, in general, of the coincidence set

$$I = \{u = \psi\} = \{x \in Q : u(x) = \psi(x)\}. \quad (4.1)$$

In the complementary set Λ of this measurable subset, from (3.1), we have

$$Hu = f \quad \text{a.e. in } \Lambda = \{u > \psi\} = Q \setminus I. \quad (4.2)$$

It is clear that, in general, I and Λ are measurable subsets defined up to a null set. This is however sufficient for our purposes in this work, since we are interested in their characteristic functions. Set

$$\chi = \chi_{(u=\psi)} = \begin{cases} 1 & \text{if } x \in \{u = \psi\}, \\ 0 & \text{if } x \in \{u > \psi\}. \end{cases} \quad (4.3)$$

As a consequence of (4.2) and property (2.11), we may conclude that the solution u of (2.10) solves the equation

$$Hu - (H\psi - f)\chi = f \quad \text{a.e. in } Q. \quad (4.4)$$

This important remark allows us to include the first order obstacle problem in the general framework of stability of the coincidence set with respect to perturbation of data (see [10, p. 204], for the elliptic theory).

Theorem 4.1. *Suppose u_n and $\chi_n = \chi_{\{u_n = \psi_n\}}$ denote the solution of (2.10) and the characteristic function of its coincidence set associated with a sequence f_n , h_n and ψ_n satisfying (2.7), (2.8) and*

$$f_n \rightarrow f \text{ in } L^2(Q), \quad h_n \rightarrow h \text{ in } L^2_\ell(\Sigma_-) \quad \text{and} \quad \psi_n \rightarrow \psi \text{ in } \tilde{L}^2_B(Q).$$

If u and χ refer to the corresponding limit problem in which we assume

$$H\psi \neq f \quad \text{a.e. in } Q, \quad (4.5)$$

then the coincidence sets converge in measure, or equivalently

$$\chi_n \rightarrow \chi \quad \text{in } L^p(Q), \quad 1 \leq p < \infty. \tag{4.6}$$

Proof. We remark $0 \leq \chi_n \leq 1$, so that there is a function $\chi_* \in L^\infty(Q)$, $0 \leq \chi_* \leq 1$, and a subsequence

$$\chi_n \rightharpoonup \chi_* \quad \text{in } L^\infty(Q)\text{-weak*}.$$

By Proposition 2.1, we know that

$$u_n \rightarrow u \quad \text{in } L^2(Q),$$

and, from remark (4.4) for u_n ,

$$Hu_n - (H\psi_n - f_n)\chi_n = f_n \quad \text{a.e. in } Q.$$

So we may pass to the limit and obtain

$$Hu - (H\psi - f)\chi_* = f \quad \text{a.e. in } Q. \tag{4.7}$$

Comparing (4.7) with (4.4) and using the assumption (4.5) we immediately conclude

$$\chi_* = \chi = \chi_{\{u=\psi\}},$$

i.e., the whole sequence converges $\chi_n \rightarrow \chi$ first weakly in $L^p(Q)$ and, since they are characteristic functions, also strongly for any $p < \infty$.

Remark 4.1. As a consequence of Theorem 4.1, we can immediately conclude also that $u_n \rightarrow u$ in $\tilde{L}^2_B(Q)$ -strong under the assumption (4.5), which however is not necessary, as we shall see in Theorem 5.1.

On the other hand, we know that u is approximated by the solution u_ε of (3.3), i.e., we have

$$Hu_\varepsilon - (H\psi - f)^+ q_\varepsilon = f \quad \text{in } Q, \tag{4.8}$$

where we set

$$0 \leq q_\varepsilon \equiv 1 - \vartheta_\varepsilon(u_\varepsilon - \psi) \leq \chi_\varepsilon \leq 1 \quad \text{a.e. in } Q. \tag{4.9}$$

Here we have introduced χ_ε as the characteristic function of the ‘‘approximating coincidence set’’

$$I_\varepsilon = \{x \in Q : \psi(x) \leq u_\varepsilon(x) < \psi(x) + \varepsilon\}.$$

To prove (4.9) it is sufficient to recall the definition of ϑ_ε : since $u_\varepsilon \geq \psi$ always, if $u_\varepsilon(x) \geq \psi(x) + \varepsilon$ (i.e. $\chi_\varepsilon(x) = 0$), then $q_\varepsilon(x) = 0$.

As $\varepsilon \rightarrow 0$, we may consider subsequences such that

$$q_\varepsilon \rightharpoonup q \quad \text{and} \quad \chi_\varepsilon \rightharpoonup \chi_* \quad \text{in } L^\infty(Q)\text{-weak*} \quad (4.10)$$

for some functions q and χ_* such that

$$0 \leq q \leq \chi_* \leq 1 \quad \text{a.e. in } Q. \quad (4.11)$$

From (4.8) we find

$$Hu - (H\psi - f)^+ q = f \quad \text{in } Q. \quad (4.12)$$

In the coincidence set $I = \{u = \psi\}$ we have $Hu = H\psi$ a.e. and, if we assume $H\psi \neq f$, from (4.12) we must have $q = 1$ in I , since we have always $Hu \geq f$ by (3.1). Therefore the nondegeneracy condition (4.5) implies

$$q \geq \chi = \chi_{\{u=\psi\}} \quad \text{a.e. in } Q. \quad (4.13)$$

But the definition of q_ε and (4.10) with the convergence of $u_\varepsilon \rightarrow u$ in $L^2(Q)$ yield as $\varepsilon \rightarrow 0$,

$$0 = (u_\varepsilon - \psi - \varepsilon)^+ q_\varepsilon \rightharpoonup (u - \psi)^+ q = 0,$$

and this implies $q = 0$ if $u > \psi$, i.e.,

$$q \leq \chi = \chi_{\{u=\psi\}} \quad \text{a.e. in } Q. \quad (4.14)$$

Then (4.11), (4.13) and (4.14) imply

$$q = \chi_* = \chi = \chi_{\{u=\psi\}}. \quad (4.15)$$

By (4.9) we remark $q_\varepsilon^2 \leq q_\varepsilon$ and from

$$\int_Q q = \lim_{\varepsilon \rightarrow 0} \int_Q q_\varepsilon \geq \liminf_{\varepsilon \rightarrow 0} \int_Q q_\varepsilon^2 \geq \int_Q q^2 = \int_Q \chi$$

we may conclude the strong convergences as $\varepsilon \rightarrow 0$

$$\chi_\varepsilon \rightarrow \chi \quad \text{and} \quad q_\varepsilon \rightarrow \chi \quad \text{in } L^p(Q)\text{-strong, } \forall p < \infty.$$

Then (4.8) implies also $Hu_\varepsilon \rightarrow Hu$ in $L^2(Q)$ -strong, and we have proven the following result on the strong approximation of the first order obstacle problem by solutions of semilinear hyperbolic problems (3.3).

Theorem 4.2. *Let u_ε and u denote the solutions of (2.10) and (3.3) respectively, under the nondegeneracy assumption (4.5).*

Then, as $\varepsilon \rightarrow 0$, we have

$$u_\varepsilon \rightarrow u \quad \text{in } \tilde{L}_B^2(Q)\text{-strong}$$

and

$$\lim_{\varepsilon \rightarrow 0} [1 - \vartheta_\varepsilon(u_\varepsilon - \psi)] = \lim_{\varepsilon \rightarrow 0} \chi_{\{\psi \leq u_\varepsilon < \psi + \varepsilon\}} = \chi_{\{u = \psi\}}$$

for the strong topologies of $L^p(Q)$, $\forall p, 1 \leq p < \infty$.

Remark 4.2. We observe that here the nondegeneracy assumption $H\psi \neq f$ a.e. in Q is required, as in Theorem 4.1 on the stability of the coincidence sets. Analogously Theorem 4.2 yields a stability in Lebesgue measure of the approximation of the coincidence set $\{u = \psi\}$, i.e., we have

$$\{\psi \leq u_\varepsilon < \psi + \varepsilon\} \rightarrow \{u = \psi\} \text{ in measure.}$$

§5. An Estimate on the Coincidence Set

Let $\beta > 0$ be the constant of (2.6) and $\alpha > 0$, such that

$$\frac{1}{2} |(\nabla \cdot \mathbf{b})(x)| \leq \alpha, \quad \forall x \in Q.$$

Denote by S the monotone graph corresponding to the sign function, i.e.,

$$S(t) = 1 \text{ if } t > 0, \quad S(t) = -1 \text{ if } t < 0 \quad \text{and} \quad S(0) = [-1, 1].$$

Lemma 5.1. For any $w \in \tilde{L}^2_B(Q)$ and any measurable function s such that $s(x) \in S(w(x))$, a.e. $x \in Q$, we have

$$\int_Q s Hw \, dx \geq (\beta - \alpha) \int_Q |w| \, dx - \int_{\Sigma_-} |w| \ell \, d\Sigma. \tag{5.1}$$

Proof. By the property (2.11), we remark that we have

$$s Hw = \text{sign}(w) Hw \quad \text{a.e. in } Q, \tag{5.2}$$

where $\text{sign}(t) = 1$ if $t > 0$, $\text{sign}(t) = -1$ if $t < 0$ and $\text{sign}(0) = 0$. Hence it is sufficient to prove (5.1) with s replaced by $\text{sign}(w)$, which can be approximated in $L^2(Q)$ by the sequence of functions $s_\delta(w) \in \tilde{L}^2_B(Q)$, where $s_\delta(t)$ are smooth functions approximating the sign, such that $|s_\delta(t)| \leq 1$, $s'_\delta \geq 0$, $s_\delta(0) = 0$ and $s_\delta(t) \rightarrow \text{sign}(t)$ as $\delta \rightarrow 0$ for $t \in \mathbb{R}$.

Integrating by parts and setting $m_\delta(t) = \int_0^t s_\delta(\tau) \, d\tau$ we have

$$\begin{aligned} \int_Q s_\delta(w) Hw \, dx &= \int_Q [b_0 s_\delta(w) w + \mathbf{b} \cdot \nabla m_\delta(w)] \, dx \\ &= \int_Q [b_0 s_\delta(w) w - (\nabla \cdot \mathbf{b}) m_\delta(w)] \, dx + \int_{\Sigma_+ \cup \Sigma_-} (\mathbf{b} \cdot \mathbf{n}) m_\delta(w) \, d\Sigma. \end{aligned} \tag{5.3}$$

Noting that $m_\delta(w) \rightarrow |w|$ in $L^2(Q)$ and in $L^2_\ell(\Sigma)$ as $\delta \rightarrow 0$, from (5.3) we obtain

$$\begin{aligned} & \int_Q \text{sign}(w) H(w) dx \\ &= \int_Q \left[b_0 - (\nabla \cdot \mathbf{b}) \right] |w| dx + \int_{\Sigma_+ \cup \Sigma_-} (\mathbf{b} \cdot \mathbf{n}) |w| d\Sigma \\ &\geq \int_Q \left[b_0 - \frac{1}{2}(\nabla \cdot \mathbf{b}) \right] |w| dx - \frac{1}{2} \int_Q (\nabla \cdot \mathbf{b}) |w| dx + \int_{\Sigma_-} (\mathbf{b} \cdot \mathbf{n}) |w| d\Sigma \\ &\geq (\beta - \alpha) \int_Q |w| dx - \int_{\Sigma_-} |w| \ell d\Sigma \end{aligned}$$

and (5.1) follows from (5.2).

Remark 5.1. If $\beta \geq \alpha$ or if we assume instead

$$b_0(x) - \nabla \cdot \mathbf{b}(x) \geq 0 \quad \text{a.e. in } Q, \tag{5.4}$$

the estimate (5.1) reduces to

$$\int_Q s Hw dx \geq - \int_{\Sigma_-} |w| \ell d\Sigma, \quad \forall w \in \tilde{L}^2_B(Q), \tag{5.5}$$

provided $s(x) \in S(w(x))$ a.e. $x \in Q$.

These estimates may be used “to measure” the stability of the coincidence set in the nondegenerate case.

We recall the Lewy-Stampacchia type inequality (3.2) in the form

$$0 \leq Hu - f \leq (H\psi - f)^+ \quad \text{a.e. in } Q$$

and, recalling that the solution u of the obstacle problem also solves the equation (4.4), we have $\zeta = \zeta(u) \geq 0$, where

$$\zeta = Hu - f = (H\psi - f) \chi_{\{u=\psi\}} = (H\psi - f)^+ \chi_{\{u=\psi\}}. \tag{5.6}$$

Lemma 5.2. *Let u_i for $i = 1, 2$ denote the solution to (2.10) for data f_i, h_i and ψ_i under the assumptions (2.6), (2.7), (2.8) respectively and set $\zeta_i = \zeta(u_i)$. Then*

$$\|\zeta_1 - \zeta_2\|_{L^1(Q)} \leq C_1 (\|f_1 - f_2\|_{L^2(Q)} + \|h_1 - h_2\|_{L^2_\ell(\Sigma_-)} + \|\psi_1 - \psi_2\|_{\sim}), \tag{5.7}$$

where $C_1 > 0$ is a constant independent of the data.

Proof. As in Proposition 2.1, by using the translation argument, we may assume $\psi_1 = \psi_2 = 0$ without loss of generality.

From (5.6) for $i = 1, 2$ we obtain

$$\zeta_1 - \zeta_2 = H(u_1 - u_2) - (f_1 - f_2) \quad \text{a.e. in } Q. \tag{5.8}$$

We define almost everywhere in Q the measurable function s by

$$s(x) = \begin{cases} -1 & \text{on } \{u_1 < u_2\} \cup \{\zeta_2 < \zeta_1\}, \\ 0 & \text{on } \{u_1 = u_2\} \cap \{\zeta_1 = \zeta_2\}, \\ 1 & \text{on } \{u_1 > u_2\} \cup \{\zeta_2 > \zeta_1\}, \end{cases} \tag{5.9}$$

and we observe that $s \in S(u_1 - u_2)$ a.e. in Q . Indeed, if $\zeta_2 > \zeta_1 \geq 0$, by (5.6) and (3.1) we have $u_2 = 0$ and the subset $\{u_2 > u_1\} \cap \{\zeta_1 < \zeta_2\}$ cannot have positive measure. Similarly the same conclusion holds for $\{u_2 < u_1\} \cap \{\zeta_1 > \zeta_2\}$ and s given by (5.9) is a.e. well-defined.

Multiplying (5.8) by s and using (5.1) with $w = u_1 - u_2$, we obtain

$$\begin{aligned} & \int_Q |\zeta_1 - \zeta_2| dx \\ &= \int_Q (\zeta_2 - \zeta_1) s dx = - \int_Q s Hw dx + \int_Q s(f_1 - f_2) dx \\ &\leq (\alpha - \beta) \int_Q |u_1 - u_2| dx + \int_{\Sigma_-} |h_1 - h_2| \ell d\Sigma + \int_Q |f_1 - f_2| dx \end{aligned}$$

and using the estimate (2.15) we easily conclude (5.7).

Remark 5.2. Under the assumption (5.4) (or if $\alpha \leq \beta$) when $\psi_1 = \psi_2 = 0$, we may improve the estimate (5.7) by exactly the simpler one

$$\|\zeta_1 - \zeta_2\|_{L^1(Q)} \leq \|f_1 - f_2\|_{L^1(Q)} + \|h_1 - h_2\|_{L^1_2(\Sigma_-)}, \tag{5.10}$$

as a simple consequence of Remark 5.1 and the above proof.

As an immediate consequence of (5.7) we have the strong continuous dependence in $\tilde{L}^2_B(Q)$ of the first order obstacle problem with respect to the data.

Theorem 5.1. *If we assume in the obstacle problem (2.10)*

$$f_n \rightarrow f \text{ in } L^2(Q), \quad h_n \rightarrow h \text{ in } L^2_B(\Sigma_-) \quad \text{and} \quad \psi_n \rightarrow \psi \text{ in } \tilde{L}^2_B(Q),$$

the respective strong solutions satisfy the strong convergence

$$u_n \rightarrow u \quad \text{in } \tilde{L}^2_B(Q).$$

Perhaps a more interesting consequence of (5.7) can be obtained, exactly as in the elliptic theory of [10], for estimating locally the Lebesgue measure of the variation of the coincidence set associated with different data, under the local nondegeneracy assumption in an arbitrary measurable subset $\mathcal{O} \subset Q$:

$$f_1 - H\psi_1 \leq -\lambda < 0 \quad \text{and} \quad f_2 - H\psi_2 \leq -\lambda < 0 \quad \text{a.e. on } \mathcal{O}. \tag{5.11}$$

For the coincidence subsets $I_1 = \{u_1 = \psi_1\}$ and $I_2 = \{u_2 = \psi_2\}$ we denote by \div the symmetric difference

$$I_1 \div I_2 = I_1 \setminus I_2 \cup I_2 \setminus I_1,$$

where $A \setminus B = A \cap B^C$ as usual.

Theorem 5.2. *Under the assumption (5.11), we have*

$$\text{meas}\left((I_1 \div I_2) \cap \mathcal{O}\right) \leq \frac{C_1}{\lambda} (\|f_1 - f_2\|_{L^2(Q)} + \|h_1 - h_2\|_{L^2_\zeta(\Sigma_-)} + \|\psi_1 - \psi_2\|_{\sim}). \tag{5.12}$$

Proof. It suffices to remark from (5.6) that (5.11) implies

$$\lambda |\chi_{\{u_1=\psi_1\}} - \chi_{\{u_2=\psi_2\}}| \leq |\zeta_1 - \zeta_2| \quad \text{a.e. in } \mathcal{O}$$

and, using (5.7), (5.12) follows from

$$\text{meas}((I_1 \div I_2) \cap \mathcal{O}) = \int_{\mathcal{O}} |\chi_{\{u_1=\psi_1\}} - \chi_{\{u_2=\psi_2\}}| dx \leq \frac{1}{\lambda} \|\zeta_1 - \zeta_2\|_{L^1(Q)}.$$

Remark 5.3. Under the additional assumptions (5.4) and $\psi_1 = \psi_2 = 0$, (5.12) reduces to

$$\text{meas}(\mathcal{O} \cap (\{u_1 = 0\} \div \{u_2 = 0\})) \leq \frac{1}{\lambda} (\|f_1 - f_2\|_{L^1(Q)} + \|h_1 - h_2\|_{L^2_\zeta(\Sigma_-)}).$$

§6. The Stability of the Evolution Problem

In this section we set $Q = \Omega \times]0, T[$, $T > 0$, and $\Sigma' = \partial\Omega \times]0, T[$, with the assumptions of Section 2 and where $\Omega \subset \mathbb{R}^n$ with $N = n + 1$.

Then $\Sigma = \partial Q = \Sigma' \cup \Omega_0 \cup \Omega_T$ ($\Omega_k = \Omega \times \{k\}$, $k = 0, T$) and we redefine $x = (x_1, \dots, x_n) \in \bar{\Omega}$, $x_N = t \in [0, T]$, $b_i = a_i$, $i = 0, 1, \dots, n$, and we set $b_N = 1$. Then the first order evolutionary operator becomes with $\partial_t = \partial/\partial t$:

$$Hu = \partial_t u + \mathbf{a} \cdot \nabla u + a_0 u = \partial_t u + Au,$$

where the coefficients a_i , $i = 1, \dots, n$ belong to $C^1(\bar{Q})$, $a_0 \in L^\infty(Q)$ and may depend on t but do not satisfy necessarily the coercivity assumption (2.6). We still define $\ell = \ell(x, t) = \left| \sum_{i=1}^n n_i a_i(x, t) \right|$ along Σ' , with the external normal \mathbf{n} to Ω , and analogously the norm of $\tilde{L}^2_B(Q)$ is given by (2.3), where now

$$\|u\|_{L^2_\zeta(\Sigma_-)}^2 = \|u(x, 0)\|_{L^2(\Omega)}^2 + \|u|_{\Sigma'_-}\|_{L^2_\zeta(\Sigma'_-)}^2$$

since $\Sigma_- = \Omega_0 \cup \Sigma'_-$, with $\Sigma'_- = \{(x, t) \in \Sigma: \mathbf{a}(x, t) \cdot \mathbf{n} < 0\}$.

The strong formulation of the evolutionary first order obstacle problem can now be rewritten in the form

$$u \in \tilde{K}_\psi, \quad u|_{\Sigma'_- \cup \Omega_0} = h: \int_Q (\partial_t u + Au - f)(v - u) dx dt \geq 0, \quad \forall v \in K_\psi. \quad (6.1)$$

Here K_ψ and \tilde{K}_ψ are given also by (2.9) and

$$h|_{\Sigma'_-} = g \in L^2_\ell(\Sigma'_-) \quad \text{and} \quad h|_{\Omega_0} = u_0 \in L^2(\Omega) \quad (6.2)$$

for $g = g(x, t)$ and $u_0 = u_0(x)$ compatible with the obstacle in the sense of (2.8).

It is clear that all the results of the preceding sections still hold for the solution of (6.1) as a consequence of the following proposition.

Proposition 6.1. *Under the preceding assumptions (2.7), (2.8), (6.2) the unique strong solution of (6.1) satisfies the estimates (2.15) and (3.2).*

Proof. If the operator $H = \partial_t + A$ does not satisfy the condition (2.6), we consider a constant $\mu > 0$ such that, for all $t \in [0, T]$,

$$\mu + a_0 - \frac{1}{2}(\nabla \cdot \mathbf{a}) \geq \beta > 0, \quad \forall x \in \Omega. \quad (6.3)$$

Setting $u = e^{\mu t} \hat{u}$ it is easy to see that \hat{u} solve the coercive problem

$$\hat{u} \in \tilde{K}_{\hat{\psi}}, \quad \hat{u}|_{\Sigma'_- \cup \Omega_0} = \hat{h}: \int_Q (\partial_t \hat{u} + A\hat{u} + \mu\hat{u} - \hat{f})(v - \hat{u}) dx dt \geq 0, \quad \forall v \in K_{\hat{\psi}}$$

with $\hat{f} = e^{-\mu t} f$, $\hat{h} = e^{-\mu t} h$ and $\hat{\psi} = e^{-\mu t} \psi$, for which all previous results apply.

When Q is a cylinder, we may use the integration by parts formula (2.4) in a subset $Q_{\sigma, t} = \Omega \times]\sigma, t[$, $0 \leq \sigma < t < T$, in the following form for any $w \in \tilde{L}^2_B(Q_{\sigma, t})$ and $\lambda \in \mathbb{R}$:

$$\begin{aligned} \int_\sigma^t \int_\Omega w(\partial_t w + Aw) e^{\lambda t} dx dt &= \int_\sigma^t \int_\Omega \left(a_0 - \frac{1}{2} \nabla \cdot \mathbf{a} - \frac{\lambda}{2} \right) w^2 e^{\lambda t} dx dt \\ &+ \frac{1}{2} \int_\sigma^t \int_{\partial\Omega} (\mathbf{a} \cdot \mathbf{n}) w^2 e^{\lambda t} d\Gamma + \frac{1}{2} \int_\Omega [w^2(\tau) e^{\lambda \tau} - w^2(\sigma) e^{\lambda \sigma}] dx. \end{aligned} \quad (6.4)$$

In order to consider the stability of the evolutionary problem as $t \rightarrow \infty$ we need to consider first the corresponding stationary problem in Ω .

We shall assume from now on that the coefficients of A are time independent, i.e.,

$$a_i = a_i(x) \in C^1(\bar{\Omega}), \quad i = 1, \dots, n, \quad \text{and} \quad a_0 = a_0(x) \in L^\infty(\Omega), \quad (6.5)$$

and coercive, i.e., satisfying (6.3) with $\mu = 0$.

Decomposing $\Gamma = \partial\Omega$ as in Section 2, we set $\Gamma_- = \{u \in \partial\Omega : \mathbf{a} \cdot \mathbf{n} < 0\}$, $\ell(x) = |\mathbf{a}(x) \cdot \mathbf{n}(x)|$ and we consider the Hilbert spaces

$$L_A^2(\Omega) = \{v \in L^2(\Omega) : Av \in L^2(\Omega)\}$$

and

$$\tilde{L}_A^2(\Omega) = \{v \in L_A^2(\Omega) : v|_{\Gamma} \in L_\ell^2(\Gamma)\}$$

and we also consider in $\tilde{L}_A^2(\Omega)$ the norm

$$\|v\|_{\#}^2 = \|Av\|_{L^2(\Omega)}^2 + \|v\|_{L_\ell^2(\Gamma_-)}^2. \tag{6.6}$$

For the stationary problem, we assume

$$f_{\#} \in L^2(\Omega), \quad g_{\#} \in L_\ell^2(\Gamma_-) \quad \text{and} \quad \psi_{\#} \in \tilde{L}_A^2(\Omega) \quad \text{with} \quad \psi_{\#} \leq g_{\#} \quad \text{on} \quad \Gamma_-, \tag{6.7}$$

and we consider the convex sets

$$K'_{\psi_{\#}} = \{v \in L^2(\Omega) : v \geq \psi_{\#} \text{ a.e. in } \Omega\} \quad \text{and} \quad \tilde{K}'_{\psi_{\#}} = K'_{\psi_{\#}} \cap \tilde{L}_A^2(\Omega). \tag{6.8}$$

Under the coercivity assumption (6.3) with $\mu = 0$ all the results in the previous sections also apply to the first order stationary problem:

$$u_{\#} \in \tilde{K}'_{\psi_{\#}}, \quad u_{\#}|_{\Gamma_-} = g_{\#} : \int_{\Omega} (Au_{\#} - f_{\#})(v - u_{\#}) dx \geq 0, \quad \forall v \in K'_{\psi_{\#}}. \tag{6.9}$$

Let, for $t \in [0, \infty[$,

$$\xi(t) = \int_t^{t+1} \left\{ \int_{\Omega} |f(\tau) - f_{\#}|^2 dx + \int_{\Gamma_-} |g(\tau) - g_{\#}|^2 \ell d\Gamma \right\} d\tau, \tag{6.10}$$

$$\eta(t) = \int_t^{t+1} \left\{ \int_{\Omega} |(\partial_t + A)\psi(t) - A\psi_{\#}|^2 dx + \int_{\Gamma_-} |\psi(t) - \psi_{\#}|^2 \ell d\Gamma \right\} d\tau. \tag{6.11}$$

Theorem 6.1. *Assuming (6.5) and (6.3) with $\mu = 0$, if $\xi(t) + \eta(t) \rightarrow 0$ as $t \rightarrow \infty$, then the solution $u(t)$ of (6.1) is asymptotically stable in the sense*

$$u(t) \rightarrow u_{\#} \text{ in } L^2(\Omega) \quad \text{as } t \rightarrow \infty, \tag{6.12}$$

where $u_{\#}$ is the unique strong solution of (6.9).

Proof. Using the translation argument we may assume that $\psi = \psi_{\#} = 0$ without loss of generality, since the assumption (6.11) reduces to (6.10) for the corresponding translated data.

Since $u(t)$ and $u_{\#}$ satisfy the complementary problem (3.1), by integration in Ω first and, afterwards multiplication by $e^{\beta t}$, we easily obtain for $w(t) = u(t) - u_{\#}$ and $\tau > \sigma \geq 0$,

$$\begin{aligned} & \int_{\sigma}^{\tau} \int_{\Omega} w(\partial_t w + Aw) e^{\beta t} dx dt \leq \int_{\sigma}^{\tau} \int_{\Omega} w(f(t) - f_{\#}) e^{\beta t} dx dt \\ & \leq \frac{\beta}{2} \int_{\sigma}^{\tau} \int_{\Omega} w^2 e^{\beta t} dx dt + \frac{1}{2\beta} \int_{\sigma}^{\tau} \int_{\Omega} |f(t) - f_{\#}|^2 e^{\beta t} dx dt. \end{aligned} \tag{6.13}$$

On the other hand, using the formula (6.4) with $\lambda = \beta$ and neglecting the nonnegative terms, by (6.3) with $\mu = 0$, we obtain

$$\begin{aligned} & \int_{\sigma}^{\tau} \int_{\Omega} w(\partial_t w + Aw) e^{\beta t} dx dt \\ & \geq \frac{\beta}{2} \int_{\sigma}^{\tau} \int_{\Omega} w^2 e^{\beta t} dx dt - \frac{1}{2} \int_{\sigma}^{\tau} \int_{\Gamma_-} w^2 e^{\beta t} \ell d\Gamma dt \\ & \quad + \frac{1}{2} \int_{\Omega} [w^2(\tau) e^{\beta \tau} - w^2(\sigma) e^{\beta \sigma}] dx. \end{aligned} \tag{6.14}$$

Hence, combining (6.13) with (6.14) we have for all $\tau > \sigma \geq 0$,

$$\begin{aligned} & e^{\beta \tau} \int_{\Omega} w^2(\tau) dx - e^{\beta \sigma} \int_{\Omega} w^2(\sigma) dx \\ & \leq \int_{\sigma}^{\tau} \left\{ \frac{1}{\beta} \int_{\Omega} |f(t) - f_{\#}|^2 dx + \int_{\Gamma_-} |g(t) - g_{\#}|^2 \ell d\Gamma \right\} e^{\beta t} dt, \end{aligned} \tag{6.15}$$

which implies the estimate for all $t > 0, t_0 \geq 0$:

$$\int_{\Omega} |u - u_{\#}|^2(t + t_0) dx \leq e^{-\beta t} \int_{\Omega} |u - u_{\#}|^2(t_0) dx + C_{\beta} \sup_{t_0 < \tau < t+t_0} \xi(\tau), \tag{6.16}$$

where $\beta > 0$ and $C_{\beta} = (1 \vee \frac{1}{\beta}) [1 + (1 - e^{-\beta})^{-1}]$, by well-known results (see Remark 6.1, below).

From (6.16) with $t_0 = 0$ we obtain first that $u - u_{\#}$ is bounded in $L^{\infty}(0, \infty; L^2(\Omega))$ and, afterwards, that $\xi(t) \rightarrow 0$ as $t \rightarrow \infty$ implies the conclusion

$$\|u(t) - u_{\#}\|_{L^2(\Omega)} \rightarrow 0 \quad \text{as } t \rightarrow +\infty.$$

Remark 6.1. This type of global behaviour of solutions is similar to other nonlinear evolution equations (see, for instance, [9]), in particular, in monotone parabolic variational inequalities. The passage of (6.15) to (6.16) follows by the elementary standard estimate (after changing variables)

$$\begin{aligned} & \int_0^t \varphi(t_0 + s) e^{\beta(s-t)} ds \\ &= \sum_{k=0}^{n-1} \int_k^{k+1} \varphi(t_0 + s) e^{-\beta(t-s)} ds + \int_n^t \varphi(t_0 + s) e^{-\beta(t-s)} ds \\ &\leq M_t \left(\sum_{k=0}^{n-1} e^{-\beta(t+k-1)} + 1 \right) \leq M_t \left(\sum_{j=1}^{n-1} e^{-j\beta} + 1 \right) = M_t [(1 - e^{-\beta})^{-1} + 1], \end{aligned}$$

where $M_t = \sup_{t < \tau < t+t_0} \xi(\tau)$ and $\xi(t) = \int_t^{t+1} \varphi(\tau) d\tau$ as in (6.10).

As in Section 4 we can also show that (6.12) with the nondegeneracy condition (4.5) for the stationary problem yields the asymptotic stability of the coincidence sets.

Theorem 6.2. Under the assumptions of Theorem 6.1, let the condition (4.5) be fulfilled for $\psi_{\#}$ and $f_{\#}$. Then, if $\chi(t) = \chi_{\{u(t)=\psi(t)\}}$ and $\chi_{\#} = \chi_{\{u_{\#}=\psi_{\#}\}}$ denote respectively the characteristic functions of the coincidence sets of the evolutionary and the stationary problems (6.1) and (6.9), we have

$$\chi(t) \rightarrow \chi_{\#} \quad \text{in } L^p(\Omega), \quad 1 \leq p < \infty \quad \text{as } t \rightarrow \infty. \tag{6.17}$$

Proof. We can argue as in the proof of Theorem 4.1, by passing to the limit $t \rightarrow \infty$ in the equation for $u = u(t)$

$$Hu - (H\psi - f)\chi = f \quad \text{a.e. } \Omega, \quad t > 0, \tag{6.18}$$

in the sense of distributions, by noting that $u(t) \rightarrow u_{\#}$ in $L^2(\Omega)$ implies $\partial_t u(t) \rightarrow 0$ in a weak sense. Here we can use the argument of Lions (see [6], p. 509) for the translated functions $w(t) = u(t) - \psi(t)$ and $w_{\#} = u_{\#} - \psi_{\#}$, by noting that $\hat{w}(t) = w(t) - w_{\#} \rightarrow 0$ as $t \rightarrow \infty$ in the sense

$$\int_t^{t+1} \|\hat{w}(\tau)\|_{L^2(\Omega)}^2 d\tau = \int_0^1 \|\hat{w}(\sigma + t)\|_{L^2(\Omega)}^2 d\sigma \rightarrow 0.$$

Then the argument of Theorem 4.1 shows first that, if we denote by $\hat{\chi} \in L^\infty(0, \infty; L^\infty(0, 1; L^\infty(\Omega)))$ the function

$$\hat{\chi}(t):]0, 1[\ni \sigma \rightarrow \chi(\sigma + t) \in L^\infty(0, \infty; L^\infty(\Omega)),$$

we obtain from (6.18) that as $t \rightarrow \infty$

$$\widehat{\chi}(t) \rightarrow \widehat{\chi}_{\#} = \chi_{\#} \quad \text{first in } L^{\infty}(0, 1; L^{\infty}(\Omega))\text{-weak}^*,$$

and, since they are characteristic functions, also strongly in $L^p(0, 1; L^p(\Omega))$ for any $1 \leq p < \infty$, which yields (6.17).

REFERENCES

- [1] Bardos, C., Problèmes aux limites pour les équations aux dérivées partielles du premier ordre, *Ann. Scient. Éc. Norm. Sup.*, 4^e série, 3(1970), 185–233.
- [2] Bensoussan, A. & Lion, J. L., Inéquations variationnelles nonlinéaires du premier et du second ordre, *C. R. Acad. Sci. Paris*, **276**(1973), 1411–1415.
- [3] Friedman, A., Variational principles and free boundary problems, Wiley, New York, 1982.
- [4] Lévi, L., Obstacle problems for scalar conservation laws, *M2AN Math. Model. Numer. Anal.*, **35**(2001), 575–593.
- [5] Lévi, L. & Vallet, G., Entropy solutions for first-order quasilinear equations related to a bilateral obstacle condition in a bounded domain, *Chin. Ann. Math.*, **22B**:1(2001), 93–114.
- [6] Lions, J. L., Quelques Méthodes de Résolution des Problèmes Aux Limites Non-Linéaires, Dunod, Gauthier-Villars, Paris, 1969.
- [7] Lions, J. L., Partial differential inequalities, *Russian Math. Surveys*, **27**(1971), 91–161.
- [8] Mignot, F. & Puel J. P., Inéquations variationnelles et quasi variationnelles hyperboliques du premier ordre, *J. Math. Pures Appl.*, **55**(1976), 353–378.
- [9] Rodrigues, J. F., Some remarks on the asymptotic behaviour of strong solutions to monotone parabolic variational inequalities, *Rendiconti Mat.*, **4**:7(1984), 458–470.
- [10] Rodrigues, J. F., Obstacle problems in mathematical physics, North-Holland, Amsterdam, 1987.

This page intentionally left blank

INSTABILITY OF TRAVELING WAVES OF THE KURAMOTO-SIVASHINSKY EQUATION*

W. STRAUSS

*Department of Mathematics and LCDS,
Brown University, Providence RI 02912, USA*

WANG GUANXIANG

*School of Mathematical Sciences, Peking
University, Beijing 100871, China*

Visiting Brown University from September 2000 to August 2001

Abstract

Consider any traveling wave solution of the Kuramoto-Sivashinsky equation that is asymptotic to a constant as $x \rightarrow +\infty$. The authors prove that it is nonlinearly unstable under H^1 perturbations. The proof is based on a general theorem in Banach spaces asserting that linear instability implies nonlinear instability.

Keywords Traveling wave, Kuramoto-Sivashinsky equation, Instability
2000 MR Subject Classification 35K55, 35K90

§0. Introduction

The Kuramoto-Sivashinsky equation

$$u_t + u_{xxxx} + u_{xx} + uu_x = 0 \tag{0.1}$$

was derived by Kuramoto [2] as a model describing phase turbulence in reaction-diffusion systems and independently by Sivashinsky [3] as a model

*Project Supported in part by NSF Grant DMS-0071838.

of flame propagation. There are many numerical and some theoretical results showing that some of its solutions engage in very complicated dynamical behavior.

A traveling wave solution $u = \varphi(x - ct)$ satisfies, after one integration, the third-order equation

$$\varphi''' + \varphi' + \frac{1}{2}(\varphi - c)^2 = k \quad (0.2)$$

where k is a constant. A special case is a steady state $c = 0$. This ordinary differential equation has been studied extensively. Numerical studies [6 – 20] indicate the existence of heteroclinic and homoclinic orbits, as well as periodic and quasiperiodic solutions. Theoretical results include the existence of periodic solutions and heteroclinic orbits [4, 5]. In particular, Troy [4] proved that if $k = 1$, there exist at least two distinct odd solutions of (0.2) such that $\varphi(x) \rightarrow c \mp \sqrt{2}$ as $x \rightarrow \pm\infty$. He conjectured that there are an infinite number of different ones. Furthermore for $k \neq 1$ there are probably many others.

In this paper we consider any traveling wave solution $\varphi(x - ct)$ of (0.1) that approaches a constant as $x \rightarrow +\infty$. Then we consider solutions $u(x, t)$ of (0.1) with initial data $u(x, 0)$ arbitrarily near $\varphi(x)$ in the $H^1(\mathbf{R})$ norm. We prove that there exist such solutions that do not remain near $\varphi(x - ct)$ in the $H^1(\mathbf{R})$ norm at some later times. The instability of the traveling waves is a hint of the complexity of the dynamics of (0.1).

Our proof is based on the principle of linearization. We prove that the essential spectrum of the linearized generator meets the right half-plane and thus generates modes $e^{\lambda t}$ with $\Re\lambda > 0$ (Lemma 2.1). Then we invoke a general theorem that asserts that linearized instability implies nonlinear instability (Theorem 1.1).

Theorem 1.1 is a slight generalization of an earlier theorem [1] concerning nonlinear semigroups in a Banach space X . In the present case we have two Banach spaces $X \subset Z$, the linear semigroup is smoothing (mapping Z into X), while the nonlinear term loses regularity (mapping X into Z). The gain and loss of regularity compensate for each other.

§1. The Abstract Theorem

Consider an evolution equation

$$\frac{du}{dt} = Lu + F(u), \tag{1.1}$$

where L is a linear operator that generates a strongly continuous semigroup e^{tL} on a Banach space X , and F is a strongly continuous operator such that $F(0) = 0$. We focus on the instability of the zero solution of equation (1.1). About such a problem, the following question was addressed in a previous article [1].

If the spectrum of L meets the right half-plane $\{\mathcal{R}\lambda > 0\}$, does it follow that the zero solution of (1.1) is nonlinearly unstable?

Here, the zero solution is called nonlinearly stable in X if for any $\epsilon > 0 \exists \delta > 0$ such that $\|u_0\|_X < \delta$ implies that the unique solution $u \in C([0, \infty); X)$ of equation (1.1) with $u(0) = u_0$ satisfies $\sup_{0 \leq t < \infty} \|u(t)\|_X < \epsilon$. Otherwise, it is called nonlinearly unstable.

In [1], the authors considered the whole problem in only one space X , that is to say, the nonlinear operator maps X into X . However, many equations possess nonlinear terms that include derivatives and therefore F maps into a larger Banach space Z . If, therefore, the linear part is smoothing, mapping Z back into X , then we can recover the nonlinear instability as before. This is the content of the following theorem.

Theorem 1.1. *Assume the following.*

- (i) X, Z are two Banach spaces with $X \subset Z$ and $\|u\|_Z \leq C_1\|u\|_X$ for $u \in X$.
- (ii) L generates a strongly continuous semigroup e^{tL} on the space Z , and the semigroup e^{tL} maps Z into X for $t > 0$, and $\int_0^1 \|e^{tL}\|_{Z \rightarrow X} dt = C_4 < \infty$.
- (iii) The spectrum of L on X meets the right half-plane, $\{\mathcal{R}\lambda > 0\}$.
- (iv) $F : X \rightarrow Z$ is continuous and $\exists \rho_0 > 0, C_3 > 0, \alpha > 1$ such that $\|F(u)\|_Z \leq C_3\|u\|_X^\alpha$ for $\|u\|_X < \rho_0$.

Then the zero solution of (1.1) is nonlinearly unstable in the space X .

Remark 1.1. If $Z = X$, the theorem reduces to the theorem in [1].

To prove the theorem, we need the following two lemmas cited from [1]. For brevity, the proofs of the lemmas are omitted. The first lemma

asserts the existence of an approximate eigenvector v corresponding to an eigenvalue of maximal growth.

Lemma 1.1. *Let the spectrum of e^L on X be denoted by $\sigma_X(e^L)$. Let $e^\lambda \in \sigma_X(e^L)$ such that $|e^\lambda|$ equals the spectral radius of e^L on X . For every $\eta > 0$ and every integer $m > 0$, there exists $v \in X$ such that*

$$\|(e^{mL} - e^{m\lambda})v\|_X < \eta\|v\|_X, \tag{1.2}$$

$$\|e^{tL}v\|_X \leq 2Ke^{t\mathcal{R}\lambda}\|v\|_X, \quad \forall t, 0 \leq t \leq m, \tag{1.3}$$

where $K = \sup\{\|e^{\theta L}\|_{X \rightarrow X} : 0 \leq \theta \leq 1\}$ and $\mathcal{R}\lambda$ means the real part of λ .

The second lemma asserts that the whole semigroup grows at approximately the same rate as the eigenvalue.

Lemma 1.2. *Under the assumption of Lemma 1.1, for all $\epsilon > 0$, there exists a constant C_ϵ so that for all $0 \leq t < \infty$ we have*

$$e^{\mathcal{R}\lambda t} \leq \|e^{tL}\|_{X \rightarrow X} \leq C_\epsilon e^{(\mathcal{R}\lambda + \epsilon)t}.$$

Proof of Theorem 1.1. If $u \in C([0, T]; X)$ ($T \leq \infty$) is a solution of (1.1) with initial data $u(0) = v \in X$, then it formally satisfies the associated integral equation

$$u(t) = e^{tL}v + \int_0^t e^{(t-\tau)L}F(u(\tau)) d\tau, \quad 0 \leq t < T. \tag{1.4}$$

We are going to prove that there exists a universal constant $\epsilon_0 > 0$ such that $\sup_{0 \leq t < T} \|u(t)\|_X > \epsilon_0$ no matter how small $\|v\|_X$ may be.

Let us first define some quantities used below. Let

$$\mu = e^\lambda \tag{1.5}$$

as in Lemma 1.1. Choose

$$\epsilon = \frac{(\alpha - 1)\mathcal{R}\lambda}{2} \tag{1.6}$$

and C_ϵ as in Lemma 1.2. Let $C_2 = \|e^L\|_{Z \rightarrow X}$. Define k by

$$k^{\alpha-1} = 2|\mu|^\alpha C_3 \left[2K + \frac{1}{2|\mu|} \right]^\alpha \left[\frac{2C_\epsilon C_2}{(\alpha - 1)\mathcal{R}\lambda} e^{-\alpha\mathcal{R}\lambda} + C_4 \right]. \tag{1.7}$$

Let δ be free to remain arbitrarily small within the interval $(0, \delta_0)$ with

$$\delta_0 \equiv \min \left\{ \frac{1}{k}, 1, \frac{\rho_0}{2} \right\}. \tag{1.8}$$

Let T^* be the integer in the interval $(b, b + 1]$ where

$$b = \ln \left(\frac{1}{\delta k} \right) / \ln |\mu| > 0. \tag{1.9}$$

Note that (1.5) and (1.9) imply that

$$\frac{1}{k} < \delta e^{T^* \mathcal{R} \lambda} \leq \frac{|\mu|}{k}, \tag{1.10}$$

and T^* is dependent on δ as well as on μ and k . We may assume that the zero solution is stable. Thus there exists $\delta' > 0$ such that if $\|v\|_X = \delta < \delta'$, then there exists a unique solution $u \in C([0, \infty); X)$ of the integral equation (1.4). Let v be given by Lemma 1.1 with $m = T^*$ and $\eta = \frac{1}{4k}$, we take $\|v\|_X = \delta$. Now define

$$T = \sup \left\{ t : \|u(\tau) - e^{\tau L} v\|_X < \frac{\delta e^{\mathcal{R} \lambda \tau}}{2|\mu|} \text{ and } \|u(\tau)\|_X < \frac{\rho_0}{2} \text{ for } 0 < \tau \leq t \right\}. \tag{1.12}$$

Clearly $T > 0$. By (1.4) we have, for $0 < t \leq \min\{T^*, T\}$,

$$\begin{aligned} \|u(t) - e^{tL} v\|_X &\leq \int_0^{t-1} \|e^{(t-\tau-1)L}\|_{X \rightarrow X} \|e^L\|_{Z \rightarrow X} \|F(u(\tau))\|_Z d\tau \\ &\quad + \int_{t-1}^t \|e^{(t-\tau)L}\|_{Z \rightarrow X} \|F(u(\tau))\|_Z d\tau. \end{aligned} \tag{1.13}$$

Taking $\epsilon = \frac{(\alpha-1)}{2} \mathcal{R} \lambda$ in Lemma 1.2 and using the assumptions of the theorem, we have

$$\begin{aligned} \|u(t) - e^{tL} v\|_X &\leq \int_0^{t-1} C_\epsilon e^{\frac{\alpha-1}{2} \mathcal{R} \lambda (t-\tau-1)} C_2 C_3 \|u(\tau)\|_X^\alpha d\tau \\ &\quad + \int_{t-1}^t \|e^{(t-\tau)L}\|_{Z \rightarrow X} C_3 \|u(\tau)\|_X^\alpha d\tau. \end{aligned} \tag{1.14}$$

Within these integrals we use (1.3) and (1.12) to obtain

$$\|u(\tau)\|_X \leq \|e^{\tau L} v\|_X + \|u(\tau) - e^{\tau L} v\|_X \leq (2K + \frac{1}{2|\mu|}) \delta e^{\tau \mathcal{R} \lambda}. \tag{1.15}$$

Substituting (1.15) into (1.14), we have

$$\begin{aligned} &\|u(t) - e^{tL} v\|_X \\ &\leq C_2 C_3 C_\epsilon \left(2K + \frac{1}{2|\mu|} \right)^\alpha \delta^\alpha e^{\frac{\alpha-1}{2} \mathcal{R} \lambda (t-1)} \int_0^{t-1} e^{\frac{\alpha-1}{2} \mathcal{R} \lambda \tau} d\tau \end{aligned}$$

$$\begin{aligned}
& + C_3 \left(2K + \frac{1}{2|\mu|}\right)^\alpha \delta^\alpha \int_{t-1}^t \|e^{(t-\tau)L}\|_{Z \rightarrow X} e^{\tau\alpha\mathcal{R}\lambda} d\tau \\
& < C_2 C_3 C_\epsilon \left(2K + \frac{1}{2|\mu|}\right)^\alpha \delta^\alpha e^{\frac{\alpha+1}{2}\mathcal{R}\lambda(t-1)} \frac{2}{(\alpha-1)\mathcal{R}\lambda} e^{\frac{(\alpha-1)}{2}\mathcal{R}\lambda(t-1)} \\
& + C_4 C_3 \left(2K + \frac{1}{2|\mu|}\right)^\alpha \delta^\alpha e^{\alpha\mathcal{R}\lambda t} \\
& = \frac{k^{\alpha-1}}{2|\mu|^\alpha} (\delta e^{\mathcal{R}\lambda t})^\alpha
\end{aligned}$$

by (1.7). Then we have

$$\|u(t) - e^{tL}v\|_X < \frac{k^{\alpha-1}}{2|\mu|^\alpha} (\delta e^{\mathcal{R}\lambda t})^\alpha \quad \text{for } 0 \leq t \leq \min(T, T^*). \quad (1.16)$$

Now if $T \leq T^*$, then we claim that $\|u(T)\|_X \geq \frac{\rho_0}{2}$. Indeed, if $T \leq T^*$ and $\|u(T)\|_X < \frac{\rho_0}{2}$, then by definition (1.12) of T , we have

$$\|u(T) - e^{TL}v\|_X = \frac{\delta e^{\mathcal{R}\lambda T}}{2|\mu|}. \quad (1.17)$$

Combining it with (1.16) for $t = T$, we obtain

$$\frac{\delta e^{\mathcal{R}\lambda T}}{2|\mu|} < \frac{k^{\alpha-1}}{2|\mu|^\alpha} (\delta e^{\mathcal{R}\lambda T})^\alpha,$$

that is,

$$\delta e^{\mathcal{R}\lambda T} > \frac{|\mu|}{k}, \quad (1.18)$$

which means $T > T^*$ by (1.11) and leads to a contradiction. Thus the claim is proven. Next, if $\|u(T)\|_X \neq \frac{\rho_0}{2}$, we have $T > T^*$. Choose $t = T^*$ so that by (1.11) we have

$$\|u(T^*) - e^{T^*L}v\|_X < \frac{k^{\alpha-1}}{2|\mu|^\alpha} (\delta e^{\mathcal{R}\lambda T^*})^\alpha \leq \frac{k^{\alpha-1}}{2|\mu|^\alpha} \left(\frac{|\mu|}{k}\right)^\alpha = \frac{1}{2k}. \quad (1.19)$$

So

$$\|u(T^*)\|_X \geq \|e^{T^*L}v\|_X - \frac{1}{2k}. \quad (1.20)$$

On the other hand, taking $m = T^*$ and $\eta = \frac{1}{4k}$ in Lemma 1.1, (1.2) implies

$$\|e^{T^*L}v\|_X \geq \|e^{T^*\lambda}v\|_X - \frac{1}{4k} \|v\|_X.$$

Since $\|e^{T^*\lambda}v\|_X = e^{T^*\mathcal{R}\lambda} \|v\|_X$ and $\|v\|_X = \delta$, we get by (1.11)

$$\|e^{T^*L}v\|_X \geq e^{\mathcal{R}\lambda T^*} \delta - \frac{\delta}{4k} > \frac{1}{k} - \frac{\delta}{4k}.$$

Hence (1.20) implies

$$\|u(T^*)\|_X \geq \frac{1}{k} - \frac{\delta}{4k} - \frac{1}{2k} > \frac{1}{4k}, \tag{1.21}$$

since $\delta < 1$.

Therefore, there exists a time t (either T or T^*) at which

$$\|u(t)\|_X \geq \min \left\{ \frac{1}{4k}, \frac{\rho}{2} \right\} \equiv \epsilon_0,$$

and ϵ_0 is a universal constant independent of the size of the initial data v .

Remark 1.2. The proof shows that there exist $C > 0$ and $\epsilon_0 > 0$ such that for all sufficiently small positive δ , there is a solution u that satisfies $\|u(0)\|_X < \delta$ but

$$\sup_{0 \leq t \leq C|\log \delta|} \|u(t)\| \geq \epsilon_0.$$

Thus the escape time occurs logarithmically soon.

§2. Application to the Kuramoto-Sivashinsky Equation

The Kuramoto-Sivashinsky equation in one dimension is

$$v_t + v_{x^4} + v_{x^2} + \frac{1}{2}(v_x)^2 = 0, \quad -\infty < x < \infty. \tag{2.1}$$

With $u = v_x$ it can be written as

$$u_t + u_{x^4} + u_{x^2} + uu_x = 0, \quad -\infty < x < \infty. \tag{2.2}$$

If $\varphi(x-ct)$ is a traveling wave solution of (2.2), then φ satisfies the ordinary differential equation

$$\varphi''' + \varphi' + \frac{1}{2}(\varphi - c)^2 = k \tag{2.3}$$

where k is a constant. If $c = 0$, then φ is a steady-state solution. Troy [4] proved that if $k = 1$, then equation (2.3) admits at least two odd solutions satisfying

$$\lim_{x \rightarrow \infty} \varphi(x) = c - \sqrt{2}, \quad \lim_{x \rightarrow -\infty} \varphi(x) = c + \sqrt{2}.$$

The goal of this section is to prove the following theorem.

Theorem 2.1. *All the traveling waves $\varphi(x-ct)$ of the Kuramoto-Sivashinsky equation satisfying $\varphi \in L^\infty(\mathbf{R})$, $\varphi_x, \varphi_{xx} \in L^2(\mathbf{R})$ and $\varphi - b_+ \in L^2([0, \infty))$ are nonlinearly unstable in the space $H^1(\mathbf{R})$.*

This means that there exist positive ϵ_0 and C_0 , a sequence $\{u_n\}$ of solutions of the K-S equation, and a sequence of times $0 \leq t_n \leq C_0 \log n$ such that $\|u_n(0) - \varphi\|_{H^1(\mathbf{R})} \rightarrow 0$ but $\|u_n(t_n) - \varphi(\cdot - ct_n)\|_{H^1(\mathbf{R})} \geq \epsilon_0$.

If $\varphi(x - ct) \in H^1(\mathbf{R})$ is a traveling-wave solution of the K-S equation (2.2), then letting $w(x, t) = u(x, t) - \varphi(x - ct)$, we have

$$w_t + w_{x^4} + w_{x^2} + \varphi w_x + \varphi_x w + w w_x = 0, \quad -\infty < x < \infty, \quad (2.5)$$

with initial value

$$w(x, 0) = w_0(x) \equiv u_0(x) - \varphi(x). \quad (2.6)$$

So the stability of traveling-wave solutions of (2.2) is translated into the stability of the zero solution of (2.5). In order to prove Theorem 2.1, taking $Z \equiv L^2(\mathbf{R})$, $X = H^1(\mathbf{R})$, we need to prove that the four conditions of Theorem 1.1 are satisfied by the associated equation (2.5).

Denote the linear partial differential operator in (2.5) by $L = -(\partial_x^4 + \partial_x^2 + \varphi \partial_x + \varphi_x) \equiv L_0 - [(\varphi - b_+) \partial_x + \varphi_x]$ with $L_0 = -(\partial_x^4 + \partial_x^2 + b_+ \partial_x)$. Then (2.5) may be rewritten in the form (1.1),

$$w_t = Lw + F(w), \quad (2.7)$$

where $F(w) = -w w_x$. Note that F maps $H^1(\mathbf{R})$ into $L^2(\mathbf{R})$ and satisfies

$$\|F(w)\|_{L^2} \leq \|w\|_{H^1}^2.$$

This proves Condition (iv) of Theorem 1.1 with $C_3 = 1$ and $\alpha = 2$.

To prove Condition (ii) in Theorem 1.1, we need the following two lemmas.

Lemma 2.1. *Let $L_0 = -(\partial_x^4 + \partial_x^2 + b_+ \partial_x)$ for any real constant b_+ . Then*

$$\|e^{tL_0}\|_{H^m \rightarrow H^m} \leq e^{\frac{t}{4}} \quad \text{for } m \in \mathbf{R}, \quad 0 \leq t < \infty, \quad (2.8)$$

$$\|e^{tL_0}\|_{L^2 \rightarrow H^1} \leq a(t) \equiv 4t^{-\frac{1}{4}} \quad \text{for } 0 < t \leq 1. \quad (2.9)$$

Proof. We write $u(x, t) = e^{tL_0} u_0(x)$. By Fourier transformation,

$$\hat{u}(\xi, t) = e^{-t(\xi^4 - \xi^2 + i\xi)} \hat{u}_0(\xi).$$

$$\begin{aligned}
 \|u(t)\|_{H^m}^2 &= \int_{-\infty}^{\infty} (1 + \xi^2)^m |\hat{u}(\xi, t)|^2 d\xi \\
 &= \int_{-\infty}^{\infty} (1 + \xi^2)^m e^{-2t(\xi^4 - \xi^2)} |\hat{u}_0(\xi)|^2 d\xi \\
 &\leq \sup_{\xi \in \mathbf{R}} e^{-2t(\xi^4 - \xi^2)} \int_{-\infty}^{\infty} (1 + \xi^2)^m |\hat{u}_0(\xi)|^2 d\xi = e^{\frac{t}{2}} \|u_0\|_{H^m}^2.
 \end{aligned}$$

Hence

$$\|e^{tL_0}\|_{H^m \rightarrow H^m} \leq e^{\frac{t}{4}}.$$

On the other hand, letting $s = \xi^2$, we have

$$\|u(t)\|_{H^1}^2 \leq \sup_{s \in \mathbf{R}^+} f(s) \int_{-\infty}^{\infty} |\hat{u}_0(\xi)|^2 d\xi$$

with $f(s) = (1 + s)e^{-2t(s^2 - s)}$, $t > 0$. Elementary computation shows that

$$\sup_{s > 0} f(s) \leq \left(\frac{3}{2} + \frac{1}{2}t^{-\frac{1}{2}}\right) e^{\frac{t}{2}}.$$

Thus

$$\|u(x, t)\|_{H^1} \leq \left(\frac{3}{2} + \frac{1}{2}t^{-\frac{1}{2}}\right)^{\frac{1}{2}} e^{\frac{t}{4}} \|u_0\|_{L^2},$$

and

$$\|e^{tL_0}\|_{L^2 \rightarrow H^1} \leq \left(\frac{3}{2} + \frac{1}{2}t^{-\frac{1}{2}}\right)^{\frac{1}{2}} e^{\frac{t}{4}} \leq 4t^{-\frac{1}{4}} \quad \text{for } 0 < t \leq 1,$$

since $e^{\frac{t}{4}} \leq e^{\frac{1}{4}} < 2$. Thus Lemma 2.1 has been proved.

The following lemma proves Condition (ii).

Lemma 2.2. *Let $L = -[\partial_x^4 + \partial_x^2 + \varphi(x)\partial_x + \varphi'(x)] \equiv L_0 - (\varphi - b_+)\partial_x - \varphi_x$ with $\varphi \in L^\infty(\mathbf{R})$, $\varphi_x \in L^2(\mathbf{R})$, $\chi_{[0, \infty)}(\varphi - b_+) \in L^2(\mathbf{R})$. Then*

$$\|e^{tL}\|_{L^2 \rightarrow H^1} \leq C_5 t^{-\frac{1}{4}} \quad \text{for } 0 < t \leq 1, \quad (2.10)$$

$$\|e^{tL}\|_{H^1 \rightarrow H^1} \leq C_6 < \infty \quad \text{for } 0 \leq t \leq 1, \quad (2.11)$$

where the constants C_5, C_6 are defined by (2.17), (2.20) below.

Proof. Consider the initial value problem

$$u_t = Lu = L_0 u - [\varphi(x) - b_+] \partial_x u - \varphi'(x) u,$$

$$u(x, 0) = u_0(x), \quad x \in \mathbf{R}.$$

Then $u(x, t) = e^{tL} u_0(x)$, $t \geq 0$, $x \in \mathbf{R}$. Thus

$$u(t) = e^{tL_0} u_0 - \int_0^t e^{(t-\tau)L_0} [(\varphi - b_+) \partial_x u + \varphi' u] d\tau.$$

Denote $A = \|\varphi - b_+\|_{L^\infty}$, $B = \|\varphi'\|_{L^2}$. Then for $0 < t \leq 1$,

$$\begin{aligned} & \|u(t)\|_{H^1} \\ & \leq \|e^{tL_0}\|_{L^2 \rightarrow H^1} \|u_0\|_{L^2} + \int_0^t \|e^{(t-\tau)L_0}\|_{L^2 \rightarrow H^1} \|\varphi - b_+\|_{L^\infty} \|\partial_x u\|_{L^2} d\tau \\ & \quad + \int_0^t \|e^{(t-\tau)L_0}\|_{L^2 \rightarrow H^1} \|\varphi'\|_{L^2} \|u\|_{L^\infty} d\tau \end{aligned} \tag{2.12}$$

$$\leq a(t)\|u_0\|_{L^2} + (A + B) \int_0^t a(t - \tau) \|u(\tau)\|_{H^1} d\tau \quad \text{for } 0 < t \leq 1. \tag{2.13}$$

By iteration,

$$\begin{aligned} \|u(t)\|_{H^1} & \leq a(t)\|u_0\|_{L^2} + (A + B) \int_0^t a(t - \tau) \left[a(\tau)\|u_0\|_{L^2} \right. \\ & \quad \left. + (A + B) \int_0^\tau a(\tau - s) \|u(s)\|_{H^1} ds \right] d\tau \\ & = a(t)\|u_0\|_{L^2} + (A + B) \int_0^t a(t - \tau) a(\tau) \|u_0\|_{L^2} d\tau \\ & \quad + (A + B)^2 \int_0^t \int_0^\tau a(t - \tau) a(\tau - s) \|u(s)\|_{H^1} ds d\tau. \end{aligned} \tag{2.14}$$

The second term on the right side of (2.14) is

$$\begin{aligned} & (A + B) \int_0^t a(t - \tau) a(\tau) \|u_0\|_{L^2} d\tau \\ & = (A + B) \|u_0\|_{L^2} \int_0^t 4(t - \tau)^{-\frac{1}{4}} 4\tau^{-\frac{1}{4}} d\tau \\ & = 16(A + B) C_8 t^{\frac{1}{2}} \|u_0\|_{L^2} \quad \text{for } 0 < t \leq 1, \end{aligned} \tag{2.15}$$

where $C_8 = \int_0^1 (1 - r)^{-\frac{1}{4}} r^{-\frac{1}{4}} dr$. By exchanging the order of integration, we get from the third term on the right side of (2.14),

$$\begin{aligned} & \int_0^t \int_0^\tau a(t - \tau) a(\tau - s) \|u(s)\|_{H^1} ds d\tau \\ & = \int_0^t \left[\int_s^t a(t - \tau) a(\tau - s) d\tau \right] \|u(s)\|_{H^1} ds. \end{aligned}$$

Now

$$\begin{aligned} & \int_s^t a(t-\tau)a(\tau-s) d\tau \\ &= 16 \int_s^t (t-\tau)^{-\frac{1}{4}}(\tau-s)^{-\frac{1}{4}} d\tau \\ &= 16C_8(t-s)^{\frac{1}{2}} \leq 16C_8 \text{ for } 0 < s \leq t \leq 1. \end{aligned} \tag{2.16}$$

Therefore (2.13)–(2.16) imply

$$\begin{aligned} \|u\|_{H^1} &\leq [a(t) + 16C_8(A + B)]\|u_0\|_{L^2} \\ &+ \int_0^t 16C_8(A + B)^2\|u(s)\|_{H^1} ds \text{ for } 0 < t \leq 1. \end{aligned}$$

By Gronwall’s inequality, we get

$$\|u\|_{H^1} \leq [4t^{-\frac{1}{4}} + 16C_8(A + B)] \exp[16C_8(A + B)^2t]\|u_0\|_{L^2} \text{ for } 0 < t \leq 1.$$

So with the constant

$$C_5 = [4 + 16C_8(A + B)]e^{16C_8(A+B)^2}, \tag{2.17}$$

we have

$$\|u\|_{H^1} \leq C_5t^{-\frac{1}{4}}\|u_0\|_{L^2} \text{ for } 0 < t \leq 1. \tag{2.18}$$

Thus (2.10) has been proven. To prove (2.11), replacing the first term on the right side of (2.12) by $\|e^{tL_0}\|_{H^1 \rightarrow H^1}\|u_0\|_{H^1}$ and using (2.8), we have

$$\|u(t)\|_{H^1} \leq e^{\frac{t}{4}}\|u_0\|_{H^1} + (A + B) \int_0^t a(t-\tau)\|u(\tau)\|_{H^1} d\tau \text{ for } 0 < t \leq 1. \tag{2.19}$$

Similarly iterating and computing as above, we obtain

$$\begin{aligned} \|u(t)\|_{H^1} &\leq [2 + 16(A + B)]e^{16C_8(A+B)^2}\|u_0\|_{H^1} \\ &\equiv C_6\|u_0\|_{H^1} \text{ for } 0 < t \leq 1. \end{aligned} \tag{2.20}$$

Hence (2.11) is proven and the proof of Lemma 2.2 is finished.

We now proceed to verify Condition (iii) of Theorem 1.1. Formula (2.11) in Lemma 2.2 means that L generates a strongly continuous semigroup on the Banach space $H^1(\mathbf{R})$ (see [22]). By Fourier transformation, the essential spectrum of L_0 on $H^1(\mathbf{R})$ is

$$\sigma_e(L_0) \supset \{-\xi^4 + \xi^2 - ib_+\xi | \xi \in \mathbf{R}\}. \tag{2.21}$$

This curve meets the vertical lines $Re\lambda = \alpha$ for $-\infty < \alpha \leq \frac{1}{4}$ because

$-\infty < -\xi^4 + \xi^2 \leq \frac{1}{4}$. We now prove that the same curve belongs to the essential spectrum of L .

Lemma 2.3. *The essential spectrum of L on $H^1(\mathbf{R})$ contains that of L_0 .*

Proof. Let $\xi \in \mathbf{R}$ and let $\lambda = P(\xi) = -\xi^4 + \xi^2 - ib_+\xi$. Following Schechter [23], $\lambda \in \sigma_e(L)$ if there exists a sequence $\{\zeta_n\} \subset H^1(\mathbf{R})$ with

$$\|\zeta_n\|_{H^1} = 1, \quad \|(L - \lambda)\zeta_n\|_{H^1} \rightarrow 0,$$

and $\{\zeta_n\}$ does not have a strongly convergent subsequence in $H^1(\mathbf{R})$. (Here we use the definition: $\lambda \notin \sigma_e(L)$ if and only if $L - \lambda$ is Fredholm with index zero.) Now let $\zeta_0 \not\equiv 0$ be a C^∞ function with compact support in $(0, \infty)$. Define

$$\zeta_n(x) = c_n e^{-i\xi x} \zeta_0(x/n) / \sqrt{n}, \quad n = 1, 2, \dots,$$

where c_n is chosen so that $\|\zeta_n\|_{H^1} = 1$. In fact,

$$\|\zeta_n\|_{L^2} = c_n \|\zeta_0\|_{L^2} \quad \text{and} \quad 1 = \|\zeta_n\|_{H^1} \leq k c_n$$

for some positive constant k . Hence $c_n \geq \frac{1}{k} > 0$. Since $\|\zeta_n\|_{L^\infty} \rightarrow 0$ but $\|\zeta_n\|_{L^2}$ is bounded away from zero, $\{\zeta_n\}$ can have no convergent subsequence in $L^2(\mathbf{R})$.

It remains to show that $\|(L - \lambda)\zeta_n\|_{H^1} \rightarrow 0$. We write

$$L - \lambda = L_0 - \lambda + (\varphi - b_+) \partial_x - \varphi_x.$$

Now elementary computations show

$$\begin{aligned} (L_0 - \lambda)\zeta_n(x) &= e^{i\xi x} \sum_{1 \leq s \leq 4} P^{(s)}(\xi) c_n \zeta_0^{(s)}\left(\frac{x}{n}\right) / (s! n^{\frac{1}{2}+s}), \\ \partial(L_0 - \lambda)\zeta_n(x) &= i\xi(L_0 - \lambda)\zeta_n(x) \\ &\quad + e^{i\xi x} \sum_{1 \leq s \leq 4} P^{(s)}(\xi) c_n \zeta_0^{(s+1)}\left(\frac{x}{n}\right) / (s! n^{\frac{3}{2}+s}). \end{aligned}$$

Thus

$$\begin{aligned} &\|(L_0 - \lambda)\zeta_n(x)\|_{H^1} \\ &\leq (1 + |\xi|) \sum_{1 \leq s \leq 4} |P^{(s)}(\xi)| c_n \|\zeta_0^{(s)}\left(\frac{x}{n}\right)\|_{L^2} / (s! n^{\frac{1}{2}+s}) \\ &\quad + \sum_{1 \leq s \leq 4} |P^{(s)}(\xi)| c_n \|\zeta_0^{(s+1)}\left(\frac{x}{n}\right)\|_{L^2} / (s! n^{\frac{3}{2}+s}) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Moreover, for any positive integer m , $\|\partial_x^m \zeta_n\|_{L^\infty} \rightarrow 0$ as $n \rightarrow \infty$, so we have

$$\|(\varphi - b_+) \partial_x \zeta_n\|_{L^2}^2 \leq \|\partial_x \zeta_n\|_{L^\infty}^2 \|\chi_{[0, \infty)}(\varphi - b_+)\|_{L^2}^2 \rightarrow 0,$$

and

$$\begin{aligned} & \|\partial_x [(\varphi - b_+) \partial_x \zeta_n]\|_{L^2}^2 \\ & \leq 2 \int_R \varphi_x^2 [\partial_x \zeta_n(x)]^2 dx + 2 \int_R (\varphi - b_+)^2 [\partial_{xx} \zeta_n(x)]^2 dx \\ & \leq 2 \|\partial_x \zeta_n\|_{L^\infty}^2 \|\varphi_x\|_{L^2}^2 + 2 \|\partial_{xx} \zeta_n\|_{L^\infty}^2 \|\chi_{[0, \infty)}(\varphi - b_+)\|_{L^2}^2 \rightarrow 0. \end{aligned}$$

In addition,

$$\|\varphi_x \zeta_n\|_{L^2} \leq \|\zeta_n\|_{L^\infty} \|\varphi_x\|_{L^2} \rightarrow 0$$

and

$$\|\varphi_x \partial_x \zeta_n + \varphi_{xx} \zeta_n\|_{L^2} \leq \|\partial_x \zeta_n\|_{L^\infty} \|\varphi_x\|_{L^2} + \|\partial_x \zeta_n\|_{L^\infty} \|\varphi_{xx}\|_{L^2} \rightarrow 0.$$

Thus

$$\|(\varphi - b_+) \partial_x \zeta_n + \varphi_x \zeta_n\|_{H^1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

So from the estimates above,

$$\|(L - \lambda)\zeta_n\|_{H^1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

The proof of Lemma 2.3 is completed.

Therefore all the four conditions of Theorem 1.1 are satisfied by the linearized equation (2.5) and Theorem 2.1 has been proved.

REFERENCES

- [1] Shatah, J. & Strauss, W., Spectral condition for instability, *Contemporary Mathematics*, **255**(2000), 189-198.
- [2] Kuramoto, Y., Instability and turbulence of wave fronts in reaction-diffusion systems, *Prog. Theoret. Phys.*, **63**(1980), 1885-1903.
- [3] Sivashinsky, G., Nonlinear analysis of hydrodynamic instability in laminar flames, Part I, Derivation of basic equations, *Acta Astronaut.*, **4**(1977), 1117-1206.
- [4] Troy, W. C., The existence of steady solutions of Kuramoto-Sivashinsky equation, *J. Diff. Eq.*, **82**(1989), 269-313.
- [5] Jones, J. & Troy, W. C., Steady solutions of the Kuramoto-Sivashinsky equation for small wave speed, *J. Diff. Eq.*, **96**(1992), 28-55.
- [6] Yang, T. S., On traveling wave solutions of the Kuramoto-Sivashinsky equation, *Physica D*, **110**:1-2(1997), 25-42.
- [7] Michelson, D., Steady solutions of the Kuramoto-Sivashinsky equation, *Physica D*, **19**(1986), 89-111.

- [8] Michelson, D., Bunsen flames as steady solutions of the Kuramoto-Sivashinsky equation, *SIAM J. Math. Anal.*, **23**:2(1992), 364–386.
- [9] Michelson, D., Stability of the bunsen flame profiles in the Kuramoto-Sivashinsky equation, *SIAM J. Math. Anal.*, **27**:3(1996), 765–781.
- [10] Green, J. M. & Kim, J. S., The steady states of the Kuramoto-Sivashinsky equation, *Progress in Chaotic dynamics, Physica D*, **33**:1–3(1988), 99–120.
- [11] Hooper, A. P. & Grimshaw, R., Traveling wave solutions of the Kuramoto-Sivashinsky equation, *Wave Motion*, **10**:5(1988), 405–420.
- [12] Kevrekidis, I. G., Nicolaenko, B. & Scovel, J. C., Back in the saddle again: a computer assisted study of the Kuramoto-Sivashinsky equation, *SIAM J. Appl. Math.*, **50**:3(1990), 760–790.
- [13] Trifonov, Yu. Ya., Two-periodical and quasi-periodical wave solutions of the Kuramoto-Sivashinsky equation and their stability and bifurcations, *Physica D*, **54**:4 (1992), 311–330.
- [14] Kent, P. & Elgin, J., Travelling waves of the Kuramoto-Sivashinsky equation: Period-multiplying bifurcations, *Nonlinearity*, **5**:4(1992), 899–919.
- [15] Elgin, J. N. & Wu, X. S., Stability of cellular states of the Kuramoto-Sivashinsky equation, *SIAM J. Appl. Math.*, **56**:6(1996), 1621–1638.
- [16] Lan, Y. T., The “cocoon” bifurcations in three-dimensional systems with two fixed points, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, **2**:3(1992), 543–558.
- [17] Feudel, F., Feudel, U. & Brandenburg, A., On the Bifurcation phenomena of the Kuramoto-Sivashinsky equation, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, **3**:5 (1993), 1299–1303.
- [18] Kukavica, I., Oscillations of solutions of the Kuramoto-Sivashinsky equation, *Physica D*, **76**:4(1994), 369–374.
- [19] Dawson, S. P. & Mancho, A. M., Collections of heteroclinic cycles in the Kuramoto-Sivashinsky equation, *Physica D*, **100**:3–4(1997), 231–256.
- [20] Amdjadi, F., Aston, P. J. & Plechac, P., Symmetry breaking Hopf bifurcations in equations with $O(2)$ symmetry with application to the Kuramoto-Sivashinsky equation, *J. Comput. Phys.*, **131**:1(1997), 181–192.
- [21] Tadmor, E., The well-posedness of the Kuramoto-Sivashinsky equation, *SIAM J. Math. Anal.*, **17**:4(1986), 884–893.
- [22] Goldstein, J. A., Semigroups of linear operators and applications, Oxford University Press, New York, 1985.
- [23] Schechter, M., Spectra of partial differential operators, American Elsevier Publishing Company, INC.-New York, 1971.

COMPLEXITY OF LARGE TIME BEHAVIOUR OF EVOLUTION EQUATIONS WITH BOUNDED DATA

J. L. VÁZQUEZ

*Departamento de Matemáticas, Universidad Autónoma,
28049 Madrid, Spain. E-mail: juanluis.vazquez@uam.es*

E. ZUAZUA

*Departamento de Matemática Aplicada, Universidad Complutense,
28040 Madrid, Spain. E-mail: zuazua@eucmax.sim.ucm.es*

Abstract

The authors study the asymptotic behaviour of solutions of the heat equation and a number of evolution equations using scaling techniques. It is proved that in the framework of bounded data stabilization need not occur and the general asymptotic behaviour is complex. This behaviour reflects for large times, even on compact sets, the complexity of the initial data at infinity.

Keywords Asymptotic behaviour, Scaling, Omega-limit,
Heat equation, Hyperbolic conservation laws

2000 MR Subject Classification 35B40, 35K60, 35K65

§1. Introduction

In recent decades much attention has been paid to the study of the long time behaviour of suitable classes of solutions of partial differential equations using the concepts of dynamical systems. Such concepts, like orbit, omega-limit and attractor, were introduced in the first half of the century for the investigation of systems of ordinary differential equations,

which are finite dimensional instances of dynamical systems, and were successfully extended to the infinite dimensional framework. Here we want to combine these ideas with the property of invariance under the scaling group to obtain results on the complex large time behaviour of a number of evolution equations with a simple structure when they are posed in the set of bounded measurable functions defined in the whole space, $L^\infty(\mathbb{R}^N)$. Invariance under the scaling group allows the equation to copy in certain way the behaviour of the initial data at infinity of the space variable into the behaviour of the solution $u = u(\cdot, t)$ for large t . The results apply to the classical heat equation, which we take as a model, but also to a large number of evolution equations like the porous medium equation, the p -Laplacian equation, the wave equation, and scalar conservation laws. Scale invariance is at the heart of our argument. However, it is really necessary only in a weak form called quasi-invariance. This will allow our theory to be extended to homogenization problems, equations involving reaction, convection and diffusion, and the equations of fluid mechanics. The space $L^\infty(\mathbb{R}^N)$ plays an important role in providing for a setting where complexity occurs in the asymptotic behaviour of the above-mentioned equations. Thus, the evolution problems under consideration are usually well-posed in other functional spaces like $L^p(\mathbb{R}^N)$, for some $1 < p < \infty$, and the asymptotic behaviour can be then rather simple, reflecting the simple structure of the equation. Consider for instance the heat equation. When we consider solutions defined in \mathbb{R}^N it is well known that every solution with initial data $u_0 \in L^1(\mathbb{R}^N)$ converges as $t \rightarrow \infty$ towards a multiple of the fundamental solution, the one which has the same integral,

$$u(x, t) \rightarrow Ct^{-N/2} \exp\left(-\frac{x^2}{4t}\right), \quad C = (4\pi)^{-N/2} \int u_0(x) dx. \quad (1.1)$$

When the space domain is a bounded set $\Omega \subset \mathbb{R}^N$ and we take homogeneous Dirichlet conditions the typical space is $L^2(\Omega)$, the convergence to a unique type of profile takes place after multiplying by $e^{\lambda_1 t}$, where λ_1 is the so-called first eigenvalue. A similar result happens for homogeneous Neumann conditions. In the setting of almost periodic functions solutions posed in \mathbb{R}^N converge to the space average of the initial data. Similar simplicity occurs for many of the equations mentioned above. We will prove below

that, on the contrary, the simple behaviour is lost when $L^\infty(\mathbb{R}^N)$ is taken as the functional setting.

The rest of this paper is organized as follows. In Section 2 we present the main results and techniques developed in this paper in the simple example of the linear heat equation. Section 3 contains the analysis of the behavior of the initial data under scaling. In Section 4 we present the extension to an abstract setting of the results of Section 2 on the heat equation. In Section 5 we return to the heat equation to make some earlier results more precise. Section 6 is devoted to nonlinear heat equations as the porous media equation and the p -Laplacian one. In Section 7 we analyze the interesting case of scalar hyperbolic conservation laws in one and several space dimensions. In Section 8 we extend the analysis done in L^∞ to other L^p spaces to cover other existing results. In Section 9 we point out that the compactness requirements may be often avoided in the context of linear evolution equations. Finally, in Section 10, we observe that most of the results of this paper can be adapted to situations where the equation under consideration is not invariant under the scaling transformation but is quasi-invariant in an appropriate sense that we define.

§2. The Heat Equation

In order to introduce the methods and results in a simple setting we consider in the first place the linear constant-coefficient heat equation posed in \mathbb{R}^N with $N \geq 1$,

$$\begin{cases} u_t = \Delta u & \text{in } Q = \mathbb{R}^N \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{in } \mathbb{R}^N. \end{cases} \quad (2.1)$$

We want to analyze the asymptotic behaviour of solutions as $t \rightarrow \infty$. It is easy to see that for suitable initial data there exists a constant $c \in \mathbb{R}$ such that $u(\cdot, t) \rightarrow c$ in $L^\infty(\mathbb{R}^N)$ as $t \rightarrow \infty$. This happens at least in two cases.

(i) When $u_0 = c + f$ with $c \in \mathbb{R}$ and $f \in L^p(\mathbb{R}^N)$ for some $p \in [1, \infty)$, and

(ii) When u_0 is periodic.

In the first case the solution converges to c with a polynomial rate, in the second one the asymptotic c is the average of the initial datum on a periodicity cell and the convergence rate is exponential. Obviously, the

constant is not the only possible asymptotic regime. This can be easily seen considering for instance

(iii) initial data in 1D with two different asymptotic values c_{\pm} at $\pm\infty$.

In this case it can be proved that the asymptotic behaviour is given by the special solution of the heat equation with piece-wise constant initial data of Heaviside type, $u_0(x) = H(x)$ with

$$H(x) = c_+ \quad \text{for } x > 0 \quad (\text{resp. } H(x) = c_- \quad \text{for } x < 0) \quad (2.2)$$

(cf. end of this section).

The first goal of this paper is to describe what are all the possible asymptotic regimes of the solutions of the heat equation and other evolution equations for bounded initial data. It is natural to address this problem by means of scaling techniques. Given a solution $u(x, t)$ of the heat equation and a constant $\lambda > 0$ we introduce the rescaled function

$$u_{\lambda}(x, t) = u(\lambda x, \lambda^2 t). \quad (2.3)$$

The key point is that u_{λ} solves the heat equation with initial data $u_{0,\lambda}(x) = u_0(\lambda x)$. Due to the Maximum Principle we have for every (x, t) that $|u(x, t)| \leq \|u_0\|_{L^{\infty}(\mathbb{R}^N)}$, and the same bound applies to u_{λ} for all $\lambda > 0$, i.e., it is uniform in λ . By standard compactness results for the solutions of the heat equation we conclude that along a subsequence $\lambda_n \rightarrow \infty$ we have

$$\begin{aligned} u_{0,\lambda} &\rightharpoonup \phi \quad \text{in } L^{\infty}(\mathbb{R}^N) \text{ weak-star,} \\ u_{\lambda} &\rightarrow v \quad \text{in } L^{\infty}_{\text{loc}}(Q). \end{aligned}$$

Moreover, the limit function v solves the heat equation with initial data ϕ . In particular, for $t = 1$ we have

$$u_{\lambda}(x, 1) \rightarrow v(x, 1) \quad \text{in } L^{\infty}(K)$$

uniformly on any compact subset $K \subset \mathbb{R}^N$. This can be re-written as follows:

$$u(\lambda x, \lambda^2) \rightarrow v(x, 1) \quad \text{in } L^{\infty}(K). \quad (2.4)$$

By setting $\lambda^2 = t$ we deduce the convergence of the rescaled orbit

$$u(\sqrt{t}x, t) \rightarrow v(x, 1) \quad \text{in } L^{\infty}(K) \quad (2.5)$$

as $t \rightarrow \infty$ along a suitable subsequence.

Let us now denote by S_t the semigroup generated by the heat equation, defined by $(S_t u_0)(x) = u(x, t)$. Then $v(x, 1) = S_1 \phi$. Summing up, we obtain the following result.

Theorem 2.1. *The set of accumulation points in $L_{loc}^\infty(\mathbb{R}^N)$ of $u(\sqrt{t}x, t)$ as $t \rightarrow \infty$ coincides with the set $\{S_1(\phi)\}$, where ϕ ranges over the set of accumulation points as $\lambda \rightarrow \infty$ of the family $\{u_{0,\lambda} : \lambda > 0\}$ in the weak-star topology of $L^\infty(\mathbb{R}^N)$.*

Proof. The direct part has already been settled. For the converse, suppose that u is a bounded solution $u(x, t)$ defined for $x \in \mathbb{R}^N$ and $t > 0$, and assume that we have a sequence $t_n \rightarrow \infty$ such that

$$u(\sqrt{t_n}x, t_n) \rightarrow v(x) \quad \text{in } L_{loc}^\infty(\mathbb{R}^N).$$

We can write this as

$$u_{\lambda_n}(x, 1) \rightarrow v(x) \quad \text{in } L_{loc}^\infty(\mathbb{R}^N)$$

with $t_n = \lambda_n^2$. On the other hand, by standard theory, a bounded solution u has a bounded initial trace u_0 at $t = 0$. But the family of rescalings of the initial data $u_{0,\lambda_n}(x) = u_0(\lambda_n x)$ is bounded in $L^\infty(\mathbb{R}^N)$, hence relatively compact in the weak-star topology. This means that, along a finer subsequence, $u_{0,\lambda_n}(x) \rightharpoonup \phi(x)$ for some $\phi \in L^\infty(\mathbb{R}^N)$. By the direct proof we know that $v = S_1(\phi)$.

Remark 2.1. The time $t = 1$ in S_1 is taken for convenience and can be replaced by S_τ for any $\tau > 0$ after changing the function $u(\sqrt{t}x, t)$ into $u((t/\tau)^{1/2}x, t)$. In any case, it must be pointed out that the map S_1 has a regularizing effect on the set of accumulation points of $u_{0,\lambda}$: all functions of the form $S_1(\phi)$ are C^∞ smooth and enjoy certain a priori bounds (cf. Section 5).

Therefore, the key point in our project is understanding the set of accumulation points (i.e., accumulation functions) of the family $u_{0,\lambda}$. The following section will be devoted to analyzing this question. We will show that in the setting of bounded data the situation can be quite complex. We advance a remarkable result, which will be proved in that section.

Theorem 2.2. *Given any bounded sequence $\{g_j : j = 1, 2, \dots\}$ in $L^\infty(\mathbb{R}^N)$ there exists a bounded function $u_0 \in L^\infty(\mathbb{R}^N)$ such that the set of accumulation points of the family $\{u_{0,\lambda}\}_\lambda$ as $\lambda \rightarrow \infty$ in the weak-star topology of $L^\infty(\mathbb{R}^N)$ contains the whole sequence $\{g_j\}$.*

As we will see, the set of accumulation points that contains an infinite sequence will be larger than just this set of points, indeed it can contain a continuum. As an immediate consequence of the two previous results we

have

Corollary 2.1. *For every bounded sequence $\{\phi_j : j = 1, 2, \dots\}$ in $L^\infty(\mathbb{R}^N)$ there exists a bounded function $u_0 \in L^\infty(\mathbb{R}^N)$ such that, if $u(x, t)$ is the corresponding solution of the heat equation, the set of accumulation points of $u(\sqrt{t}x, t)$ as $t \rightarrow \infty$ in $L^\infty_{loc}(\mathbb{R}^N)$ contains the family $\{S_1(\phi_j)\}$.*

To conclude this section we note that the particular results on the asymptotic behaviour mentioned at the beginning of the section can be easily understood in this framework. In the case (i) when $u_0 = c + f$ with $f \in L^1(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$ the unique limit of $u_{0,\lambda}$ is the constant c , and the corresponding solution of the rescaled problem in the limit is $v = c$. On the other hand, when u_0 is periodic, the unique accumulation point of $u_{0,\lambda}$ is the average of the initial data on the periodicity cell and the corresponding solution v is this constant. Finally, in example (iii) the asymptotic limit has the form $v(x, t) = V(xt^{-1/2})$ where $V = S_1H$, H being the Heaviside function (2.2).

§3. The Scaling of Bounded Data

Let us now return to general theory. We are interested in describing the possible behaviour of bounded functions at infinity using as a tool the action of the scaling group $G = \{G_\lambda\}_{\lambda>0}$. G acts on the space $X = L^\infty(\mathbb{R}^N)$ by the formula

$$G_\lambda(f)(x) = f(\lambda x). \tag{3.1}$$

Definition 3.1. *We define the scaling omega-limit set of any function f , or G -omega-limit set, as the set $G_\infty(f)$ of all accumulation points of the family $G_\lambda(f)$ as $\lambda \rightarrow \infty$ in the weak-star topology of $L^\infty(\mathbb{R}^N)$, $\sigma(L^\infty, L^1)$. In the usual notation of dynamical systems we write*

$$G_\infty(f) = \bigcap_{r>1} \overline{\gamma_r^+(f)}, \tag{3.2}$$

where $\gamma_r^+(f)$ is the forward G -orbit of f starting at $\lambda = r$, $\gamma_r^+(f) = \{G_\lambda(f) : \lambda \geq r\}$ and the closure is taken in the weak-star topology.

Since the space $L^1(\mathbb{R}^N)$ is separable, the above topology restricted to a ball $B_c = \{f \in L^\infty(\mathbb{R}^N) : |g| \leq c\}$ is metrizable and we can characterize the G_∞ -limit sequentially. Thus, $G_\infty(f)$ is the set of functions $g \in L^\infty(\mathbb{R}^N)$

such that there exists a sequence $\lambda_n \rightarrow \infty$ and $G_{\lambda_n} f \rightarrow g$ weakly-star. We also have

Lemma 3.1. *The group acts continuously on X endowed with the weak star convergence. It follows easily that the G -omega-limit of a function $f \in X$ is never empty. It is also closed and connected in the weak-star topology. It is also invariant under the action of the group.*

Proof. The first assertion means that the map $G : X \times \mathbb{R}_+ \rightarrow X$ is continuous when X is endowed with the weak-star topology. Since weak-star topologies are not so frequent in the literature on asymptotic phenomena, we are giving a proof for the reader's convenience. Firstly, if we restrict G to be defined on a bounded subset of X (as is the case in the application of the present paper) the topology is metrizable and we only need to check that for a bounded sequence f_n converging weakly-star to f , a sequence of numbers λ_n converging to $\lambda > 0$, and a test function $\phi \in L^1(\mathbb{R}^N)$ we have $\int G_{\lambda_n}(f_n) \phi dx \rightarrow \int G_\lambda(f) \phi dx$. We write the difference as

$$\int G_{\lambda_n}(f_n) \phi dx - \int G_\lambda(f) \phi dx = \int f_n (G_{1/\lambda_n} \phi - \phi) dx + \int (f_n - f) \phi dx. \quad (3.3)$$

The first term tends to zero because ϕ is integrable and the f_n are uniformly bounded, while the second converges because $f_n \rightharpoonup f$.

When G is defined on the whole space X , we have to replace the sequence f_n by a net f_α converging to f weakly-star. But then the Banach-Steinhaus theorem implies that the family f_α is uniformly bounded and we are back in the previous situation. In particular, the decomposition (3.3) with f_α instead of f_n allows to prove just in the same way that $\int G_{\lambda_n}(f_\alpha) \phi dx \rightarrow \int G_\lambda(f) \phi dx$ as $f_\alpha \rightharpoonup f$ and $\lambda_n \rightarrow \lambda$.

The consequences are standard in the dynamical systems literature (cf. [21, 22]). The last assertion means that for every $\lambda > 0$ we have $G_\lambda(G_\infty(f)) = G_\infty(f)$.

Remark 3.1. Though we will use by default the topology of weak-star convergence in $L^\infty(\mathbb{R}^N)$, we will use the term bounded subset of $L^\infty(\mathbb{R}^N)$ in the usual uniform sense unless mention to the contrary.

3.1. Simple G -Omega-Limits

We say that the omega-limit is simple if it contains only one function. Particular examples of simple omega-limits have been mentioned in the

previous section. For case (i) it is clear that all functions $f \in L^p(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$, $1 \leq p < \infty$, have a simple omega-limit, $G_\infty(f) = \{0\}$. In case (ii) all periodic functions have a constant as simple omega-limit. This is also true for the more general class of almost periodic functions \mathcal{A} characterized as the closure in $L^\infty(\mathbb{R}^N)$ (with uniform converge) of set of trigonometric polynomial with arbitrary frequencies,

$$P(x) = \sum_1^n a_k e^{i\omega_k \cdot x} \quad (3.4)$$

with $\omega_k \in \mathbb{R}^N$ (cf. [4]). For a function $f \in \mathcal{A}$ the G_∞ -set is given by the limit of the averages

$$G_\infty(f) = \lim_{R \rightarrow \infty} \oint_{B_R} f dx. \quad (3.5)$$

On the other hand, a non-constant omega-limit is given by the piecewise-constant initial data in one dimension.

A consequence of the invariance of the omega-limit is the following characterization of simple omega-limits.

Lemma 3.1. *If an omega-limit is simple, $G_\infty(f) = \{g\}$, then g has the form $g(x) = g(x/|x|)$.*

It is immediate that this g is its own ω -limit, $g = G_\infty(g)$, though any bounded function f which coincides with g for large x could be used to obtain $G_\infty(f) = g$.

Simple omega-limits for the initial data of an evolution process of the type considered in this paper give rise to evolutions which stabilize to a unique asymptotic profile. We are interested in addressing quite the opposite situation.

3.2. General G -Omega-Limits

At the next level of complexity we construct an example that we call log-periodic. It is based on any function $f \in L^\infty(\mathbb{R}^N)$ such that for some real number $a > 1$,

$$f(ax) = f(x), \quad x \in \mathbb{R}^N.$$

In that case the G -omega-limit is given by the log-periodic orbit

$$G_\infty(f) = \{G_\lambda(f) : 1 \leq \lambda \leq a\},$$

that contains f and its scalings, and it is topologically equivalent to a cir-

cle, S^1 . An easy way of constructing such functions is to consider spherical coordinates $x = (r, \phi)$, take a bounded function $f_1(\phi)$ defined on the sphere S^{N-1} and put $f(r, \phi) = f_1(R(r)(\phi))$, where $R(r)$ is a family of transformations of the sphere such that $R(1) = id$ and $R(ar) = R(r)$ for every $r > 0$, i.e., R is log-periodic.

As we announced in Theorem 2.2, we can construct very large omega-limits. We recast the result in the present notation

Lemma 3.2. *Given any sequence $\{g_j\}$ of functions which is bounded in $L^\infty(\mathbb{R}^N)$ we can construct a function $f \in L^\infty(\mathbb{R}^N)$ such that $G_\infty(f)$ contains the closure of that sequence in the weak-star topology, $G_\infty(f) \supset \text{clo}\{g_1, g_2, \dots\}$.*

Proof. We use a zooming method. Given the function g_i , $i \geq 1$, we restrict it to the annulus $A_j = \{2^{-j} < |x| < 2^j\}$, $j \geq 1$ integer, and transfer this image to a far away distance by means of a scaling with factor $1/\lambda_{ij}$,

$$g_{ij}(x) = g_i(x/\lambda_{ij}), \quad (3.6)$$

which is defined in the annulus $A_{ij} = \{2^{-j}\lambda_{ij} < |x| < 2^j\lambda_{ij}\}$. Next, we select the factors λ_{ij} in such a way that all these sets are disjoint. This can be obtained as follows: we first arrange the indexes (i, j) in a sequence by standard diagonal process, and then choose the sequence λ_{ij} iteratively so that each annulus lies immediately outside of the preceding one. We can also leave gaps between successive annuli if we wish, only this makes the λ_{ij} even larger. We then define the desired f on each annulus A_{ij} by

$$f(x) = g_i(x/\lambda_{ij}), \quad x \in A_{ij}. \quad (3.7)$$

This formula gives then a unique value for f on $\bigcup_{ij} A_{ij}$ since the annuli A_{ij} are disjoint. If there exist gaps between successive annuli, we define f on the gaps as zero or in any other bounded way.

It is then clear that applying the group action G_λ to f with the sequence of factors λ_{ij} , for fixed i and variable $j = 1, 2, \dots$, we obtain in the limit exactly the function g_i on the annulus A_j . Since the sequence A_j expands as $j \rightarrow \infty$, we get along this sequence $G_\lambda(f) \rightarrow g_i$ with uniform convergence on compact subsets of $\mathbb{R}^N \setminus \{0\}$. This in particular implies weak-star convergence in \mathbb{R}^N .

3.3. Maximal Omega-Limits

As a corollary of this result and the fact that the unit ball in $L^\infty(\mathbb{R}^N)$ is separable in the weak-star topology we have the following consequence.

Corollary 3.1. *Given $C > 0$ there exists a function $f \in L^\infty(\mathbb{R}^N)$ with $\|f\|_\infty = C$ such that $G_\infty(f) = \{f \in L^\infty(\mathbb{R}^N) : \|f\|_\infty \leq C\}$.*

Let us look a bit closer at this kind of functions. We define the set

$$\mathcal{M} = \{f \in L^\infty(\mathbb{R}^N) : \|f\| = 1, G_\infty(f) = B_1(L^\infty(\mathbb{R}^N))\},$$

where $B_1(L^\infty(\mathbb{R}^N))$ is the unit ball in $L^\infty(\mathbb{R}^N)$. Then we have

Theorem 3.1. *The set \mathcal{M} is dense with empty interior in $B_1(L^\infty(\mathbb{R}^N))$ with the weak-star topology.*

Proof. A basis of open neighbourhoods of a function $f \in L^\infty(\mathbb{R}^N)$ has the form

$$B(\phi_1, \dots, \phi_n, \epsilon) = \left\{ g \in L^\infty(\mathbb{R}^N) : \left| \int (f - g)\phi_i dx \right| \leq \epsilon, \quad \forall i \right\},$$

where $\phi_i \in L^1(\mathbb{R}^N)$, $i = 1, \dots, N$, and $\epsilon > 0$. This condition does not depend on the behaviour of g at far away distances as long as we have a bound like $|g| \leq C$. But the behaviour for large x is precisely what determines $G_\infty(f)$.

§4. General Evolution Setting. Main Asymptotic Formula

If we consider again the argument used in Section 2 for the classical heat equation, we can see that there are two types of properties of the equation on which the proof is based, namely (I) the heat equation defines a semigroup with some good properties, and (II) this semigroup is invariant under the action of the scaling group G .

In order to set up a framework that applies to a number of linear or non-linear evolution models we introduce the following abstract setting. Instead of concentrating on the evolution equation

$$u_t = A u, \tag{4.1}$$

we look rather to the fact that it generates a semigroup. We assume that

(I.i) The equation generates a point-wise continuous semigroup in $X = L^\infty(\mathbb{R}^N)$, i.e., given any initial data $u_0 \in X$ there exists a unique solution

$u \in Y = C([0, \infty) : X)$ that we write as $S_t(u_0)$ to denote the dependence on the initial data. Then

$$S : X \times [0, \infty) \rightarrow X, \quad S(u_0, t) = S_t(u_0)$$

is the semigroup map. When we want to stress the dependence on x we write $(S_t u_0)(x) = u(x, t)$, and sometimes we write $u(t)$ when no confusion arises. X is endowed with the weak-star topology $\sigma(L^\infty, L^1)$. We impose still an additional condition

(I.ii) For every $t > 0$ the map S_t is continuous in the weak-star topology of $X = L^\infty(\mathbb{R}^N)$.

Conditions (I.i) and (I.ii) imply that the map S is separately continuous in both variables.

(II) We impose another type of condition on the semigroup, namely scale invariance. By this we mean that there is a power function $\sigma(\lambda) = \lambda^\alpha$, $\alpha > 0$, such that whenever $u(x, t)$ is a solution in Y with initial data u_0 then

$$\tilde{u} = u(\lambda x, \sigma(\lambda)t) \tag{4.2}$$

is also a solution in Y , with data $\tilde{u}_0 = G_\lambda u_0$. In the case of the heat equation $\sigma(\lambda) = \lambda^2$. This property is expressed in functional terms as $S_t(G_\lambda(u_0)) = G_\lambda(S_{\sigma(\lambda)t}(u_0))$, or in abridged form as the commutation rule:

$$S_t G_\lambda = G_\lambda S_{\sigma(\lambda)t}. \tag{4.3}$$

The next stage is to note that we are interested in the convergence on expanding sets of the form $\{|x| \leq c\psi(t)\}$. This makes it convenient to define the rescaled orbit corresponding to a standard orbit $u(\cdot, t) = S_t u_0$ as

$$U(y, t) = u(\psi(t)y, t), \tag{4.4}$$

where $\psi = \sigma^{-1}$ is the inverse function of σ , $\psi(t) = t^{1/\alpha}$. We denote the “renormalized semigroup” as

$$(R_t u_0)(y) = U(y, t), \tag{4.5}$$

which implies the identity: $R_t = G_{\psi(t)} S_t$. Notice that $R_1 = S_1$. The invariance under G_λ takes a much simpler form in rescaled form. If \tilde{U} is the rescaled orbit corresponding to initial data $\tilde{u}_0(x) = u_0(\lambda x)$ then, since

$\tilde{U}(y, t) = \tilde{u}(\psi(t)y, t) = u(\lambda\psi(t)y, \sigma(\lambda)t)$, we have

$$\tilde{U}(y, t) = U(y, \sigma(\lambda)t), \quad \text{i.e.,} \quad R_t G_\lambda u_0 = R_{\sigma(\lambda)t} u_0. \tag{4.6}$$

We define the omega-limit of the R -evolution orbit in a similar way to what we did for the scaling group G , but now we use R_t instead of G_λ .

Definition 4.1. *The R -omega-limit set of the evolution orbit $\{R_t(u_0) = U(\cdot, t)t > 0\}$ starting at $u_0 \in X$ is the set*

$$R_\infty(u_0) = \{g \in X : \exists t_n \rightarrow \infty \text{ and } U(\cdot, t_n) \rightarrow g\}, \tag{4.7}$$

with convergence in the topology of X .

The basic properties are easily established: the R -omega-limit set is a nonvoid, closed, connected subset of X . Our result relates the omega-limits of the scaling group G_λ and the scaled evolution R_t .

Theorem 4.1 (Main Asymptotic Formula). *If S_t is a semigroup fulfilling properties (I) and (II) above, then for every $u_0 \in X$ we have*

$$R_\infty(u_0) = S_1(G_\infty(u_0)). \tag{4.8}$$

Proof. It repeats the lines of Theorem 2.1. Let $g \in G_\infty(u_0)$ and let $\lambda_n \rightarrow \infty$ be such that $G_{\lambda_n}(u_0) \rightarrow g$ weak-star. We note that $R_{\sigma(\lambda)}u_0 = R_1(G_\lambda u_0) = S_1(G_\lambda u_0)$. Using the property (I.ii) we conclude that along the sequence $t_n = \sigma(\lambda_n)$,

$$\lim_{n \rightarrow \infty} R_{t_n} u_0 = \lim_{n \rightarrow \infty} S_1(G_{\lambda_n} u_0) = S_1(g).$$

Thus, $R_\infty(u_0) \subset S_1(G_\infty(u_0))$. The converse is similar.

We conclude that the orbit $R_t(u_0)$ stabilizes to a limit only if $G_\infty(u_0)$ is simple. Using the fact that $G_\infty(u_0)$ is closed under the action of G and that the renormalized semigroup R_t commutes with G in the sense of (4.6) we have

Corollary 4.1. *Given $u_0 \in X$ and given $g \in G_\infty(u_0)$ then $R_t(g) \in R_\infty(u_0)$ for every $t > 0$.*

Thus, the whole renormalized orbit of g is contained in the omega-limit of u_0 , proving in this way that the asymptotic behaviour of $u(x, t)$ can oscillate strongly even at a fixed point, say, $x = 0$. Summing up, the results show that the evolution equation does not stabilize to a certain asymptotic profile for general initial data in X , but it rather copies the complexity of the G -omega-limit which, as we have seen, can be quite wild.

4.1. Regularizing Semigroups

In many cases we have the extra condition on the semigroup:

(I.iii) The semigroup is regularizing. By this we mean that S_t maps bounded sets of X into relatively compact subsets in a better space $X_1 \subset X$ for $t > 0$. In the case of the heat equation X_1 is the set $BC(\mathbb{R}^N)$ of continuous and bounded functions with the topology of uniform convergence on compact sets. Then the convergence in the definition of $R_\infty(u_0)$ is uniform over compact subsets of \mathbb{R}^N .

4.2. Local and Global Convergence

We insist that the local convergences that we have established for the orbit of the semigroup are uniform on compact sets of the variables $y = \sqrt{t}x$. In particular, when the semigroup is regularizing we may consider the space variable x ranging over a fixed bounded set $A \subset \mathbb{R}^N$, and then we obtain the following uniform convergence result.

Theorem 4.2. *Let S_t be a regularizing semigroup under the above hypotheses and let $u(x, t)$ be an orbit with initial data $u_0 \in L^\infty(\mathbb{R}^N)$. Then for every $g \in G_\infty(u_0)$ there exists a sequence $t_n \rightarrow \infty$ such that*

$$u(x, t_j) \rightarrow v(0, 1) \quad \text{as } t \rightarrow \infty \quad (4.9)$$

uniformly on compact sets $A \subset \mathbb{R}^N$, where v is the solution of the heat equation with initial data g .

In particular, there exist bounded initial data for which the set of accumulation points of $\{u(0, t)\}$ as $t \rightarrow \infty$ is the closed interval $[a, b] \subset \mathbb{R}$. It suffices to take as set g_j in Lemma 3.2 the set of constant functions $g_j(x) = r_j$, where r_j ranges over a dense set in $[a, b]$. The same applies to the omega-limit of $\{u(x, t)\}$ for fixed $x \in \mathbb{R}^N$.

On the other hand, the orbits with complex asymptotics cannot have uniform convergence since the behaviour that will appear at later stages of the evolution must be present at any smaller t , typically as $|x| \rightarrow \infty$.

4.3. Log-Periodic Solutions

Corresponding to log-periodic data, $u_0(ax) = u_0(x)$ for some $a > 0$, we have a solution $u(x, t)$ such that $u(ax, \sigma(a)t) = u(x, t)$. In terms of the rescaled orbits $U(x, t) = u(\sqrt{t}x, t)$ we have $U(x, \sigma(a)t) = U(x, t)$. This means that the rescaled orbit is log-periodic in time with log-period $T = \sigma(a)$. Recall that this is the same as saying that it is periodic in logarithmic

time $\tau = \log(t)$. In this case we can control the speed of recurrence of the orbit around the different points of the omega-limit, which is not true for more complex situations.

§5. The Heat Equation Revisited

Since the heat equation generates a regularizing semigroup, the results of the last sections apply to say that the set of accumulation points of the renormalized orbit $U(x, t) = u(\sqrt{tx}, t)$, which we will denote by $R_\infty(u_0)$, is the image by S_1 of the set of accumulation points $G_\infty(u_0)$. The convergence to the omega-limit $R_\infty(u_0)$ takes place locally uniformly in x along subsequences $t_n \rightarrow \infty$.

Range Constraints. We note that $R_\infty(u_0)$ can be a very large set. The maximum extension of such a set is the image of a ball in $L^\infty(\mathbb{R}^N)$. We have to add that, since the evolution copies the set $G_\infty(u_0)$ in a regularized way given by the map S_1 , this reduces the range of $R_\infty(u_0)$ since S_1 has strong regularizing properties. In other words, there are some constraints for a function to belong to $R_\infty(u_0)$ for some u_0 . In the first place, the solutions of the heat equation are C^∞ functions. On the other hand, there are quantitative aspects: thus, for every $u_0 \geq 0$ there is an a priori second-order estimate $\Delta \log(u) \geq -\frac{N}{2t}$ (cf. [1]). This implies a universal constraint on the range of any omega-limit set $R_\infty(u_0)$.

Corollary 5.1. *For every $g \in R_\infty(u_0)$ with $u_0 \geq 0$ we have*

$$\Delta \log(g) \geq -\frac{N}{2}. \quad (5.1)$$

There is a version of this result for any bounded set in $L^\infty(\mathbb{R}^N)$, not just for positive solutions, since a lower bound of the form $u \geq -c$ can be converted into $v \geq 0$ by defining $v = u + c$ which is again a solution of the heat equation. In this way, we obtain a lower bound for $\Delta \log(g+c)$ instead of $\Delta \log(g)$.

Let us remark that the extent of this regularization depends on the amount of renormalization. Thus, a renormalization with larger compression, using a function of the form

$$u(cyt^{1/2}, t), \quad c > 1,$$

instead of Definition 4.1, produces a renormalized semigroup whose omega-

limit satisfies

$$R'_\infty(u_0) = S_{1/c^2}(G_\infty(u_0))$$

which is just a rescaled version of the previous one, but this time the evolution is shorter for $c > 1$ and the regularization smaller. Thus, the a priori bound from below on the Laplacian of $\log(u)$ is now $-c^2N/2$. In the limit $c \rightarrow \infty$ we recover increasingly the whole range of G_∞ .

§6. Nonlinear Heat Equations

The above general setting applies perfectly not only to the heat equation, but also to the most popular models of quasilinear heat equations, namely the porous medium equation and the p -Laplacian equation.

6.1. The Porous Medium and Related Equations

We will restrict here consideration to the subset of nonnegative solutions, hence $X = L^\infty_+(\mathbb{R}^N)$, which is the case treated in detail in the literature because of its applications, and avoids technical difficulties. It is well-known (cf. [2]) that the equation

$$u_t = \Delta(u^m), \quad m > 0, \tag{6.1}$$

defines a semigroup of contractions in $L^1(\mathbb{R}^N)$ and can be extended to a much wider class of initial data, in particular to $L^\infty_+(\mathbb{R}^N)$. The equation is called the porous medium equation for $m > 1$, the fast diffusion equation for $0 < m < 1$ and it reduces to the classical heat equation for $m = 1$.

Let us examine the properties of the evolution process. In all cases $m > 0$, for every $u_0 \in L^\infty(\mathbb{R}^N)$, $u_0 \geq 0$ there exists a unique continuous weak solution $u(x, t)$, i.e., $u \in BC(Q) \cap C([0, \infty) : L^1_{loc}(\mathbb{R}^N))$ and

$$\int \int \{u\eta_t + u^m \Delta\eta\} dxdt + \int u_0(x)\eta(x, 0) dx = 0 \tag{6.2}$$

holds for every test function $\eta \in C^{2,1}(Q)$, $\eta \geq 0$, which is compactly supported in $\mathbb{R}^N \times [0, \infty)$. Hence, the equation generates a semigroup in this class of solutions and the property (I.i) holds. Also, the Maximum Principle holds. The closure property (I.ii) holds easily: the limits of bounded weak solutions are still bounded weak solutions and weak-star converging initial data pass to the limit in formula (6.2). This uses the fact that the map S_t is regularizing, property (I.iii), since the images of bounded sets are

relatively compact subsets of C^α for some $\alpha > 0$ in compact sets of \mathbb{R}^N (cf. [10]). Finally, scale invariance holds with $\sigma(\lambda) = \lambda^2$. We then have

Theorem 6.1. *The results of Theorems 4.1 and 4.2 apply to this problem and the convergence in the definition of R_∞ , formula (4.7), takes place uniformly on compact sets.*

The possible range of the omega-limit set $R_\infty(u_0)$ is now controlled by the constraints

$$\Delta(u^{m-1}) \geq -\frac{C}{t}, \quad u_t \geq -\frac{Cu}{t},$$

which hold for $m > 1$, while for $(N-2)/N < m < 1$ we have even stronger constraints

$$|\Delta(u^{m-1})| \leq \frac{C}{t}, \quad |u_t| \leq \frac{Cu}{t}.$$

These a priori bounds have universal constants $C = C(m, N) > 0$ independent of the solution under consideration (under the only restriction that $u \geq 0$) (cf. [1,6]). For instance, when $0 < m < 1$ we have $\Delta(u^{m-1}) \geq \frac{m}{t}$. Let us finally remark that some of the results can be extended to the very fast diffusion range $m < 0$, where the equation is written as $u_t = \nabla \cdot (u^{m-1} \nabla u)$ (cf. for instance [33] and references therein). More generally, we can consider the general filtration equation $u_t = \Delta \Phi(u)$, where Φ is a continuous increasing real function, or even a maximal monotone graph in \mathbb{R}^2 (cf. [2]). In order not to lengthen the presentation we restrain from entering into the specific details of these extensions. Let us only assert that the general theory of Section 4 applies under suitable conditions on Φ .

6.2. The p -Laplacian Equation

The equation reads

$$u_t = \text{Div}(|\nabla u|^{p-2} \nabla u), \quad p > 1. \quad (6.3)$$

The properties of the evolution are similar [11], and a regularizing semi-group is obtained. Scale-invariance holds with $\sigma(\lambda) = \lambda^p$. The regularizing effect holds: S_t takes bounded initial data into solutions which are $C^{1+\alpha}$ in x , C^β in time for $t \geq \tau > 0$. Regularizing estimates have been proved in [16] and take the form for $p > 2$, $\Delta_p v \geq -\frac{C}{t}$, $u_t \geq -\frac{Cu}{t}$, where $v = u^{(p-2)/(p-1)}$ and $C > 0$ is a universal constant that depends only on p and N . The estimate is valid for all nonnegative solutions $u = u(x, t)$. But the restriction to

positive data is here irrelevant since the equation is invariant under vertical displacements $u \rightarrow u + c$.

§7. Scalar Conservation Laws. The Compact Case

We consider in this section the asymptotics of scalar conservation laws in one and more space variables.

We start with the one-dimensional theory which relies on better known results and requires less assumptions on the non-linearity.

7.1. One Space Dimension

We study the first-order equation

$$u_t + f(u)_x = 0, \quad (7.1)$$

where $u(x, t)$ is a scalar function and f is a locally Lipschitz continuous scalar function. We may assume that $f(0) = 0$. The proper concept of solution is given by the entropy solutions that can be formulated as those bounded distributional solutions that satisfy a family of entropy inequalities. We take from [19, 20] the precise definition of solution.

Definition 7.1. *An entropy solution of equation (7.1) is a function $u \in L^\infty_{\text{loc}}(Q)$ satisfying*

$$\frac{\partial}{\partial t}|u - k| + \frac{\partial}{\partial x} \{ \text{sign}(u - k)(f(u) - f(k)) \} \leq 0 \quad (7.2)$$

in $\mathcal{D}'(Q)$ for any $k \in \mathbb{R}$. An entropy solution of the Cauchy problem with initial data $u(x, 0) = u_0(x) \in L^\infty(\mathbb{R})$ is an entropy solution such that

$$u(\cdot, t) \rightarrow u_0 \quad \text{in } L^1_{\text{loc}}(\mathbb{R}) \quad \text{as } t \rightarrow 0 \text{ essentially.}$$

The Kruzhkov condition (7.2) can be equivalently formulated as two conditions: (i) the equation is satisfied in the sense of distributions, (ii) for every entropy pair (ϕ, ψ) where ϕ is a convex function and $\psi'(u) = f'(u)\phi'(u)$ we have

$$\frac{\partial}{\partial t}\phi(u) + \frac{\partial}{\partial x}\psi(u) \leq 0$$

in the sense of distributions (cf. [28]).

Kruzhkov proves that the bounded entropy solution of the initial problem exists and is unique, and we have the local estimate for any two such

solutions

$$\int_{B_t} |u(x, t) - v(x, t)| dx \leq \int_{B_0} |u_0(x) - v_0(x)| dx, \tag{7.3}$$

where $B_0 = B(0, R)$ is the ball of center 0 and radius $R > 0$ and $B_t = B(0, R(t))$ is the ball of radius $R(t) = R - Nt$, where N is the Lipschitz constant of f in the interval $[-M, M]$ with $M = \max\{\|u_0\|_\infty, \|v_0\|_\infty\}$. The Maximum Principle applies:

$$u_0 \leq v_0 \quad \text{implies} \quad u(x, t) \leq v(x, t) \quad \text{in } Q.$$

Moreover, it is proved that the solution belongs to the class $u \in C([0, \infty) : L^1_{loc}(\mathbb{R})) \cap L^\infty(Q)$. We have therefore a semigroup of solutions in $L^\infty(\mathbb{R})$ (cf. [9]).

On the other hand, there is a large literature on the asymptotic behavior of solutions of scalar conservation laws. For instance, in the context of periodic solutions, it is by now well known even for 1D systems of hyperbolic conservation laws, that solutions stabilize around a constant (cf. [7]). In what concerns L^1 -solutions, the asymptotic profiles of solutions are also well known (see for instance [24]). A lot is also known about the stability of shock waves (see for instance [17]).

However, here, we are interested in the possible behavior of all solutions with bounded initial data.

One of the main difficulties of applying the scaling techniques as above is the obtainment of compactness. Here we shall use Tartar’s results that hold under a suitable non-degeneracy condition on the nonlinearity. The problem of identifying the initial data of the limiting solutions will require also special care.

Theorem 7.1. *Assume that f is not affine on any interval of the real line. Then the equation (7.1) generates a regularizing semigroup S_t in $L^\infty(\mathbb{R})$ satisfying the conditions of Section 4, so that Theorems 4.1 and 4.2 apply. The scaling law is now linear, $\sigma(\lambda) = \lambda$.*

Proof. Properties (I.i) and (II) are already established. Under the stated conditions on f , also called weakly genuine nonlinearity, Tartar [29, 30] proves that bounded sequences of entropy solutions are relatively compact in the strong topology of $L^p_{loc}(Q)$ for any $p < \infty$. Using the local estimate we conclude compactness in the space $C([\tau, T] : L^1_{loc}(\mathbb{R}))$ for every

$0 < \tau < T < \infty$. It is immediate to see that the limit $v(x, t)$ of such a sequence is an entropy solution in any time interval $t \geq \tau$ with $\tau > 0$.

We still have to examine property (I.ii), i.e., the weak-star continuity of the semigroup map, in the present setting. This is a delicate but essential point. Indeed, the family of rescaled solutions $u_\lambda(x, t)$ of Theorem 4.1 are such that their initial data converge weak-star in $L^\infty(\mathbb{R})$ along subsequences, $u_{0, \lambda_n} \rightarrow \phi$ in $L^\infty(\mathbb{R})$ -ws, and we want to conclude that $u_{\lambda_n}(x, 1)$ converges to $v(x, 1) = S_1(\phi)$. We state the general result independently

Lemma 7.1. *Let $\{u_n\}$ be a sequence of bounded entropy solutions of equation (7.1) with initial data $\{u_{0,n}\}$ which converge to a certain ϕ in $L^\infty(\mathbb{R})$ weak-star. Then $u_n \rightarrow v = S_t\phi$ in the sense of $C([\tau, T] : L^p_{\text{loc}}(\mathbb{R}))$ for every $0 < \tau < T < \infty$ and every $1 \leq p < \infty$.*

Proof. (i) By the results of [29] the sequence u_n converges in $L^p_{\text{loc}}(Q)$ and also a.e. after passing if necessary to a subsequence. The limit v is an entropy solution of the equation. The remaining problem consists of identifying the solution in terms of the limit ϕ of the initial data. The classical uniqueness result by Kruzhkov is not enough for our purposes since we do not have strong convergence of $v(t)$ in $L^1_{\text{loc}}(\mathbb{R})$ as $t \rightarrow 0$. In other words, we have to discard the possible occurrence of an initial layer of discontinuity. This difficulty has been recently solved by Chen and Rascle [8] who proved uniqueness of entropy solutions of equation (7.1) assuming that the initial data are taken in the sense of measures.

(ii) Here we give an alternative proof which easily follows from the uniqueness result of Liu and Pierre [26].

Proposition 7.1. *Under the assumption that $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is Lipschitz continuous and $f(0) = 0$ a nonnegative entropy of equation (7.1) such that $u \in L^\infty((0, T) : L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times (\tau, T))$, for all $\tau > 0$, is uniquely determined by its initial data u_0 taken in the narrow sense, i.e.,*

$$\lim_{t \rightarrow 0} \int u(x, t)\eta(x) dx = \int u_0(x)\eta(x) dx$$

for every continuous and bounded real function η .

Indeed, [26] allows the initial data to be finite measures, but this is no concern for us here. Let us note that such a generality forces them to work with nonnegative f and nonnegative solutions, two restrictions that we can dispense with, since we can always displace u into $u + c$ and conserve the

equation after a corresponding change in f to satisfy both restrictions.

Let us check the initial data of our limit solution v in the case when it has compact support: passing to the limit in the distributional identity

$$\int (u_n(x, t) - u_{0,n}(x))\eta(x) dx = \int_0^t \int f(u_n(x, t))\eta(x)_x dx dt$$

valid for every $\eta \in C_0^1(\mathbb{R})$ we get

$$\int (v(x, t) - \phi(x))\eta(x) dx = \int_0^t \int f(v(x, t))\eta(x)_x dx dt.$$

By density and using the fact that the support of v is uniformly bounded we prove that

$$\int u(x, t)\eta(x) dx \rightarrow \int u_0(x)\eta(x) dx$$

for every $\eta \in C(\mathbb{R})$. Under these conditions the result of Liu and Pierre [26] and the remark allow us to identify v as the unique entropy solution with initial data ϕ , $v(t) = S_t\phi$.

(iii) Unfortunately, the condition of belonging to $L^1(\mathbb{R})$ is not fulfilled in our case. But the above uniqueness result can still be used, arguing as follows. Let us take a radius $R > 0$, a time $T > 0$ and work in the cylinder $Z = B(0, R) \times (0, T)$. By the local estimates (7.3) the values of the functions u_n on Z are only determined by the initial data in $B(0, R_1)$ with $R_1 = R + NT$. Therefore, we may replace the data $u_{0,n}$ by functions $u'_{0,n}$ which coincide with $u_{0,n}$ on $B(0, R_1)$ and are zero for $|x| > R_1$. The solutions $u'_n(x, t)$ will vanish outside an expanding cone $K = \{(x, t) : |x| > R_1 + Nt\}$ according again to estimate (7.3). Summing up, we may assume that our functions u_n belong to the space $C([0, \infty) : L^1(\mathbb{R}))$ without altering their value on Z .

Remark About Riemann Problems. A typical problem in scalar conservation laws is the Riemann Problem, which consists of taking initial data of the step or Heaviside type, (2.2), in which case the solution has the form of a shock or a rarefaction wave, or combinations of these two types. We observe that for data u_0 whose G_∞ limit is a Heaviside function, $G_\infty(u_0) = \{H\}$, we obtain convergence of the rescaled solution to the profile at time $t = 1$ of the solution of the Riemann problem with data $H(x)$. This is well-known, but we can also consider initial data whose G_∞

limit contains any number of Heaviside functions, so that the rescaled orbit oscillates infinitely many times among the corresponding profiles.

7.2. Several Space Dimensions

We look now at the scalar conservation law in several dimensions,

$$\frac{\partial u}{\partial t} + \sum_{i=1}^N \frac{\partial}{\partial x_i} f_i(u) = 0, \quad (7.4)$$

where $u(x, t)$ is a scalar function and $f = (f_1, \dots, f_N)$ is a locally Lipschitz continuous vector function. We may assume that $f(0) = 0$. The proper concept of solution is given again by the entropy solutions with a formula similar to (7.2) (cf. [19, 20]). The above theory applies almost literally with the following changes:

(i) The compactness property has been established by Lions, Perthame and Tadmor [25] under the following non-degeneracy condition on the non-linearity: for every $(\tau, \zeta) \in \mathbb{R} \times \mathbb{R}^N$ which is not $(0, 0)$ there holds

$$\text{meas} \{ \xi : \tau + \zeta \cdot f'(\xi) = 0 \} = 0, \quad (7.5)$$

where meas refers to the Lebesgue measure in \mathbb{R} . Their method relies on the kinetic formulation of the equation, a powerful tool for the several dimensional problem introduced by the authors.

(ii) The uniqueness of entropy solutions taking bounded data in the weak sense has been recently proved by A. Vasseur [32] using also the kinetic formulation of the equation. The result needs the following regularity on the nonlinearity: $f \in (C^3(\mathbb{R}))^N$. We conjecture that this requirement is technical.

Theorem 7.2. *With these two requirements on f , Theorem 7.1 holds with \mathbb{R} replaced by \mathbb{R}^N and equation (7.1) by (7.4).*

§8. Lack of Compactness

The property of continuity in the weak topology needed for the semigroup is difficult to establish in the case of nonlinear equations unless we also have compactness, i.e., a regularizing semigroup. This last property holds for the PME and the PLE and we imposed conditions of non-degeneracy to have it in the scalar conservation law. However, the Main Asymptotic Formula holds for many equations without compactness. We list below two instances.

I. Linear Equations. One of the simplest examples is the case of linear operators, like the scalar conservation law $u_t + a u_x = 0$, even in N dimensions,

$$u_t + \sum_{i=1}^N a_i u_{x_i} = 0.$$

By means of the explicit representation of solutions it is easy to see that no regularization happens but the semigroup has the desired properties, hence $R_\infty = S_1 G_\infty$.

The argument applies of course in the heat equation, but also for the wave equation, where the representation is also explicit and no regularization holds.

II. Scalar Conservation Laws. But the result also holds for some nonlinear evolution equations without compactness. Thus, we can study the scalar conservation law (7.1) of previous section with the assumption of local Lipschitz continuity on f but no condition of non-degeneracy. In this case a result of Tartar [30, p.202] states that

Weak Continuity. Whenever u_n is a sequence of entropy solutions of equation (7.1) which are uniformly bounded ($|u_n(x, t)| \leq C$ for every n, x, t) and $u_n \rightharpoonup u$ in $L^\infty\text{-}w^*$, then $f(u_n) \rightharpoonup f(u)$ in the same sense. On the other hand, $f'(u_n)$ converges to $f'(u)$ locally in L^p strong.

Using this result we can perform the asymptotic study of the scalar conservation law in one dimension for any locally Lipschitz continuous real function f . The scaling law is linear, $\sigma(\lambda) = \lambda$. But, as far as we know, the weak-continuity result is not known in several space dimensions, and the asymptotic analysis cannot proceed for equation (7.4) with general Lipschitz f_i . The question seems difficult and needs further study.

§9. Extending the Scope of the Theory

The general setting outlined above for the application to the heat equations and scalar conservation laws in $L^\infty(\mathbb{R}^N)$ admits three kinds of extensions where the Main Asymptotic Formula $R_\infty(u_0) = S_1(G_\infty(u_0))$ still holds when suitably interpreted:

- (i) it can be applied to a number of other equations,

(ii) it can be modified to include cases where scale invariance is not strictly respected,

(iii) it can be extended so as to include asymptotic results in other functional settings under the cover of the same formula.

The L^1 Setting. Leaving aside for the moment points (i) and (ii), we will devote this section to showing how to adapt the setting of Section 4 to recover the well-known results on asymptotic behaviour of heat equations and conservation laws in the framework of $L^1(\mathbb{R}^N)$. Indeed, we can keep the semigroup setting of Sections 3, 4 with the following modifications.

- The proper concept of scaling of the data in $L^1(\mathbb{R}^N)$ is given by the group $G^{(1)} = \{G_\lambda^{(1)}\}$ acting by $G_\lambda^{(1)}(f) = \lambda^N f(\lambda x)$. Again, this group is continuous in the weak topology $\sigma(L^1, L^\infty)$. We must now note that $L^1(\mathbb{R}^N)$ is not closed in this topology, its closure being the space of bounded measures $\mathcal{M}(\mathbb{R}^N)$, which is endowed with the so-called narrow convergence. Indeed, with a definition of $G^{(1)}$ -omega-limit as in Section 3, it is easy to see that

Lemma 9.1. *For every $f \in L^\infty(\mathbb{R}^N)$ we have a simple omega-limit, $G_\infty^{(1)}(f) = \{c\delta\}$, where $c = \int f(x) dx \in \mathbb{R}$ and δ is the Dirac mass located at $x = 0$.*

- The equation generates a semigroup acting on $X = L^1(\mathbb{R}^N)$. We ask that

(I) The map S is separately continuous in both variables. X is endowed with the weak topology. We also need the semigroup to extend continuously to include the space of measures, or at least, the Dirac deltas.

(II) Scale invariance is formulated in the new setting as

$$u_\lambda(x, t) = \lambda^N u(\lambda x, \lambda^\alpha t),$$

with suitable exponent α . We want u_λ to be a solution of the equation if u is, i.e., we need the commutation rule $S_t G_\lambda u_0 = G_\lambda S_{\lambda^\alpha t}$.

Under these conditions we define the rescaled orbit corresponding to a standard orbit $u(\cdot, t) = S_t u_0$ as

$$U(y, t) = t^{N/\alpha} u(t^{1/\alpha} y, t), \tag{9.1}$$

and the renormalized semigroup as $(R_t u_0)(y) = U(y, t)$, which implies the identity

$$R_t u_0 = G_{t^{1/\alpha}} S_t u_0. \quad (9.2)$$

The R -omega-limit is defined as in Section 4, and it stays in $L^1(\mathbb{R}^N)$ if we make sure that convergence takes place in the strong topology (otherwise, it might become a measure).

Theorem 9.1. *Under the above assumptions the Main Asymptotic Formula holds true. It means that the R -omega-limit of the orbits of the semigroup is given by the unique profile at $t = 1$ of the fundamental solution of the equation.*

If the semigroup is regularizing, i.e., if the map S_t sends bounded subsets of $L^1(\mathbb{R}^N)$ into relatively compact subsets in the strong topology, then the convergence of the rescaled orbit to the R_∞ -limit holds in the strong topology.

Applications. Indeed, the four types of equations already considered generate a regularizing semigroup acting on $L^1(\mathbb{R}^N)$.

(i) In the case of the heat equation the solution with a Dirac delta exists, is unique, and has the necessary continuity property. It is called the fundamental solution and is given by formula (1.1). Let us remark that the L^1 norm is conserved by the semigroup, $\int u(x, t) dx = u_0(x) dx$. The exponent α in the scaling law is $1/2$ as in the L^∞ case. The regularizing effect takes $L^1(\mathbb{R}^N)$ into bounded subsets of C^k for all k (regularity both in space and time). The theorem states the well-known fact that for every $u_0 \in L^1(\mathbb{R}^N)$ we have

$$\lim_{t \rightarrow \infty} t^{N/2} u(\sqrt{t}x, t) = c(4\pi)^{-N/2} \exp\left(-\frac{x^2}{4}\right),$$

and the convergence is uniform in \mathbb{R}^N .

(ii) For the Porous Medium Equation we have a fundamental solution, usually called Barenblatt solution, if $m > m_* = (N-1)/N$, and the scaling exponent is $\alpha = N(m-1) + 2$. Again the semigroup is regularizing into C^β for some $0 < \beta < 1$ (regularity both in space and time). In this parameter range the theorem applies (cf. [34]).

(iii) For the p -Laplacian equation we have $\alpha = N(p-2) + p$, and we impose the restriction $p > 2N/(N+1)$ in order to obtain a fundamental solution, also called Barenblatt solution. Regularization occurs into $C^{1,\alpha}$

in space, C^β in time.

(iv) For the scalar conservation law we must take $f(u) = u^p$ with $p \geq 1$ and then $\alpha = N(p - 1) + 1$.

In all cases the L^1 norm is conserved

$$\int u_\lambda(x, t) dx = \int u(x, t) dx = \int u_0(x) dx.$$

Note however that the existence of fundamental entropy solutions for scalar hyperbolic conservation laws is only known in one space dimension. It is unique in the class of constant sign solutions, but the occurrence of N -waves may produce non-uniqueness phenomena, depending on whether the nonlinearity f is even or odd (cf. [26]).

§10. Quasi-Invariance

We may go back to the abstract setting of Section 4 and relax both aspects of the assumption list and still follow the argument advanced in Section 2. We introduce next the concepts of quasi-invariance and limit equation and arrive at a generalized version of the Main Asymptotic Formula.

Quasi-Invariance and Limit Equation. A simple case of this project is exemplified with the viscous approximation to a conservation law, governed by equation

$$u_t + f(u)_x = \varepsilon u_{xx}, \tag{10.1}$$

with $\varepsilon > 0$ which approximates and regularizes the solutions of the scalar conservation law (7.1). When we perform the scaling $u_\lambda(x, t) = u(\lambda x, \lambda t)$, we see that u_λ does not satisfy the same equation, it satisfies instead

$$u_{\lambda,t} + f(u_\lambda)_x = \frac{\varepsilon}{\lambda} u_{\lambda,xx}.$$

Using the same definition of rescaled orbit and passing to the limit $\lambda \rightarrow \infty$ along a subsequence we get as in Section 7 a limit function v that satisfies the limit equation

$$v_t + f(v)_x = 0. \tag{10.2}$$

This is called quasi-invariance, a concept that relates the scaling of equation (10.1) to (7.1). The compactness argument of Tartar applies to the family of equations and we get in the limit an entropy solution with initial data $\phi = \lim_n u_{0,\lambda_n}$. Summing up, we have

Theorem 10.1. *The main asymptotic formula $R_\infty(u_0) = S_1G_\infty(u_0)$ holds if R_∞ is the omega-limit of the rescaled orbit and S_t is the semigroup generated by the entropy solution of the limit equation (10.2).*

Limit Families of Equations. The situation can be more involved when the limit equation is not unique. Such a situation arises in Homogenization Problems. We consider next the heat equation

$$\rho(x)u_t = \Delta u, \tag{10.3}$$

where $\rho \in L^\infty(\mathbb{R}^N)$ is such that $0 < c_1 \leq \rho(x) \leq c_2$. Let $S_t(u_0; \rho)$ denote the heat semigroup with weight ρ . When we re-scale the solutions by means of formula (2.3),

$$u_\lambda(x, t) = u(\lambda x, \lambda^2 t),$$

we get the equation satisfied by u_λ , $\rho_\lambda(x)u_{\lambda,t} = \Delta u_\lambda$, where $\rho_\lambda(x) = \rho(\lambda x)$. Let us now pass to a subsequence where

$$u_0(\lambda x) \rightarrow \phi, \quad \rho_\lambda \rightarrow a, \quad u_\lambda(x, t) \rightarrow v(x, t).$$

We easily see that such a limit v satisfies the homogenized equation

$$a(x)v_t = \Delta v(x), \tag{10.4}$$

where a ranges over all the elements in $G_\infty(\rho)$. In this case we have a situation where, instead of a limit equation, there exists in general a family of limit equations. Note however that in the important case of periodic coefficients the limit equation is unique since a is a constant, the average of ρ . This happens also for almost periodic coefficients.

We have then the following theorem.

Theorem 10.2. *The set of accumulation points of the rescaled orbit satisfies*

$$R_\infty(u_0) \subset \{S_1(\phi; a) : \phi \in G_\infty(u_0), a \in G_\infty(\rho)\}. \tag{10.5}$$

Equality of the sets in the theorem is not necessarily true since ρ and u_0 may be coupled. If we take for instance $u_0 = \rho$ then we only obtain R -omega-limits of the form $S_1(a; a)$. Depending on how independent are the oscillations of u_0 and ρ we may obtain larger R -omega-limit sets, even equality in the formula.

About an Asymptotic Theory with Quasi-Invariance. The general evolution setting introduced in Section 4 for invariant semigroups can

be easily modified to cover the case of quasi-invariant semi-groups and their limits according to the above two examples. Since there are no special novelties we leave the details to the reader and delay the detailed study of this aspect until other relevant examples have been examined. Actually, the ideas of quasi-invariance can be applied to study the asymptotic behavior of numerous processes appearing in the physical sciences. This is the case for equations combining nonlinear diffusion, reaction and convection.

Diffusion-convection equations of the form

$$u_t = \Delta u + \frac{\partial}{\partial x_1} f(u), \quad (10.6)$$

where f is a scalar function, are studied in [13, 14] in the setting of L^1 data. Convergence of the rescaled solutions is proved towards the solutions of the equation with partial diffusivity

$$u_t = \Delta' u + \frac{\partial}{\partial x_1} f(u), \quad (10.7)$$

where Δ' denotes the Laplacian with respect to the variables $x' = (x_2, \dots, x_n)$. In order to extend the results of this paper to this model for $L^\infty(\mathbb{R}^N)$ data we need a different scaling

$$u_\lambda(x, t) = u(\lambda x_1, \sqrt{\lambda} x', \lambda t).$$

Proving the asymptotic result should offer only difficulties related to compactness and identification of the limit solution (cf. Section 7). Besides, the scaling of the initial data in Section 3 has to be adapted to this anisotropic formula. In any case, Theorem 2.2 still holds in this setting, which guarantees the occurrence of the complexity we want to describe.

Further interesting lines of application concern the equations of fluid mechanics and the kinetic equations, but the technical difficulties are greater.

Acknowledgments. The authors are grateful to P. L. Lions, M. Rascle, D. Serre and L. Tartar for useful information about conservation laws. The work of the first author was supported by DGES Grant PB94-0153 and EU Programme TMR FMRX-CT98-0201, and the second author by DGES Grant PB96-0663 and TMR Programme Homogenization and Multiple Scales. The first author also acknowledges partial support from TICAM, Univ. of Texas at Austin.

REFERENCES

- [1] Aronson, D. & Ph. Bénilan, A nonlinear heat equation with singular diffusivity, *Commun. in Partial Differ. Equations*, **13**(1979), 985–1039.
- [2] Ph. Bénilan, Equations d'évolution dans un espace de Banach quelconque et applications, Thesis, Orsay, 1972.
- [3] Ph. Bénilan & Kruzhkov, S. N., Conservation laws with continuous flux functions, *Nonlinear Diff. Eqns. Applic. (NODEA)*, **3:4**(1996), 395–419.
- [4] Bohr, H., Almost periodic functions, Chelsea, 1947.
- [5] Brezis, H., Analyse fonctionnelle, Masson, Paris, 1983.
- [6] Chasseigne, E. & Vazquez, J. L., Theory of extended solutions for fast diffusion equations in optimal classes of data, radiation from singularities, *Archive Rat., Mech. Anal.* (to appear).
- [7] Chen, G. Q. & Frid, H., Initial layers and uniqueness of weak entropy solutions to hyperbolic conservation laws, *Arch. Rat. Mech. Anal.*, **153**(2000), 205–220.
- [8] Chen, G. Q. & Rasle, M., Initial layers and uniqueness of weak entropy solutions to hyperbolic laws, *Arch. Rat. Mech. Anal.*, **153**(2000), 205–220.
- [9] Crandall, M. G., The semigroup approach to first-order quasilinear equations in several space variables, *Israel J. Math.*, **12**(1972), 108–132.
- [10] Di Benedetto, E., Degenerate parabolic equations, Series Universitext, Springer-Verlag, New York, 1993.
- [11] Di Benedetto, E. & Friedman, A., Hölder estimates for non-linear degenerate parabolic systems, *J. Reine Angew. Math.*, **357**(1985), 1–22; Addendum, *J. Reine Angew. Math.*, **363**(1985), 217–220.
- [12] Duoandikoetxea, J. & Zuazua, E., Moments, masses de Dirac et décomposition de fonctions, *C. R. Acad. Sci. Paris.*, **315**(1992), 693–698.
- [13] Escobedo, M., Vazquez, J. L. & Zuazua, E., A diffusion-convection equation in several space variables, *Indiana Univ. Math. J.*, **42**(1993), 1413–1440.
- [14] Escobedo, M., Vazquez, J. L. & Zuazua, E., Entropy solutions for diffusion-convection equations with partial diffusivity, *Transactions Amer. Math. Soc.*, **343**(1994), 829–842.
- [15] Escobedo, M. & Zuazua, E., Long time behavior for convection-diffusion equations, *SIAM J. Math. Anal.*, **28**(1997), 570–594.
- [16] Esteban, J. R. & Vazquez, J. L., Régularité des solutions positives de l'équation parabolique p -laplacienne, *Comptes Rendus Acad. Sciences Paris I*, **310**(1990), 105–110.
- [17] Freistuhler, H. & Serre, D., L^1 -stability of shock waves in scalar viscous conservation laws, *Comm. Pure Appl. Math.*, **51**(1998), 291–301.
- [18] Godlewski, E. & Raviart, P. A., Hyperbolic systems of conservat laws, *Mathématiques & Appl. 3–4*, Ellipses, Paris, 1991.
- [19] Kruzhkov, S. N., Generalized solutions of the Cauchy problem in the large for nonlinear equations of first order, *Dokl. Akad. Nauk SSSR*, **187:1**(1969), 29–32 (English Transl. in *Soviet Math. Dokl.*, **10**, 1969).
- [20] Kruzhkov, S. N., First-order quasilinear equations in several independent variables, *Mat. Sbornik*, **81:2**(1970), 228–255 (English Transl. *Math. USSR Sb.*, **10**(1970), 217–243).
- [21] Hale, J. K., Asymptotic behavior of dissipative systems, *Math. Surveys and Monographs* **25**, Amer. Math. Soc., Providence, 1988.

- [22] Ladyzhenskaya, O. A., A dynamical system generated by the Navier-Stokes equation, *Zap. Nauchn. Sem. Leningrad Otdel. Mat. Inst. Steklov*, **27**(1972), 91–115.
- [23] Lax, P. D., Hyperbolic systems of conservation laws II, *Comm. Pure Appl. Math.*, **10**(1957), 537–566.
- [24] Liu, T. P., Hyperbolic and viscous conservation laws, CBMSS-NSF Regional Conference Series in Applied Mathematics, **72**, SIAM, Philadelphia, 1999.
- [25] Lions, P. L., Perthame, B. & Tadmor, E., A kinetic formulation of multidimensional scalar conservation laws and related equations, *J. Am. Math. Soc.*, **7**(1994), 169–191.
- [26] Liu, T. P. & Pierre, M., Source solutions and asymptotic behaviour in conservation laws, *J. Diff. Eqns.*, **51**(1984), 419–441.
- [27] Ortega, J. & Zuazua, E., Large time behavior in \mathbb{R}^N for linear parabolic equations with periodic coefficients, *Asymptotic Analysis*, **22**(2000), 51–85.
- [28] Serre, D., Systèmes de lois de conservation, II, Diderot Ed., Paris, 1996.
- [29] Tartar, L., Une nouvelle méthode de résolution d'équations aux dérivées partielles non linéaires, Lecture Notes in Math. [M], **665**, Springer, Berlin, 1977, 228–241.
- [30] Tartar, L., Compensated compactness and applications to partial differential equations [A], in “Nonlinear Analysis and Mechanics: Heriot-Watt Symposium”, vol. IV, Pitman, San Francisco, 1979. 136–212.
- [31] Témam, R., Infinite dimensional dynamical systems in mechanics and physics, Springer Verlag, 1988.
- [32] Vasseur, A., Strong traces for solutions to the multi-dimensional conservation laws, Preprint, Univ. Nice, 2000.
- [33] Vazquez, J. L., Nonexistence of solutions for nonlinear heat equations of fast-diffusion type, *Journal Math. Pures et Appliquées*, **71**(1992), 503–526.
- [34] Vazquez, J. L., Asymptotic behaviour for the Porous Medium Equation in the whole space, Ph. D. Course, UAM 1996/97 (see UAM address <http://www.adi.uam.es/~jvazquez>).