HABILITATION A DIRIGER LES RECHERCHES EN SCIENCES

présentée à

L'UNIVERSITE DE MONTPELLIER II

Spécialité

MATHEMATIQUES

par

Jérôme DRONIOU

Sujet du mémoire:

Etude théorique et numérique d'équations aux dérivées partielles elliptiques, paraboliques et non-locales

présentée et soutenue publiquement le 26 novembre 2004 devant le jury composé de :

M. **Hédy ATTOUCH**	Président du jury
M. **Pierre FABRIE**	Rapporteur
M. **Thierry GALLOUET**	Examinateur
M. **Benoit PERTHAME**	Rapporteur
M. **Lionel THIBAULT**	Examinateur
M. **Juan-Luis VAZQUEZ**	Rapporteur

Table des Matières

I Présentation de l'activité scientifique **6**

1 Synthèse des travaux **7**
1.1 Activités scientifiques avant et pendant la thèse de doctorat 8
1.1.1 Equations elliptiques non-coercitives 8
1.1.2 Equations paraboliques à données mesures 8
1.1.3 Etude des espaces de Sobolev 9
1.2 Equations elliptiques à données mesures 10
1.2.1 Problématique 10
1.2.2 Apports 11
1.3 Schémas volumes finis pour équations elliptiques 13
1.3.1 Problématique 13
1.3.2 Apports 14
1.4 Approximation parabolique en domaine borné 18
1.4.1 Problématique 18
1.4.2 Apports 19
1.5 Régularisation non-locale de lois de conservation scalaires 20
1.5.1 Problématique 20
1.5.2 Apports 21

2 Liste des Travaux **26**
2.1 Publications dans des revues internationales avec comité de lecture 26
2.2 Publications dans des actes de congrès 27
2.3 Autres publications 27

II Régularité locale de solutions d'équations elliptiques avec mesures **28**

3 Global and local estimates for nonlinear noncoercive elliptic equations with measure data **29**
3.1 Introduction and main results 29
3.1.1 The problem 29
3.1.2 Hypotheses and notations 30
3.1.3 Main results 31
3.2 Global Estimates 32
3.2.1 Estimate on $\ln(1+|u|)$ 32
3.2.2 Proof of the global estimates 33
3.3 Local Estimates 35
3.3.1 Preliminary results 35
3.3.2 Proof of the local estimates 39
3.4 Existence and regularity result for an equation with measure data 44

3.5 Appendix . . . 45

III Schémas volumes finis pour équations elliptiques 47

4 Finite volume methods for convection-diffusion equations with right-hand side in H^{-1} 48
4.1 Introduction . . . 48
4.2 Definition of the scheme and main result . . . 49
4.2.1 Technical results . . . 51
4.3 A Priori Estimates . . . 52
4.3.1 Estimate on $\ln(1+|u_{\mathcal{T}}|)$ . . . 52
4.3.2 Estimate on $||u_{\mathcal{T}}||_{1,\mathcal{M}}$ . . . 57
4.4 Proof of the existence, uniqueness and convergence result . . . 60
4.5 Another scheme . . . 66

5 Error estimates for the convergence of a finite volume discretization of convection-diffusion equations 67
5.1 Introduction . . . 67
5.1.1 The problem . . . 67
5.1.2 Definition of the scheme . . . 68
5.2 Statement of the main result . . . 70
5.3 The H^1 framework . . . 71
5.4 The H^2 framework . . . 73
5.5 Proof of the main result . . . 80
5.6 Numerical results . . . 81
5.7 Appendix . . . 82
5.7.1 Technical lemmas . . . 82
5.7.2 Interpolation . . . 87

6 A finite volume scheme for a noncoercive elliptic equation with measure data 89
6.1 Introduction . . . 89
6.2 Conservative finite volume discretization and convergence result . . . 90
6.3 A Priori Estimates . . . 95
6.3.1 Estimate on $\ln(1+|u_{\mathcal{T}}|)$ . . . 96
6.3.2 Estimate on $||u_{\mathcal{T}}||_{1,q,\mathcal{M}}$ . . . 99
6.3.3 Proof of Theorem 6.2.2 . . . 106
6.4 Proof of Theorem 6.2.1 . . . 106
6.5 A scheme with jump of the fluxes . . . 111
6.6 Numerical results . . . 112
6.6.1 Comparison of the two finite volume schemes . . . 112
6.6.2 Two and three-dimensional tests on a Cartesian mesh . . . 113
6.6.3 Two-dimensional tests on an unstructured mesh . . . 115
6.6.4 Spherical domain and mesh . . . 115
6.7 Appendix . . . 117

IV Approximation parabolique d'une équation hyperbolique en domaine borné 120

7 An error estimate for the parabolic approximation of multidimensional scalar conservation laws with boundary conditions 121
7.1 Introduction . . . 121
7.2 Preliminaries . . . 123

7.2.1 Known estimates on u and v . . . 123
7.2.2 Notations . . . 123
7.2.3 Kinetic formulations of (7.1.1) and (7.1.2) . . . 124
7.2.4 Main ideas of the proof . . . 125
7.3 Estimate in the interior of the domain . . . 125
7.4 Transport and regularization of the kinetic equations . . . 127
7.4.1 Transport of the kinetic equations . . . 127
7.4.2 Transport of the BLN condition . . . 128
7.4.3 Regularization of the transported equations . . . 129
7.5 Combination of the equations and new estimates . . . 130
7.5.1 Passing to the limit in β and ν . . . 130
7.5.2 Choice of ϕ and continuation of the estimates . . . 132
7.6 Estimate for the boundary term . . . 136
7.6.1 Introduction of f_+^b . . . 136
7.6.2 Apparition of Ha . . . 136
7.6.3 Regularization of f_+^b . . . 136
7.6.4 Estimate of $S^{\alpha,\varepsilon}$ and conclusion concerning the boundary term . . . 138
7.7 Conclusion . . . 139
7.8 Appendix . . . 140
7.8.1 Estimate of $T_5^{\alpha,\varepsilon}$ using boundary layers . . . 140
7.8.2 Technical results . . . 141

V Diffusion non-locale **144**

8 Global solution and smoothing effect for a non-local regularization of an hyperbolic equation **145**
8.1 Introduction . . . 145
8.2 Properties of the kernel of g . . . 146
8.3 Definition and first properties of the solutions . . . 148
8.4 Uniqueness of the solution . . . 150
8.5 Regularizing effect . . . 150
8.5.1 Spatial regularity . . . 150
8.5.2 Temporal regularity . . . 153
8.6 L^∞ estimate and global existence . . . 156
8.6.1 Construction of an approximate solution by a splitting method . . . 156
8.6.2 Compactness result on the sequence $(u^\delta)_{\delta>0}$ . . . 157
8.6.3 Passing to the limit $\delta \to 0$ . . . 158
8.6.4 Conclusion . . . 160

9 Généralisations du chapitre 8 **162**
9.1 En dimension supérieure à 1 . . . 162
9.1.1 Positivité du noyau . . . 162
9.1.2 Intégrabilité de transformées de Fourier . . . 164
9.1.3 Modifications à apporter aux preuves . . . 166
9.2 Equations hyperboliques plus générales . . . 167
9.2.1 Introduction, hypothèses . . . 167
9.2.2 Premières constatations . . . 168
9.2.3 Partie hyperbolique . . . 169
9.2.4 Partie non-locale . . . 176
9.2.5 Conclusion . . . 176

10 Vanishing non-local regularization of a scalar conservation law **182**
10.1 Introduction 182
10.2 Approximate entropy inequalities for (10.1.1) 184
10.2.1 Construction of u^ε 184
10.2.2 The approximate entropy inequalities 184
10.3 Proof of the convergence 189
10.4 Proof of the error estimate 193
10.5 Appendix 196
10.5.1 An expression and an estimate of $g[\varphi]$ 196
10.5.2 Technical lemmas on the kernel of g 197

Partie I

Présentation de l'activité scientifique

Chapitre 1

Synthèse des travaux

Je présente ici mes travaux de recherche, qui peuvent essentiellement se classer en 5 thèmes:

1. Etude d'**équations elliptiques et paraboliques**, tout particulièrement les équations de convection-diffusion générales, ainsi que le cas des données mesures.
2. **Analyse fonctionnelle dans les espaces de Sobolev**, qui donne des outils utiles au thème 1.
3. **Etude numérique des équations de convection-diffusion avec seconds membres peu réguliers** (dans des espaces de type H^{-s} ou mesures), extension naturelle du thème 1.
4. Obtention de vitesse de convergence pour la **régularisation parabolique d'une équation hyperbolique** en domaine borné.
5. Etude de **régularisations non-locales de lois de conservation scalaires**, en particulier le cas où l'on remplace le laplacien d'une équation parabolique par une puissance fractionnaire de ce laplacien; ce sujet a quelques liens avec le thème 4 (il en est une généralisation, du moins — pour l'instant — lorsque l'on considère le sujet 4 posé dans l'espace entier et non dans un ouvert borné).

Dans une première section, je vais brièvement rappeler les travaux effectués avant et pendant ma thèse de doctorat, sans entrer dans les détails des sujets ou des preuves.
Dans les sections suivantes, je présenterai les travaux réalisés depuis la fin de ma thèse (principalement dans les thèmes 1, 3, 4 et 5, bien que le thème 2 soit omniprésent dans le thème 1 et, sous une forme discrète, dans le thème 3), en m'attachant à chaque fois à situer dans un premier temps, le plus simplement possible, la problématique et les résultats existants, avant d'indiquer de manière plus précise les résultats que j'ai obtenus dans le sujet. Toutefois, pour éviter des énoncés trop complexes, qui cacheraient le coeur des résultats ou demanderaient d'introduire beaucoup de notations (en particulier dans les parties numériques), certains des théorèmes ci-après sont écrits dans une forme affaiblie par rapport aux articles dont ils sont tirés, ou de manière un peu imprécise. Je renvoie le lecteur intéressé aux articles correspondants, reproduits dans les chapitres suivants.

Dans ce chapitre, les références numériques (comme [18]) renvoient à la bibliographie en fin de mémoire, et les références comprenant une lettre D (comme [D10]) renvoient à la liste de publications de l'auteur située au chapitre 2.

1.1 Activités scientifiques avant et pendant la thèse de doctorat

1.1.1 Equations elliptiques non-coercitives

Le principal apport de ma thèse est probablement le résultat d'existence et d'unicité concernant les équations elliptiques non-coercitives. On sait depuis longtemps résoudre le problème

$$\left\{ \begin{array}{ll} -\Delta u + \mathrm{div}(\mathbf{v}u) = f & \text{dans } \Omega\,, \\ u = 0 & \text{sur } \partial\Omega \end{array} \right. \tag{1.1.1}$$

où Ω est un ouvert borné de $\mathbb{R}^N$, $f \in H^{-1}(\Omega) = (H_0^1(\Omega))'$ et $\mathbf{v}$ est un champ de vecteurs à divergence positive; le théorème de Lax-Milgram donne l'existence et l'unicité d'une solution dans $H_0^1(\Omega)$. Cependant, en général, ce théorème ne s'applique plus si on ne suppose rien sur la divergence de $\mathbf{v}$: la forme bilinéaire associée à (1.1.1) peut ne plus être coercitive.

[D4] prouve que, sans aucune autre hypothèse sur $\mathbf{v}$ que des propriétés d'intégrabilité minimales, on peut obtenir des estimations *a priori* sur les solutions de (1.1.1) et donc montrer l'existence (puis l'unicité par une technique de dualité) d'une solution à ce problème. Outre le fait qu'il exprime que, contrairement à ce que l'on semblait penser, les termes d'ordre 1 ne perturbent pas sensiblement la théorie des équations elliptiques, ce résultat a aussi une application directe: grâce à la résolution de (1.1.1) sans hypothèse sur $\mathrm{div}(\mathbf{v})$, [D3] (article co-écrit avec T. Gallouët) étend au cas de certains opérateurs elliptiques non-linéaires non-monotones le résultat de [10] sur l'unicité pour des problèmes elliptiques avec seconds membres mesures (voir la section 1.2 pour plus de détails sur ces problématiques).

La méthode utilisée pour obtenir des estimations sur les solutions de (1.1.1) s'est révélée très générale et a été adaptée au cas d'opérateurs non-linéaires, ainsi qu'aux discrétisations numériques d'équations de convection-diffusion (voir les sections 1.2 et 1.3 ci-dessous).

1.1.2 Equations paraboliques à données mesures

La capacité est un outil essentiel dans l'étude des espaces fonctionnels associés aux équations elliptiques ou paraboliques; elle permet de voir que les fonctions dans des espaces de Sobolev sont définies mieux que "presque partout". De manière précise: dans le cas elliptique, la $W^{1,p}$-capacité d'un compact $K \subset \Omega$ est définie par

$$\mathrm{cap}(K) = \inf \left\{ \int_\Omega |\nabla v|^p \,;\; v \in C_c^\infty(\Omega)\,,\; v \geq \mathbf{1}_K \right\}$$

($\mathbf{1}_K$ est la fonction caractéristique de K); on définit ensuite la capacité d'un ouvert $U \subset \Omega$ par $\mathrm{cap}(U) = \sup_{K \subset U} \mathrm{cap}(K)$, puis celle d'un ensemble quelconque $A \subset \Omega$ par $\mathrm{cap}(A) = \inf_{U \supset A} \mathrm{cap}(U)$. Une des propriétés importantes et bien connues de la capacité elliptique est la suivante: si $v \in W_0^{1,p}(\Omega)$, alors il existe un représentant cap-quasi-continu de v, c'est à dire une fonction $\widetilde{v}$ qui coincide presque partout (au sens de la mesure de Lebesgue) avec v et telle que pour tout $\varepsilon > 0$, il existe un ouvert $U \subset \Omega$ tel que $\mathrm{cap}(U) < \varepsilon$ et $\widetilde{v}_{|\Omega \setminus U}$ est continue. On peut aussi prouver que si $(u, v) \in W_0^{1,p}(\Omega)$ et $u = v$ presque partout, alors leurs représentants cap-quasi-continus coincident en dehors d'un ensemble de capacité nulle: c'est en ce sens que l'on peut dire "les fonctions de $W_0^{1,p}$ sont définies mieux que presque partout" (des ensembles peuvent être de mesure de Lebesgue nulle, mais de capacité non-nulle).
La capacité permet aussi de caractériser certaines mesures pour lesquelles on sait prouver l'existence et l'unicité de la solution à des problèmes elliptiques ou paraboliques ayant ces mesures comme seconds membres. La capacité associée à l'équation de la chaleur a été étudiée dans [75]; dans [D6] (co-écrit avec A. Prignet et A. Porretta), nous avons introduit et étudié la capacité associée à des opérateurs paraboliques non-linéaires plus généraux, dont le p-laplacien instationnaire donne un exemple:

$$\left\{ \begin{array}{ll} \partial_t u - \mathrm{div}(|\nabla u|^{p-2}\nabla u) = \mu & \text{dans }]0,T[\times\Omega\,, \\ u(0,\cdot) = u_0 & \text{sur } \Omega\,, \\ u = 0 & \text{sur }]0,T[\times\partial\Omega \end{array} \right. \tag{1.1.2}$$

(la page 11 montre une forme plus générale pour la partie spatiale des opérateurs paraboliques que l'on peut considérer ici).
Cette capacité est un peu plus délicate à manipuler que celle correspondant à l'équation de la chaleur ($p = 2$), à cause de problèmes de dualité entre les espaces considérés et de l'analyse fonctionnelle nécessaire concernant ces espaces. Néanmoins, nous prouvons dans ce cadre les résultats usuels de la théorie des capacités, et surtout nous étendons au cas parabolique le résultat de décomposition de mesure de [13]; il est alors possible de donner une structure pour les mesures ne chargeant pas les ensembles de capacité parabolique nulle, structure qui permet de résoudre (1.1.2) à l'aide des solutions renormalisées (l'existence et l'unicité d'une solution renormalisée sont prouvées).
Dans la continuité de ce sujet, A. Prignet et moi-même collaborons en ce moment à un travail dans lequel nous étendons la notion de solution "entropique" pour (1.1.2), notion connue lorsque $\mu \in L^1$, au cas de mesures ne chargeant pas les ensembles de capacité nulle, et nous prouvons que cette notion coïncide avec celle de solution renormalisée.

Dans le cadre des équations paraboliques à données mesures, je peux aussi mentionner [D1] (réalisé en collaboration avec J.-P. Raymond); ce travail, effectué avant ma thèse, s'intéresse au contrôle optimal d'une équation parabolique semi-linéaire de la forme

$$\begin{cases} \partial_t u(t,x) - \Delta u(t,x) + |u|^{\gamma-1}u(t,x) = f(t)\delta_{x_0} & t \in]0,T[\,,\ x \in \Omega\,, \\ \frac{\partial u}{\partial \mathbf{n}} = 0 & \text{sur }]0,T[\times\partial\Omega\,, \\ u(0,\cdot) = u_0 & \text{sur } \Omega \end{cases} \tag{1.1.3}$$

où δ_{x_0} est la masse de Dirac en $x_0 \in \Omega \subset \mathbb{R}^N$ et $\gamma \in [1, N/(N-2)[$. Le contrôle optimal de cette équation consiste à se donner une fonction u_d sur Ω et à ajuster f pour que, en un temps T final, la solution de (1.1.3) soit le plus proche possible de u_d; généralement, on demande aussi que ce contrôle f ne soit pas trop grand (dans une norme choisie). Plus précisément, on étudie le problème

$$\inf\left\{ \int_\Omega |u_f(T) - u_d|^s + \int_0^T |f|^q\,;\ f \in K\,,\ u_f \text{ vérifie (1.1.3)} \right\} \tag{1.1.4}$$

où K est un sous-ensemble de $L^q(0,T)$ et u_d est une fonction fixée.
En utilisant une technique d'équations duales (différente de celle qu'utilise P. Baras dans sa thèse), nous avons obtenu des estimations nouvelles sur les solutions de (1.1.3), et nous avons prouvé qu'il existe des minima à (1.1.4), minima pour lesquels nous donnons une condition nécessaire au travers d'une équation duale à (1.1.3).

1.1.3 Etude des espaces de Sobolev

Les espaces de Sobolev sont un outil omniprésent dans l'étude des EDP elliptiques et paraboliques. Leur compréhension est donc une étape nécessaire avant d'aborder les équations en question.

Si les espaces de Sobolev sur un ouvert Ω ayant la "propriété du segment" (une manière de dire que Ω est localement situé d'un seul côté de son bord) sont généralement bien connus (voir par exemple [1]), le cas d'ouverts dont le bord est une variété lipschitzienne générale ne semble pas très étudié ([72] traite surtout le cas d'ouverts que l'on peut localement, avec un bon choix de coordonnées, écrire comme des épigraphes de fonctions lipschitziennes, ce qui est plus fort). Dans le polycopié [Dp2], dont la rédaction a débuté avant ma thèse, j'analyse les propriétés des espaces de Sobolev sur des ouverts qui sont localement transportables, par des homéomorphismes bilipschitziens, sur des demi-espaces; ces ouverts ne vérifient pas forcément la propriété du segment, mais les résultats classiques des espaces de Sobolev (prolongement, densité, trace, intégration par parties...) sont quand même établis; les preuves font parfois appel à des techniques de géométrie différentielle, mais qui doivent être adaptées à des variétés lipschitziennes et non C^1.

Ce polycopié a directement servi dans [D7] (co-écrit avec R. Eymard, D. Hilhorst et X.D. Zhou); en effet, des propriétés fines sur les transports des espaces de Sobolev par homéomorphismes bilipschitziens se sont révélées nécessaires pour étudier le maillage très général utilisé dans cet article.
Un autre résultat concernant l'analyse des espaces de Sobolev est donné dans [D5]: il s'agit de la densité dans $W^{1,p}(\Omega)$, lorsque Ω est régulier ou polygonal ([1]), de

$$\{\varphi \in C^\infty(\overline{\Omega}) \mid \nabla\varphi \cdot \mathbf{n} = 0 \text{ sur } \partial\Omega\}$$

(i.e. des fonctions régulières satisfaisant une condition de Neumann). La technique de preuve est relativement classique dans le cas où l'ouvert est régulier, et nettement moins dans le cas où il est polygonal: dans cette dernière situation, l'idée est d'approcher (grâce à une récurrence sur la dimension des singularités du bord de Ω) une fonction régulière fixée par des fonctions régulières qui ne dépendent, au voisinage d'une partie linéaire de $\partial\Omega$ (face, arête, sommet...), que de la coordonnée le long de cette partie.

Dans le cas des équations paraboliques, ce sont surtout les espaces de Lebesgue ou de Sobolev à valeurs dans des Banach qui sont utiles. J'ai essayé, dans [Dp1], de présenter tous les outils (parfois pas forcément bien connus) nécessaires à l'étude des équations paraboliques à l'aide de ces espaces. En particulier, ce polycopié a été grandement enrichi lors de la rédaction de [D6] (par exemple, le résultat de densité en annexe de ce dernier article découle de théorèmes énoncés de manière *ad hoc* dans [Dp1]).

Je vais maintenant aborder les sujets que j'ai étudiés depuis la fin de ma thèse; certains sont bien sûrs en continuité des travaux que je viens de mentionner.

1.2 Equations elliptiques à données mesures

1.2.1 Problématique

L'étude des équations elliptiques avec seconds membres mesures est un thème très actif depuis une quinzaine d'années. Le premier résultat à ce sujet est probablement celui de G. Stampacchia [77], qui concerne les équations linéaires de la forme:

$$\begin{cases} -\mathrm{div}(A\nabla u) = \nu & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega \end{cases} \tag{1.2.1}$$

où Ω est un ouvert borné de $\mathbb{R}^N$ et $A : \Omega \to M_N(\mathbb{R})$ est une application matricielle bornée qui vérifie le critère d'ellipticité usuel $(A(x)\xi \cdot \xi \geq \alpha|\xi|^2$ pour presque tout $x \in \Omega$ et tout $\xi \in \mathbb{R}^N$, avec $X \cdot Y$ désignant le produit scalaire de deux vecteurs de $\mathbb{R}^N$ et $|\cdot|$ la norme associée). Stampacchia établit, par une méthode de dualité qui utilise fortement le caractère linéaire de l'équation (1.2.1), une formulation qui permet d'avoir l'existence et l'unicité d'une solution lorsque ν est une mesure bornée sur Ω.
Le cas linéaire est donc entièrement résolu depuis 1965 (le résultat complet dans [77] concerne des équations linéaires plus générales).

Le cas des équations semi-linéaires, de la forme

$$-\Delta u + g(u) = f \quad \text{avec } f \in L^1(\Omega)$$

a été abordé peu après, dans [19] et [5] par exemple. Si l'on considère des équations elliptiques non-linéaires plus générales, comme

$$\begin{cases} -\mathrm{div}(a(x,\nabla u))) = \nu & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega \end{cases} \tag{1.2.2}$$

[1] Ce dernier cas ayant un intérêt dans l'étude des discrétisations numériques d'équations elliptiques.

dont un exemple modèle est le p-laplacien: $a(x, \nabla u) = |\nabla u|^{p-2}\nabla u$ (avec $1 < p < \infty$), le premier résultat d'existence général d'une solution lorsque ν est une mesure est dû à T. Gallouët et L. Boccardo [11], et date de 1989; la formulation pour laquelle une solution de (1.2.2) était obtenue est classique: on multiplie l'équation par une fonction-test assez régulière et on intègre par parties; la solution est alors dans les espaces de Sobolev $W_0^{1,q}(\Omega)$ pour tout $q < \frac{N(p-1)}{N-1}$ (ceci pour $p > 2 - \frac{1}{N}$, afin que les espaces en question soient bien définis).
Cependant, on sait que, même dans le cas linéaire, cette formulation ne permet pas d'avoir l'unicité de la solution (voir [76]), et ce problème d'unicité reste encore majoritairement ouvert aujourd'hui.

Des formulations plus fortes ont alors vu le jour, avec l'espoir d'arriver à une définition de solution qui permettrait d'obtenir l'unicité. On peut par exemple citer la formulation entropique, introduite dans [13], qui donne l'unicité mais n'est adaptée qu'aux seconds membres dans $L^1(\Omega) + W^{-1,p'}(\Omega)$ (c'est à dire, si on considère des mesures, les seconds membres qui ne chargent pas les ensembles de $W_0^{1,p}$-capacité nulle); la "formulation" SOLA venue de [29], qui consiste à considérer uniquement des solutions de (1.2.2) approchables par les solutions de problèmes avec seconds membres plus réguliers ([2]), permet de prouver l'unicité pour des mesures générales, mais avec de fortes restrictions sur l'opérateur considéré (en particulier, p doit être égal à 2): voir [10] ou [D3].
Il est fort probable que 1999 ait vu un grand pas s'accomplir dans la recherche d'une formulation qui donne l'existence et l'unicité de la solution à (1.2.2), avec l'introduction dans [28] d'une notion de solution renormalisée qui accepte toute mesure au second membre et tout opérateur elliptique non-linéaire; l'existence d'une solution renormalisée est prouvé dans cet article, et quelques résultats d'unicité partiels (lorsque deux solutions sont "comparables") sont aussi donnés. Mais un résultat d'unicité vraiment satisfaisant se fait toujours attendre, même s'il parait clair à beaucoup de spécialistes du domaine que la formulation de [28] a éliminé toutes les causes de non-unicité.

1.2.2 Apports

Malgrè les avancées mentionnées ci-dessus dans la recherche d'une bonne formulation, il semblerait que nous manquions encore de connaissances précises sur la solution du problème avec donnée mesure; il faut probablement chercher à établir des propriétés supplémentaires que cette solution vérifie et qui permettraient de la cerner davantage pour en prouver l'unicité.

Le travail [D8], reproduit dans le chapitre 3 de ce mémoire, donne une telle propriété; cette dernière est aisée à énoncer et tout à fait naturelle, mais sa démonstration demande une bonne dose de technique et des astuces d'estimation.
Considérons un problème non-linéaire simple, de la forme de (1.2.2) où l'on suppose que a vérifie

$$\begin{aligned} &|a(x,\xi)| \le C(b(x) + |\xi|^{p-1}) \quad \text{avec } b \in L^{p'}(\Omega)\,, \\ &a(x,\xi)\cdot\xi \ge \alpha|\xi|^p \quad \text{avec } \alpha > 0\,, \\ &(a(x,\xi) - a(x,\eta))\cdot(\xi-\eta) > 0 \quad \text{dès que } \xi \ne \eta \end{aligned}$$

(ce sont les hypothèses usuelles sur les opérateurs de Leray-Lions de la forme (1.2.2)). Lorsque le second membre est dans le dual $W^{-1,p'}(\Omega)$ de $W_0^{1,p}(\Omega)$, on sait qu'il existe une (unique) solution à (1.2.2) dans $W_0^{1,p}(\Omega)$; cet espace apparait comme un espace d'énergie naturel associé à l'opérateur $-\mathrm{div}(a(x,\nabla u))$. Lorsque le second membre est une mesure, on sait que la solution ne vit pas en général dans cet espace, mais dans $W_0^{1,q}(\Omega)$ avec $q < \frac{N(p-1)}{N-1}$ lorsque $p > 2 - \frac{1}{N}$ (la solution a un "gradient" dans des espaces de Marcinkiewitz lorsque $p \le 2 - \frac{1}{N}$); c'est bien sûr le peu de régularité du second membre qui provoque cette perte de régularité sur la solution. Cependant, on obtient le résultat suivant:

[2] C'est à dire une formulation qui s'attache à préserver en premier lieu la stabilité de la solution par rapport à de petites perturbations du second membre ν.

Théorème 1.2.1 *Supposons que $\nu = \mu + f$ où μ est une mesure bornée dont le support est contenu dans un compact K de $\mathbb{R}^N$ et $f \in W^{-1,p'}(\Omega)$. Alors il existe une solution u à (1.2.2) qui soit dans $W^{1,p}(U)$ pour tout ouvert U de Ω qui ne rencontre pas K.*

Autrement dit, dès que l'on s'éloigne de la zone où la mesure n'est effectivement pas dans le dual de l'espace d'énergie associé à l'opérateur, on retrouve toute la régularité permise par l'opérateur. La motivation première de ce résultat n'était pas d'aborder directement le problème de l'unicité pour (1.2.2) (et si le théorème 1.2.1 peut aider à prouver cette unicité, il n'est pas évident qu'il en soit la clef cependant), mais de résoudre une difficulté qui se posait lors de la rédaction de [D12] (difficulté qui a depuis été résolue d'une autre manière); dans le cadre linéaire de [D12], le théorème 1.2.1 est de toutes façons très facile à montrer: si θ est une fonction qui s'annule sur le support de μ, alors on peut écrire grâce à la linéarité une équation sur θu qui montre rapidement, par une technique de bootstrap, que cette fonction est dans l'espace d'énergie H^1 associé à l'opérateur.
Dans le cas non-linéaire, il faut aussi employer des fonctions de troncature et une technique de bootstrap, mais il est hors de question d'écrire une équation sur θu qui montre directement que cette fonction a la régularité recherchée. L'idée de la preuve est la suivante:

i) En multipliant l'équation sur u par $\theta^{p-1} T_k(\theta u)$, où θ est une fonction bien choisie qui s'annule au voisinage de K et $T_k(s) = \min(k, \max(s, -k))$ est la troncature usuelle (voir la figure 6.2 page 96), on obtient une estimation
$$\int_\Omega |\nabla(T_k(\theta u))|^p \leq C(u,\alpha) k^\alpha, \tag{1.2.3}$$
où $C(u,\alpha)$ ne dépend que de la norme de u dans $W^{1,p-\alpha}(\mathrm{supp}(\theta))$ (on peut prendre $\alpha \in]0,1]$ quelconque). On sait de plus que, pour α assez proche de 1, u est dans $W^{1,p-\alpha}(\Omega)$ (ce sont des estimations globales connues pour les équations elliptiques à seconds membres mesures).

ii) Une estimation de la forme (1.2.3) donne, par des méthodes classiques, une estimation sur $\nabla(\theta u)$ dans l'espace de Marcinkiewitz d'exposant $\frac{N(p-\alpha)}{N-\alpha}$, i.e.
$$\mathrm{meas}(\{x \in \Omega \mid |\nabla(\theta u)(x)| \geq k\}) \leq M(u,\alpha) k^{-\frac{N(p-\alpha)}{N-\alpha}}$$
où $M(u,\alpha)$ ne dépend que de $C(u,\alpha)$. Ainsi, pour α assez proche de 1, l'étape i) montre que θu est estimée dans $W^{1,q}(\Omega)$ ([3]) dès que $q < \frac{N(p-\alpha)}{N-\alpha}$ (l'espace de Marcinkiewitz d'exposant r s'injecte dans tous les espaces de Lebesgue d'exposants $< r$). Cela améliore, quitte à rester à l'intérieur du support de θ, les estimations que l'on connaissait sur u.

iii) En utilisant une technique de bootstrap basée sur la répétition des deux étapes précédentes (l'estimation meilleure de la deuxième étape permet de baisser le α de la première étape, donc d'améliorer encore les estimations de la deuxième étape, etc...), on obtient une borne sur u dans tous les $W^{1,q}$ pour $q < p$, pourvu que l'on reste loin de K.

iv) On utilise alors une fonction test de la forme $\theta^p u$ avec θ bien choisie qui s'annule au voisinage de K, et les estimations de iii) avec q assez proche de p permettent directement de borner θu dans $W^{1,p}(\Omega)$.

Ceci prouve des estimations *a priori* sur la solution de (1.2.2) dans $W^{1,p}$ loin du support de μ; une technique d'approximation aisée donne alors l'existence d'une solution dans ces espaces.

Le théorème 1.2.1 n'est pas limité aux équations de la forme (1.2.2); il est aussi valable lorsque l'opérateur a une dépendance (sous une forme usuelle) par rapport à u. Mais surtout, il s'applique également dans le cadre des équations non-coercitives, de la forme
$$\begin{cases} -\mathrm{div}(a(x,\nabla u)) + \mathrm{div}(\Phi(x,u)) = \mu & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega \end{cases} \tag{1.2.4}$$

[3] On suppose pour simplifier que $N > 2$ et que $p > 2 - \frac{1}{N}$.

avec $|\Phi(x,u)| \leq g(x)(1+|u|^{p-1})$ et $g \in L^{N/(p-1)}(\Omega)$.
Comme signalé précédemment, la difficulté de ces équations est que l'on ne peut obtenir, par les méthodes usuelles, des estimations *a priori* sur les solutions (même lorsque le second membre est régulier). Une deuxième caractéristique notable de [D8] est de montrer que la technique d'estimation développée dans [D4] pour le problème linéaire non-coercitif avec données régulières s'adapte (et semble donc véritablement robuste et générale, puisqu'elle a aussi été appliquée avec succès aux schémas numériques) au cas des équations non-linéaires et des seconds membres mesures. On obtient alors la forme générale suivante du théorème 1.2.1:

Théorème 1.2.2 *Soit $p > 2 - \frac{1}{N}$. Supposons que $\nu = \mu + f$ où μ est une mesure bornée dont le support est contenu dans un compact K de $\mathbb{R}^N$ et $f \in W^{-1,p'}(\Omega)$. Alors il existe une solution u à (1.2.4) qui est dans $W_0^{1,q}(\Omega)$ pour tout $q < \frac{N(p-1)}{N-1}$ et dans $W^{1,p}(U)$ pour tout ouvert U qui ne rencontre pas K.*

Dans l'énoncé des théorèmes 1.2.1 et 1.2.2, je n'ai pas précisé en quel sens les solutions doivent être prises; c'était en fait volontaire: le sens importe peu. En effet, les solutions à ces problèmes non-linéaires avec données mesures sont toujours construites par des méthodes d'approximation (que ce soit les solutions au sens faible de [11], les solutions entropiques de [13], les SOLA, ou les solutions renormalisées de [28]); or la preuve des régularités locales dans les théorèmes 1.2.1 et 1.2.2 consiste justement à prouver que des solutions correspondant à des problèmes approchés sont bornées dans les bons espaces. On en déduit qu'essentiellement toute solution construite par une méthode d'approximation est dans $W^{1,p}$ lorsque l'on est loin de la partie "délicate" de la mesure.

1.3 Schémas volumes finis pour équations elliptiques

1.3.1 Problématique

L'approximation numérique d'équations aux dérivées partielles est un problème extrèmement important en pratique. En effet, les équations aux dérivées partielles que le mathématicien étudie, même si elles correspondent parfois à des modèles très simplifiés de la physique ou de la mécanique, apparaissent naturellement dans des problèmes concrets; dans ces cas-là, pouvoir calculer de manière approchée la solution (si une résolution analytique est inconnue) est crucial: bien souvent, le physicien ou le mécanicien n'est pas tant intéressé par le fait de savoir qu'une solution existe que par la calculer explicitement.
Dans la famille des méthodes numériques, les Volumes Finis ont émergé depuis peu mais semblent se tailler une part de plus en plus importante des applications concrètes, car ils ont des propriétés très intéressantes vis-a-vis des problèmes physiques sous-jacents: en particulier, la conservation des flux. Une équation venant d'une loi de conservation a donc de bonnes chances de se trouver correctement discrétisée par une méthode de Volumes Finis.

Considérons le cas simple d'une équation de la forme

$$\operatorname{div}(q) = f \text{ dans } \Omega \tag{1.3.1}$$

où Ω est un ouvert borné de $\mathbb{R}^N$ (dans la pratique, $N = 2$ ou 3 en stationnaire, et $N = 2$, 3 ou 4 en instationnaire — une des coordonnées est alors le temps), $q : \Omega \to \mathbb{R}^N$ et $f : \Omega \to \mathbb{R}$; cette équation correspond à un bilan effectué sur une quantité dont $q \cdot \mathbf{n}\, dS$ est le flux au travers d'une surface infinitésimale dS de normale $\mathbf{n}$; f est la source volumique de la quantité en question. Le principe de la discrétisation par Volumes Finis consiste à faire le travail inverse du physicien qui a obtenu (1.3.1) par un bilan: on prend cette équation et on l'intègre sur un petit volume non-infinitésimal.
Plus précisément, si on a pavé l'ouvert Ω par des polygônes K_i (par exemple), on intègre (1.3.1) sur chaque polygone, et on trouve, par Stokes:

$$\sum_{\sigma \in \mathcal{E}_i} \int_\sigma q \cdot \mathbf{n}_i = \int_{\partial K_i} q \cdot \mathbf{n}_i = \int_{K_i} f \tag{1.3.2}$$

où l'on a noté $\mathcal{E}_i$ l'ensemble des arêtes du polygône K_i et $\mathbf{n}_i$ est la normale extérieure à K_i. Il faut ensuite trouver une valeur approchée *ad hoc*, en fonction de la quantité qu'on souhaite calculer, de $q \cdot \mathbf{n}_i$ sur chaque arête $\sigma \in \mathcal{E}_i$; dans le cas elliptique, si on note u la quantité dont q est le flux, on a $q = -\nabla u$ (loi de Fourier adimensionnée, par exemple): en partant de l'idée qu'on cherche une approximation de u constante sur chaque maille K_i, un choix simple pour $q \cdot \mathbf{n}_i$ consiste à poser $q \cdot \mathbf{n}_i = -\nabla u \cdot \mathbf{n}_i = \frac{u_i - u_j}{d_{i,j}}$ sur l'arête σ située entre les polygônes K_i et K_j, où $d_{i,j}$ représente une longueur bien choisie (typiquement, si on imagine que les u_i sont des valeurs approchées de la quantité cherchée "aux centres des K_i", $d_{i,j}$ sera la distance entre ces "centres").
L'intérêt est bien sûr que l'on s'est ramené à une famille d'équations (une équation (1.3.2) par maille K_i; les équations sont de plus linéaires dans l'exemple considéré) portant sur un nombre fini de valeurs $(u_i)_i$; si le système est sympathique, il va exister une solution que l'on peut espérer calculer de manière informatique, et si l'on peut prouver que ces valeurs $(u_i)_i$ approchent effectivement les valeurs de la solution exacte sur les mailles $(K_i)_i$, on a obtenu une méthode de calcul approchée effective de la quantité qu'on cherche à estimer.
Ceci donne l'idée de base du principe des Volumes Finis; comme leur définition précise est délicate et dépend de l'équation étudiée (la manière dont on partitionne Ω en petits volumes dépend de la nature de l'équation: elliptique, parabolique, hyperbolique...), nous n'irons pas plus en détail dans cette définition. Une référence générale sur les Volumes Finis est le livre [38] de R. Eymard, T. Gallouët et R. Herbin.

1.3.2 Apports

Dans le cadre des équations elliptiques, les études de Volumes Finis existantes se limitaient souvent à des seconds membres $f \in L^2(\Omega)$, ou au moins des fonctions avec une intégrabilité suffisante (voir par exemple [38], [63], [71]...). Or les problèmes réels font parfois intervenir des seconds membres qui n'ont pas cette régularité.
Dans tout ce qui suit, la dimension N de l'espace est 2 ou 3.

Second membre dans le dual de l'espace d'énergie naturel

Considérons le problème modèle:

$$\begin{cases} -\Delta u = f & \text{dans } \Omega\,, \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \tag{1.3.3}$$

Il est bien connu (voir par exemple la résolution de ce problème par le théorème de Lax-Milgram) qu'un espace naturel pour f est le dual $H^{-1}(\Omega)$ de l'espace d'énergie $H^1_0(\Omega)$ associé au laplacien avec conditions au bord de Dirichlet. Qui plus est, de tels seconds membres peuvent naturellement apparaitre dans des problèmes concrets: dans [44] est présentée une équation venant d'un problème électrochimique avec, en particulier, un second membre pouvant se mettre sous la forme $f = \mathrm{div}(G)$ ([4]). Il est donc naturel de chercher à discrétiser (1.3.3) lorsque $f \in H^{-1}(\Omega)$.

Ce fut le but de l'article [D9], reproduit dans le chapitre 4 et écrit en collaboration avec T. Gallouët. Contrairement à [44] qui utilise la forme particulière du G dans $f = \mathrm{div}(G)$, nous présentons dans ce travail une méthode totalement générale pour discrétiser ces équations avec un $G \in L^2(\Omega)^N$ quelconque. L'idée est la suivante: dans le cas $f = \mathrm{div}(G)$ (f n'est donc pas une fonction a priori), pour donner un sens au membre de droite de (1.3.2), on l'écrit comme

$$\int_{K_i} f = \int_{\partial K_i} G \cdot \mathbf{n}_i = \sum_{\sigma \in \mathcal{E}_i} \int_\sigma G \cdot \mathbf{n}_i.$$

Ceci a un sens si G est suffisamment régulière pour que son intégrale sur σ (pour la mesure $(N-1)$-dimensionnelle) soit définie. Dans le cas contraire, il faut remplacer $G \cdot \mathbf{n}_i$ par une approximation adaptée;

[4] Ω étant borné, on sait que $f \in H^{-1}(\Omega)$ si et seulement si $f = \mathrm{div}(G)$ avec $G \in L^2(\Omega)^N$.

l'approximation proposée dans [D9] consiste à prendre

$$G \approx \frac{1}{\mathrm{mes}(\triangle_\sigma)} \int_{\triangle_\sigma} G \quad \text{sur } \sigma,$$

où $\triangle_\sigma$ est un "épaississement" de σ bien choisi et de mesure de Lebesgue non-nulle (il s'agit d'un polygône en forme de diamant autour de σ). Le terme $\int_\sigma G \cdot \mathbf{n}_i$ est alors remplacé par

$$|\sigma| \left(\frac{1}{\mathrm{mes}(\triangle_\sigma)} \int_{\triangle_\sigma} G \right) \cdot \mathbf{n}_i$$

(où $|\sigma|$ est la mesure $(N-1)$-dimensionnelle de σ), ce dernier terme étant bien défini dès que G est une fonction intégrable. Si on considère des maillages et une discrétisation du laplacien comme dans [38], on obtient alors le résultat suivant:

Théorème 1.3.1 *Si $f = \mathrm{div}(G)$ avec $G \in L^2(\Omega)^N$, le problème discrétisé a une unique solution et cette dernière converge dans $L^2(\Omega)$ vers la solution de (1.3.3), lorsque le pas de discrétisation (=le maximum des diamètres des polygônes qui pavent Ω) tend vers 0.*

Ce résultat permet de combler un vide qu'avait la théorie des Volumes Finis par rapport à celle des Eléments Finis: dans cette dernière, le traitement de seconds membres dans $H^{-1}(\Omega)$ est connu depuis toujours (ils ne se traitent pas différemment des seconds membres dans $L^2(\Omega)$).

J'ai annoncé dans la section précédente que la méthode de [D4] pour obtenir des estimations *a priori* sur les solutions d'équations elliptiques non-coercitives s'adapte aussi aux volumes finis; outre traiter des seconds membres plus généraux que d'habitude, c'est ce que montre [D9]. En effet, le résultat du théorème 1.3.1 est aussi valable pour des discrétisations d'équations de la forme

$$\begin{cases} -\Delta u + \mathrm{div}(\mathbf{v}u) = \mathrm{div}(G) & \text{dans } \Omega\,, \\ u = 0 & \text{sur } \partial\Omega \end{cases} \tag{1.3.4}$$

où $\mathbf{v} \in L^p(\Omega)^N$ pour $p > N$ (le caractère non-coercitif de (1.3.4) vient bien sûr du fait que l'on ne suppose pas $\mathrm{div}(\mathbf{v}) \geq 0$). Lors de la construction du schéma Volumes Finis sur ce problème général, on se retrouve à devoir définir une valeur de $\mathbf{v}$ sur chaque arête σ du maillage: ceci se fait comme pour G en moyennant $\mathbf{v}$ sur un "diamant" autour de cette arête.
Pour discrétiser un terme convectif comme $\mathrm{div}(\mathbf{v}u)$ ci-dessus, il est bien connu qu'un "décentrement" doit être effectué (il faut utiliser une valeur approchée pour u qui correspond à des mailles en amont du flux $\mathbf{v}$), si l'on veut s'assurer d'estimations valables pour un maillage de n'importe quelle taille; l'équation non-coercitive (1.3.4) n'échappe pas à cette règle et c'est grâce à ce décentrement ainsi qu'à une adaptation astucieuse des méthodes d'estimation de [D4] qu'on arrive à prouver des estimations *a priori* sur la solution du problème discrétisé (ces estimations permettent de voir que cette solution existe, et sont la clef pour montrer qu'elle converge vers la solution du problème continu).

Second membre avec une régularité intermédiaire

Comme indiqué plus haut, la discrétisation par Volumes Finis d'équations elliptiques avec seconds membres dans $L^2(\Omega)$ est bien connue. En particulier, comme en Eléments Finis, il est possible de montrer un ordre de convergence de la solution du problème discrétisé vers la solution du problème continu: [48] montre que cette convergence se fait dans $L^2(\Omega)$ (et même pour une norme un peu plus forte: une version discrète de la norme H^1) en $\mathcal{O}(h)$, où h est le pas du maillage (le diamètre de la plus grande maille).
Puisque [D9] construit un schéma Volumes Finis pour des seconds membres plus généraux, il peut paraitre naturel de vouloir établir des vitesses de convergence pour ces seconds membres; si l'on se place dans le cadre uniquement H^{-1}, il est à peu près certain que l'on ne pourra pas établir un ordre de convergence particulier: tout comme en Eléments Finis, la vitesse de convergence est directement liée à la régularité

de la solution du problème continu, et si on considère un second membre qui n'est que dans $H^{-1}(\Omega)$, la solution n'est que dans $H_0^1(\Omega)$, ce qui est insuffisant pour établir des erreurs de consistance ayant un ordre par rapport à h. Cependant, si l'on augmente la régularité du second membre, la solution peut voir sa régularité augmenter en parallèle, ce qui peut devenir suffisant pour établir des vitesses de convergence précises.

C'est le propos de [D11], reproduit dans le chapitre 5. Précisément, dans ce travail, on considère des équations de la forme

$$\begin{cases} -\Delta u + \mathrm{div}(\mathbf{v}u) + bu = f + \mathrm{div}(G) & \text{dans } \Omega\,, \\ u = 0 & \text{sur } \partial\Omega\,, \end{cases} \tag{1.3.5}$$

où $\mathbf{v} : \overline{\Omega} \to \mathbb{R}^N$ est continue, $b \in L^\infty(\Omega)$ est positive, $f \in L^2(\Omega)$ et $G \in H^s(\Omega)^N$ avec $s \in [0,1]$ ($s = 0$ correspond à $G \in L^2(\Omega)^N$ et second membre dans $H^{-1}(\Omega)$, et $s = 1$ correspond à un second membre dans $L^2(\Omega)$), que l'on discrétise par la méthode présentée dans [D9] ([5]). Le résultat obtenu est le suivant:

Théorème 1.3.2 *On suppose que $N = 2$ ou que Ω est convexe, et que* $\mathrm{div}(\mathbf{v}) \in L^2(\Omega)$. *Sous des hypothèses de non-dégénérescence du maillage, si $G \in H^s(\Omega)^N$ et la solution de (1.3.5) est dans $H^{1+s}(\Omega)$, alors l'erreur en norme $L^2(\Omega)$ entre la solution du problème discrétisé et la solution de (1.3.5) est en $\mathcal{O}(h^s)$.*

L'idée de la preuve est simple: on prouve l'estimation correspondante quand $s = 0$ (i.e. une estimation en $\mathcal{O}(1)$... bref, il faut montrer que la solution du problème discrétisé reste bornée dans la norme considérée) et quand $s = 1$ (i.e. une estimation en $\mathcal{O}(h)$); le cas des s entre 0 et 1 s'obtient par application de méthodes d'interpolation.
En fait, la vitesse de convergence est prouvée dans une norme plus forte que $L^2(\Omega)$: une forme discrétisée de la norme de $H_0^1(\Omega)$; cette norme est couramment employée dans les Volumes Finis pour équations elliptiques, car c'est celle qui est naturellement adaptée à la formulation discrétisée de l'équation (tout comme la norme de $H_0^1(\Omega)$ est adaptée à la formulation variationnelle de ces mêmes équations). Afin d'utiliser cette norme "H_0^1-discrète", on est obligé d'introduire une discrétisation de la solution de (1.3.5) sur le maillage considéré; contrairement à [48], puisque l'on ne suppose pas que cette solution est dans H^2 et donc continue, on ne peut prendre sa valeur en un point de chaque maille comme discrétisation: on doit introduire une moyenne de la solution sur un ensemble de mesure non-nulle dans la maille. Il est facile de voir que, sauf cas géométrique particulier, prendre la moyenne sur toute la maille n'est pas un bon choix, car les erreurs de consistance ne seront pas contrôlables; il faut donc prendre une moyenne sur une boule contenue dans la maille, et centrée en un point adapté au maillage considéré. De plus, même lorsque $G \in H^1(\Omega)^N$, le schéma que l'on utilise n'est pas le même que dans [48] (on utilise le fait qu'une partie du second membre s'écrit $\mathrm{div}(G)$, tandis que [48] profite directement du fait que le second membre est dans $L^2(\Omega)$, ce qui donne un schéma différent). Cela signifie que, même dans le cas $s = 1$, on doit re-montrer que le schéma qu'on a choisi donne aussi une vitesse de convergence en $\mathcal{O}(h)$; les méthodes employées sont classiques: il faut estimer les erreurs de consistance sur la solution de (1.3.5) lorsqu'elle est injectée dans l'équation discrétisée; l'obtention de ces erreurs demande cependant quelques astuces techniques, car la discrétisation forcée par moyennes sur des boules n'est pas standard et rajoute quelques termes supplémentaires. Les hypothèses de "non-dégénérescence" ([6]) interviennent lorsque l'on somme les erreurs de consistance sur chaque maille.
L'estimation dans le cas $s = 0$ utilise sensiblement les mêmes techniques, mais est plus simple. Les outils d'interpolation sont ensuite usuels: en notant

$$E_s = \{(f, G, u) \in L^2(\Omega) \times H^s(\Omega)^N \times H^{1+s}(\Omega) \mid (f, G, u) \text{ vérifie (1.3.5)}\}$$

[5] A noter que, $\mathbf{v}$ étant continue, on n'utilise en fait pas sa moyenne sur un diamant autour de chaque arête, mais simplement sa moyenne sur chaque arête, qui est bien définie.

[6] Elles demandent essentiellement à ce que le nombre d'arêtes de chaque maille soit borné par une constante, et à ce que les mailles ne s'écrasent pas trop.

et T l'application qui à (f, G, u) (où u est solution de (1.3.5)) associe la différence entre la discrétisation de u et la solution du problème discrétisé, le cas $s = 1$ affirme que T a une norme de l'ordre $\mathcal{O}(h)$ lorsque l'espace de départ est E_1, et de l'ordre de $\mathcal{O}(1)$ lorsque l'espace de départ est E_0 (l'espace d'arrivée étant toujours le même, celui des fonctions constantes sur chaque maille muni de la norme H_0^1-discrète); les théorèmes d'interpolation disent alors que T a une norme d'ordre $\mathcal{O}(h^s)$ lorsque l'espace de départ est l'interpolé d'ordre s entre E_1 et E_0, et les hypothèses sur N ou la convexité de Ω et la divergence de $\mathbf{v}$ ne servent qu'à s'assurer que cet espace interpolé est bien E_s.

Des résultats numériques sont aussi présentés, qui illustrent et confirment les estimations théoriques prouvées. J'avais en fait déjà effectué ces tests numériques (ils ont été présentés lors du congrès international "Finite Volumes for Complex Applications III" qui s'est tenu à Porquerolles en 2002), et c'est en constatant qu'on voyait apparaitre un ordre de convergence naturel qu'il m'a paru important d'arriver à prouver cet ordre.

Second membre mesure

Une des motivations de l'étude des équations elliptiques avec seconds membres mesures est que ces dernières interviennent naturellement dans certains modèles physiques: par exemple dans les modèles de Thomas-Fermi en physique atomique [5], ou dans des modélisations d'écoulements poreux en réservoir [42] (dans ce deuxième cas, la mesure représente des puits qui, vu les échelles en jeu, peuvent être considérés comme des sources ponctuelles en dimension 2 et linéaires en dimension 3). Il parait donc important de pouvoir calculer des approximations des solutions à ces équations.
L'étude d'une discrétisation par Volumes Finis pour le problème modèle (1.3.3) lorsque f est une mesure bornée sur Ω a été effectuée dans [46]; le principe est d'adapter les techniques de [14] afin de montrer des estimations sur les solutions de la discrétisation de (1.3.3). Il est à noter que ces estimations ne semblent pas être directement faisables dans le cadre des Eléments Finis et que ce n'est qu'en comparant les schémas Eléments Finis aux schémas Volumes Finis qu'on parvient, lorsque f est une mesure, à étudier (1.3.3) par la méthode des Eléments Finis (voir [47]).

Dans [D12], reproduit en chapitre 6 et co-écrit avec T. Gallouët et R. Herbin, nous présentons un schéma Volumes Finis pour une équation de convection-diffusion plus générale:

$$\begin{cases} -\Delta u + \operatorname{div}(\mathbf{v}u) + bu = \nu & \text{dans } \Omega\,, \\ u = 0 & \text{sur } \partial\Omega\,, \end{cases} \tag{1.3.6}$$

où $b \in L^2(\Omega)$ est positive, $\mathbf{v} \in C(\overline{\Omega})^N$ et ν est une mesure bornée sur Ω. Cette équation mélange donc deux difficultés: elle est non-coercitive (on n'a rien supposé sur $\operatorname{div}(\mathbf{v})$), et son second membre est une mesure.
Comme dans les autres résultats numériques mentionnés ici, le principal travail pour prouver la convergence vers la solution du problème continu réside dans l'obtention de bornes ([7]) sur les solutions du problème approché. Afin d'obtenir ces bornes dans le cas présent, nous mélangeons les techniques d'estimation des équations non-coercitives avec les techniques d'estimation lorsque le second membre est une mesure, en adaptant tout ceci au cadre discret du schéma numérique; nous obtenons alors le

Théorème 1.3.3 *Sous les hypothèses usuelles concernant le maillage de Ω, la solution de la discrétisation de (1.3.6) converge, lorsque le pas du maillage tend vers 0 et dans $L^q(\Omega)$ pour tout $q < \frac{N}{N-2}$, vers l'unique solution faible de (1.3.6).*

A noter que nous ne considérons pas une solution par dualité à (1.3.6) comme dans [77], mais uniquement une solution faible; son unicité est cependant prouvée, en utilisant les résultats de [52] concernant la régularité, en domaine polygonal, de la solution du problème dual.

[7] Pour des normes $W^{1,q}$ discrètes qui donnent de la compacité dans des espaces de Lebesgue.

On peut aussi remarquer que, comme c'est souvent le cas dans la discrétisation par Volumes Finis, on ne suppose pas que le problème continu (1.3.6) a une solution: la preuve de la convergence des solutions des problèmes approchés donne l'existence d'une solution au problème continu.

Nous exposons aussi quelques tests numériques, lorsque le second membre est une masse de Dirac en un point de Ω. Plusieurs phénomènes émergent à la lecture des ordres de convergence observés lors de ces tests:

- La vitesse de convergence dépend fortement du respect ou non de la symétrie du maillage par rapport à la masse de Dirac considérée; si le maillage est symétrique, on obtient un ordre de convergence (en norme L^1) en $\mathcal{O}(h^{1,77})$ tandis que, si cette symétrie est brisée, on descent à $\mathcal{O}(h)$.
- Si on calcule l'erreur loin de l'endroit où la mesure est concentrée, on retrouve l'ordre de convergence connu en $\mathcal{O}(h^2)$ lorsque le maillage est symétrique par rapport à la mesure, mais uniquement une convergence en $\mathcal{O}(h)$ dans le cas d'une dissymétrie du maillage. Les liens entre le maillage et la mesure du second membre jouent donc un grand rôle y compris sur les zones où la solution est régulière (en dehors du support de la mesure).

Ces résultats semblent suggérer qu'on pourrait établir des ordres de convergence pour la discrétisation par Volumes Finis de (1.3.6). Dans le cas d'une mesure ν totalement générale (très dispersée dans l'ouvert), cela n'est probablement pas réalisable: la solution manque clairement de régularité pour qu'on puisse estimer les erreurs de consistance du schéma. Mais si, comme ci-dessus, on considère le cas d'une mesure concentrée sur une petite zone, alors il devrait être possible d'établir un ordre de convergence loin de cette zone et, si le support de la mesure est effectivement très restreint, de contrôler également les erreurs autour de ce support.

1.4 Approximation parabolique en domaine borné

1.4.1 Problématique

Les équations hyperboliques sont omniprésentes en mécanique des fluides: dès qu'une quantité (comme la masse, la quantité de mouvement, etc...) est conservée, un bilan donne souvent une équation (ou un système d'équations) aux dérivées partielles hyperbolique vérifée par cette quantité. La forme de base en mathématiques, dans ce domaine, est la loi de conservation scalaire suivante:

$$\left\{ \begin{array}{ll} \partial_t u(t,x) + \mathrm{div}(f(u))(t,x) = 0 & t>0\,,\ x\in\mathbb{R}^N\,, \\ u(0,x) = u_0(x) & x\in\mathbb{R}^N\,, \end{array} \right. \tag{1.4.1}$$

où $f \in C^1(\mathbb{R};\mathbb{R}^N)$ et $u_0 \in L^\infty(\mathbb{R}^N)$ sont données.

Le résultat essentiel d'existence et d'unicité pour (1.4.1) est dû à S.N. Kružkov [60], qui résoud en particulier le problème de l'unicité en introduisant une notion de "solution entropique" (on se rend vite compte sur des exemples qu'une simple notion de solution faible, consistant à multiplier (1.4.1) par une fonction régulière et à intégrer par parties, ne donne pas l'unicité). L'idée pour prouver l'existence d'une solution, et trouver une formulation suffisamment forte afin de récupérer l'unicité de la solution, est d'approcher la loi de conservation scalaire par une équation comportant un peu de diffusion, à savoir

$$\left\{ \begin{array}{ll} \partial_t u^\varepsilon(t,x) + \mathrm{div}(f(u^\varepsilon))(t,x) - \varepsilon\Delta u^\varepsilon(t,x) = 0 & t>0\,,\ x\in\mathbb{R}^N\,, \\ u^\varepsilon(0,x) = u_0(x) & x\in\mathbb{R}^N. \end{array} \right. \tag{1.4.2}$$

C'est une idée qui s'inspire de phénomènes physiques réels: dans de nombreuses situations (pas la conservation de la masse, bien sûr, mais par exemple des lois sur la saturation d'une phase dans un écoulement diphasique), une quantité vérifie une équation de la forme (1.4.2) et ce n'est que parce que ε est jugé

assez petit que la diffusion est négligée et (1.4.1) est considérée. Il parait donc naturel de considérer ce dernier problème comme une limite de (1.4.2) lorsque $\varepsilon \to 0$.
Kružkov montre que la solution de (1.4.2) (que l'on sait exister et être unique) converge vers une solution de (1.4.1) au sens suivant: pour tout $\kappa \in \mathbb{R}$, en notant $\phi_\kappa(s) = \mathrm{sgn}(s-\kappa)(f(s)-f(\kappa))$,

$$\partial_t |u(t,x)-\kappa| + \mathrm{div}(\phi_\kappa(u(t,x))) \leq 0 \quad \text{ au sens des distributions dans }]0,\infty[\times\mathbb{R}^N \tag{1.4.3}$$

(à ceci doit aussi être ajoutée la condition initiale, mais nous ne rentrerons pas dans ces questions). La fonction $s \to |s-\kappa|$ est appelée entropie, et ϕ_κ est le flux entropique associé.

Une question qui se pose alors naturellement (tant d'un point de vue mathématique que d'un point de vue physique, afin de comprendre dans quelle mesure le modèle (1.4.1) est effectivement une bonne approximation de (1.4.2) lorsque la diffusion est négligée) est de savoir s'il existe une vitesse de convergence, par rapport à ε, de u^ε vers u. La réponse est donnée par N.N. Kuznecov dans [61]: si la condition initiale est de plus intégrable et a une variation bornée, alors la différence entre u^ε et u dans $C([0,T];L^1(\mathbb{R}^N))$ est en $\mathcal{O}(\sqrt{\varepsilon})$ (pour tout $T>0$).

Ceci concerne des lois de conservation posées sur tout l'espace $\mathbb{R}^N$; cependant, de nombreux problèmes surgissent dans des domaines bornés $\Omega \subset \mathbb{R}^N$. Une des difficultés essentielles des lois de conservations en domaine borné est la manière dont les conditions au bord doivent être comprises; on ne peut, contrairement aux équations paraboliques, imposer une condition sur tout le bord pour la solution: le cas d'école

$$\left\{ \begin{array}{ll} \partial_t u(t,x) + \partial_x u(t,x) = 0 & t>0\,,\ 0<x<1\,, \\ u(0,x) = u_0(x) & 0<x<1\,, \end{array} \right. \tag{1.4.4}$$

dont la solution dans la zone $\{0 \leq t \leq x \leq 1\}$ est $u(t,x) = u_0(x-t)$, montre que l'on ne peut imposer la valeur de u en $x=1$.
Le premier résultat général concernant les conditions au bord pour les équations hyperboliques vient de [3]; lorsque les données sont régulières, la condition au bord doit être formulée de la manière suivante:

$$\forall t>0\,,\ \forall x \in \partial\Omega\,,\ \forall \kappa \in [u(t,x),u_b(t,x)]\,,\ \mathrm{sgn}(u(t,x)-u^b(t,x))(f(u(t,x))-f(\kappa))\cdot \mathbf{n}(x) \geq 0$$

où $\mathbf{n}$ est la normale extérieure à Ω et u_b est la condition au bord "voulue". La méthode pour obtenir cette formulation et prouver l'existence et l'unicité d'une solution est similaire à celle de Kružkov: le problème est approché par une équation parabolique.
Le résultat de [3] a ensuite été amélioré dans [73], où les termes de bord sont inclus dans une formulation purement intégrale qui donne existence et unicité de la solution avec des données uniquement bornées. L'idée essentielle est de ne prendre que des "demi-entropies" de la forme $s \to (s-\kappa)^+$ et $s \to (s-\kappa)^-$ dans la formulation entropique de Kružkov, et non les entropies $s \to |s-\kappa|$.

1.4.2 Apports

Etant donné qu'une bonne formulation au problème de Cauchy-Dirichlet

$$\left\{ \begin{array}{ll} \partial_t u(t,x) + \mathrm{div}(f(u))(t,x) = 0 & t>0\,,\ x\in\Omega\,, \\ u(0,x) = u_0(x) & x\in\Omega\,, \\ u(t,x) = u_b(t,x) & t>0\,,\ x\in\partial\Omega \end{array} \right. \tag{1.4.5}$$

a été obtenue, comme pour le problème de Cauchy (1.4.1), par approximation parabolique

$$\left\{ \begin{array}{ll} \partial_t u^\varepsilon(t,x) + \mathrm{div}(f(u^\varepsilon))(t,x) - \varepsilon\Delta u^\varepsilon(t,x) = 0 & t>0\,,\ x\in\Omega\,, \\ u^\varepsilon(0,x) = u_0(x) & x\in\Omega\,, \\ u^\varepsilon(t,x) = u_b(t,x) & t>0\,,\ x\in\partial\Omega\,, \end{array} \right. \tag{1.4.6}$$

on peut légitimement se poser la question de la vitesse de convergence de u^ε vers u: le résultat de Kuznecov est-il encore vrai en domaine borné? Le problème vient bien sûr des couches limites au bord

du domaine: (1.4.6) impose la valeur de u^ε sur tout $\partial\Omega$ (car le problème est parabolique) alors que l'on a vu que la solution de (1.4.5) peut être loin de cette valeur dans certaines zones de $\partial\Omega$...

Dans le cas où la solution de (1.4.5) est régulière (ce qui demande en particulier de ne regarder la solution qu'en temps petit), des techniques d'étude asymptotique de couches limites permettent de prouver des vitesses de convergence en $\mathcal{O}(\sqrt{\varepsilon})$ ou $\mathcal{O}(\varepsilon)$ selon que le bord de l'ouvert est caractéristique ou non, i.e. selon qu'il existe effectivement une couche limite ou non (ces résultats sont aussi valables pour les systèmes): voir par exemple [53], [50], [20]... . Mais dans le cas où l'on suppose que des chocs peuvent apparaitre dans la solution de (1.4.5) (i.e. si on regarde en temps plus long), aucun ordre de convergence ne semblait avoir été prouvé jusqu'à [D14], qui établit le résultat suivant.

Théorème 1.4.1 *On suppose que Ω, f, u_0 et u_b sont de classe C^2. Si u^ε est la solution de (1.4.6) et u est la solution entropique de (1.4.5), alors pour tout $T > 0$ on a*

$$||u^\varepsilon - u||_{C([0,T];L^1(\Omega))} = \mathcal{O}(\varepsilon^{1/3}).$$

La convergence est donc moins forte que dans le cas du problème de Cauchy, et cela reste une question ouverte que de savoir si cet ordre 1/3 est optimal (la couche limite du bord provoquant donc un véritable ralentissement de la convergence lorsqu'on regarde en temps long) ou non.

Pour prouver ce résultat, nous utilisons une formulation "cinétique" de (1.4.5), qui est équivalente à la formulation entropique; cette formulation a été introduite pour le problème de Cauchy par P.-L. Lions, B. Perthame et E. Tadmor dans [67] et adaptée au problème de Cauchy-Dirichlet dans [58]. L'obtention de ces formulations cinétiques est aisé à comprendre: il suffit d'écrire l'équivalent de (1.4.3) avec la demi-entropie $s \to (s-\kappa)^+$ et de dériver cette inéquation par rapport à κ; on obtient alors une équation sur la fonction $(\mathrm{sgn}(u(t,x)-\kappa))^+$ dont le second membre est la dérivée d'une mesure. La difficulté réside dans la manière dont les conditions initiales et au bord doivent être considérées.
L'idée de la preuve du théorème 1.4.1 est alors de mettre $\mathrm{sgn}(u-\kappa)^+$ dans la formulation cinétique vérifiée par $\mathrm{sgn}(u^\varepsilon-\kappa)^-$, de mettre $\mathrm{sgn}(u^\varepsilon-\kappa)^-$ dans la formulation cinétique vérifée par $\mathrm{sgn}(u-\kappa)^+$ et de sommer. Tout serait très simple si l'on pouvait effectivement faire cela, mais $\mathrm{sgn}(u-\kappa)^+$ et $\mathrm{sgn}(u^\varepsilon-\kappa)^-$ sont loin d'avoir assez de régularité pour être utilisées comme fonctions tests dans les formulations cinétiques; il faut donc régulariser ces fonctions avant de les injecter dans les équations. Le choix des régularisations est crucial et permet de se débarasser des termes de bord génants (problème des couches limites) par un bon décentrement lorsqu'on se rapproche de $\partial\Omega$; il faut ensuite estimer les erreurs commises lors de ces régularisations (qui étalent les fonctions), et une optimisation des paramètres de convolution donne l'ordre 1/3 du théorème.

1.5 Régularisation non-locale de lois de conservation scalaires

1.5.1 Problématique

Le dernier sujet que j'aborde dans ce mémoire vient d'une question directe de P. Clavin, mécanicien à l'IRPHE de Marseille. Lors de l'étude de certaines détonations de gaz (voir [24], [25]), et après simplifications, si on écrit le front du choc comme le graphe $y = v(t,x)$ d'une fonction, on se rend compte que $u = \partial_x v(t,x)$ vérifie une équation de Bürgers modifiée:

$$\partial_t u + \partial_x(u^2/2) + g[u] = 0 \tag{1.5.1}$$

où g est définie en variables de Fourier spatiales par

$$\widehat{g[u]}(\xi) = r(\xi)\widehat{u}(\xi) \quad \text{avec} \quad r(\xi) \sim |\xi| \text{ au voisinage de l'infini.}$$

Les tests effectués montrent que, pour ce modèle, (1.5.1) se comporte à peu près comme Bürgers: des chocs peuvent apparaitre dans la solution.
La question de Clavin était la suivante: si on suppose que

$$\widehat{g[u]}(\xi) = |\xi|^\lambda \widehat{u}(\xi) \quad \text{avec} \quad 1 < \lambda \leq 2 \qquad (\text{i.e. } g[u] = (-\Delta)^{\lambda/2} u), \tag{1.5.2}$$

autrement dit si on met "un peu plus de dérivées" dans g, est-ce que (1.5.1) a un effet régularisant sur la solution? La question est naturelle quand on se rend compte que, lorsque $\lambda = 2$, on a $g[u] = -\Delta u$ ([8]), auquel cas (1.5.1) n'est autre que la régularisation parabolique d'une loi de conservation (cas dans lequel on sait dire que la solution est effectivement régulière).

L'équation sous sa forme générale est la suivante:

$$\begin{cases} \partial_t u(t,x) + \operatorname{div}(f(u))(t,x) + g[u](t,x) = 0 & t > 0\,,\ x \in \mathbb{R}^N\,, \\ u(0,x) = u_0(x) & x \in \mathbb{R}^N \end{cases} \tag{1.5.3}$$

avec g définie par (1.5.2). Ces équations n'interviennent pas que dans les détonations de gaz; le terme $g[u]$ correspond à ce qu'on appelle parfois une "diffusion anormale" (ce terme diffuse beaucoup plus que le laplacien) et il apparait dans des problèmes de construction de semi-conducteur par déposition de vapeur chimique, ainsi que dans d'autres phénomènes physiques (voir [81]).
Les quelques résultats connus sur cette équation sont principalement dûs à P. Biler, T. Funaki et W.A. Woyczinsky [7], et sont les suivants:

- Si $N = 1$, $u_0 \in H^1(\mathbb{R})$, $f(u) = u^2$ et $\lambda > 3/2$, une unique solution existe dans $L^\infty(0,T;H^1(\mathbb{R})) \cap L^2(0,T;H^{1+\lambda/2}(\mathbb{R}))$ pour tout $T > 0$.
- Si $N = 1$, $u_0 \in H^1(\mathbb{R}) \cap W^{1,1}(\mathbb{R})$, $f(u) = u^2$ et $1/2 < \lambda \leq 2$, une unique solution existe dans $L^\infty(]0,\infty[\times\mathbb{R}) \cap L^\infty(0,\infty;BV(\mathbb{R}))$.
- Si $N \geq 1$, $f(u) = u^r$, alors il existe une unique solution *locale en temps* dans des espaces de Morrey et, si u_0 est assez petit dans certains espaces, cette solution est globale.

Les techniques employées sont des estimations d'énergie par des méthodes usuelles: multiplication de l'équation par u ou ses dérivées et intégrations par parties; les limites sur la puissance apparaissant dans f ou la dimension de l'espace viennent des injections de Sobolev employées pour contrôler le terme donné par f dans ces estimations.

Pour palier le problème de l'existence uniquement locale en temps, [8] et [9] rajoutent un terme $-\Delta u$ à l'équation et considèrent donc

$$\partial_t u + \operatorname{div}(f(u)) + g[u] - \Delta u = 0\,,$$

ce qui permet d'avoir des estimations suffisantes pour montrer que la solution est globale en temps; ces travaux étudient en effet surtout le comportement asymptotique de la solution.

1.5.2 Apports

Existence, unicité et régularité

Les résultats d'existence et de régularité pour (1.5.3) n'etaient donc pas très concluants; mais lorsque T. Gallouët, J. Vovelle et moi-même avons abordé cette équation suite à la question de P. Clavin, nous ne connaissions pas ces références et nous avons développé dans [D10] (voir chapitre 8) une méthode totalement différente de celles décrites ci-dessus.

[8] A une constante positive multiplicative près, dépendant de la définition choisie pour la transformée de Fourier.

La première idée que nous avons eue pour aborder ce problème (et qui figure en fait aussi dans [7]) est de considérer $\mathrm{div}(f(u))$ comme une perturbation de $\partial_t u + g[u] = 0$. On sait en effet bien décrire le semi-groupe d'évolution correspondant à l'opérateur g: en passant en Fourier, on voit immédiatement que la solution de $\partial_t v + g[v] = 0$ avec condition initiale v_0 est donnée par

$$v(t,x) = K(t,\cdot) * v_0(x) \quad \text{avec} \quad K(t,x) = \mathcal{F}^{-1}(\xi \to e^{-t|\xi|^\lambda})$$

($\mathcal{F}^{-1}$ est la transformée de Fourier inverse). On applique alors une formule de Duhamel sur $\partial_t u + g[u] = -\mathrm{div}(f(u))$ et on obtient formellement

$$\begin{aligned} u(t,x) &= K(t,\cdot) * u_0(x) - \int_0^t K(t-s,\cdot) * \mathrm{div}(f(u(s,\cdot)))(x)\, ds \\ &= K(t,\cdot) * u_0(x) - \int_0^t \nabla K(t-s,\cdot) * (f(u(s,\cdot)))(x)\, ds. \end{aligned} \tag{1.5.4}$$

Le noyau K n'est pas explicitement calculable, mais on sait donner assez d'estimations sur lui et sa dérivée pour s'assurer que (1.5.4) a un sens quand u est bornée.
On peut donc prendre comme définition de solution à (1.5.3) toute fonction bornée qui vérifie (1.5.4). Cela permet, par un point fixe contractant, de prouver l'unicité de la solution sur n'importe quel intervalle de temps, son existence en temps petit et sa régularité, pour tout u_0 bornée, dès que $t > 0$. Cependant, cela ne donne aucun moyen de s'assurer que la solution est globale en temps: cette formulation ne permet pas de voir que la norme infinie de u reste bornée et n'explose pas en un certain $T > 0$.

Comme l'équation (1.5.3) ne permet pas non plus d'obtenir des estimations *a priori* sur la norme infinie de u (à cause du caractère non-local de g, les techniques usuelles — basées sur des intégrations par parties — qui donnent ces estimations dans le cas parabolique ne peuvent s'adapter ici), il a fallu trouver une autre méthode pour obtenir une solution globale en temps. La méthode en question est celle du splitting; c'est une technique couramment employée en analyse numérique, mais qui, à notre connaissance, n'avait jamais été utilisée pour construire directement des solutions à un problème continu. Pour construire une solution globale, on procède comme suit.

- Pendant un petit intervalle de temps $[0,\delta]$, on laisse évoluer u selon l'opérateur $\partial_t u + 2g[u] = 0$ (la solution est bien connue et donnée par la convolution par K).
- Puis sur l'intervalle de temps $[\delta, 2\delta]$, on part de la valeur trouvée en δ par l'étape précédente et on laisse évoluer u selon la loi de conservation $\partial_t u + 2\mathrm{div}(f(u)) = 0$ (là aussi, on connait de nombreuses propriétés pour cette équation).
- On recommence ensuite sur $[2\delta, 3\delta]$ en résolvant $\partial_t u + 2g[u] = 0$.
- Puis on résoud $\partial_t u + 2\mathrm{div}(f(u)) = 0$ sur $[3\delta, 4\delta]$, etc...

On a ainsi construit une fonction u^δ qui consiste à laisser chaque opérateur $\partial_t + 2g$ et $\partial_t + 2\mathrm{div}(f(\cdot))$ évoluer séparément sur des intervalles de temps distincts. L'intérêt majeur est le suivant: comme le noyau K est positif (voir [65]) et d'intégrale spatiale 1, l'opérateur $\partial_t + 2g$ ne fait pas croitre les normes L^∞, L^1 et BV; la loi de conservation scalaire a la même propriété de non-croissance de ces normes. Ainsi, des estimations sur u^δ, indépendantes de δ, sont obtenues dans L^∞, L^1 et BV. On peut alors espérer, comme pour le splitting en analyse numérique, que u^δ se rapproche, lorsque $\delta \to 0$, d'une solution u de (1.5.3) ([9]); c'est ce qui est prouvé, et on peut résumer les résultats obtenus dans le théorème suivant.

Théorème 1.5.1 *Si $u_0 \in L^\infty(\mathbb{R}^N)$, il existe une unique solution $u \in L^\infty(]0,\infty[\times\mathbb{R}^N)$ à (1.5.3) au sens (1.5.4). De plus, cette solution vérifie*

[9] Le facteur "2" qui a été mis devant g s'explique de la manière suivante: au total, $\partial_t + 2g$ n'aura évolué que sur la moitié du temps quand on construit u^δ, et il est donc nécessaire de lui donner un poids double pour retrouver effectivement g et non $\frac{1}{2}g$ à la fin; idem pour $2\mathrm{div}(f(\cdot))$.

i) $u \in C^\infty(]0,\infty[\times\mathbb{R}^N)$ *et a toutes ses dérivées bornées sur* $]t_0,\infty[\times\mathbb{R}^N$ *(pour tout* $t_0 > 0$*),*

ii) pour tout $t \geq 0$, $||u(t)||_{L^\infty(\mathbb{R}^N)} \leq ||u_0||_{L^\infty(\mathbb{R}^N)}$.

Outre la borne L^∞ donnée dans le point ii) du théorème, la construction d'une solution approchée par splitting permet d'obtenir de nombreuses autres propriétés: toute propriété qui est commune aux opérateurs $\partial_t + g$ et $\partial_t + \mathrm{div}(f(\cdot))$ est vérifiée par (1.5.3); par exemple, on peut montrer un principe de contraction L^1 des solutions, ou un principe du maximum.
Les résultats d'existence et de régularité de [7] se trouvent donc nettement améliorés par ce théorème. Qui plus est, les méthodes employées ne sont pas limitées aux opérateurs décrits ci-dessus: essentiellement, elles s'appliquent dès que le noyau de l'opérateur est positif et suffisamment intégrable; par exemple, on peut sans problème considérer des sommes d'opérateurs g pour des λ différents (pourvu qu'ils restent entre 0 et 2 — ceci afin que leurs noyaux soient positifs — et que l'un d'eux au moins soit strictement supérieur à 1 — pour que le noyau de l'ensemble ait de bonnes propriétés d'intégrabilité, i.e. qu'au moins un des opérateurs régularise effectivement l'équation). En fait, le principe de la méthode de splitting est très général: il consiste à étudier chaque opérateur d'évolution (ici, la loi de conservation scalaire et $\partial_t + g$) séparément et à vérifier qu'ils ont suffisamment de propriétés communes (ici, qu'ils donnent des estimations de compacité dans les mêmes espaces) pour être cumulés.

Le chapitre 9 n'est pas tiré d'un article; il s'agit d'une part de l'adaptation du chapitre 8 au cas d'une dimension d'espace quelconque ([D10] a en effet été écrit, pour des raisons de facilité de lecture, dans le cas $N = 1$), et d'autre part d'une généralisation relativement immédiate (mais assez technique) de ce qui précède au cas

$$\begin{cases} \partial_t u(t,x) + \mathrm{div}(f(t,x,u(t,x))) + g[u](t,x) = h(t,x,u(t,x)) & t>0\,,\ x\in\mathbb{R}^N\,, \\ u(0,x) = u_0(x) & x\in\mathbb{R}^N\,. \end{cases} \tag{1.5.5}$$

Il m'a semblé intéressant de l'inclure dans ce mémoire.

Diffusion evanescente

Si l'on considère effectivement, au vu du théorème 1.5.1, l'opérateur g comme régularisant la loi de conservation scalaire, une question naturelle est alors de savoir si cette régularisation donne une solution qui est proche ou non de celle la loi de conservation originelle. Autrement dit, si on considère

$$\begin{cases} \partial_t u^\varepsilon(t,x) + \mathrm{div}(f(u^\varepsilon))(t,x) + \varepsilon g[u^\varepsilon](t,x) = 0 & t>0\,,\ x\in\mathbb{R}^N\,, \\ u^\varepsilon(0,x) = u_0(x) & x\in\mathbb{R}^N\,, \end{cases} \tag{1.5.6}$$

est-ce que u^ε se rapproche de la solution entropique de (1.4.1) lorsque $\varepsilon \to 0$?

La réponse n'est pas forcément évidente si on pense au cas de la limite dispersive décrite dans [30]: lorsque l'on regarde l'équation de KdV (Bürgers régularisée par un opérateur d'ordre 3)

$$\partial_t v^\varepsilon + \partial_x((v^\varepsilon)^2/2) = \varepsilon\partial_{xxx} v^\varepsilon\,,$$

alors v^ε oscille beaucoup trop lorsque $\varepsilon \to 0$ et ne converge effectivement pas vers la solution entropique de l'équation de Bürgers.

Quand on cherche à montrer la convergence de l'approximation parabolique

$$\partial_t v^\varepsilon + \mathrm{div}(f(v^\varepsilon)) - \varepsilon\Delta v^\varepsilon = 0 \tag{1.5.7}$$

vers le problème hyperbolique, un point clef est d'établir une inégalité d'entropie pour le problème parabolique, i.e. en prenant η une fonction convexe et $\phi' = \eta' f'$, à montrer que

$$\partial_t \eta(v^\varepsilon) + \mathrm{div}(\phi(v^\varepsilon)) - \varepsilon\Delta(\eta(v^\varepsilon)) \leq 0.$$

Cela s'établit très simplement dans ce cas précis: il suffit de multiplier (1.5.7) par $\eta'(v^\varepsilon)$ et d'utiliser la convexité de η pour voir que

$$\Delta(\eta(v^\varepsilon)) = \eta''(v^\varepsilon)|\nabla v^\varepsilon|^2 + \eta'(v^\varepsilon)\Delta v^\varepsilon \geq \eta'(v^\varepsilon)\Delta v^\varepsilon. \tag{1.5.8}$$

Dans le cas de (1.5.3), cette méthode échoue car (1.5.8) est basé sur la dérivation de fonctions composées et aucune formule ne semble exister *a priori* qui nous dise comment $g[\eta(u^\varepsilon)]$ se comporte par rapport à $\eta'(u^\varepsilon)g[u^\varepsilon]$.

Cependant, en replongeant dans la méthode de splitting, on est capable d'établir effectivement une inégalité d'entropie pour (1.5.3), comme le montre [D13], reproduit en chapitre 10. L'idée est de regarder les pertes d'entropie sur chaque intervalle $[k\delta,(k+1)\delta]$ apparaissant dans la construction de l'approximation; sur les intervalles où on laisse évoluer $\partial_t + 2\mathrm{div}(f(\cdot))$, on sait comment ces pertes se font ([10]); sur les intervalles où c'est $\partial_t + 2g$ qui évolue, on a une représentation par convolution de la solution: on utilise sur cette expression la formule de Jensen (la convolution se faisant contre un noyau d'intégrale 1) qui nous permet de voir que l'entropie a bien plutôt tendance à diminuer, modulo un terme qui est petit quand δ est petit.
On peut alors, en laissant $\delta \to 0$, établir l'inégalité d'entropie suivante pour la solution de (1.5.6)

$$\partial_t \eta(u^\varepsilon) + \mathrm{div}(\phi(u^\varepsilon)) + \varepsilon g[\eta(u^\varepsilon)] \leq 0.$$

Il est alors possible d'appliquer la méthode de dédoublement des variables entre cette inégalité entropique et celle vérifiée par la solution de (1.4.1) pour prouver le

Théorème 1.5.2 *Si $u_0 \in L^\infty(\mathbb{R}^N)$, u^ε est la solution de (1.5.6) et u est la solution entropique de (1.4.1), alors, lorsque $\varepsilon \to 0$, $u^\varepsilon \to u$ dans $C([0,T];L^1_{\mathrm{loc}}(\mathbb{R}^N))$ pour tout $T>0$.*

Ensuite, comme dans la section précédente, il parait normal de s'interroger sur une éventuelle vitesse de convergence.

Théorème 1.5.3 *Si $u_0 \in L^\infty(\mathbb{R}^N)\cap L^1(\mathbb{R}^N)\cap BV(\mathbb{R}^N)$, u^ε est la solution de (1.5.6) et u est la solution entropique de (1.4.1), alors pour tout $T>0$ on a*

$$||u^\varepsilon - u||_{C([0,T];L^1(\mathbb{R}^N))} = \mathcal{O}(\varepsilon^{1/\lambda}).$$

La méthode employée est la même que dans [61] et dans le théorème 1.5.2: on utilise la technique de dédoublement des variables de Kružkov et on estime la taille de chaque terme venant de l'étalement provoqué par cette méthode. Pour obtenir l'ordre en $1/\lambda$, on établit que $g[\varphi] = C|\cdot|^{-N-(\lambda-2)} * \Delta\varphi$ (lorsque φ est dans la classe de Schwartz) et on découpe, selon un paramètre à optimiser, les intégrales qui apparaissent pour profiter des comportements de $|\cdot|^{-N-(\lambda-2)}$ près de 0 et à l'infini.
Le théorème 1.5.3 indique donc que g diffuse moins que le laplacien (la convergence vers l'équation hyperbolique est plus rapide avec g pour $\lambda<2$ qu'avec une régularisation parabolique); cela semble en contradiction avec ce que j'ai dit plus haut, à savoir que g diffuse plus que le laplacien. En fait, il n'y a pas de contradiction une fois qu'on précise les choses: *en temps long*, g diffuse effectivement plus que le laplacien, mais *en temps court*, g diffuse moins; cela se comprend aisément quand on regarde les propriétés d'homogénéité du noyau K_λ associé à g et du noyau K_2 (la gaussienne) associé au laplacien. On a en effet

$$K_\lambda(t,x) = t^{-N/\lambda}K(1,t^{-1/\lambda}x).$$

[10] En fait, quand δ est petit, la solution reste régulière lors de l'évolution de l'opérateur hyperbolique et il n'y a donc pas de perte d'entropie.

Si l_λ est l'étalement de $K_\lambda(1,\cdot)$ (i.e. sa largeur caractéristique: par exemple, $K_\lambda(1,x) \geq \frac{1}{2}K_\lambda(1,0)$ dès que $|x| \leq l_\lambda$), alors la formule précédente montre que $t^{1/\lambda}l_\lambda$ est l'étalement de $K_\lambda(t,\cdot)$; pour $\lambda < 2$ fixé, on voit donc que $K_\lambda(t,\cdot)$ s'étale moins que $K_2(t,\cdot)$ pour t petit, et plus pour t grand.

Une continuation naturelle dans l'étude de ces problèmes non-locaux est de considérer le cas où l'équation est posée sur un ouvert borné de $\mathbb{R}^N$. La bonne définition pour g sera alors certainement par une puissance fractionnaire du laplacien avec conditions au bord de Dirichlet (par exemple), et il faudra comprendre précisément les propriétés du semi-groupe associé à cet opérateur pour pouvoir adapter la méthode qui précède.

Chapitre 2

Liste des Travaux

2.1 Publications dans des revues internationales avec comité de lecture

[D1] *Optimal Pointwise Control of a Semilinear Parabolic Equation*, J. Droniou and J.-P. Raymond. Nonlinear Anal. **39** (2000), no. 2, Ser. A: Theory Methods, 135-156.

[D2] *Solving convection-diffusion equations with mixed, Neumann and Fourier boundary conditions and measures as data, by a duality method*, J. Droniou. Adv. Differential Equations **5** (2000), no. 10-12, 1341-1396.

[D3] *A uniqueness result for quasilinear elliptic equations with measures as data*, J. Droniou and T. Gallouët. Rend. Mat. Appl. (7) **21** (2001), no. 1-4, 57-86.

[D4] *Non-coercive Linear Elliptic Problems*, J. Droniou. Potential Anal. **17** (2002), no. 2, 181-203.

[D5] *A density result in Sobolev spaces*, J. Droniou. J. Math. Pures Appl. (9) **81** (2002), no. 7, 697-714.

[D6] *Parabolic capacity and soft measures for nonlinear equations*, J. Droniou, A. Porretta and A. Prignet. Potential Anal. **19** (2003), no. 2, 99-161.

[D7] *Convergence of a finite volume – mixed finite element method for a system of a hyperbolic and an elliptic equations*, J. Droniou, R. Eymard, D. Hilhorst and X. D. Zhou. IMA J. Numer. Anal. **23** (2003), no. 3, 507-538.

[D8] *Global and local estimates for nonlinear noncoercive elliptic equations with measure data*, J. Droniou. Comm. Partial Differential Equations **28** (2003), no. 1-2, 129-153.

[D9] *Finite volume methods for convection-diffusion equations with right-hand side in H^{-1}*, J. Droniou and T. Gallouët. M2AN Math. Model. Numer. Anal. **36** (2002), no. 4, 705-724.

[D10] *Global solution and smoothing effect for a non-local regularization of an hyperbolic equation*, J. Droniou, T. Gallouët and J. Vovelle. J. Evol. Equ. **3** (2003), no. 3, 499-521.

[D11] *Error estimates for the convergence of a finite volume discretization of convection-diffusion equations*, J. Droniou. J. Numer. Math. **11** (2003), no. 1, 1-32.

[D12] *A finite volume scheme for a noncoercive elliptic equation with measure data*, J. Droniou, T. Gallouët and R. Herbin. SIAM J. Numer. Anal. **41** (2003), no. 6, 1997-2031.

[D13] *Vanishing non-local regularization of a scalar conservation law*, J. DRONIOU. Electron. J. Differential Equations **2003** (2003), no. 117, 1-20.

[D14] *An error estimate for the parabolic approximation of multidimensional scalar conservation laws with boundary conditions*, J. DRONIOU, C. IMBERT AND J. VOVELLE. Ann. Inst. H. Poincaré Anal. Non Linéaire **21** (2004), no. 5, 689-714.

Dans la suite de ce document, je n'ai reproduit que les articles qui correspondent à des travaux effectués après ma thèse ([D8] à [D14]). Tous mes articles sont cependant disponibles sur la page web `http://www-gm3.univ-mrs.fr/~droniou/travaux.html`.

2.2 Publications dans des actes de congrès

[Dc1] *Contrôle de l'architecture et des représentations internes dans les réseaux de neurones multicouches*, J. Droniou, A. Elisseeff, H. Paugam-Moisy et O. Teytaud, Actes de la Conférence sur l'Apprentissage, CAP'99, 185–194. Palaiseau, France, 1999.

[Dc2] *A finite volume scheme for noncoercive Dirichlet problems with right-hand sides in* H^{-1}, J. Droniou and T. Gallouët, "Finite Volumes for Complex Applications III", Hermes Penton Science, 195–202. Porquerolles, France, 2002.

2.3 Autres publications

[Dp1] *Intégration et Espaces de Sobolev à Valeurs Vectorielles*, J. Droniou, Polycopiés de l'Ecole Doctorale de Maths-Info de Marseille,
disponible sur `http://www-gm3.univ-mrs.fr/polys/gm3-02/`.

[Dp2] *Quelques Résultats sur les Espaces de Sobolev*, J. Droniou, Polycopiés de l'Ecole Doctorale de Maths-Info de Marseille,
disponible sur `http://www-gm3.univ-mrs.fr/polys/gm3-03/`.

Partie II

Régularité locale de solutions d'équations elliptiques avec mesures

Chapitre 3

Global and local estimates for nonlinear noncoercive elliptic equations with measure data

Reference: J. Droniou. *Comm. Partial Differential Equations* **28** (2003), no. 1-2, 129-153.

Abstract We study nonlinear noncoercive elliptic problems with measure data, proving first that the global estimates already known when the problem is coercive are still true for noncoercive problems. We then prove new estimates, on sets far from the support of the singular part of the right-hand side, in the energy space associated to the operator, which entails additional regularity results on the solutions.

3.1 Introduction and main results

3.1.1 The problem

Let Ω be a bounded open subset of $\mathbb{R}^N$. The purpose of this work is to obtain global and local estimates on the weak solutions to:

$$\begin{cases} -\mathrm{div}(a(x,u,\nabla u)) - \mathrm{div}(\Phi(x,u)) = \mu + f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{3.1.1}$$

where $-\mathrm{div}(a(x,u,\nabla u))$ is a Leray-Lions operator acting on $W_0^{1,p}(\Omega)$ $(1 < p \leq N)$, $\Phi(x,u)$ is a convection term with growth properties, $f \in W^{-1,p'}(\Omega)$ and $\mu \in \mathcal{M}(\Omega)$ (see the precise hypotheses on these data in Subsection 3.1.2).

Nonlinear elliptic problems with measure data have been studied in a number of papers. A quite efficient way to prove the existence of a solution to such problems is, as first shown in [11], to use an approximation method: taking a sequence of regular data which converges to the measure of the right-hand side, one can prove that the solutions corresponding to these regular data converge, in a sense, to a solution of the equation with measure data.
This method has been widely used to obtain different kinds of solutions to elliptic problems with measure data: see for example [12], [13], [4] or [28]. In each of these works, the first step to prove the convergence of the solutions corresponding to regular right-hand sides is, of course, to obtain estimates on these solutions in adequate spaces.
We intend here to obtain new estimates on elliptic problems with measure data.

One of the novelty of this paper is the presence of the convective term defined by Φ; because of this term, the elliptic equation (3.1.1) is not coercive, and obtaining estimates on the solutions to this problem (with regular or singular right-hand side) is thus quite difficult. [54] give estimates for a noncoercive problem with right-hand side in $L^1(\Omega)$, using the tools of renormalized solutions. Here, we will rather use the method developed in [31] for linear noncoercive elliptic problems; this method was then adapted in [32] to nonlinear variational noncoercive elliptic problems and to finite volume schemes for linear noncoercive elliptic equations. We will see that, mixing the methods of [31] (to handle the noncoercive feature of the equation) and of [14] (to handle the singularity of the right-hand side), estimates on solutions to (3.1.1), similar to the ones already known when the problem is coercive, are easy to prove.

But the main originality of the work we present here is certainly the local estimate outside the support of the singular part of the right-hand side.
It is well-known that the solutions to elliptic problems with right-hand side measures do not belong, in general, to the energy space $W_0^{1,p}(\Omega)$ associated to the equation (the global estimates we obtain on solutions to (3.1.1) are, roughly speaking, in $W_0^{1,q}(\Omega)$ spaces, with $q < \frac{N(p-1)}{N-1}$). This is quite obvious, since the right-hand side does not belong to the dual space of $W_0^{1,p}(\Omega)$. But if the singular part of the right-hand side is concentrated on some subset of Ω, one can hope that, outside this subset, the solution is as regular as the operator allows.
We will indeed prove estimates (and thus regularity results), far from the support of μ, on solutions to (3.1.1) in $W^{1,p}$. These estimates are not straightforward because, to study the solutions far from the support of μ, we must introduce cut-off functions that not only have to satisfy some special properties but also entail the apparition of terms which are not easily bounded; we thus first use a bootstrap technique to reach estimates in $W^{1,q}$ for all $q < p$, and then, thanks to these estimates, we prove the desired bound in $W^{1,p}$.

In the rest of this section, we state the precise hypotheses on the data and the main results of this paper (global and local estimates, as well as their consequences on the regularity of solutions corresponding to right-hand side measures). In Section 2, we prove the global estimates on solutions to (3.1.1). Section 3, the biggest part of this paper, is devoted to the proof of the local estimates, far from the support of μ. We then quickly show, in Section 4, how these estimates allow to obtain solutions with regularity properties when the right-hand side is a measure. Section 5 is an appendix with two easy technical results useful in the rest of the paper.

3.1.2 Hypotheses and notations

Ω is a bounded open subset of $\mathbb{R}^N$ ($N \geq 2$). $|\cdot|$ is the Euclidean norm in $\mathbb{R}^N$; $B(e,r)$ and $\overline{B}(e,r)$ denote the open and closed ball in $\mathbb{R}^N$ of center e and radius r. meas(A) is the Lebesgue measure of a measurable set $A \subset \mathbb{R}^N$.
We take $p \in]1, N]$; if $p < N$, we denote $N_* = N$ and, if $p = N$, we take $N_* > N$. We let $p^* = \frac{N_* p}{N_* - p}$. $\overline{p}$ is a real number in $[\frac{N_* p}{N_* - p + 1}, \frac{(N_* - 1)p}{N_* - p}[$ ([1]).

$W_0^{1,r}(\Omega)$ denotes the usual Sobolev space, endowed with the norm $||u||_{W_0^{1,r}(\Omega)} = || \, |\nabla u| \, ||_{L^r(\Omega)}$; $W^{-1,r'}(\Omega)$ is the dual space of $W_0^{1,r}(\Omega)$. $\mathcal{M}(\Omega)$ is the space of bounded measures on Ω, identified, through the Riesz theorem, to the dual space of $C_c(\Omega)$ (this last space is endowed with the supremum norm).

[1] The choice $\overline{p} \geq \frac{N_* p}{N_* - p + 1}$ is only made to avoid the introduction of a new notation in the following proofs; this restriction on $\overline{p}$ is not a problem: if a satisfies (3.1.4) for $0 < \overline{p} < \frac{(N_* - 1)p}{N_* - p}$, then it also satisfies this hypotheses for some $\overline{p} \in [\frac{N_* p}{N_* - p + 1}, \frac{(N_* - 1)p}{N_* - p}[$ (notice that, indeed, $\frac{N_* p}{N_* - p + 1} < \frac{(N_* - 1)p}{N_* - p}$).

The hypotheses on the data of (3.1.1) are:

$$a : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N \text{ is a Caratheodory function,} \tag{3.1.2}$$

$$\begin{gathered}\text{there exists } \nu > 0 \text{ and } \Theta \in L^1(\Omega) \text{ such that } a(x,s,\xi)\cdot\xi \geq \nu|\xi|^p - \Theta(x) \\ \text{for a.e. } x \in \Omega, \text{ for all } (s,\xi) \in \mathbb{R} \times \mathbb{R}^N,\end{gathered} \tag{3.1.3}$$

$$\begin{gathered}\text{there exists } \beta > 0 \text{ and } h \in L^{p'}(\Omega) \text{ such that} \\ |a(x,s,\xi)| \leq h(x) + \beta|s|^{\overline{p}-1} + \beta|\xi|^{p-1} \text{ for a.e. } x \in \Omega, \text{ for all } (s,\xi) \in \mathbb{R} \times \mathbb{R}^N,\end{gathered} \tag{3.1.4}$$

$$\begin{gathered}(a(x,s,\xi) - a(x,s,\eta))\cdot(\xi - \eta) > 0 \\ \text{for a.e. } x \in \Omega, \text{ for all } (s,\xi,\eta) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N \text{ such that } \xi \neq \eta.\end{gathered} \tag{3.1.5}$$

$$\begin{gathered}\Phi : \Omega \times \mathbb{R} \to \mathbb{R}^N \text{ is a Caratheodory function such that there exists } g \in L^{\frac{N_*}{p-1}}(\Omega) \text{ satisfying} \\ |\Phi(x,s)| \leq g(x)(1+|s|^{p-1}) \text{ for a.e. } x \in \Omega, \text{ for all } s \in \mathbb{R},\end{gathered} \tag{3.1.6}$$

$$\begin{gathered}\mu \in \mathcal{M}(\Omega)\,,\ f \in W^{-1,p'}(\Omega)\,,\ ||\mu||_{\mathcal{M}(\Omega)} + ||f||_{W^{-1,p'}(\Omega)} \leq \Lambda \text{ and} \\ K \text{ is a compact subset of } \mathbb{R}^N \text{ such that } \mathrm{supp}(\mu) \subset K.\end{gathered} \tag{3.1.7}$$

We know (see [32]) that, under Hypotheses (3.1.2)—(3.1.7), if μ belongs to $W^{-1,p'}(\Omega)$, there exists at least a solution to (3.1.1) in the sense

$$\left\{\begin{array}{l} u \in W_0^{1,p}(\Omega), \\ \displaystyle\int_\Omega a(x,u,\nabla u)\cdot\nabla\varphi + \int_\Omega \Phi(x,u)\cdot\nabla\varphi = \langle\mu,\varphi\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} + \langle f,\varphi\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}\,, \\ \qquad \forall\varphi \in W_0^{1,p}(\Omega). \end{array}\right. \tag{3.1.8}$$

3.1.3 Main results

Our main results are the following.

Theorem 3.1.1 (Global Estimates) *Under Hypotheses (3.1.2)—(3.1.7), if $\mu \in W^{-1,p'}(\Omega)$ then, for all $0 < r < \frac{N_*(p-1)}{N_*-p}$ and all $0 < s < \frac{N_*(p-1)}{N_*-1}$, there exists $C > 0$ only depending on (Λ, r, s) such that, for any solution u of (3.1.8), we have*

$$\int_\Omega |u|^r \leq C \quad \text{and} \quad \int_\Omega |\nabla u|^s \leq C.$$

Remark 3.1.1 *Of course, this C and all the constants in the sequel also depend on the other numerous data involved in the hypotheses, i.e. on $(\Omega, p, N_*, \nu, \Theta, g, h, \beta, \overline{p})$, but we have chosen to emphasize the dependance only on the important data, that is to say the space in which we obtain the estimates (through the exponents (r,s) in the global estimates), the norm of μ in $\mathcal{M}(\Omega)$ (through Λ) and, in the local estimates, the support of μ (through K). In fact, a close examination of the proofs also shows that the constants appearing in the global estimates do not depend on $(h, \beta, \overline{p})$.*

Remark 3.1.2 *Notice that we can have $\frac{N_*(p-1)}{N_*-p} \leq 1$ and $\frac{N_*(p-1)}{N_*-1} \leq 1$, in which case Theorem 3.1.1 does not give estimates in Lebesgue or Sobolev spaces.*
One can also notice that, once we have the result of Theorem 3.1.1, it is possible to obtain estimates on u in the Marcinkiewicz space of exponent $\frac{N_*(p-1)}{N_*-p}$ and on ∇u in the Marcinkiewicz space of exponent $\frac{N_*(p-1)}{N_*-1}$ (but this does not improve the integrability properties of u or ∇u).

Theorem 3.1.2 (Local Estimates) *Under Hypotheses (3.1.2)—(3.1.7), if $\mu \in W^{-1,p'}(\Omega)$ then, for all $\varepsilon > 0$, there exists $C > 0$ only depending on $(\Lambda, K, \varepsilon)$ such that, for any solution u of (3.1.8), denoting $F = K + \overline{B}(0,\varepsilon)$, we have*

$$||u||_{L^{p^*}(\Omega\setminus F)} \leq C \quad \text{and} \quad ||u||_{W^{1,p}(\Omega\setminus F)} \leq C.$$

As a consequence of these estimates, we obtain the following existence and regularity result on a nonlinear noncoercive elliptic equation with measure data.

Theorem 3.1.3 (Existence result for a right-hand side measure) *Suppose that $p > 2 - \frac{1}{N_*}$. Under Hypotheses (3.1.2)—(3.1.7), there exists a solution to (3.1.1) in the sense*

$$\left\{ \begin{array}{l} u \in \bigcap_{q<\frac{N_*(p-1)}{N_*-1}} W_0^{1,q}(\Omega)\,, \\ \forall \varepsilon > 0\,, \ \text{denoting } F_\varepsilon = \overline{\mathrm{supp}(\mu) + B(0,\varepsilon)},\ u \in L^{p^*}(\Omega\setminus F_\varepsilon) \cap W^{1,p}(\Omega\setminus F_\varepsilon)\,, \\ \displaystyle\int_\Omega a(x,u,\nabla u)\cdot\nabla\varphi + \int_\Omega \Phi(x,u)\cdot\nabla\varphi = \int_\Omega \varphi\, d\mu + \langle f,\varphi\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}\,, \\ \qquad \forall \varphi \in \bigcup_{s>N_*} W_0^{1,s}(\Omega). \end{array} \right. \tag{3.1.9}$$

Remark 3.1.3 *We have defined F_ε as the closure of* $\mathrm{supp}(\mu) + B(0,\varepsilon)$ *to make sure that $W^{1,p}(\Omega\setminus F_\varepsilon)$ is a Sobolev space on an open set.*

3.2 Global Estimates

3.2.1 Estimate on $\ln(1+|u|)$

The following proposition is a nonlinear form of a proposition in [31].

Proposition 3.2.1 *Under Hypotheses (3.1.2)—(3.1.7), if $\mu \in W^{-1,p'}(\Omega)$ then there exists C only depending on Λ such that, for any solution u of (3.1.8), we have $||\ln(1+|u|)||_{W_0^{1,p}(\Omega)} \leq C$.*

Proof of Proposition 3.2.1
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^p}$. Using $\varphi(u)$ as a test function in (3.1.8), we find

$$\begin{aligned} &\int_\Omega a(x,u,\nabla u)\cdot\frac{\nabla u}{(1+|u|)^p} \\ &\quad\leq \int_\Omega g(1+|u|^{p-1})\frac{|\nabla u|}{(1+|u|)^p} + \langle \mu,\varphi(u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} + \langle f,\varphi(u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}. \end{aligned} \tag{3.2.1}$$

But

$$a(x,u,\nabla u)\cdot\frac{\nabla u}{(1+|u|)^p} \geq \nu\frac{|\nabla u|^p}{(1+|u|)^p} - \frac{\Theta(x)}{(1+|u|)^p} \geq \nu|\nabla(\ln(1+|u|))|^p - \Theta(x). \tag{3.2.2}$$

Since $1+|u|^{p-1} \leq 2(1+|u|)^{p-1}$, we can write

$$\int_\Omega g(1+|u|^{p-1})\frac{|\nabla u|}{(1+|u|)^p} \leq 2\int_\Omega g\frac{|\nabla u|}{(1+|u|)} \leq 2||g||_{L^{p'}(\Omega)}||\,|\nabla(\ln(1+|u|))|\,||_{L^p(\Omega)} \tag{3.2.3}$$

(notice that $g \in L^{p'}(\Omega)$, since $p' \leq \frac{N_*}{p-1}$).
Since $|\varphi'(u)| \leq \frac{1}{(1+|u|)}$, we have $|\nabla(\varphi(u))| = |\varphi'(u)|\,|\nabla u| \leq \frac{|\nabla u|}{(1+|u|)} = |\nabla(\ln(1+|u|))|$. Thus,

$$||\varphi(u)||_{W_0^{1,p}(\Omega)} = ||\,|\nabla(\varphi(u))|\,||_{L^p(\Omega)} \leq ||\,|\nabla(\ln(1+|u|))|\,||_{L^p(\Omega)}$$

and, φ being bounded by $\frac{1}{p-1}$, we obtain

$$|\langle\mu,\varphi(u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}+\langle f,\varphi(u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}|\leq\frac{\Lambda}{p-1}+\Lambda||\,|\nabla(\ln(1+|u|))|\,||_{L^p(\Omega)}. \quad (3.2.4)$$

Injecting (3.2.2), (3.2.3) and (3.2.4) in (3.2.1), we obtain

$$\nu||\,|\nabla(\ln(1+|u|))|\,||_{L^p(\Omega)}^p\leq C+C||\,|\nabla(\ln(1+|u|))|\,||_{L^p(\Omega)},$$

with C only depending on Λ, which concludes the proof ("$X^p\leq C+CX$" implies that X has to be bounded, since $p>1$). ■

3.2.2 Proof of the global estimates

In the proof of the next proposition, we mix the ideas of [31] (or [32]) — to handle the noncoercive characteristic of the equation — with the ideas of [14] — to handle the measure on the right-hand side of the equation.

Proposition 3.2.2 *Let $\alpha>1$. Under Hypotheses (3.1.2)—(3.1.7), if $\mu\in W^{-1,p'}(\Omega)$ then there exists $C>0$ only depending on (Λ,α) such that, for any solution u of (3.1.8),*

$$\int_\Omega\frac{|\nabla u|^p}{(1+|u|)^\alpha}\leq C.$$

Proof of Proposition 3.2.2
We define $\varphi(s)=\int_0^s\frac{dt}{(1+|t|)^\alpha}$, $T_k(s)=\max(-k,\min(s,k))$ and $S_k(s)=s-T_k(s)$.
Step 1: estimate on $S_k(u)$.
Let $k>0$. Using $\varphi(S_k(u))$ as a test function in (3.1.8), we get, since φ is bounded by $\frac{1}{\alpha-1}$,

$$\begin{aligned}&\int_\Omega a(x,u,\nabla u)\cdot\frac{\nabla(S_k(u))}{(1+|S_k(u)|)^\alpha}\\&\quad\leq\int_\Omega g(1+|u|^{p-1})\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^\alpha}+\frac{\Lambda}{\alpha-1}+\Lambda||\,|\nabla(\varphi(S_k(u)))|\,||_{L^p(\Omega)}.\end{aligned} \quad (3.2.5)$$

Since $\nabla(S_k(u))=\mathbf{1}_{E_k}\nabla u$, where $E_k=\{|u|>k\}$ and $\mathbf{1}_{E_k}$ is the characteristic function of E_k, we have

$$a(x,u,\nabla u)\cdot\frac{\nabla(S_k(u))}{(1+|S_k(u)|)^\alpha}\geq\nu\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}-\frac{\Theta}{(1+|S_k(u)|)^\alpha}\geq\nu\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}-\Theta. \quad (3.2.6)$$

Since $|\varphi'(s)|^p=\frac{1}{(1+|s|)^{\alpha p}}\leq\frac{1}{(1+|s|)^\alpha}$ for all $s\in\mathbb{R}$ (because $\alpha p\geq\alpha$), we can write

$$||\,|\nabla(\varphi(S_k(u)))|\,||_{L^p(\Omega)}\leq\left(\int_\Omega\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}\right)^{1/p}. \quad (3.2.7)$$

We have $|u|\leq k+|S_k(u)|$, which implies

$$\begin{aligned}&\int_\Omega g(1+|u|^{p-1})\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^\alpha}\\&\quad\leq\int_\Omega g(1+2^{p-1}k^{p-1}+2^{p-1}|S_k(u)|^{p-1})\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^\alpha}\\&\quad\leq(1+2^{p-1}k^{p-1})||g||_{L^{p'}(\Omega)}\left(\int_\Omega\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^{\alpha p}}\right)^{1/p}+2^{p-1}\int_\Omega g\frac{|S_k(u)|^{p-1}}{(1+|S_k(u)|)^{\frac{\alpha}{p'}}}\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^{\frac{\alpha}{p}}}\\&\quad\leq C_1(1+k^{p-1})\left(\int_\Omega\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}\right)^{1/p}+C_1\left(\int_\Omega g^{p'}\frac{|S_k(u)|^p}{(1+|S_k(u)|)^\alpha}\right)^{1/p'}\left(\int_\Omega\frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}\right)^{1/p}\end{aligned}$$

(we have used $(1+|S_k(u)|)^{\alpha p} \geq (1+|S_k(u)|)^{\alpha}$).
Denoting

$$A_k = \int_\Omega \frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha} \quad \text{and} \quad \psi(s) = \frac{|s|}{(1+|s|)^{\frac{\alpha}{p}}},$$

we have just proved that

$$\int_\Omega g(1+|u|^{p-1})\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^\alpha} \leq C_1(1+k^{p-1})A_k^{1/p} + C_1 A_k^{1/p}\left(\int_\Omega g^{p'}\psi(S_k(u))^p\right)^{1/p'}. \tag{3.2.8}$$

Let $\delta > 0$ (fixed later on) and write $g = g_1 + g_2$ with $g_1 \in L^\infty(\Omega)$ and $g_2 \in L^{\frac{N_*}{p-1}}(\Omega)$ such that $||g_2||_{L^{\frac{N_*}{p-1}}(\Omega)} \leq \delta$ (the choice of (g_1, g_2) only depend on δ). Thanks to Hölder's inequality with exponents $(\frac{N_*}{p}, \frac{N_*}{N_*-p})$ and to the Sobolev injections, since $\psi(S_k(u)) = 0$ outside E_k, we have

$$\begin{aligned}
\int_\Omega g^{p'}\psi(S_k(u))^p &\leq ||g||^{p'}_{L^{\frac{N_*}{p-1}}(E_k)} ||\psi(S_k(u))||^p_{L^{\frac{N_* p}{N_*-p}}(\Omega)} \\
&\leq C_2\left(||g_1||^{p'}_{L^{\frac{N_*}{p-1}}(E_k)} + ||g_2||^{p'}_{L^{\frac{N_*}{p-1}}(E_k)}\right) || \, |\nabla(\psi(S_k(u)))| \, ||^p_{L^p(\Omega)} \\
&\leq C_2\left(||g_1||^{p'}_{L^\infty(\Omega)}\mathrm{meas}(E_k)^{\frac{p}{N_*}} + \delta^{p'}\right) || \, |\nabla(\psi(S_k(u)))| \, ||^p_{L^p(\Omega)}.
\end{aligned}$$

Moreover, $|\psi'(s)| \leq \frac{1+\frac{\alpha}{p}}{(1+|s|)^{\frac{\alpha}{p}}}$ for all $s \in \mathbb{R}$, so that $|\nabla(\psi(S_k(u)))|^p \leq (1+\frac{\alpha}{p})^p \frac{|\nabla(S_k(u))|^p}{(1+|S_k(u)|)^\alpha}$ and

$$\int_\Omega g^{p'}\psi(S_k(u))^p \leq C_3\left(||g_1||^{p'}_{L^\infty(\Omega)}\mathrm{meas}(E_k)^{\frac{p}{N_*}} + \delta^{p'}\right) A_k$$

with C_3 only depending on α. Used in (3.2.8), this inequality allows us to write

$$\int_\Omega g(1+|u|^{p-1})\frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^\alpha} \leq C_4(1+k^{p-1})A_k^{1/p} + C_4\left(||g_1||^{p'}_{L^\infty(\Omega)}\mathrm{meas}(E_k)^{\frac{p}{N_*}} + \delta^{p'}\right)^{1/p'} A_k \tag{3.2.9}$$

where C_4 only depends on α.
(3.2.5), (3.2.6), (3.2.7) and (3.2.9) give

$$\begin{aligned}
A_k &\leq C_5 + C_5(1+k^{p-1})A_k^{1/p} + C_5\left(||g_1||^{p'}_{L^\infty(\Omega)}\mathrm{meas}(E_k)^{\frac{p}{N_*}} + \delta^{p'}\right)^{1/p'} A_k \\
&\leq C_5 + C_5(1+k^{p-1})A_k^{1/p} + C_5\left(||g_1||_{L^\infty(\Omega)}\mathrm{meas}(E_k)^{\frac{p-1}{N_*}} + \delta\right) A_k
\end{aligned} \tag{3.2.10}$$

where C_5 only depends on (Λ, α) (we have used the fact that, for $(s,t) \in \mathbb{R}^+$, $(s+t)^{1/p'} \leq s^{1/p'} + t^{1/p'}$). Set $\delta = 1/(4C_5)$, which only depends on (Λ, α). By Proposition 3.2.1, Tchebychev's inequality and Poincaré's inequality, we have $\mathrm{meas}(E_k) \leq \frac{||\ln(1+|u|)||^p_{L^p(\Omega)}}{(\ln(1+k))^p} \leq \frac{C_6}{(\ln(1+k))^p}$ with C_6 only depending on Λ. There exists thus k_0 only depending on (Λ, α), such that $C_5||g_1||_{L^\infty(\Omega)}\mathrm{meas}(E_{k_0})^{\frac{p-1}{N_*}} \leq \frac{1}{4}$.
With these choices of δ and k_0, (3.2.10) becomes $A_{k_0} \leq C_5 + C_5(1+k_0^{p-1})A_{k_0}^{1/p} + \frac{1}{2}A_{k_0}$, which leads to $A_{k_0} \leq 2C_5 + 2C_5(1+k_0^{p-1})A_{k_0}^{1/p}$. By Young's inequality, we obtain thus C_7 only depending on (Λ, α), such that $A_{k_0} \leq C_7$, that is to say

$$\int_\Omega \frac{|\nabla(S_{k_0}(u))|^p}{(1+|S_{k_0}(u)|)^\alpha} \leq C_7. \tag{3.2.11}$$

Step 2: conclusion.

By Proposition 3.2.1, we have $||\ln(1+|u|)||^p_{W_0^{1,p}(\Omega)} \leq C_8$ with C_8 only depending on Λ, thus

$$\int_\Omega |\nabla(T_{k_0}(u))|^p = \int_{\{|u|\leq k_0\}} |\nabla u|^p \leq (1+k_0)^p \int_{\{|u|\leq k_0\}} \frac{|\nabla u|^p}{(1+|u|)^p} \leq (1+k_0)^p C_8. \tag{3.2.12}$$

Since $u = T_{k_0}(u) + S_{k_0}(u)$ and $1+|S_{k_0}(u)| \leq 1+|u|$, (3.2.11) and (3.2.12) show that

$$\begin{aligned}\int_\Omega \frac{|\nabla u|^p}{(1+|u|)^\alpha} &\leq 2^p \int_\Omega \frac{|\nabla(T_{k_0}(u))|^p}{(1+|u|)^\alpha} + 2^p \int_\Omega \frac{|\nabla(S_{k_0}(u))|^p}{(1+|u|)^\alpha} \\ &\leq 2^p \int_\Omega |\nabla(T_{k_0}(u))|^p + 2^p \int_\Omega \frac{|\nabla(S_{k_0}(u))|^p}{(1+|S_{k_0}(u)|)^\alpha} \\ &\leq 2^p (1+k_0)^p C_8 + 2^p C_7,\end{aligned}$$

which concludes the proof. ■

We can now prove the global estimates theorem.
Proof of Theorem 3.1.1
Since $r < \frac{N_*(p-1)}{N_*-p}$ and $s < \frac{N_*(p-1)}{N_*-1}$, there exists $\alpha \in]1,p[$ only depending on (r,s) such that $r < \frac{N_*(p-\alpha)}{N_*-p}$ and $s < \frac{N_*(p-\alpha)}{N_*-\alpha}$.
Let C_1 only depending on (Λ,α) (i.e. on (Λ,r,s)) given by Proposition 3.2.2. We have, for all $k \geq 1$,

$$\int_\Omega |\nabla(T_k(u))|^p = \int_{\{|u|\leq k\}} |\nabla u|^p \leq (1+k)^\alpha \int_\Omega \frac{|\nabla u|^p}{(1+|u|)^\alpha} \leq C_1(1+k)^\alpha \leq 2^\alpha C_1 k^\alpha. \tag{3.2.13}$$

Since $0 < r < \frac{N_*(p-\alpha)}{N_*-p}$ and $0 < s < \frac{N_*(p-\alpha)}{N_*-\alpha}$, the proof is then an easy consequence of Lemmas 3.5.2 and 3.5.1. ■

3.3 Local Estimates

3.3.1 Preliminary results

Proposition 3.3.1 *Suppose Hypotheses (3.1.2)—(3.1.7) and $\mu \in W^{-1,p'}(\Omega)$. Let $\theta \in C^\infty(\mathbb{R}^N;\mathbb{R}^+)$ be such that*

$$\begin{gathered}\forall m > 0\,,\ \theta^m \in C^\infty(\mathbb{R}^N;\mathbb{R}^+)\,, \\ \forall m \in]0,1[\,,\ \exists Q_m \text{ such that } |\nabla\theta| \leq Q_m \theta^m\,, \\ \theta = 0 \text{ on a neighborhood of } K.\end{gathered} \tag{3.3.1}$$

Then there exists C only depending on (Λ,θ) such that, if u is a solution to (3.1.8), we have, for all $\alpha \in]0,1]$ and all $k \geq 0$,

$$\int_\Omega |\nabla(T_k(\theta u))|^p \leq C(1+k^\alpha) + Ck^\alpha \int_{\Omega\cap \mathrm{supp}(\theta)} \left(|u|^{\overline{p}-\alpha} + |\nabla u|^{p-\alpha}\right).$$

Remark 3.3.1 *Functions satisfying (3.3.1) exist and will be constructed in the proof of Proposition 3.3.2.*

Proof of Proposition 3.3.1

θ and θ^{p-1} being regular functions, we can take $\theta^{p-1}T_k(\theta u)$ as a test function in (3.1.8); this gives, since $\theta = 0$ on a neighborhood of $\mathrm{supp}(\mu)$,

$$\begin{aligned}
&\langle f, \theta^{p-1}T_k(\theta u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} \\
&= \int_\Omega a(x,u,\nabla u)\cdot\nabla(\theta^{p-1}T_k(\theta u)) + \int_\Omega \Phi(x,u)\cdot\nabla(\theta^{p-1}T_k(\theta u)) \\
&= \int_\Omega \theta^{p-1}a(x,u,\nabla u)\cdot\nabla(T_k(\theta u)) + \int_\Omega T_k(\theta u)a(x,u,\nabla u)\cdot\nabla(\theta^{p-1}) \\
&\quad + \int_\Omega \theta^{p-1}\Phi(x,u)\cdot\nabla(T_k(\theta u)) + \int_\Omega T_k(\theta u)\Phi(x,u)\cdot\nabla(\theta^{p-1}) \\
&= \int_\Omega \theta^{p}\mathbf{1}_{\{|\theta u|\le k\}}a(x,u,\nabla u)\cdot\nabla u + \int_\Omega \theta^{p-1}u\mathbf{1}_{\{|\theta u|\le k\}}a(x,u,\nabla u)\cdot\nabla\theta \\
&\quad + \int_\Omega T_k(\theta u)a(x,u,\nabla u)\cdot\nabla(\theta^{p-1}) + \int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\theta^{p}\Phi(x,u)\cdot\nabla u \\
&\quad + \int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\theta^{p-1}u\Phi(x,u)\cdot\nabla\theta + \int_\Omega T_k(\theta u)\Phi(x,u)\cdot\nabla(\theta^{p-1}). \qquad (3.3.2)
\end{aligned}$$

We have $\theta^{p-1}|\nabla\theta| \le C_1\theta$ on Ω, with C_1 only depending on θ (indeed, if $p \ge 2$, we just use the fact that θ^{p-2} and $\nabla\theta$ are bounded on Ω and, if $p \in]1,2[$, we use (3.3.1) to get $|\nabla\theta| \le Q_{2-p}\theta^{2-p}$); moreover, $\nabla(\theta^{p-1})$ is bounded on Ω (say by C_2). Thus, thanks to Hypotheses (3.1.3) and (3.1.4), we have

$$\begin{aligned}
&\int_\Omega \theta^{p}\mathbf{1}_{\{|\theta u|\le k\}}a(x,u,\nabla u)\cdot\nabla u + \int_\Omega \theta^{p-1}u\mathbf{1}_{\{|\theta u|\le k\}}a(x,u,\nabla u)\cdot\nabla\theta \\
&+ \int_\Omega T_k(\theta u)a(x,u,\nabla u)\cdot\nabla(\theta^{p-1}) \\
&\quad\ge \nu\int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\theta^{p}|\nabla u|^{p} - \int_\Omega \theta^{p}\Theta - \int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}\right)\theta^{p-1}|\nabla\theta|\,|u| \\
&\qquad - \int_\Omega \left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}\right)|\nabla(\theta^{p-1})|\,|T_k(\theta u)| \\
&\quad\ge \nu\int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\theta^{p}|\nabla u|^{p} - \int_\Omega \theta^{p}\Theta - C_1\int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}\right)|\theta u| \\
&\qquad - C_2\int_\Omega \left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}\right)|T_k(\theta u)| \\
&\quad\ge \nu\int_\Omega \mathbf{1}_{\{|\theta u|\le k\}}\theta^{p}|\nabla u|^{p} - \int_\Omega \theta^{p}\Theta - (C_1+C_2)\int_\Omega \left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}\right)|T_k(\theta u)|. \qquad (3.3.3)
\end{aligned}$$

By Hypothesis (3.1.6), we also have, with the same (C_1, C_2) as before and using the fact that $st \leq s^r + t^{r'}$ if $1 < r < \infty$ and $(s,t) \in \mathbb{R}^+$ (simplified Young's inequality),

$$\begin{aligned}
&\int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^p \Phi(x,u) \cdot \nabla u + \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^{p-1} u \Phi(x,u) \cdot \nabla \theta + \int_\Omega T_k(\theta u) \Phi(x,u) \cdot \nabla(\theta^{p-1}) \\
&\geq -\int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^p g(1+|u|^{p-1}) |\nabla u| - \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^{p-1} |\nabla \theta| \, |u| g(1+|u|^{p-1}) \\
&\quad -C_2 \int_\Omega g(1+|u|^{p-1}) |T_k(\theta u)| \\
&\geq -\int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} |\theta \nabla u| \theta^{p-1} g(1+|u|^{p-1}) - C_1 \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} |\theta u| g(1+|u|^{p-1}) \\
&\quad -C_2 \int_\Omega g(1+|u|^{p-1}) |T_k(\theta u)| \\
&\geq -\frac{\nu}{2} \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} |\theta \nabla u|^p - \left(\frac{2}{\nu}\right)^{p'/p} \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^p g^{p'} (1+|u|^{p-1})^{p'} \\
&\quad -(C_1 + C_2) \int_\Omega |T_k(\theta u)| g(1+|u|^{p-1}) \\
&\geq -\frac{\nu}{2} \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^p |\nabla u|^p - 2^{p'} \left(\frac{2}{\nu}\right)^{p'/p} \int_\Omega \mathbf{1}_{\{|\theta u| \leq k\}} \theta^p g^{p'} (1+|u|^p) \\
&\quad -(C_1 + C_2) \int_\Omega |T_k(\theta u)| g(1+|u|^{p-1}). \qquad (3.3.4)
\end{aligned}$$

Gathering (3.3.2), (3.3.3) and (3.3.4), we obtain C_3 only depending on θ such that

$$\begin{aligned}
&\langle f, \theta^{p-1} T_k(\theta u) \rangle_{W^{-1,p'}(\Omega), W_0^{1,p}(\Omega)} \\
&\geq \frac{\nu}{2} \int_{\{|\theta u| \leq k\}} \theta^p |\nabla u|^p - C_3 \int_\Omega \left(h + |u|^{\overline{p}-1} + |\nabla u|^{p-1} + g + g|u|^{p-1}\right) |T_k(\theta u)| \\
&\quad -C_3 \int_{\{|\theta u| \leq k\}} (\theta^p g^{p'} + |\theta u|^p g^{p'}) - C_3.
\end{aligned}$$

This inequality allows us to write

$$\begin{aligned}
\int_\Omega |\nabla(T_k(\theta u))|^p &= \int_{\{|\theta u| \leq k\}} |\nabla(\theta u)|^p \\
&\leq 2^p \int_{\{|\theta u| \leq k\}} |\nabla \theta|^p |u|^p + 2^p \int_{\{|\theta u| \leq k\}} \theta^p |\nabla u|^p \\
&\leq C_4 \int_{\{|\theta u| \leq k\}} |\nabla \theta|^p |u|^p + C_4 \langle f, \theta^{p-1} T_k(\theta u) \rangle_{W^{-1,p'}(\Omega), W_0^{1,p}(\Omega)} \\
&\quad +C_4 \int_\Omega \left(h + g + g|u|^{p-1} + |u|^{\overline{p}-1} + |\nabla u|^{p-1}\right) |T_k(\theta u)| \\
&\quad +C_4 \int_\Omega g^{p'} |T_k(\theta u)|^p + C_4, \qquad (3.3.5)
\end{aligned}$$

where C_4 only depends on θ (notice that $\theta^p g^{p'} \in L^1(\Omega)$ since $p' \leq \frac{N_*}{p-1}$ and θ is bounded on Ω).

Take $\delta > 0$ (to be fixed later on); there exists $g_1 \in L^\infty(\Omega)$ and $g_2 \in L^{\frac{N_*}{p-1}}(\Omega)$, only depending on δ, such that $g = g_1 + g_2$ and $||g_2||_{L^{\frac{N_*}{p-1}}(\Omega)} \leq \delta$.

We have $|\nabla\theta|^p \leq C_5\theta$ on Ω (this is (3.3.1) with $m = 1/p$), where C_5 only depends on θ. Thus, (3.3.5) gives C_6 only depending on θ such that

$$\begin{aligned}
\int_\Omega |\nabla(T_k(\theta u))|^p &\leq C_6 \int_{\{|\theta u|\leq k\}} |\theta u||u|^{p-1} + C_6\langle f, \theta^{p-1}T_k(\theta u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} \\
&\quad + C_6 \int_{\Omega\cap \mathrm{supp}(\theta)} \left(h + g + g|u|^{p-1} + |u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)| \\
&\quad + C_6||g_1||^{p'}_{L^\infty(\Omega)} \int_\Omega |T_k(\theta u)||\theta u|^{p-1} + C_6\int_\Omega g_2^{p'}|T_k(\theta u)|^p + C_6 \\
&\leq C_6\langle f, \theta^{p-1}T_k(\theta u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} \\
&\quad + C_6 \int_{\Omega\cap \mathrm{supp}(\theta)} \left(h + g + (1+g)|u|^{p-1} + |u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)| \\
&\quad + C_6||g_1||^{p'}_{L^\infty(\Omega)}||\theta||^{p-1}_{L^\infty(\Omega)} \int_\Omega |u|^{p-1}|T_k(\theta u)| + C_6\int_\Omega g_2^{p'}|T_k(\theta u)|^p + C_6. \quad (3.3.6)
\end{aligned}$$

But, θ^{p-1} being regular, by the Hölder inequality, the Sobolev injection and the Poincaré inequality, we have

$$\begin{aligned}
&C_6\langle f, \theta^{p-1}T_k(\theta u)\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)} + C_6\int_{\Omega\cap\mathrm{supp}(\theta)} (h+g)|T_k(\theta u)| + C_6\int_\Omega g_2^{p'}|T_k(\theta u)|^p \\
&\leq C_6\Lambda||\theta^{p-1}T_k(\theta u)||_{W_0^{1,p}(\Omega)} + C_6||h+g||_{L^{p'}(\Omega)}||T_k(\theta u)||_{L^p(\Omega)} + C_6||g_2||^{\frac{p}{p-1}}_{L^{\frac{N_*}{p-1}}(\Omega)}||T_k(\theta u)||^p_{L^{\frac{N_*p}{N_*-p}}(\Omega)} \\
&\leq (C_7\Lambda + C_7||h+g||_{L^{p'}(\Omega)})||\,|\nabla(T_k(\theta u))|\,||_{L^p(\Omega)} + C_7\delta^{p'}||\,|\nabla(T_k(\theta u))|\,||^p_{L^p(\Omega)} \\
&\leq \frac{1}{p'}\left(C_7\Lambda + C_7||h+g||_{L^{p'}(\Omega)}\right)^{p'} + \left(\frac{1}{p} + C_7\delta^{p'}\right)||\,|\nabla(T_k(\theta u))|\,||^p_{L^p(\Omega)},
\end{aligned}$$

where C_7 only depends on θ.
Fix now $\delta > 0$ such that $\frac{1}{p} + C_7\delta^{p'} < 1$ (such a choice of δ only depends on θ). Returning to (3.3.6), we find C_8 only depending on (Λ, θ) such that

$$\int_\Omega |\nabla(T_k(\theta u))|^p \leq C_8 + C_8\int_{\Omega\cap\mathrm{supp}(\theta)} \left((1+g)|u|^{p-1} + |u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)|. \quad (3.3.7)$$

Let $\alpha \in]0,1]$. We have, if $\alpha < 1$,

$$|T_k(\theta u)| = |T_k(\theta u)|^\alpha |T_k(\theta u)|^{1-\alpha} \leq k^\alpha\theta^{1-\alpha}|u|^{1-\alpha} \leq (1 + ||\theta||_{L^\infty(\Omega)})k^\alpha|u|^{1-\alpha}, \quad (3.3.8)$$

so that

$$\left(|u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)| \leq (1 + ||\theta||_{L^\infty(\Omega)})k^\alpha\left(|u|^{\overline{p}-\alpha} + |\nabla u|^{p-1}|u|^{1-\alpha}\right).$$

Using the simplified Young's inequality with $\frac{p-\alpha}{p-1} > 1$, we find

$$|\nabla u|^{p-1}|u|^{1-\alpha} \leq |\nabla u|^{p-\alpha} + |u|^{p-\alpha},$$

which gives

$$\left(|u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)| \leq (1 + ||\theta||_{L^\infty(\Omega)})k^\alpha\left(|u|^{\overline{p}-\alpha} + |u|^{p-\alpha} + |\nabla u|^{p-\alpha}\right).$$

Notice that $\frac{N_*}{N_*-p+1} > 1$ (because $p > 1$), so that $\overline{p} > p$. In particular, $|u|^{p-\alpha} \leq 1 + |u|^{\overline{p}-\alpha}$ and

$$\left(|u|^{\overline{p}-1} + |\nabla u|^{p-1}\right)|T_k(\theta u)| \leq (1 + ||\theta||_{L^\infty(\Omega)})k^\alpha\left(1 + 2|u|^{\overline{p}-\alpha} + |\nabla u|^{p-\alpha}\right). \quad (3.3.9)$$

This inequality is still valid if $\alpha = 1$ (we simply bound $T_k(\theta u)$ by k).
Thanks again to (3.3.8) if $\alpha < 1$, or bounding $T_k(\theta u)$ by k if $\alpha = 1$, we can also write

$$(1+g)|u|^{p-1}|T_k(\theta u)| \leq (1+||\theta||_{L^\infty(\Omega)})k^\alpha(1+g)|u|^{p-\alpha}$$

which gives, by the simplified Young's inequality with $\frac{\overline{p}-\alpha}{p-\alpha} > 1$,

$$(1+g)|u|^{p-1}|T_k(\theta u)| \quad \leq \quad (1+||\theta||_{L^\infty(\Omega)})k^\alpha\left((1+g)^{\frac{\overline{p}-\alpha}{\overline{p}-p}} + |u|^{\overline{p}-\alpha}\right). \tag{3.3.10}$$

We have $\overline{p} \geq \frac{N_* p}{N_*-p+1}$, which implies $\frac{\overline{p}-\alpha}{\overline{p}-p} \leq \frac{\overline{p}}{\overline{p}-p} = \frac{1}{1-p/\overline{p}} \leq \frac{1}{1-(N_*-p+1)/N_*} = \frac{N_*}{p-1}$ and $(1+g)^{\frac{\overline{p}-\alpha}{\overline{p}-p}} \leq (1+g)^{\frac{N_*}{p-1}} \in L^1(\Omega)$. Using this last inequality in (3.3.10) and injecting the result in (3.3.7) along with (3.3.9), we obtain

$$\int_\Omega |\nabla(T_k(\theta u))|^p \leq C_8 + C_9 k^\alpha + C_9 k^\alpha \int_{\Omega\cap\mathrm{supp}(\theta)} \left(|u|^{\overline{p}-\alpha} + |\nabla u|^{p-\alpha}\right),$$

with C_9 only depending on (Λ, θ), which concludes the proof. ■

Corollary 3.3.1 *Let $M \geq 0$ and $\alpha \in]0,1]$. Under Hypotheses (3.1.2)—(3.1.7), if $\mu \in W^{-1,p'}(\Omega)$, θ satisfies (3.3.1) and u is a solution to (3.1.8) which satisfies*

$$\int_{\Omega\cap\mathrm{supp}(\theta)} \left(|u|^{\overline{p}-\alpha} + |\nabla u|^{p-\alpha}\right) \leq M,$$

then, for all $0 < r < \frac{N_*(p-\alpha)}{N_*-p}$ and all $0 < s < \frac{N_*(p-\alpha)}{N_*-\alpha}$, there exists C only depending on $(\Lambda, \theta, M, \alpha, r, s)$ such that

$$\int_\Omega \left(|\theta u|^r + |\nabla(\theta u)|^s\right) \leq C.$$

Proof of Corollary 3.3.1.
By Proposition 3.3.1, we have C_1 only depending on (Λ, θ) such that, for all $k \geq 1$,

$$\int_\Omega |\nabla(T_k(\theta u))|^p \leq C_1 + C_1 k^\alpha + C_1 M k^\alpha \leq (2C_1 + C_1 M)k^\alpha.$$

The corollary is then an easy consequence of Lemmas 3.5.2 and 3.5.1. ■

3.3.2 Proof of the local estimates

We first prove, thanks to a bootstrap technique based on Corollary 3.3.1, local estimates in $W^{1,q}$ for all $q < p$ and then, using these estimates for q close enough to p, we deduce Theorem 3.1.2.

Proposition 3.3.2 *Suppose Hypotheses (3.1.2)—(3.1.7). If $\varepsilon > 0$, $F = K + \overline{B}(0,\varepsilon)$ and $1 \leq q < p$, then there exists C only depending on $(\Lambda, K, \varepsilon, q)$ such that, if u is a solution to (3.1.8),*

$$||u||_{L^{\frac{N_* q}{N_*-q}}(\Omega\setminus F)} + ||u||_{W^{1,q}(\Omega\setminus F)} \leq C.$$

Proof of Proposition 3.3.2
The proof is based on an induction reasoning, which uses the following sequence: $\alpha_1 = 1$ and, for $n \geq 2$,

$$\alpha_n = \max\left(\overline{p} - \frac{N_*(p-\alpha_{n-1})}{N_*-p}; \frac{(N_*-p)\alpha_{n-1}}{N_*-\alpha_{n-1}}\right). \tag{3.3.11}$$

Step 1: study of $(\alpha_n)_{n\geq 1}$.
Let us prove by induction that $(\alpha_n)_{n\geq 1}$ is a decreasing sequence of numbers in $]0,1]$. Indeed, suppose that, for $n \geq 2$, $\alpha_{n-1} \in]0,1]$; then $\frac{N_*-p}{N_*-\alpha_{n-1}} \in]0,1[$ (because $p > 1 \geq \alpha_{n-1}$), so that $\frac{(N_*-p)\alpha_{n-1}}{N_*-\alpha_{n-1}} \in]0,\alpha_{n-1}[$; moreover, by definition of $\overline{p}$,

$$\overline{p} - \frac{N_*(p-\alpha_{n-1})}{N_*-p} < \frac{(N_*-1)p}{N_*-p} - \frac{N_*(p-\alpha_{n-1})}{N_*-p} = \frac{N_*\alpha_{n-1}-p}{N_*-p} = \alpha_{n-1} - (1-\alpha_{n-1})\frac{p}{N_*-p}, \quad (3.3.12)$$

and this last quantity belongs to $]-\infty,\alpha_{n-1}]$ (because $1-\alpha_{n-1} \geq 0$); thus, α_n belongs to $]0,\alpha_{n-1}[$.
Denote $\alpha_\infty \in [0,\alpha_1[$ the limit of the decreasing sequence $(\alpha_n)_{n\geq 1}$. By passing to the limit $n \to \infty$ in (3.3.11) (the right-hand side of this equality is a continuous function of α_{n-1} on $[0,1]$), we obtain

$$\alpha_\infty = \max\left(\overline{p} - \frac{N_*(p-\alpha_\infty)}{N_*-p}; \frac{(N_*-p)\alpha_\infty}{N_*-\alpha_\infty}\right).$$

This maximum cannot be $\overline{p} - \frac{N_*(p-\alpha_\infty)}{N_*-p}$, because this would lead (thanks to the computations of (3.3.12) applied to α_∞ instead of α_{n-1}) to $\alpha_\infty < \alpha_\infty - (1-\alpha_\infty)\frac{p}{N_*-p}$, which is impossible since $\alpha_\infty < \alpha_1 = 1$. Thus, we have $\alpha_\infty = \frac{N_*-p}{N_*-\alpha_\infty}\alpha_\infty$ and, since $\frac{N_*-p}{N_*-\alpha_\infty} \in]0,1[$, this shows that $\alpha_\infty = 0$. Thus, the sequence $(\alpha_n)_{n\geq 1}$ tends to 0 as $n \to \infty$.

We conclude this step by proving that, for all $n \geq 2$,

$$\overline{p} - \alpha_n \leq \frac{N_*(p-\alpha_{n-1})}{N_*-p} \quad \text{and} \quad p - \alpha_n \leq \frac{N_*(p-\alpha_{n-1})}{N_*-\alpha_{n-1}}. \quad (3.3.13)$$

The first inequality of (3.3.13) is an immediate consequence of the definition of α_n. To obtain the second inequality, we write $\alpha_n \geq \frac{(N_*-p)\alpha_{n-1}}{N_*-\alpha_{n-1}}$, which implies $p - \alpha_n \leq \frac{N_*p-\alpha_{n-1}p-N_*\alpha_{n-1}+p\alpha_{n-1}}{N_*-\alpha_{n-1}} = \frac{N_*(p-\alpha_{n-1})}{N_*-\alpha_{n-1}}$.

Step 2: a set of functions satisfying (3.3.1).
Let $\varphi : \mathbb{R} \to \mathbb{R}^+$ be defined by $\varphi(s) = \exp(-\frac{1}{1-s})$ if $s < 1$ and $\varphi(s) = 0$ if $s \geq 1$; for all $m > 0$, $\varphi^m \in C^\infty(\mathbb{R};\mathbb{R}^+)$ and, for all $m \in]0,1[$, there exists Q_m such that $|\varphi'| \leq Q_m\varphi^m$ on $\mathbb{R}^+$ (indeed, this inequality is satisfied on $[1,\infty[$ where $\varphi' = 0$ and, on $[0,1[$, we have $|\varphi'|\varphi^{-m} = \frac{1}{(1-s)^2}\exp(-\frac{1-m}{1-s})$, which is bounded since $1-m > 0$).
If $e \in \mathbb{R}^N$ and $\eta > 0$, we define $\theta_{e,\eta}(x) = \varphi\left(\frac{|x-e|^2}{\eta^2}\right)$. For all $m > 0$, $\theta_{e,\eta}^m = \varphi^m\left(\frac{|\cdot-e|^2}{\eta^2}\right)$ is in $C^\infty(\mathbb{R}^N;\mathbb{R}^+)$ by composition. Moreover, if $m \in]0,1[$, we have, on $B(e,\eta)$,

$$|\nabla\theta_{e,\eta}| = \frac{2|\cdot-e|}{\eta^2}|\varphi'|\left(\frac{|\cdot-e|^2}{\eta^2}\right) \leq \frac{2}{\eta}Q_m\varphi^m\left(\frac{|\cdot-e|^2}{\eta^2}\right) = \frac{2Q_m}{\eta}\theta_{e,\eta}^m.$$

Since this inequality is also satisfied outside $B(e,\eta)$ (because $\nabla\theta_{e,\eta} = 0$ outside this ball), we deduce that $|\nabla\theta_{e,\eta}| \leq 2\eta^{-1}Q_m\theta_{e,\eta}^m$ on $\mathbb{R}^N$.
$\theta_{e,\eta}$ being null outside $B(e,\eta)$, we conclude that, if $\overline{B}(e,\eta) \cap K = \emptyset$, then $\theta_{e,\eta}$ satisfies (3.3.1).

Step 3: the bootstrap.
Let $e \in \Omega\backslash K$ and $0 < \eta < \mathrm{dist}(e,K)$. Define $\eta_n = \frac{\eta}{2} + \frac{\eta}{2^n}$.
In this step, we want to prove by induction that, for all $n \geq 1$, for all $0 < r < \frac{N_*(p-\alpha_n)}{N_*-p}$ and all $0 < s < \frac{N_*(p-\alpha_n)}{N_*-\alpha_n}$, there exists $C_{n,r,s}$ only depending on $(\Lambda, e, \eta, n, r, s)$ such that

$$\int_{\Omega\cap B(e,\eta_n)} (|u|^r + |\nabla u|^s) \leq C_{n,r,s}.$$

The case $n = 1$ is an immediate consequence of Theorem 3.1.1, since $\alpha_1 = 1$
Take now $n \geq 2$ and suppose that, for all $0 < r_0 < \frac{N_*(p-\alpha_{n-1})}{N_*-p}$ and all $0 < s_0 < \frac{N_*(p-\alpha_{n-1})}{N_*-\alpha_{n-1}}$, there exists C_{n-1,r_0,s_0} only depending on $(\Lambda, e, \eta, n-1, r_0, s_0)$ such that

$$\int_{\Omega\cap B(e,\eta_{n-1})} (|u|^{r_0} + |\nabla u|^{s_0}) \leq C_{n-1,r_0,s_0}.$$

Let $0 < r < \frac{N_*(p-\alpha_n)}{N_*-p}$ and $0 < s < \frac{N_*(p-\alpha_n)}{N_*-\alpha_n}$; we can find $\gamma \in]\alpha_n, 1]$, only depending on (n, r, s), such that $r < \frac{N_*(p-\gamma)}{N_*-p}$ and $s < \frac{N_*(p-\gamma)}{N_*-\gamma}$. Let $r_0 = \overline{p} - \gamma$ and $s_0 = p - \gamma$; by (3.3.13), since $\gamma > \alpha_n$, we have $0 < r_0 < \frac{N_*(p-\alpha_{n-1})}{N_*-p}$ and $0 < s_0 < \frac{N_*(p-\alpha_{n-1})}{N_*-\alpha_{n-1}}$, thus, for C_{n-1,r_0,s_0} as above,

$$\int_{\Omega\cap\overline{B}(e,\eta_{n-1})} \left(|u|^{\overline{p}-\gamma} + |\nabla u|^{p-\gamma}\right) = \int_{\Omega\cap B(e,\eta_{n-1})} \left(|u|^{\overline{p}-\gamma} + |\nabla u|^{p-\gamma}\right) \leq C_{n-1,r_0,s_0}$$

(notice that $\mathrm{meas}(\Omega \cap \partial(B(e,\eta_{n-1}))) = 0$). Since $r < \frac{N_*(p-\gamma)}{N_*-p}$ and $s < \frac{N_*(p-\gamma)}{N_*-\gamma}$, Corollary 3.3.1 applied to $\alpha = \gamma$ and $\theta = \theta_{e,\eta_{n-1}}$ (the support of which is included in $\overline{B}(e,\eta_{n-1})$) gives then C_1 only depending on $(\Lambda, \theta_{e,\eta_{n-1}}, C_{n-1,r_0,s_0}, \gamma, r, s)$, i.e. only depending on $(\Lambda, e, \eta, n, r, s)$, such that

$$\int_\Omega \left(|\theta_{e,\eta_{n-1}} u|^r + |\nabla(\theta_{e,\eta_{n-1}} u)|^s\right) \leq C_1.$$

Since $\theta_{e,\eta_{n-1}} \geq C_2 > 0$ on $B(e,\eta_n)$, where C_2 only depends on (η, n), we deduce that

$$\int_{\Omega\cap B(e,\eta_n)} |u|^r \leq \frac{C_1}{C_2^r} \tag{3.3.14}$$

and that

$$\begin{aligned}
\int_{\Omega\cap B(e,\eta_n)} |\nabla u|^s &\leq \frac{1}{C_2^s} \int_{\Omega\cap B(e,\eta_n)} |\theta_{e,\eta_{n-1}} \nabla u|^s \\
&\leq \frac{1}{C_2^s} \int_{\Omega\cap B(e,\eta_n)} 2^s |\nabla(\theta_{e,\eta_{n-1}} u)|^s + 2^s |u \nabla \theta_{e,\eta_{n-1}}|^s \\
&\leq \frac{2^s C_1}{C_2^s} + \frac{2^s ||\,|\nabla \theta_{e,\eta_{n-1}}|\,||^s_{L^\infty(\mathbb{R}^N)}}{C_2^s} \int_{\Omega\cap B(e,\eta_n)} |u|^s.
\end{aligned}$$

Since $\frac{N_*(p-\alpha_n)}{N_*-p} \geq \frac{N_*(p-\alpha_n)}{N_*-\alpha_n}$, we can always suppose that $r \geq s$, and we obtain thus

$$\begin{aligned}
\int_{\Omega\cap B(e,\eta_n)} |\nabla u|^s &\leq \frac{2^s C_1}{C_2^s} + \frac{2^s ||\,|\nabla \theta_{e,\eta_{n-1}}|\,||^s_{L^\infty(\mathbb{R}^N)}}{C_2^s} \int_{\Omega\cap B(e,\eta_n)} (1 + |u|^r) \\
&\leq \frac{2^s C_1}{C_2^s} + \frac{2^s ||\,|\nabla \theta_{e,\eta_{n-1}}|\,||^s_{L^\infty(\mathbb{R}^N)}}{C_2^s} \left(\mathrm{meas}(\Omega) + \frac{C_1}{C_2^r}\right).
\end{aligned} \tag{3.3.15}$$

(3.3.14) and (3.3.15) conclude the induction.

Step 4: conclusion.
Let $q \in [1, p[$; we have then $\frac{N_* q}{N_*-q} < \frac{N_* p}{N_*-p}$. Since $\alpha_n \to 0$, there exists $n_0 \geq 1$ only depending on q such that $\frac{N_* q}{N_*-q} < \frac{N_*(p-\alpha_{n_0})}{N_*-p}$ and $q < \frac{N_*(p-\alpha_{n_0})}{N_*-\alpha_{n_0}}$.
Take $e \in \Omega\backslash F$; we notice that $\varepsilon < \mathrm{dist}(e, K)$. By Step 3, there exists thus $C(e)$ only depending on $(\Lambda, e, \varepsilon, n_0, \frac{N_* q}{N_*-q}, q)$, i.e. on $(\Lambda, e, \varepsilon, q)$, such that

$$\int_{\Omega\cap B(e,\frac{\varepsilon}{2})} \left(|u|^{\frac{N_* q}{N_*-q}} + |\nabla u|^q\right) \leq \int_{\Omega\cap B(e,\frac{\varepsilon}{2}+\frac{\varepsilon}{2^{n_0}})} \left(|u|^{\frac{N_* q}{N_*-q}} + |\nabla u|^q\right) \leq C(e). \tag{3.3.16}$$

Since $\overline{\Omega\backslash F}$ is a compact subset of $\mathbb{R}^N$ which is covered by $(B(e,\varepsilon/2))_{e\in\Omega\backslash F}$ (the points of $\partial(\Omega\backslash F)$ are in the union of these balls because the radius is fixed), we can find $(e_1,\dots,e_l)\in\Omega\backslash F$ — only depending on (F,ε), i.e. on (K,ε) — such that $\Omega\backslash F\subset\cup_{i=1}^l B(e_i,\varepsilon/2)$. Writing (3.3.16) for $e_1,\dots,e_l$ and summing these inequalities, we find

$$\int_{\Omega\backslash F}\left(|u|^{\frac{N_*q}{N_*-q}}+|\nabla u|^q\right)\le\sum_{i=1}^l C(e_i),$$

which is the desired estimate. ■

We can now prove the local estimates theorem.

Proof of Theorem 3.1.2.

Let $\psi\in C_c^\infty(\mathbb{R}^N)$ be such that $\psi=1$ on a neighborhood of $K+\overline{B}(0,\frac{\varepsilon}{3})$ and $\psi=0$ outside $K+B(0,\frac{\varepsilon}{2})$ (the choice of ψ only depends on (K,ε)). We take $\theta=|1-\psi|^\zeta$, with $\zeta=\max(2,\frac{2}{p-1})$, and notice that θ, θ^{p-1} and θ^p are C^1 on $\mathbb{R}^N$ (they all can be written as $|1-\psi|^l$ with $l>1$, and, if $l>1$, $s\to|s|^l$ is C^1 on $\mathbb{R}$). Moreover, $\theta=0$ on a neighborhood of K. Thus, using $\theta^p u$ as a test function in (3.1.8), we obtain

$$\begin{aligned}
\Lambda||\theta^p u||_{W_0^{1,p}(\Omega)} &\ge \langle f,\theta^p u\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}\\
&= \int_\Omega a(x,u,\nabla u)\cdot\nabla(\theta^p u)+\int_\Omega\Phi(x,u)\cdot\nabla(\theta^p u)\\
&= \int_\Omega\theta^p a(x,u,\nabla u)\cdot\nabla u+\int_\Omega u a(x,u,\nabla u)\nabla(\theta^p)\\
&\quad+\int_\Omega\theta^p\Phi(x,u)\cdot\nabla u+\int_\Omega u\Phi(x,u)\cdot\nabla(\theta^p)\\
&\ge \nu\int_\Omega|\theta\nabla u|^p-\int_\Omega\theta^p\Theta-\int_{\Omega\backslash(K+\overline{B}(0,\frac{\varepsilon}{3}))}\theta^p g(1+|u|^{p-1})|\nabla u|\\
&\quad-\int_{\Omega\backslash(K+\overline{B}(0,\frac{\varepsilon}{3}))}|u|\left(h+\beta|u|^{\overline{p}-1}+\beta|\nabla u|^{p-1}+g+g|u|^{p-1}\right)|\nabla(\theta^p)| \qquad (3.3.17)
\end{aligned}$$

(notice that $\nabla(\theta^p)$ and θ are null on $K+\overline{B}(0,\frac{\varepsilon}{3})$).

Since θ^{p-1} is C^1, we have

$$||\theta^p u||_{W_0^{1,p}(\Omega)}=||\theta^{p-1}\theta u||_{W_0^{1,p}(\Omega)}\le C_1||\theta u||_{W_0^{1,p}(\Omega)}$$

where C_1 only depends on θ^{p-1}, i.e. on (K,ε).

Take $\delta>0$ and write $g=g_1+g_2$ with $g_1\in L^\infty(\Omega)$ and $g_2\in L^{\frac{N_*}{p-1}}(\Omega)$ such that $||g_2||_{L^{\frac{N_*}{p-1}}(\Omega)}\le\delta$ (the choice of such a decomposition only depends on δ). Since θ, g_1 and $\nabla(\theta^p)$ are bounded on Ω, we deduce from (3.3.17), by Young's inequality (and using the fact that $\overline{p}\ge p$, which implies $|u|^p\le 1+|u|^{\overline{p}}$),

$$\begin{aligned}
&C_1\Lambda||\theta u||_{W_0^{1,p}(\Omega)}\\
&\ge \nu\int_\Omega|\theta\nabla u|^p-C_2-\int_{\Omega\backslash(K+\overline{B}(0,\frac{\varepsilon}{3}))}g_2|\theta u|^{p-1}|\theta\nabla u|\\
&\quad-C_2\int_{\Omega\backslash(K+\overline{B}(0,\frac{\varepsilon}{3}))}\left(g|\nabla u|+|u|^{p-1}|\nabla u|+h^{p'}+g^{p'}+|u|^p+|u|^{\overline{p}}+|\nabla u|^{p-1}|u|+g|u|^p\right)\\
&\ge \frac{\nu}{2}\int_\Omega|\theta\nabla u|^p-C_3-C_3\int_\Omega g_2^{p'}|\theta u|^p\\
&\quad-C_3\int_{\Omega\backslash(K+\overline{B}(0,\frac{\varepsilon}{3}))}\left(g|\nabla u|+|u|^{p-1}|\nabla u|+|u|^{\overline{p}}+|\nabla u|^{p-1}|u|+g|u|^p\right) \qquad (3.3.18)
\end{aligned}$$

where C_2, C_3 only depend on (K,ε,δ) — recall that g_1 only depends on δ and that θ only depends on (K,ε).

Since, by the simplified Young's inequality with $\frac{N_*}{p-1} > 1$, we have

$$g|\nabla u| \leq g^{\frac{N_*}{p-1}} + |\nabla u|^{\frac{N_*}{N_*-p+1}} \quad \text{and} \quad g|u|^p \leq g^{\frac{N_*}{p-1}} + |u|^{\frac{N_*p}{N_*-p+1}},$$

we obtain, from (3.3.18), C_4 only depending on $(\Lambda, K, \varepsilon, \delta)$ such that

$$\begin{aligned}
&\int_\Omega |\nabla(\theta u)|^p \\
&\leq \quad 2^p \int_\Omega |\theta \nabla u|^p + 2^p \int_\Omega |u \nabla \theta|^p \\
&\leq \quad C_4 ||\theta u||_{W_0^{1,p}(\Omega)} + C_4 + C_4 \int_\Omega g_2^{p'} |\theta u|^p \\
&\quad + C_4 \int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} \left(|\nabla u|^{\frac{N_*}{N_*-p+1}} + |u|^{\frac{N_*p}{N_*-p+1}} + |u|^{\overline{p}} + |\nabla u|^{p-1}|u| + |u|^{p-1}|\nabla u| \right) \qquad (3.3.19)
\end{aligned}$$

(we have used the fact that $\nabla\theta$ is bounded on Ω and null on $K + \overline{B}(0, \frac{\varepsilon}{3})$, and we have bounded $|u|^p$ by $1 + |u|^{\overline{p}}$).
Let $q \in]\max(p-1, 1), p[$. By the simplified Young's inequality (with $\frac{q}{p-1} > 1$ and q), we have

$$|\nabla u|^{p-1}|u| \leq |\nabla u|^q + |u|^{\frac{q}{q-p+1}} \quad \text{and} \quad |u|^{p-1}|\nabla u| \leq |u|^{\frac{q(p-1)}{q-1}} + |\nabla u|^q. \qquad (3.3.20)$$

Moreover, by Hölder's inequality (with $\frac{N_*}{p}$) and the Sobolev injections,

$$\int_\Omega g_2^{p'} |\theta u|^p \leq ||g_2||^{p'}_{L^{\frac{N_*}{p-1}}(\Omega)} ||\theta u||^p_{L^{\frac{N_*p}{N_*-p}}(\Omega)} \leq C_5 \delta^{p'} ||\, |\nabla(\theta u)|\, ||^p_{L^p(\Omega)}. \qquad (3.3.21)$$

Setting $\delta > 0$ such that $1 - \frac{1}{p} - C_5\delta^{p'} > 0$ and writing $C_4||\theta u||_{W_0^{1,p}(\Omega)} \leq \frac{1}{p'} C_4^{p'} + \frac{1}{p}||\, |\nabla(\theta u)|\, ||^p_{L^p(\Omega)}$, (3.3.21) and (3.3.20) injected in (3.3.19) give us C_6 only depending on $(\Lambda, K, \varepsilon)$ such that, for all $q \in]\max(p-1, 1), p[$,

$$\begin{aligned}
&\int_\Omega |\nabla(\theta u)|^p \\
&\leq \quad C_6 + C_6 \int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} \left(|u|^{\overline{p}} + |u|^{\frac{N_*p}{N_*-p+1}} + |u|^{\frac{q}{q-p+1}} + |u|^{\frac{q(p-1)}{q-1}} + |\nabla u|^{\frac{N_*}{N_*-p+1}} + |\nabla u|^q \right) \qquad (3.3.22)
\end{aligned}$$

We notice that $\overline{p} < \frac{(N_*-1)p}{N_*-p} < \frac{N_*p}{N_*-p}$, that $\frac{N_*p}{N_*-p+1} < \frac{N_*p}{N_*-p}$ and that $\frac{N_*}{N_*-p+1} < p$ (this last inequality comes down to $(p - N_*)(p-1) < 0$, which is true since $p \in]1, N_*[$). Moreover, as $q \to p$, we have $\frac{q}{q-p+1} \to p < \frac{N_*p}{N_*-p}$, $\frac{q(p-1)}{q-1} \to p < \frac{N_*p}{N_*-p}$ and $\frac{N_*q}{N_*-q} \to \frac{N_*p}{N_*-p}$. We can thus find $q \in]\max(p-1, 1), p[$ such that $\overline{p} \leq \frac{N_*q}{N_*-q}$, $\frac{N_*p}{N_*-p+1} \leq \frac{N_*q}{N_*-q}$, $\frac{N_*}{N_*-p+1} \leq q$, $\frac{q}{q-p+1} \leq \frac{N_*q}{N_*-q}$ and $\frac{q(p-1)}{q-1} \leq \frac{N_*q}{N_*-q}$. Applying then Proposition 3.3.2 to this q (and with $\varepsilon/3$ instead of ε), we find C_7 only depending on $(\Lambda, K, \varepsilon)$ such that

$$\begin{aligned}
\int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} \left(|u|^{\overline{p}} + |u|^{\frac{N_*p}{N_*-p+1}} + |u|^{\frac{q}{q-p+1}} + |u|^{\frac{q(p-1)}{q-1}} \right) &\leq 4 \int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} \left(1 + |u|^{\frac{N_*q}{N_*-q}} \right) \\
&\leq 4\mathrm{meas}(\Omega) + 4C_7,
\end{aligned}$$

and

$$\int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} \left(|\nabla u|^{\frac{N_*}{N_*-p+1}} + |\nabla u|^q \right) \leq \int_{\Omega \setminus (K + \overline{B}(0, \frac{\varepsilon}{3}))} (1 + 2|\nabla u|^q) \leq \mathrm{meas}(\Omega) + 2C_7.$$

Returning to (3.3.22), we obtain

$$\int_\Omega |\nabla(\theta u)|^p \leq C_6 + C_6(5\mathrm{meas}(\Omega) + 6C_7) := C_8$$

with C_8 only depending on $(\Lambda, K, \varepsilon)$, that is to say $||\theta u||_{W_0^{1,p}(\Omega)} \leq C_8^{1/p}$. By the Sobolev injection, we get $||\theta u||_{L^{p^*}(\Omega)} \leq C_9 C_8^{1/p}$. Since $\theta \equiv 1$ on the open set $\Omega \backslash (K + \overline{B}(0,\varepsilon))$, these last two estimates give the result of the theorem. ■

3.4 Existence and regularity result for an equation with measure data

We can now prove Theorem 3.1.3. In fact, thanks to the previous estimates and to the technique of [12], the proof is very simple.

Proof of Theorem 3.1.3

We can find $(\mu_n)_{n\geq 1} \in W^{-1,p'}(\Omega) \cap \mathcal{M}(\Omega)$ converging to μ in $\mathcal{M}(\Omega)$ weak-$*$ and such that, for all $\varepsilon > 0$, $\mathrm{supp}(\mu_n) \subset \mathrm{supp}(\mu) + B(0,\varepsilon)$ for n large enough (in fact, most of the classical ways to approximate μ by regular data — for example through a convolution method, or by discretizing μ on a grid — satisfy this property on the supports of the approximations). Take u_n a solution to (3.1.8) with μ_n instead of μ. Since $(\mu_n)_{n\geq 1}$ converges in $\mathcal{M}(\Omega)$ weak-$*$, it is bounded in this space; thus, by Theorem 3.1.1,

$$(u_n)_{n\geq 1} \text{ is bounded in } W_0^{1,q}(\Omega) \text{ for all } q < \tfrac{N_*(p-1)}{N_*-1} \tag{3.4.1}$$

(notice that, by choice of $p > 2 - \frac{1}{N_*}$, we have $\frac{N_*(p-1)}{N_*-1} > 1$). Moreover, for all $\varepsilon > 0$, denoting $K = \overline{\mathrm{supp}(\mu) + B(0,\varepsilon/2)}$, we have $\mathrm{supp}(\mu_n) \subset K$ for n large enough; by Theorem 3.1.2, $(u_n)_{n\geq 1}$ is thus bounded in $L^{p^*}(\Omega \backslash (K + \overline{B}(0,\varepsilon/2))$ and in $W^{1,p}(\Omega \backslash (K + \overline{B}(0,\varepsilon/2))$; since $K + \overline{B}(0,\varepsilon/2) \subset F_\varepsilon := \overline{\mathrm{supp}(\mu) + B(0,\varepsilon)}$, this implies that

$$(u_n)_{n\geq 1} \text{ is bounded in } L^{p^*}(\Omega \backslash F_\varepsilon) \text{ and in } W^{1,p}(\Omega \backslash F_\varepsilon) \text{ for all } \varepsilon > 0. \tag{3.4.2}$$

By (3.4.1), (3.4.2), a diagonal process and Rellich's theorem, we can extract a sequence of $(u_n)_{n\geq 1}$, still denoted $(u_n)_{n\geq 1}$, such that $u_n \to u$ a.e. on Ω, weakly in $W_0^{1,q}(\Omega)$ for all $q < \frac{N_*(p-1)}{N_*-1}$ and, for all $\varepsilon > 0$, weakly in $L^{p^*}(\Omega \backslash F_\varepsilon)$ and in $W^{1,p}(\Omega \backslash F_\varepsilon)$.

Using then the technique of [12], we can prove that $\nabla u_n \to \nabla u$ a.e. on Ω (there is no Φ in [12], but the technique to prove the a.e. convergence of the gradients works fine even with this additional term). This allows to pass to the limit in the equation satisfied by u_n and to see that u is a solution to (3.1.9). ■

If $p \leq 2 - \frac{1}{N_*}$ (in this case, $N_* = N$), it is well-known that a solution to (3.1.1) with μ measure is not to be sought in a Sobolev space (one can notice that, in this case, Theorem 3.1.1 does not give an estimate in a Sobolev space). To solve this problem, two main notions of solutions have been introduced: entropy solutions (see [4]) or renormalized solutions (see [28]).

For each of these notions, the existence of a solution is proved thanks to an approximation method; thus, Theorem 3.1.2 also allows to obtain entropy or renormalized solutions with better local regularity results than usual. For example, under Hypotheses (3.1.2)—(3.1.6), if $\mu \in L^1(\Omega)$ and $f \in W^{-1,p'}(\Omega)$, the technique of [4], associated to Estimate (3.2.13) and to Theorem 3.1.2, allows to prove the existence of a solution to (3.1.1) in the sense:

$$\left\{ \begin{array}{l} u : \Omega \to \mathbb{R} \text{ is a measurable function,} \\ \forall k \geq 0\,,\ T_k(u) \in W_0^{1,p}(\Omega)\,, \\ \forall \varepsilon > 0\,,\ \text{denoting } F_\varepsilon = \overline{\mathrm{supp}(\mu) + B(0,\varepsilon)},\ u \in L^{p^*}(\Omega \backslash F_\varepsilon) \cap W^{1,p}(\Omega \backslash F_\varepsilon)\,, \\ \displaystyle\int_\Omega a(x,u,\nabla u) \cdot \nabla (T_k(u-\varphi)) + \int_\Omega \Phi(x,u) \cdot \nabla (T_k(u-\varphi)) = \int_\Omega \mu T_k(u-\varphi) \\ \qquad + \langle f, T_k(u-\varphi) \rangle_{W^{-1,p'}(\Omega), W_0^{1,p}(\Omega)}\,,\ \forall \varphi \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega). \end{array} \right. \tag{3.4.3}$$

3.5 Appendix

The first lemma is a well-known result concerning the integrability properties of functions in Marcinkiewicz spaces.

Lemma 3.5.1 *Let $v : \Omega \to \mathbb{R}$ be a measurable function and $r > 0$. Suppose that there exists M such that, for all $k \geq 1$, $\mathrm{meas}(\{|v| \geq k\}) \leq Mk^{-r}$; then, for all $s \in]0, r[$, there exists C only depending on (Ω, M, r, s) such that*

$$\int_\Omega |v|^s \leq C.$$

Proof of Lemma 3.5.1
We know (this is a simple application of the Fubini-Tonelli theorem) that

$$\int_\Omega |v|^s = \int_0^\infty \mathrm{meas}(\{|v|^s \geq t\})\, dt.$$

Bounding $\mathrm{meas}(\{|v|^s \geq t\})$ by $\mathrm{meas}(\Omega)$ if $t \leq 1$ and by $Mt^{-r/s}$ if $t \geq 1$ (because $\{|v|^s \geq t\} = \{|v| \geq t^{1/s}\}$), and using the fact that $\frac{r}{s} > 1$, we find

$$\int_\Omega |v|^s \leq \mathrm{meas}(\Omega) + M \int_1^\infty t^{-r/s}\, dt = \mathrm{meas}(\Omega) + \frac{M}{\frac{r}{s} - 1}\,,$$

which concludes the proof. ■

The following result is a very simple generalization of a lemma in [4].

Lemma 3.5.2 *Let $\alpha \in]0, p[$ and $v \in W_0^{1,p}(\Omega)$. If there exists M such that, for all $k \geq 1$,*

$$\int_\Omega |\nabla(T_k(v))|^p \leq Mk^\alpha,$$

then there exists C only depending on (Ω, p, N_*, M) such that, for all $k > 0$,

$$\mathrm{meas}(\{|v| \geq k\}) \leq Ck^{-\frac{N_*(p-\alpha)}{N_*-p}} \quad \text{and} \quad \mathrm{meas}(\{|\nabla v| \geq k\}) \leq Ck^{-\frac{N_*(p-\alpha)}{N_*-\alpha}}.$$

Proof of Lemma 3.5.2.
Let $k \geq 1$. Thanks to the Sobolev injection, we can find C_1 only depending on (Ω, p, N_*) such that

$$||T_k(v)||_{L^{p^*}(\Omega)} \leq C_1 || \, |\nabla(T_k(v))| \, ||_{L^p(\Omega)} \leq C_1 M^{\frac{1}{p}} k^{\frac{\alpha}{p}}.$$

Since $\mathrm{meas}(\{|v| \geq k\}) = \mathrm{meas}(\{|T_k(v)| \geq k\}) \leq k^{-p^*} ||T_k(v)||^{p^*}_{L^{p^*}(\Omega)}$, we deduce that

$$\mathrm{meas}(\{|v| \geq k\}) \leq C_1^{p^*} M^{\frac{p^*}{p}} k^{-p^*} k^{\frac{p^*\alpha}{p}},$$

which concludes the first estimate for $k \geq 1$, since $p^*(1 - \frac{\alpha}{p}) = \frac{N_*(p-\alpha)}{N_*-p}$. If $k \leq 1$, we simply write $\mathrm{meas}(\{|v| \geq k\}) \leq \mathrm{meas}(\Omega) \leq \mathrm{meas}(\Omega) k^{-\frac{N_*(p-\alpha)}{N_*-p}}$.
Let $k \geq 1$ and take $\lambda \geq 1$. We have

$$\{|\nabla v| \geq k\} \subset \{|\nabla v| \geq k\,,\ |v| \geq \lambda\} \cup \{|\nabla v| \geq k\,,\ |v| \leq \lambda\} \subset \{|v| \geq \lambda\} \cup \{|\nabla(T_\lambda(v))| \geq k\} \cup A,$$

where $\mathrm{meas}(A) = 0$ (because $\nabla(T_\lambda(v)) = \mathbf{1}_{\{|v|\leq\lambda\}}\nabla v$ a.e. on Ω). Thus, using the first estimate and the hypothesis of the lemma,

$$\mathrm{meas}(\{|\nabla v| \geq k\}) \leq C_2\lambda^{-\frac{N_*(p-\alpha)}{N_*-p}} + k^{-p}\int_\Omega |\nabla(T_\lambda(v))|^p \leq C_2\lambda^{-\frac{N_*(p-\alpha)}{N_*-p}} + M\lambda^\alpha k^{-p} \tag{3.5.1}$$

where C_2 only depends on (Ω, p, N_*, M).
Choose now $\lambda = k^{\frac{p}{p_1+\alpha}}$, where $p_1 = \frac{N_*(p-\alpha)}{N_*-p}$ (this choice comes down to taking $\lambda = k^\beta$ for some β such that the powers of k in the right-hand side of (3.5.1) are the same; this also comes down — up to a multiplicative constant — to minimizing the right-hand side of (3.5.1) on λ). We have $\lambda \geq 1$ and

$$\lambda^{-\frac{N_*(p-\alpha)}{N_*-p}} = k^{-\frac{pp_1}{p_1+\alpha}} \quad \text{and} \quad \lambda^\alpha k^{-p} = k^{-\left(p-\frac{\alpha p}{p_1+\alpha}\right)} = k^{-\frac{pp_1}{p_1+\alpha}}.$$

Since $\frac{pp_1}{p_1+\alpha} = \frac{pN_*(p-\alpha)}{N_*(p-\alpha)+\alpha(N_*-p)} = \frac{N_*(p-\alpha)}{N_*-\alpha}$, (3.5.1) concludes the second inequality if $k \geq 1$; for $k \leq 1$, we simply bound $\mathrm{meas}(\{|\nabla v| \geq k\})$ by $\mathrm{meas}(\Omega) \leq \mathrm{meas}(\Omega)k^{-\frac{N_*(p-\alpha)}{N_*-\alpha}}$. ■

Partie III

Schémas volumes finis pour équations elliptiques

Chapitre 4

Finite volume methods for convection-diffusion equations with right-hand side in H^{-1}

Reference: J. Droniou and T. Gallouët. *M2AN Math. Model. Numer. Anal.* **36** (2002), no. 4, 705-724.

Abstract We prove the convergence of a finite volume method for a noncoercive linear elliptic problem, with right-hand side in the dual space of the natural energy space of the problem.

4.1 Introduction

We take Ω a polygonal open subset of $\mathbb{R}^d$ ($d = 2$ or 3), and we study the problem

$$\left\{ \begin{array}{ll} -\Delta u + \mathrm{div}(\mathbf{v}u) + bu = L & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{array} \right. \tag{4.1.1}$$

with the following hypotheses on the datas:

$$\begin{array}{c} \exists p > d \text{ such that } \mathbf{v} \in (L^p(\Omega))^d\,, \\ b \in L^r(\Omega) \text{ with } r > 1 \text{ if } d = 2 \text{ and } r = \frac{3}{2} \text{ if } d = 3\,,\ b \geq 0 \text{ a.e. on } \Omega\,, \\ L \in H^{-1}(\Omega). \end{array} \tag{4.1.2}$$

Of course, solutions to (4.1.1) are taken in a weak sense, that is to say

$$\left\{ \begin{array}{l} u \in H_0^1(\Omega)\,, \\ \displaystyle\int_\Omega \nabla u \cdot \nabla \varphi - \int_\Omega u\mathbf{v} \cdot \nabla \varphi + \int_\Omega bu\varphi = \langle L, \varphi \rangle_{H^{-1}(\Omega), H_0^1(\Omega)}\,,\ \forall \varphi \in H_0^1(\Omega). \end{array} \right. \tag{4.1.3}$$

Existence and uniqueness of a solution to (4.1.3) have already been proved in [31] (see also [32] for nonlinear problems).

Our purpose is to prove the convergence of a finite volume discretization of (4.1.1). Finite volume methods have been widely used to approximate solutions to convection-diffusion equations, either using structured or unstructured grids (see for example [27], [48], [38], [39], [45]). The grids we consider here are the same as in [38], that is to say grids made of convex polygonal control volumes with some geometrical properties (see the next section).

There are two main originalities in the work we present here. First, we consider elliptic problems which are not necessarily coercive, because it is not supposed that $\frac{1}{2}\mathrm{div}(\mathbf{v}) + b$ is nonnegative. Moreover, the regularity we have taken on the velocity $\mathbf{v}$ is minimal (that is, just enough for (4.1.3) to make sense — in previous papers on the finite volume discretization of convection-diffusion equations, the convection velocity is in general C^1-continuous, see e.g. [38] or [48]); considering a non-regular convection velocity is a first step toward the treatment of coupled systems, in which $\mathbf{v}$ comes from the resolution of another partial differential equation.
The second originality concerns the right-hand side: here too, we consider a datum with minimal regularity (that is, in the dual space of the energy space associated to the equation — previous papers take in general a right-hand side in $L^2(\Omega)$); in fact, $H^{-1}(\Omega)$ is a natural space for right-hand sides of convection-diffusion equations.

In the next section, we define the finite volume scheme used to discretize (4.1.1), and we state the main convergence result of this paper; since we consider data $\mathbf{v}$ and L which lack of regularity (with respect to previous works), we present a new way to discretize them, using what we call "half-diamonds". We also give, in this section, technical results useful to the rest of the paper. In Section 4.3, we prove *a priori* estimates on the solutions to our finite volume discretization of (4.1.1); the problem being noncoercive, obtaining estimates on these solutions is not straightforward: we must adapt the techniques of [31] to the discrete setting. Along with the compactness results of [38], these *a priori* estimates allow us, in Section 4.4, to prove our main result, that is to say existence and uniqueness of the approximate solutions and their convergence toward the solution of (4.1.3); to prove the convergence result with our irregular data, we approximate them by regular data and adapt then known techniques (see [38], for example). In the last section, we present a modified scheme which consists in discretizing the data $\mathbf{v}$ and L using another method (based on the "full-diamonds"); comparing this scheme to the one of Section 4.2, we easily obtain the convergence of the associated approximate solutions.

4.2 Definition of the scheme and main result

Definition 4.2.1 *An admissible mesh $\mathcal{T}$ of Ω is a finite family of polygonal open convex subsets of Ω (the "control volumes"), together with a finite family $\mathcal{E}$ of disjoint subsets of $\overline{\Omega}$ contained in affine hyperplanes (the "edges") and a family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points in Ω such that*

i) $\overline{\Omega} = \bigcup_{K\in\mathcal{T}} \overline{K}$,

ii) each $\sigma \in \mathcal{E}$ is a non-empty open subset of ∂K for some $K \in \mathcal{T}$,

iii) by denoting $\mathcal{E}_K = \{\sigma \in \mathcal{E} \mid \sigma \subset \partial K\}$, $\partial K = \cup_{\sigma\in\mathcal{E}_K} \overline{\sigma}$ for all $K \in \mathcal{T}$,

iv) for all $K \neq L$ in $\mathcal{T}$, either the $(d-1)$-dimensional measure of $\overline{K} \cap \overline{L}$ is null, or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, that we denote then $\sigma = K|L$,

v) for all $K \in \mathcal{T}$, $x_K \in K$,

vi) for all $\sigma = K|L \in \mathcal{E}$, the line (x_K, x_L) intersects and is orthogonal to σ,

vii) for all $\sigma \in \mathcal{E}$, $\sigma \subset \partial\Omega \cap \partial K$, the line which is orthogonal to σ and going through x_K intersects σ.

The size of the mesh is then defined by $\mathrm{size}(\mathcal{T}) = \sup_{K\in\mathcal{T}} \mathrm{diam}(K)$ (where $\mathrm{diam}(K)$ is the diameter of K). We denote by $\mathrm{meas}(K)$ the Lebesgue measure of $K \in \mathcal{T}$. The unit normal to $\sigma \in \mathcal{E}_K$ outward to K is denoted by $\mathbf{n}_{K,\sigma}$.
We define $\mathcal{E}_{\mathrm{int}} = \{\sigma \in \mathcal{E} \mid \sigma \not\subset \partial\Omega\}$ and $\mathcal{E}_{\mathrm{ext}} = \mathcal{E}\backslash\mathcal{E}_{\mathrm{int}}$. If $\sigma \in \mathcal{E}$, $m(\sigma)$ is the $(d-1)$-dimensional measure of σ; if $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$, d_σ is the Euclidean distance between the points (x_K, x_L) and $d_{K,\sigma}$ denotes

the distance between x_K and σ; if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, $d_\sigma = d_{K,\sigma}$ is the distance between x_K and σ. The transmissivity through an edge σ is $\tau_\sigma = \frac{m(\sigma)}{d_\sigma}$. We denote by γ the $(d-1)$-dimensional measure on the edges of the mesh.
If $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, the "half-diamond" $\triangle_{K,\sigma}$ is defined by $\triangle_{K,\sigma} = \{tx_K + (1-t)x\,,\ t \in [0,1]\,,\ x \in \sigma\}$. It will be useful to notice that $\mathrm{meas}(\triangle_{K,\sigma}) = \frac{m(\sigma)d_{K,\sigma}}{d}$.
The following quantity measures the "regularity" of the mesh:

$$\mathrm{reg}(\mathcal{T}) = \inf_{K\in\mathcal{T}} \left(\inf_{\sigma\in\mathcal{E}_K} \frac{d_{K,\sigma}}{d_\sigma} \right).$$

If $\mathcal{T}$ is an admissible mesh, and under Hypothesis (4.1.2), we can define the finite volume discretization of (4.1.1).
We first write

$$L = f + \mathrm{div}(G)\,, \quad \text{with } f \in L^2(\Omega) \text{ and } G \in (L^2(\Omega))^d.$$

It is well-known that any element of $H^{-1}(\Omega)$ can be written this way; in fact, in models of physical problems, the right-hand side naturally appears in this form, see e.g. [44], and there is thus no trouble to define the following scheme (this is also why we have kept f, which can be taken, from a theoretical point of view, null).
The finite volume discretization consists in integrating the equation $-\Delta u + \mathrm{div}(\mathbf{v}u) + bu = f + \mathrm{div}(G)$ on a control volume K: with some integrates by parts, we formally obtain

$$\sum_{\sigma\in\mathcal{E}_K} -\int_\sigma \nabla u\cdot\mathbf{n}_{K,\sigma}\,d\gamma + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma u\mathbf{v}\cdot\mathbf{n}_{K,\sigma}\,d\gamma + \int_K bu = \int_K f + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma G\cdot\mathbf{n}_{K,\sigma}\,d\gamma.$$

By letting u_K be an approximate value of u on the control volume K, we must then discretize each term of this relation. To this aim, we denote, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$\begin{aligned} &v_{K,\sigma} = \left(\frac{1}{\mathrm{meas}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} \mathbf{v}\right)\cdot\mathbf{n}_{K,\sigma}\,, \quad b_K = \frac{1}{\mathrm{meas}(K)}\int_K b\,,\\ &f_K = \frac{1}{\mathrm{meas}(K)}\int_K f \quad \text{and} \quad G_{K,\sigma} = \left(\frac{1}{\mathrm{meas}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} G\right)\cdot\mathbf{n}_{K,\sigma} \end{aligned} \tag{4.2.1}$$

(these are, respectively, approximate values of $\mathbf{v}\cdot\mathbf{n}_{K,\sigma}$ on σ, of b on K, of f on K and of $G\cdot\mathbf{n}_{K,\sigma}$ on σ), and the finite volume scheme is written

$$\forall K\in\mathcal{T}\,,\ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)v_{K,\sigma}u_{K,\sigma,+} + \mathrm{meas}(K)b_Ku_K = \mathrm{meas}(K)f_K + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)G_{K,\sigma}\,, \tag{4.2.2}$$

$$\forall K\in\mathcal{T}\,,\forall\sigma\in\mathcal{E}_K\,,\quad F_{K,\sigma} = -\tfrac{m(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)\,, \tag{4.2.3}$$

$$\begin{aligned} \forall\sigma = K|L\in\mathcal{E}_{\rm int}\,,\quad & F_{K,\sigma} + m(\sigma)v_{K,\sigma}u_{K,\sigma,+} - m(\sigma)G_{K,\sigma}\\ &= -(F_{L,\sigma} + m(\sigma)v_{L,\sigma}u_{L,\sigma,+} - m(\sigma)G_{L,\sigma})\,, \end{aligned} \tag{4.2.4}$$

$$\begin{aligned} &\forall\sigma\in\mathcal{E}_{\rm ext}\,, && u_\sigma = 0\,,\\ &\forall\sigma = K|L\in\mathcal{E}_{\rm int}\,, && u_{K,\sigma,+} = u_K \ \text{ if } \ v_{K,\sigma}\geq 0\,,\quad u_{K,\sigma,+} = u_L \ \text{ otherwise,}\\ &\forall\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K\,, && u_{K,\sigma,+} = u_K \ \text{ if } \ v_{K,\sigma}\geq 0\,,\quad u_{K,\sigma,+} = 0 \ \text{ otherwise.} \end{aligned} \tag{4.2.5}$$

Equations (4.2.2)—(4.2.5) are a linear system in $(u_K)_{K\in\mathcal{T}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$, but thanks to (4.2.4) (which describes the conservativity of the fluxes), we can eliminate the unknowns $(u_\sigma)_{\sigma\in\mathcal{E}}$, so that (4.2.2)—(4.2.5) can be considered as a linear system of size $\mathrm{Card}(\mathcal{T})$, with unknowns $(u_K)_{K\in\mathcal{T}}$.
We naturally identify the set $\mathbb{R}^{\mathrm{Card}(\mathcal{T})}$ to the set $X(\mathcal{T})$ of functions defined a.e. on Ω and constant on each control volume $K \in \mathcal{T}$.

Our main result is the following.

Theorem 4.2.1 *If $\mathcal{T}$ is an admissible mesh, then there exists a unique solution to (4.2.2)—(4.2.5). Moreover, let $\alpha > 0$; denoting by $u_\mathcal{T} \in X(\mathcal{T})$ the solution to (4.2.2)—(4.2.5), $u_\mathcal{T}$ converges in $L^q(\Omega)$, for all $q < \frac{2d}{d-2}$, to the unique solution of (4.1.3), as* $\mathrm{size}(\mathcal{T}) \to 0$ *with* $\mathrm{reg}(\mathcal{T}) \geq \alpha$.

Remark 4.2.1 *We will not use, to prove this theorem, the existence of a solution to (4.1.3). The finite volume method allows, as usual, to prove the existence of a solution to the continuous problem.*

Remark 4.2.2 *In dimension $d = 2$, the regularity we suppose on $\mathbf{v}$ is minimal in order for all the terms in (4.1.3) to make sense (see the Sobolev imbeddings in [1]). But, if $d = 3$, the minimal regularity on the convection velocity would be: $\mathbf{v} \in (L^3(\Omega))^3$; in fact, cutting $\mathbf{v}$ in two parts (one small in $(L^3(\Omega))^3$, the other in $(L^\infty(\Omega))^3$ — see [31] for the reasoning in the continuous case), we could also prove Theorem 4.2.1 under this minimal hypothesis on $\mathbf{v}$. However, for the legibility of the following proofs, we prefer to suppose Hypothesis (4.1.2).*

4.2.1 Technical results

To prove this existence, uniqueness and convergence result, we first search for *a priori* estimates on the solutions to (4.2.2)—(4.2.5). These estimates are obtained via the following discrete H_0^1 norm.

Definition 4.2.2 *If $\mathcal{T}$ is an admissible mesh and $v_\mathcal{T} = (v_K)_{K\in\mathcal{T}} \in X(\mathcal{T})$, we define*

$$||v_\mathcal{T}||_{1,\mathcal{T}} = \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (D_\sigma v_\mathcal{T})^2\right)^{1/2},$$

where $D_\sigma v_\mathcal{T} = |v_K - v_L|$ if $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$ and $D_\sigma v_\mathcal{T} = |v_K|$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$.

Notice that this norm takes into account a boundary condition "$v_\mathcal{T} = 0$ on $\partial\Omega$", since we have defined $D_\sigma v_\mathcal{T} = |v_K|$ if $\sigma \subset \partial\Omega$ (this comes down to consider that functions of $X(\mathcal{T})$ are defined on $\mathbb{R}^N$ and are null outside Ω).
The following proposition sums up a few useful properties of the norm $||\cdot||_{1,\mathcal{T}}$.

Proposition 4.2.1 *i) (Discrete Poincaré inequality) If $\mathcal{T}$ is an admissible mesh and $v_\mathcal{T} \in X(\mathcal{T})$, then $||v_\mathcal{T}||_{L^2(\Omega)} \leq \mathrm{diam}(\Omega)||v_\mathcal{T}||_{1,\mathcal{T}}$ (where $\mathrm{diam}(\Omega)$ is the diameter of Ω).*

ii) (Discrete Sobolev inequality) If $\mathcal{T}$ is an admissible mesh and $0 < \zeta \leq \mathrm{reg}(\mathcal{T})$, then there exists C only depending on (Ω, ζ) such that, for all $q \in [1, \frac{2d}{d-2}[$, for all $v_\mathcal{T} \in X(\mathcal{T})$, $||v_\mathcal{T}||_{L^q(\Omega)} \leq Cq||v_\mathcal{T}||_{1,\mathcal{T}}$.

iii) (Discrete Rellich Theorem) If $(\mathcal{T}_n)_{n\geq 1}$ is a sequence of admissible meshes such that $\mathrm{size}(\mathcal{T}_n) \to 0$ and if $v_n \in X(\mathcal{T}_n)$ is such that $(||v_n||_{1,\mathcal{T}_n})_{n\geq 1}$ is bounded, then $(v_n)_{n\geq 1}$ is relatively compact in $L^2(\Omega)$ and any adherence value in $L^2(\Omega)$ of $(v_n)_{n\geq 1}$ belongs to $H_0^1(\Omega)$.

For a proof of these properties, see [38].

The following discrete integrate by parts formula will be quite useful in the sequel.

Lemma 4.2.1 *Let $\mathcal{T}$ be an admissible mesh and $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ satisfy (4.2.2)—(4.2.5). Then, for all $\varphi_\mathcal{T} = (\varphi_K)_{K\in\mathcal{T}} \in X(\mathcal{T})$, we have*

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi_K - \varphi_L) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi_K \\
&\quad = \sum_{K\in\mathcal{T}} \mathrm{meas}(K) f_K \varphi_K + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi_K - \varphi_L) \\
&\quad + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma}\right)(\varphi_K - \varphi_L),
\end{aligned} \tag{4.2.6}$$

where we have denoted $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\mathrm{int}}$ and $u_L = u_{L,\sigma,+} = v_{L,\sigma} = d_{L,\sigma} = G_{L,\sigma} = \varphi_L = 0$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$.

Proof of Lemma 4.2.1
We notice that, thanks to (4.2.4), the quantity $a_{K,\sigma} = F_{K,\sigma} + m(\sigma)v_{K,\sigma}u_{K,\sigma,+} - m(\sigma)G_{K,\sigma}$ is conservative, that is to say, if $\sigma = K|L \in \mathcal{E}_{\rm int}$, then $a_{K,\sigma} = -a_{L,\sigma}$.
Multiplying (4.2.2) by φ_K and summing on the control volumes $K \in \mathcal{T}$, we have

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} a_{K,\sigma}\varphi_K + \sum_{K\in\mathcal{T}} {\rm meas}(K) b_K u_K \varphi_K = \sum_{K\in\mathcal{T}} {\rm meas}(K) f_K \varphi_K.$$

Using the conservativity of $a_{K,\sigma}$ and gathering by edges, we deduce

$$\sum_{\sigma\in\mathcal{E}} a_{K,\sigma}(\varphi_K - \varphi_L) + \sum_{K\in\mathcal{T}} {\rm meas}(K) b_K u_K \varphi_K = \sum_{K\in\mathcal{T}} {\rm meas}(K) f_K \varphi_K \tag{4.2.7}$$

where $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $\varphi_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.
Let us now compute the $(a_{K,\sigma})_{K\in\mathcal{T},\ \sigma\in\mathcal{E}_K}$. If $\sigma = K|L \in \mathcal{E}_{\rm int}$, then (4.2.3) and (4.2.4) give u_σ; indeed, dividing (4.2.4) by $m(\sigma)$, we have

$$-\frac{u_\sigma}{d_{K,\sigma}} + \frac{u_K}{d_{K,\sigma}} + v_{K,\sigma}u_{K,\sigma,+} - G_{K,\sigma} = \frac{u_\sigma}{d_{L,\sigma}} - \frac{u_L}{d_{L,\sigma}} - v_{L,\sigma}u_{L,\sigma,+} + G_{L,\sigma},$$

that is, noticing that $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$,

$$\frac{d_\sigma}{d_{K,\sigma}d_{L,\sigma}}u_\sigma = \frac{u_K}{d_{K,\sigma}} + \frac{u_L}{d_{L,\sigma}} + v_{K,\sigma}u_{K,\sigma,+} + v_{L,\sigma}u_{L,\sigma,+} - G_{K,\sigma} - G_{L,\sigma},$$

which gives

$$u_\sigma = \frac{d_{L,\sigma}}{d_\sigma}u_K + \frac{d_{K,\sigma}}{d_\sigma}u_L + \frac{d_{K,\sigma}d_{L,\sigma}}{d_\sigma}\left(v_{K,\sigma}u_{K,\sigma,+} + v_{L,\sigma}u_{L,\sigma,+} - G_{K,\sigma} - G_{L,\sigma}\right).$$

With this value of u_σ, we obtain

$$\begin{aligned} a_{K,\sigma} &= -\frac{m(\sigma)}{d_{K,\sigma}}\left(\frac{d_{K,\sigma}}{d_\sigma}u_L - \frac{d_{K,\sigma}}{d_\sigma}u_K\right) - \frac{m(\sigma)d_{L,\sigma}}{d_\sigma}\left(v_{K,\sigma}u_{K,\sigma,+} + v_{L,\sigma}u_{L,\sigma,+} - G_{K,\sigma} - G_{L,\sigma}\right) \\ &\quad + m(\sigma)v_{K,\sigma}u_{K,\sigma,+} - m(\sigma)G_{K,\sigma} \\ &= \tau_\sigma(u_K - u_L) + m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+} - \frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}\right) \\ &\quad - m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right). \end{aligned}$$

Note that this equality is also valid if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, providing that we define $u_L = u_{L,\sigma,+} = v_{L,\sigma} = G_{L,\sigma} = \varphi_L = 0$ in this case.
Using this expression in (4.2.7), we obtain the desired formula. ■

4.3 A Priori Estimates

We prove here some *a priori* estimates on the solution to (4.2.2)—(4.2.5). As already said, we adapt the methods of [31] to the discrete setting; however, the estimation of the convection term (the noncoercive part of the equation) requires new ideas, to take advantage of the upwind choice in (4.2.5).

4.3.1 Estimate on $\ln(1 + |u_\mathcal{T}|)$

Proposition 4.3.1 *Let $\mathcal{T}$ be an admissible mesh. If $(u_K)_{K\in\mathcal{T}}$ is a solution to (4.2.2)—(4.2.5), then*

$$||\ln(1 + |u_\mathcal{T}|)||^2_{1,\mathcal{T}} \le 2||f||_{L^1(\Omega)} + 2d\left(||\,|G|\,||_{L^2(\Omega)} + ||\,|\mathbf{v}|\,||_{L^2(\Omega)}\right)^2,$$

where $|X|$ denotes the Euclidean norm of a vector $X \in \mathbb{R}^d$.

Proof of Proposition 4.3.1
Step 1: A preliminary estimate.
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. Applying Formula (4.2.6) to $(\varphi_K)_{K\in\mathcal{T}} = (\varphi(u_K))_{K\in\mathcal{T}}$, and since φ is bounded by 1 and $b_K u_K \varphi(u_K) \geq 0$ for all $K \in \mathcal{T}$, we have

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \\
&\leq \sum_{K\in\mathcal{T}} \mathrm{meas}(K)|f_K| + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+} \right) (\varphi(u_K) - \varphi(u_L)) \\
&\quad + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right) (\varphi(u_K) - \varphi(u_L)) \qquad (4.3.1)
\end{aligned}$$

(with the notation $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\mathrm{int}}$ and $u_L = u_{L,\sigma,+} = v_{L,\sigma} = d_{L,\sigma} = G_{L,\sigma} = \varphi(u_L) = 0$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$).

We have

$$\sum_{K\in\mathcal{T}} \mathrm{meas}(K)|f_K| \leq \sum_{K\in\mathcal{T}} \int_K |f| = ||f||_{L^1(\Omega)}. \qquad (4.3.2)$$

Using the Cauchy-Schwarz inequality, we can write

$$\begin{aligned}
&\left| \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right) (\varphi(u_K) - \varphi(u_L)) \right| \\
&\leq \left(\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right)^2 \right)^{1/2} \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (\varphi(u_K) - \varphi(u_L))^2 \right)^{1/2}. \qquad (4.3.3)
\end{aligned}$$

Since $\left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right)^2 \leq 2 \frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma}^2 + 2 \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma}^2$ (we have used the fact that $\frac{d_{K,\sigma}}{d_\sigma}$ and $\frac{d_{L,\sigma}}{d_\sigma}$ are bounded by 1), gathering by control volumes, we have

$$\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right)^2 \leq 2 \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} G_{K,\sigma}^2.$$

But, for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$, by Jensen's inequality and since $\mathrm{meas}(\triangle_{K,\sigma}) = \frac{m(\sigma) d_{K\sigma}}{d}$, we have $m(\sigma) d_{K,\sigma} G_{K,\sigma}^2 \leq d \int_{\triangle_{K,\sigma}} |G|^2$. Using the fact that $\{\triangle_{K,\sigma},\ K \in \mathcal{T},\ \sigma \in \mathcal{E}_K\}$ is (up to a set of null Lebesgue measure) a partition of Ω, we deduce

$$\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right)^2 \leq 2d \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \int_{\triangle_{K,\sigma}} |G|^2 = 2d ||\,|G|\,||_{L^2(\Omega)}^2. \qquad (4.3.4)$$

φ being nondecreasing and Lipschitz-continuous with Lipschitz constant 1, we have $(\varphi(u_K) - \varphi(u_L))^2 \leq (u_K - u_L)(\varphi(u_K) - \varphi(u_L))$; (4.3.3) and (4.3.4) give then

$$\begin{aligned}
&\left| \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right) (\varphi(u_K) - \varphi(u_L)) \right| \\
&\leq \sqrt{2d}\, ||\,|G|\,||_{L^2(\Omega)} \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \right)^{1/2}. \qquad (4.3.5)
\end{aligned}$$

Now, we need to estimate the terms of (4.3.1) coming from the discretization of the convection part of (4.1.1). We first notice that, if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$,

$$\left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)) = -\frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+} \varphi(u_K) \le 0.$$

Indeed, if $v_{K,\sigma} \ge 0$, this last term is $-\frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_K \varphi(u_K)$, which is nonpositive since $s\varphi(s) \ge 0$ for all $s \in \mathbb{R}$; if $v_{K,\sigma} < 0$, this last term is null (because $u_{K,\sigma,+} = 0$ in this case). Thus,

$$\begin{aligned}
&\sum_{\sigma \in \mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)) \\
&\quad \le \sum_{\sigma \in \mathcal{E}_{\rm int}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)). \qquad (4.3.6)
\end{aligned}$$

Let $\sigma = K|L \in \mathcal{E}_{\rm int}$ and denote

$$\Lambda = \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)), \quad A = \left|\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma}\right| \text{ and } B = \left|\frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma}\right|.$$

We separate the cases.

- If $v_{K,\sigma}$ and $v_{L,\sigma}$ are nonnegative, then $\Lambda = (Au_L - Bu_K)(\varphi(u_K) - \varphi(u_L))$ is, by item i) of Lemma 4.3.1 below, bounded from above by 0 if $u_K u_L \le 0$ and by $|A - B| \inf(|u_L|, |u_K|)|\varphi(u_K) - \varphi(u_L)|$ otherwise.
- If $v_{K,\sigma}$ and $v_{L,\sigma}$ are negative, then $\Lambda = (-Au_K + Bu_L)(\varphi(u_K) - \varphi(u_L)) = (Bu_L - Au_K)(\varphi(u_K) - \varphi(u_L))$ is once again bounded from above by 0 if $u_K u_L \le 0$ and by $|A - B| \inf(|u_L|, |u_K|)|\varphi(u_K) - \varphi(u_L)|$ otherwise.
- If $v_{K,\sigma} \ge 0$ and $v_{L,\sigma} < 0$, then $\Lambda = -(A + B)u_K(\varphi(u_K) - \varphi(u_L))$ is, by item ii) of Lemma 4.3.1 below, bounded from above by 0 if $u_K u_L \le 0$ and by $(A + B) \inf(|u_L|, |u_K|)|\varphi(u_K) - \varphi(u_L)|$ otherwise.
- If $v_{K,\sigma} < 0$ and $v_{L,\sigma} \ge 0$, then $\Lambda = (A + B)u_L(\varphi(u_K) - \varphi(u_L)) = -(A + B)u_L(\varphi(u_L) - \varphi(u_K))$ is, as before, bounded from above by 0 if $u_K u_L \le 0$ and by $(A + B) \inf(|u_K|, |u_L|)|\varphi(u_L) - \varphi(u_K)|$ otherwise.

In either case, we notice that $\Lambda \le 0$ if $u_K u_L \le 0$ and that $\Lambda \le (A + B) \inf(|u_K|, |u_L|)|\varphi(u_K) - \varphi(u_L)|$ otherwise; thus, by denoting $\mathcal{A} = \{\sigma = K|L \in \mathcal{E}_{\rm int} \mid u_K u_L > 0\}$, (4.3.6) gives

$$\begin{aligned}
&\sum_{\sigma \in \mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)) \\
&\quad \le \sum_{\sigma \in \mathcal{A}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} |v_{L,\sigma}| + \frac{d_{K,\sigma}}{d_\sigma} |v_{K,\sigma}|\right) \inf(|u_K|, |u_L|)|\varphi(u_K) - \varphi(u_L)|.
\end{aligned}$$

Since $\frac{d_{K,\sigma}}{d_\sigma}$ and $\frac{d_{L,\sigma}}{d_\sigma}$ are bounded by 1, we obtain

$$\begin{aligned}
&\sum_{\sigma \in \mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)(\varphi(u_K) - \varphi(u_L)) \\
&\quad \le \left(2 \sum_{\sigma \in \mathcal{A}} m(\sigma) d_\sigma \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma}^2 + \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma}^2\right)\right)^{1/2} \left(\sum_{\sigma \in \mathcal{A}} \tau_\sigma \inf(|u_K|, |u_L|)^2 (\varphi(u_K) - \varphi(u_L))^2\right)^{1/2} \qquad (4.3.7)
\end{aligned}$$

Gathering by control volumes, and using Jensen's inequality, we can write

$$\sum_{\sigma\in\mathcal{A}} m(\sigma)d_\sigma\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}^2+\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}^2\right)\le\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\frac{m(\sigma)d_{K,\sigma}}{\mathrm{meas}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}|\mathbf{v}|^2.$$

Since $\frac{m(\sigma)d_{K,\sigma}}{\mathrm{meas}(\triangle_{K,\sigma})}=d$ and $\{\triangle_{K,\sigma}\,,\ K\in\mathcal{T}\,,\ \sigma\in\mathcal{E}_K\}$ is (up to a set of null Lebesgue measure) a partition of Ω, we deduce

$$\sum_{\sigma\in\mathcal{A}} m(\sigma)d_\sigma\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}^2+\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}^2\right)\le d\,||\,|\mathbf{v}|\,||_{L^2(\Omega)}^2. \tag{4.3.8}$$

For all $\sigma=K|L\in\mathcal{A}$, since $u_Ku_L>0$, item iii) of Lemma 4.3.1 gives

$$\inf(|u_K|,|u_L|)^2(\varphi(u_K)-\varphi(u_L))^2\le(u_K-u_L)(\varphi(u_K)-\varphi(u_L)).$$

Using this and (4.3.8) in (4.3.7), we finally obtain

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(u_K)-\varphi(u_L))\\ &\quad\le\quad\sqrt{2d}\,||\,|\mathbf{v}|\,||_{L^2(\Omega)}\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi(u_K)-\varphi(u_L))\right)^{1/2}.\end{aligned} \tag{4.3.9}$$

Gathering (4.3.2), (4.3.5) and (4.3.9) in (4.3.1), we get

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi(u_K)-\varphi(u_L))\\ &\quad\le\quad||f||_{L^1(\Omega)}+\sqrt{2d}\left(||\,|G|\,||_{L^2(\Omega)}+||\,|\mathbf{v}|\,||_{L^2(\Omega)}\right)\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi(u_K)-\varphi(u_L))\right)^{1/2},\end{aligned}$$

which gives, thanks to Young's inequality,

$$\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi(u_K)-\varphi(u_L))\le 2||f||_{L^1(\Omega)}+2d\left(||\,|G|\,||_{L^2(\Omega)}+||\,|\mathbf{v}|\,||_{L^2(\Omega)}\right)^2. \tag{4.3.10}$$

Step 2: Estimate on $\ln(1+|u_\mathcal{T}|)$.
We notice that, for all $s\in\mathbb{R}$, $\ln(1+|s|)=\int_0^s\frac{\mathrm{sgn}(t)\,dt}{1+|t|}$. Thus, for all $(x,y)\in\mathbb{R}^2$, by the Cauchy-Schwarz inequality and since φ is nondecreasing,

$$\begin{aligned}(\ln(1+|x|)-\ln(1+|y|))^2&=\left(\int_y^x\frac{\mathrm{sgn}(t)\,dt}{1+|t|}\right)^2\\ &\le\quad|x-y|\left|\int_y^x\frac{dt}{(1+|t|)^2}\right|=|x-y||\varphi(x)-\varphi(y)|=(x-y)(\varphi(x)-\varphi(y)).\end{aligned}$$

Using this bound in (4.3.10), we deduce the desired estimate on $\ln(1+|u_\mathcal{T}|)$. ■

It remains to state and prove the following technical result, which has been used in the course of the preceding proof. This lemma shows the usefulness of the upwind choice in (4.2.5): thanks to the first two items of the lemma, the upwind choice allows to reduce the estimate on the discrete convection term to the cases $u_Ku_L>0$; these cases are then, thanks to item iii), bounded by the discrete diffusion term.

Lemma 4.3.1 *Let* $\varphi(s)=\int_0^s\frac{dt}{(1+|t|)^2}$.

i) Let A and B be nonnegative real numbers and $(x,y) \in \mathbb{R}^2$. If $xy \leq 0$, then

$$(Ax - By)(\varphi(y) - \varphi(x)) \leq 0$$

and, if $xy > 0$, then

$$(Ax - By)(\varphi(y) - \varphi(x)) \leq |A - B| \inf(|x|, |y|)|\varphi(y) - \varphi(x)|.$$

ii) Let $(x,y) \in \mathbb{R}^2$. If $xy \leq 0$, then

$$-y(\varphi(y) - \varphi(x)) \leq 0$$

and, if $xy > 0$, then

$$-y(\varphi(y) - \varphi(x)) \leq \inf(|x|, |y|)|\varphi(y) - \varphi(x)|.$$

iii) Let $(x,y) \in \mathbb{R}^2$. If $xy > 0$, then

$$\inf(|x|, |y|)^2(\varphi(y) - \varphi(x))^2 \leq (y - x)(\varphi(y) - \varphi(x)).$$

Proof of Lemma 4.3.1
The first two items are only consequences of the nondecreasingness of φ and of the fact that $s\varphi(s) \geq 0$ for all $s \in \mathbb{R}$.
Consider i). Suppose first that $xy \leq 0$. Up to a permutation of x and y, there is no loss of generality if we assume that $x \leq 0$. If $x = 0$, then $(Ax - By)(\varphi(y) - \varphi(x)) = -By\varphi(y) \leq 0$. If $x < 0$, then $y \geq 0 > x$ and, A and B being nonnegative, we have $By \geq 0 \geq Ax$, thus $Ax - By \leq 0$; φ being nondecreasing, we deduce that $(Ax - By)(\varphi(y) - \varphi(x)) \leq 0$.
Suppose now that $xy > 0$. Up to a permutation of x and y, we can suppose that $|x| \leq |y|$. We have then

$$(Ax - By)(\varphi(y) - \varphi(x)) = (A - B)x(\varphi(y) - \varphi(x)) + B(x - y)(\varphi(y) - \varphi(x)).$$

φ being nondecreasing, the second term of the right-hand side of this equality is nonpositive, and we obtain thus $(Ax - By)(\varphi(y) - \varphi(x)) \leq (A - B)x(\varphi(y) - \varphi(x)) \leq |A - B|\,|x|\,|\varphi(y) - \varphi(x)|$ as desired.

Let us now study the second item. If $xy \leq 0$, then either $x = 0$, or $y = 0$, or $x < 0 < y$ or $y < 0 < x$. In the first case, $-y(\varphi(y) - \varphi(x)) = -y\varphi(y) \leq 0$; in the second case, $-y(\varphi(y) - \varphi(x)) = 0$; in the third case, $-y \leq 0$ and $\varphi(y) - \varphi(x) \geq 0$ so that the result holds; in the fourth case, $-y \geq 0$ but $\varphi(y) - \varphi(x) \leq 0$ and the result still holds. Assume now that $xy > 0$; the result is obvious if $|y| = \inf(|x|, |y|)$, so that we can take $|y| \geq |x|$; then either $0 < x \leq y$ or $y \leq x < 0$. In both cases, the nondecreasingness of φ easily gives $-y(\varphi(y) - \varphi(x)) \leq 0$, and the desired inequality is thus satisfied.

To prove the third item, we notice that, since φ is C^1-continuous on $\mathbb{R}$, there exists $\theta \in [x,y]$ such that $\varphi(y) - \varphi(x) = \varphi'(\theta)(y - x)$. Using the fact that φ is nondecreasing, we obtain

$$\begin{aligned} \inf(|x|, |y|)^2(\varphi(y) - \varphi(x))^2 &\leq \frac{\inf(|x|, |y|)^2}{(1 + |\theta|)^2}|y - x|\,|\varphi(y) - \varphi(x)| \\ &\leq \frac{\inf(|x|, |y|)^2}{(1 + |\theta|)^2}(y - x)(\varphi(y) - \varphi(x)). \end{aligned}$$

But, since x and y have the same sign and $\theta \in [x,y]$, we have $\inf(|x|, |y|) \leq |\theta|$, and the result is thus a consequence of the previous inequality. ■

4.3.2 Estimate on $||u_{\mathcal{T}}||_{1,\mathcal{M}}$

Theorem 4.3.1 *Let $\mathcal{T}$ be an admissible mesh, $0 < \zeta \leq \mathrm{reg}(\mathcal{T})$ and M be an upper bound of $|||\mathbf{v}|||_{L^p(\Omega)}$. There exists $C > 0$ only depending on (Ω, p, M, ζ) such that, if $u_{\mathcal{T}}$ is a solution to (4.2.2)—(4.2.5), then*

$$||u_{\mathcal{T}}||_{1,\mathcal{T}} \leq C(||f||_{L^2(\Omega)} + |||G|||_{L^2(\Omega)}).$$

Proof of Theorem 4.3.1
(4.2.2)—(4.2.5) being a linear system, proving a bound on $u_{\mathcal{T}}$ whenever

$$||f||_{L^2(\Omega)} + |||G|||_{L^2(\Omega)} \leq 1 \tag{4.3.11}$$

is enough to prove the theorem in the general case.
We denote, for $k > 0$, $T_k(s) = \max(-k, \min(s, k))$ and $S_k(s) = s - T_k(s)$.

Step 1: estimate on $S_k(u_{\mathcal{T}})$.
Let $k > 0$. We use (4.2.6) with $\varphi_K = S_k(u_K)$; since $(S_k(u_K) - S_k(u_L))^2 \leq (u_K - u_L)(S_k(u_K) - S_k(u_L))$ (S_k is nondecreasing and Lipschitz-continuous with 1 as Lipschitz constant) and $b_K u_K S_k(u_K) \geq 0$ ($S_k(s)$ has the same sign as s), we get

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_k(u_K) - S_k(u_L))^2 \\
&\quad\leq \sum_{K\in\mathcal{T}} \mathrm{meas}(K) f_K S_k(u_K) + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right) (S_k(u_K) - S_k(u_L)) \\
&\qquad + \sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma}\right) (S_k(u_K) - S_k(u_L)).
\end{aligned} \tag{4.3.12}$$

By means of the Cauchy-Schwarz inequality, the discrete Poincaré inequality and (4.3.11), we have

$$\begin{aligned}
\left|\sum_{K\in\mathcal{T}} \mathrm{meas}(K) f_K S_k(u_K)\right| &\leq \left(\sum_{K\in\mathcal{T}} \mathrm{meas}(K) f_K^2\right)^{1/2} \left(\sum_{K\in\mathcal{T}} \mathrm{meas}(K) (S_k(u_K))^2\right)^{1/2} \\
&\leq ||f||_{L^2(\Omega)} ||S_k(u_{\mathcal{T}})||_{L^2(\Omega)} \\
&\leq \mathrm{diam}(\Omega) ||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}}.
\end{aligned} \tag{4.3.13}$$

The Cauchy-Schwarz inequality, associated to (4.3.4) and (4.3.11), gives

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma}\right) (S_k(u_K) - S_k(u_L)) \\
&\quad\leq \left(\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma}\right)^2\right)^{1/2} \left(\sum_{\sigma\in\mathcal{E}} \tau(\sigma) (S_k(u_K) - S_k(u_L))^2\right)^{1/2} \\
&\quad\leq \sqrt{2d} ||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}}.
\end{aligned} \tag{4.3.14}$$

We bound now the convection term, beginning with

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right) (S_k(u_K) - S_k(u_L)) \\
&\quad\leq \left(\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+}\right)^2\right)^{1/2} ||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}}.
\end{aligned} \tag{4.3.15}$$

Since $d_{K,\sigma}/d_\sigma \le 1$ and $d_{L,\sigma}/d_\sigma \le 1$, gathering by control volumes and using Hölder's inequality (with $p/2 > 1$ and $p/(p-2)$), we find

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+} \right)^2 \\
&\quad \le \; 2 \sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma}^2 u_{L,\sigma,+}^2 + \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma}^2 u_{K,\sigma,+}^2 \right) \\
&\quad \le \; 2 \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} v_{K,\sigma}^2 u_{K,\sigma,+}^2 \\
&\quad \le \; 2 \left(\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} |v_{K,\sigma}|^p \right)^{\frac{2}{p}} \left(\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} |u_{K,\sigma,+}|^{\frac{2p}{p-2}} \right)^{\frac{p-2}{p}}. \qquad (4.3.16)
\end{aligned}$$

But, by Jensen's inequality,

$$m(\sigma) d_{K,\sigma} |v_{K,\sigma}|^p \le \frac{m(\sigma) d_{K,\sigma}}{\mathrm{meas}(\triangle_{K,\sigma})} \int_{\triangle_{K,\sigma}} |\mathbf{v}|^p = d \int_{\triangle_{K,\sigma}} |\mathbf{v}|^p$$

so that, since $\{\triangle_{K,\sigma}\,,\ K\in\mathcal{T}\,,\ \sigma\in\mathcal{E}_K\}$ is (up to a set of null Lebesgue measure) a partition of Ω,

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} |v_{K,\sigma}|^p \le d || \, |\mathbf{v}| \, ||_{L^p(\Omega)}^p \le d M^p. \qquad (4.3.17)$$

On the other hand,

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} |u_{K,\sigma,+}|^{\frac{2p}{p-2}} = \sum_{K\in\mathcal{T}} |u_K|^{\frac{2p}{p-2}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d(K,\sigma)$$

where

- $d(K,\sigma) = d_{K,\sigma}$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$ satisfies $v_{K,\sigma} \ge 0$ and $v_{L,\sigma} \ge 0$,
- $d(K,\sigma) = d_{K,\sigma} + d_{L,\sigma} = d_\sigma$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$ satisfies $v_{K,\sigma} \ge 0$ and $v_{L,\sigma} < 0$,
- $d(K,\sigma) = d_{L,\sigma}$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$ satisfies $v_{K,\sigma} < 0$ and $v_{L,\sigma} < 0$,
- $d(K,\sigma) = 0$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$ satisfies $v_{K,\sigma} < 0$ and $v_{L,\sigma} \ge 0$,
- $d(K,\sigma) = d_{K,\sigma}$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$ satisfies $v_{K,\sigma} \ge 0$,
- $d(K,\sigma) = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$ satisfies $v_{K,\sigma} < 0$.

In either case, we have $d(K,\sigma) \le d_\sigma \le \frac{d_{K,\sigma}}{\zeta}$, so that

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} |u_{K,\sigma,+}|^{\frac{2p}{p-2}} \le \frac{1}{\zeta} \sum_{K\in\mathcal{T}} |u_K|^{\frac{2p}{p-2}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} \le \frac{d}{\zeta} ||u_\mathcal{T}||_{L^{\frac{2p}{p-2}}(\Omega)}^{\frac{2p}{p-2}} \qquad (4.3.18)$$

(we have used $\sum_{\sigma\in\mathcal{E}_K} m(\sigma) d_{K,\sigma} = d\mathrm{meas}(K)$).
(4.3.16), (4.3.17) and (4.3.18) together give

$$\left(\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma \left(\frac{d_{L,\sigma}}{d_\sigma} v_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} v_{K,\sigma} u_{K,\sigma,+} \right)^2 \right)^{1/2} \le \frac{\sqrt{2} d^{\frac{1}{p}+\frac{p-2}{2p}}}{\zeta^{\frac{p-2}{2p}}} M ||u_\mathcal{T}||_{L^{\frac{2p}{p-2}}(\Omega)}. \qquad (4.3.19)$$

Since $|u_{\mathcal{T}}| \le k + |S_k(u_{\mathcal{T}})|$, (4.3.15) entails

$$\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(S_k(u_K) - S_k(u_L)) \\ \le \quad C_1 k||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}} + C_1||S_k(u_{\mathcal{T}})||_{L^{\frac{2p}{p-2}}(\Omega)}||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}} \tag{4.3.20}$$

where C_1 only depends on (Ω, p, M, ζ) (a dependence on Ω takes into account a dependence on d). But $p > d$, so that $\frac{2p}{p-2} < \frac{2d}{d-2}$. Let $q \in]\frac{2p}{p-2}, \frac{2d}{d-2}[$ (the choice of such a q only depends on (d,p)). Since $S_k(u_{\mathcal{T}}) = 0$ outside $E_k = \{x \in \Omega \mid |u_{\mathcal{T}}(x)| \ge k\}$, the Hölder inequality and the discrete Sobolev inequality give

$$||S_k(u_{\mathcal{T}})||_{L^{\frac{2p}{p-2}}(\Omega)} \le \mathrm{meas}(E_k)^{\frac{p-2}{2p}-\frac{1}{q}}||S_k(u_{\mathcal{T}})||_{L^q(\Omega)} \le C_2\mathrm{meas}(E_k)^{\frac{p-2}{2p}-\frac{1}{q}}||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}}$$

where C_2 only depends on (Ω, q, ζ) (i.e. on (Ω, p, ζ)). (4.3.20) leads then to

$$\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(S_k(u_K) - S_k(u_L)) \\ \le \quad C_3 k||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}} + C_3\mathrm{meas}(E_k)^{\frac{p-2}{2p}-\frac{1}{q}}||S_k(u_{\mathcal{T}})||^2_{1,\mathcal{T}} \tag{4.3.21}$$

where C_3 only depends on (Ω, p, M, ζ).

Gathering (4.3.13), (4.3.14) and (4.3.21) in (4.3.12), we obtain

$$||S_k(u_{\mathcal{T}})||^2_{1,\mathcal{T}} \le (\mathrm{diam}(\Omega) + \sqrt{2d} + C_3 k)||S_k(u_{\mathcal{T}})||_{1,\mathcal{T}} + C_3\mathrm{meas}(E_k)^{\frac{p-2}{2p}-\frac{1}{q}}||S_k(u_{\mathcal{T}})||^2_{1,\mathcal{T}}. \tag{4.3.22}$$

But, by Tchebycheff's inequality, the discrete Poincaré inequality and Proposition 4.3.1, we have

$$\mathrm{meas}(E_k) = \mathrm{meas}(\{x \in \Omega \mid \ln(1+|u_{\mathcal{T}}(x)|) \ge \ln(1+k)\}) \le \frac{||\ln(1+|u_{\mathcal{T}}|)||^2_{L^2(\Omega)}}{(\ln(1+k))^2} \le \frac{C_4}{(\ln(1+k))^2},$$

where C_4 only depends on (Ω, p, M). Thus, since $\frac{p-2}{2p} - \frac{1}{q} > 0$, we can find k_0 only depending on (Ω, p, M, ζ) such that $C_3\mathrm{meas}(E_{k_0})^{\frac{p-2}{2p}-\frac{1}{q}} < \frac{1}{2}$ and (4.3.22) allows to write

$$||S_{k_0}(u_{\mathcal{T}})||_{1,\mathcal{T}} \le 2(\mathrm{diam}(\Omega) + \sqrt{2d} + C_3 k_0) = C_5 \tag{4.3.23}$$

with C_5 only depending on (Ω, p, M, ζ).

Step 2: estimate on $T_{k_0}(u_{\mathcal{T}})$ and conclusion.
With the k_0 obtained in the previous step, using $\varphi_K = T_{k_0}(u_K)$ in (4.2.6), the fact that $(T_{k_0}(u_K) - T_{k_0}(u_L))^2 \le (u_K - u_L)(T_{k_0}(u_K) - T_{k_0}(u_L))$, that $b_K u_K T_{k_0}(u_K) \ge 0$ and that $|T_{k_0}(u_K)| \le k_0$, we find

$$\begin{aligned} &||T_{k_0}(u_{\mathcal{T}})||^2_{1,\mathcal{T}} \\ &\le \quad k_0||f||_{L^1(\Omega)} + \sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(T_{k_0}(u_K) - T_{k_0}(u_L)) \\ &\quad + \sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(T_{k_0}(u_K) - T_{k_0}(u_L)). \end{aligned} \tag{4.3.24}$$

The Cauchy-Schwarz inequality, (4.3.4) and (4.3.11) lead to

$$\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(T_{k_0}(u_K) - T_{k_0}(u_L)) \le \sqrt{2d}||T_{k_0}(u_{\mathcal{T}})||_{1,\mathcal{T}}. \tag{4.3.25}$$

Thanks to the Cauchy-Schwarz inequality and to (4.3.19), we also have

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(T_{k_0}(u_K) - T_{k_0}(u_L)) \\ \leq\ C_6||u_\mathcal{T}||_{L^{\frac{2p}{p-2}}(\Omega)}||T_{k_0}(u_\mathcal{T})||_{1,\mathcal{T}} \\ \leq\ \left(C_7 + C_7||S_{k_0}(u_\mathcal{T})||_{L^{\frac{2p}{p-2}}(\Omega)}\right)||T_{k_0}(u_\mathcal{T})||_{1,\mathcal{T}}\end{aligned}$$

where C_6 and C_7 only depend on (Ω, p, M, ζ) (we have used $|u_\mathcal{T}| \leq k_0 + |S_{k_0}(u_\mathcal{T})|$). Thanks to the discrete Sobolev inequality (recall that $\frac{2p}{p-2} < \frac{2d}{d-2}$) and to (4.3.23), we deduce that there exists C_8 only depending on (Ω, p, M, ζ) such that

$$\sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(T_{k_0}(u_K) - T_{k_0}(u_L)) \leq C_8||T_{k_0}(u_\mathcal{T})||_{1,\mathcal{T}}.$$

This inequality, injected in (4.3.24) together with (4.3.25), gives $||T_{k_0}(u_\mathcal{T})||_{1,\mathcal{T}} \leq C_9$ with C_9 only depending on (Ω, p, M, ζ).

Since $u_\mathcal{T} = T_{k_0}(u_\mathcal{T}) + S_{k_0}(u_\mathcal{T})$, we deduce that $||u_\mathcal{T}||_{1,\mathcal{T}} \leq C_5 + C_9$, which concludes this proof. ■

4.4 Proof of the existence, uniqueness and convergence result

Proof of Theorem 4.2.1
The existence of a unique solution to (4.2.2)—(4.2.5) is an immediate consequence of the estimate of Theorem 4.3.1: indeed, if $f = G = 0$, then this theorem shows that any solution to (4.2.2)—(4.2.5) is null, that is to say that the square matrix defining this linear system is invertible.

Let us now prove the convergence result.
Since the solution to (4.1.3) is unique (see [31]), it is sufficient to prove that, for any sequence of admissible meshes $(\mathcal{T}_n)_{n\geq 1}$ such that $\mathrm{size}(\mathcal{T}_n) \to 0$ and $\mathrm{reg}(\mathcal{T}_n) \geq \alpha$, we can extract a subsequence (still denoted $(\mathcal{T}_n)_{n\geq 1}$) such that the solution $u_{\mathcal{T}_n}$ to (4.2.2)—(4.2.5) (with $\mathcal{T}_n$ instead of $\mathcal{T}$) converges to the solution of (4.1.3).
Take such a sequence $(\mathcal{T}_n)_{n\geq 1}$. Thanks to Theorem 4.3.1 and to item iii) of Proposition 4.2.1, we see that, up to a subsequence, we can suppose that $u_{\mathcal{T}_n} \to u$ in $L^2(\Omega)$, for some $u \in H_0^1(\Omega)$; by the discrete Sobolev inequality, $(u_{\mathcal{T}_n})_{n\geq 1}$ is also bounded in $L^q(\Omega)$ for all $q < \frac{2d}{d-2}$, so that Vitali's Theorem gives the convergence of $(u_{\mathcal{T}_n})_{n\geq 1}$ to u in $L^q(\Omega)$ for all $q < \frac{2d}{d-2}$.
We are now going to prove that u is a solution to (4.1.3), which is enough, as noticed above, to conclude the proof of the theorem.

To simplify the notation, we forget the index n.
Of course, it is sufficient to prove that u satisfies the equation of (4.1.3) for all $\varphi \in C_c^\infty(\Omega)$. Take such a φ. Using (4.2.6) with $\varphi_K = \varphi(x_K)$, we have

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(x_K) - \varphi(x_L)) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(x_K) \\ &= \sum_{K\in\mathcal{T}} \mathrm{meas}(K) f_K \varphi(x_K) + \sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K) - \varphi(x_L)) \\ &\quad + \sum_{\sigma\in\mathcal{E}} m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(\varphi(x_K) - \varphi(x_L)) \qquad (4.4.26)\end{aligned}$$

(with $\varphi(x_L)=0$ whenever $\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K$).

Step 1: convergence of the diffusion and the lower order terms.
The convergence proof in [38] immediately gives

$$\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi(x_K)-\varphi(x_L))\to\int_\Omega\nabla u\cdot\nabla\varphi\,,\quad\sum_{K\in\mathcal{T}}\mathrm{meas}(K)b_Ku_K\varphi(x_K)\to\int_\Omega bu\varphi$$
$$\text{and}\quad\sum_{K\in\mathcal{T}}\mathrm{meas}(K)f_K\varphi(x_K)\to\int_\Omega f\varphi \tag{4.4.27}$$

as $\mathrm{size}(\mathcal{T})\to 0$ (in fact, to prove the convergence of $\sum_{K\in\mathcal{T}}\mathrm{meas}(K)b_Ku_K\varphi(x_K)$, we must slightly adapt the method of [38], since b is constant in this reference).

Step 2: convergence of the term involving G.
Let us study the convergence of $\sum_{\sigma\in\mathcal{E}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))$. We first notice that, for $\mathrm{size}(\mathcal{T})$ small enough, since φ has a compact support in Ω, this sum is reduced to $\mathcal{E}_{\rm int}$; we take, from now on, $\mathrm{size}(\mathcal{T})$ satisfying this property.
Fix $\varepsilon>0$ and take $H\in(C^1(\overline{\Omega}))^d$ such that $\||G-H|\|_{L^1(\Omega)}\le\varepsilon$; let, for $K\in\mathcal{T}$ and $\sigma\in\mathcal{E}_K$, $H_{K,\sigma}=(\frac{1}{\mathrm{meas}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}H)\cdot\mathbf{n}_{K,\sigma}$.
By regularity of φ and gathering by control volumes, we write

$$\begin{aligned}
&\left|\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))\right.\\
&\qquad\left.-\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}H_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}H_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))\right|\\
&\le C_1\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)d_\sigma\left(\frac{d_{K,\sigma}}{d_\sigma}|G_{K,\sigma}-H_{K,\sigma}|+\frac{d_{L,\sigma}}{d_\sigma}|G_{L,\sigma}-H_{L,\sigma}|\right)\\
&\le C_1\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}m(\sigma)d_{K,\sigma}|G_{K,\sigma}-H_{K,\sigma}|\\
&\le C_1d\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\int_{\triangle_{K,\sigma}}|G-H|\le C_1d\varepsilon
\end{aligned}\tag{4.4.28}$$

where C_1 only depends on φ.
By regularity of H and φ, we have

$$\begin{aligned}
&\left|\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}H_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}H_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))\right.\\
&\qquad\left.-\sum_{\sigma\in\mathcal{E}_{\rm int}}\left(\frac{d_{K,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{K,\sigma}\,d\gamma-\frac{d_{L,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{L,\sigma}\,d\gamma\right)(\varphi(x_K)-\varphi(x_L))\right|\\
&\le C_2\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)d_\sigma\left(\frac{d_{K,\sigma}}{d_\sigma}\mathrm{size}(\mathcal{T})+\frac{d_{L,\sigma}}{d_\sigma}\mathrm{size}(\mathcal{T})\right)\le C_2d\mathrm{meas}(\Omega)\mathrm{size}(\mathcal{T})
\end{aligned}\tag{4.4.29}$$

where C_2 only depends on (H,φ).
Gathering by control volumes and noticing that $\mathbf{n}_{K,\sigma}=-\mathbf{n}_{L,\sigma}$ whenever $\sigma=K|L\in\mathcal{E}_{\rm int}$, we can moreover

write

$$\sum_{\sigma\in\mathcal{E}_{\rm int}}\left(\frac{d_{K,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{K,\sigma}\,d\gamma\ -\frac{d_{L,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{L,\sigma}\,d\gamma\right)(\varphi(x_K)-\varphi(x_L))$$
$$=\sum_{K\in\mathcal{T}}\varphi(x_K)\sum_{\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm int}}\int_\sigma H\cdot\mathbf{n}_{K,\sigma}\,d\gamma$$
$$=\sum_{K\in\mathcal{T}}\varphi(x_K)\int_{\partial K\setminus\partial\Omega}H\cdot\mathbf{n}_{K,\sigma}\,d\gamma.$$

Since $\varphi=0$ on the control volumes $K\in\mathcal{T}$ such that $\partial K\cap\partial\Omega\neq\emptyset$, we have in fact

$$\sum_{\sigma\in\mathcal{E}_{\rm int}}\left(\frac{d_{K,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{K,\sigma}\,d\gamma-\frac{d_{L,\sigma}}{d_\sigma}\int_\sigma H\cdot\mathbf{n}_{L,\sigma}\,d\gamma\right)(\varphi(x_K)-\varphi(x_L))$$
$$=\sum_{K\in\mathcal{T}}\varphi(x_K)\int_{\partial K}H\cdot\mathbf{n}_{K,\sigma}\,d\gamma$$
$$=\sum_{K\in\mathcal{T}}\varphi(x_K)\int_K\mathrm{div}(H)\overset{\mathrm{size}(\mathcal{T})\to0}{\longrightarrow}\int_\Omega\varphi\mathrm{div}(H)=-\int_\Omega H\cdot\nabla\varphi,\tag{4.4.30}$$

the convergence being a consequence of the regularity of φ and H.
We also remark that

$$\left|\int_\Omega H\cdot\nabla\varphi-\int_\Omega G\cdot\nabla\varphi\right|\leq C_3\varepsilon,\tag{4.4.31}$$

where C_3 only depends on φ.
Gathering (4.4.28), (4.4.29), (4.4.30) and (4.4.31), we deduce that

$$\limsup_{\mathrm{size}(\mathcal{T})\to0}\left|\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))-\left(-\int_\Omega G\cdot\nabla\varphi\right)\right|\leq(C_1d+C_3)\varepsilon$$

for all $\varepsilon>0$ and, since C_1 and C_3 only depend on φ, this gives

$$\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)(\varphi(x_K)-\varphi(x_L))\to-\int_\Omega G\cdot\nabla\varphi\tag{4.4.32}$$

as $\mathrm{size}(\mathcal{T})\to0$.

Step 3: convergence of the convective term.
It remains to study the convergence of the term in (4.4.26) coming from the convection, that is to say $\sum_{\sigma\in\mathcal{E}_{\rm int}}m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))$ (the sum is reduced to $\mathcal{E}_{\rm int}$ because $\mathrm{size}(\mathcal{T})$ has been chosen small enough).
Take $\varepsilon>0$ and $\mathbf{w}\in(C^1(\overline{\Omega}))^d$ such that $||\,|\mathbf{v}-\mathbf{w}|\,||_{L^2(\Omega)}\leq\varepsilon$. Let, if $K\in\mathcal{T}$ and $\sigma\in\mathcal{E}_K$, $\mathbf{w}_{K,\sigma}=$

$(\frac{1}{\mathrm{meas}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}\mathbf{w})\cdot\mathbf{n}_{K,\sigma}$. We have, by regularity of φ,

$$\begin{aligned}
&\left|\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right.\\
&\qquad\left.-\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}w_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}w_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right|\\
&\le\ C_1\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)d_\sigma\left(\frac{d_{L,\sigma}}{d_\sigma}|v_{L,\sigma}-w_{L,\sigma}|\,|u_{L,\sigma,+}|+\frac{d_{K,\sigma}}{d_\sigma}|v_{K,\sigma}-w_{K,\sigma}|\,|u_{K,\sigma,+}|\right)\\
&\le\ C_1\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}|v_{K,\sigma}-w_{K,\sigma}|\,|u_{K,\sigma,+}|\\
&\le\ C_1\left(\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}(v_{K,\sigma}-w_{K,\sigma})^2\right)^{1/2}\left(\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}(u_{K,\sigma,+})^2\right)^{1/2}
\end{aligned}$$

(C_1, which only depends on φ, is the same constant as before). The same way we have obtained (4.3.18), we can prove that

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}(u_{K,\sigma,+})^2\le\frac{d}{\zeta}||u_\mathcal{T}||^2_{L^2(\Omega)}\le C_4$$

where C_4 only depends on $(\Omega,p,||\,|\mathbf{v}|\,||_{L^p(\Omega)},\zeta)$ (we use here Theorem 4.3.1 and the discrete Poincaré inequality to obtain a bound on $u_\mathcal{T}$ in $L^2(\Omega)$). Moreover, by Jensen's inequality,

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}(v_{K,\sigma}-w_{K,\sigma})^2\le\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\frac{m(\sigma)d_{K,\sigma}}{\triangle_{K,\sigma}}\int_{\triangle_{K,\sigma}}|\mathbf{v}-\mathbf{w}|^2=d||\,|\mathbf{v}-\mathbf{w}|\,||^2_{L^2(\Omega)}\le d\varepsilon^2.$$

Thus, we have

$$\begin{aligned}
&\left|\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right.\\
&\qquad\left.-\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}w_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}w_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right|\\
&\le\ \varepsilon C_1\sqrt{C_4 d}
\end{aligned}\tag{4.4.33}$$

By regularity of $\mathbf{w}$ and φ, and gathering by control volumes, we find C_5 only depending on $\mathbf{w}$ and φ such that

$$\begin{aligned}
&\left|\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}w_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}w_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right.\\
&\qquad\left.-\sum_{\sigma\in\mathcal{E}_{\rm int}}\left(\frac{d_{L,\sigma}}{d_\sigma}\int_\sigma\mathbf{w}\cdot\mathbf{n}_{L,\sigma}\,d\gamma\,u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}\int_\sigma\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma\,u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))\right|\\
&\le\ C_5\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)d_\sigma\left(\frac{d_{L,\sigma}}{d_\sigma}\mathrm{size}(\mathcal{T})|u_{L,\sigma,+}|+\frac{d_{K,\sigma}}{d_\sigma}\mathrm{size}(\mathcal{T})|u_{K,\sigma,+}|\right)\\
&\le\ C_5\mathrm{size}(\mathcal{T})\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}|u_{K,\sigma,+}|.
\end{aligned}$$

Once again we can prove, the same way we have obtained (4.3.18), that

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} m(\sigma)d_{K,\sigma}|u_{K,\sigma,+}|\le\frac{d}{\zeta}||u_\mathcal{T}||_{L^1(\Omega)},$$

which is bounded by C_6 only depending on $(\Omega, p, ||\,|\mathbf{v}|\,||_{L^p(\Omega)}, \zeta)$. Thus we get

$$\begin{aligned}
&\left| \sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma) \left(\frac{d_{L,\sigma}}{d_\sigma} w_{L,\sigma} u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} w_{K,\sigma} u_{K,\sigma,+} \right) (\varphi(x_K) - \varphi(x_L)) \right. \\
&\qquad \left. - \sum_{\sigma\in\mathcal{E}_{\rm int}} \left(\frac{d_{L,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{L,\sigma}\, d\gamma\, u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\, d\gamma\, u_{K,\sigma,+} \right) (\varphi(x_K) - \varphi(x_L)) \right| \\
&\quad \le \quad C_5 C_6 \mathrm{size}(\mathcal{T}). \qquad (4.4.34)
\end{aligned}$$

Denoting by $\overline{w}_{K,\sigma} = \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\, d\gamma$, $\overline{u}_\sigma = \frac{d_{L,\sigma} u_{L,\sigma,+} + d_{K,\sigma} u_{K,\sigma,+}}{d_\sigma}$ and noticing that $\overline{w}_{K,\sigma} = -\overline{w}_{L,\sigma}$ whenever $\sigma = K|L \in \mathcal{E}_{\rm int}$, we can write

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}_{\rm int}} \left(\frac{d_{L,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{L,\sigma}\, d\gamma\, u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\, d\gamma\, u_{K,\sigma,+} \right) (\varphi(x_K) - \varphi(x_L)) \\
&\qquad = \quad - \sum_{\sigma\in\mathcal{E}_{\rm int}} \overline{w}_{K,\sigma} \overline{u}_\sigma (\varphi(x_K) - \varphi(x_L)).
\end{aligned}$$

Gathering by control volumes (and since $\overline{w}_{K,\sigma} = -\overline{w}_{L,\sigma}$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$), this gives

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}_{\rm int}} \left(\frac{d_{L,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{L,\sigma}\, d\gamma\, u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma} \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\, d\gamma\, u_{K,\sigma,+} \right) (\varphi(x_K) - \varphi(x_L)) \\
&\qquad = \quad - \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm int}} \overline{w}_{K,\sigma} \overline{u}_\sigma \varphi(x_K) \\
&\qquad = \quad - \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} \overline{u}_\sigma \varphi(x_K) \qquad (4.4.35)
\end{aligned}$$

(recall that $\mathrm{size}(\mathcal{T})$ is small enough so that $\varphi(x_K) = 0$ whenever $\mathcal{E}_{\rm ext} \cap \mathcal{E}_K \neq \emptyset$).
The technique is then the same as in [38]: we decompose

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} \overline{u}_\sigma \varphi(x_K) = \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} (\overline{u}_\sigma - u_K) \varphi(x_K) + \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} u_K \varphi(x_K). \qquad (4.4.36)$$

Since $\sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} = \int_K \mathrm{div}(\mathbf{w})$, by convergence of $u_\mathcal{T}$ to u in $L^2(\Omega)$ and by regularity of φ, we have

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} u_K \varphi(x_K) = \sum_{K\in\mathcal{T}} u_K \varphi(x_K) \int_K \mathrm{div}(\mathbf{w}) \longrightarrow \int_\Omega u\varphi \mathrm{div}(\mathbf{w}) \qquad (4.4.37)$$

as $\mathrm{size}(\mathcal{T}) \to 0$. We also have

$$\begin{aligned}
&\left| \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \overline{w}_{K,\sigma} (\overline{u}_\sigma - u_K) \varphi(x_K) \; - \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} (\overline{u}_\sigma - u_K) \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma} \varphi\, d\gamma \right| \\
&\qquad\qquad \le \quad C_1 \mathrm{size}(\mathcal{T}) ||\,|\mathbf{w}|\,||_{C(\overline{\Omega})} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} m(\sigma) |\overline{u}_\sigma - u_K|.
\end{aligned}$$

But $\overline{u}_\sigma$ is a convex combinaison of (u_K, u_L) if $\sigma = K|L \in \mathcal{E}_{\rm int}$, and $\overline{u}_\sigma \in \{0, u_K\}$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, so that,

in either case, $|\overline{u}_\sigma - u_K| \leq D_\sigma u_\mathcal{T}$ and

$$\begin{aligned}\left|\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\overline{w}_{K,\sigma}(\overline{u}_\sigma-u_K)\varphi(x_K) - \sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}(\overline{u}_\sigma-u_K)\int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\varphi\,d\gamma\right| \\ \leq\ C_1\mathrm{size}(\mathcal{T})||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}m(\sigma)D_\sigma u_\mathcal{T} \\ \leq\ 2C_1\mathrm{size}(\mathcal{T})||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\sum_{\sigma\in\mathcal{E}}m(\sigma)D_\sigma u_\mathcal{T} \\ \leq\ 2C_1\mathrm{size}(\mathcal{T})||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\left(\sum_{\sigma\in\mathcal{E}}m(\sigma)d_\sigma\right)^{1/2}||u_\mathcal{T}||_{1,\mathcal{T}}.\end{aligned}$$

$||u_\mathcal{T}||_{1,\mathcal{T}}$ being bounded as $\mathrm{size}(\mathcal{T})\to 0$ and $\sum_{\sigma\in\mathcal{E}}m(\sigma)d_\sigma$ being constant (it is $d\mathrm{meas}(\Omega)$), we deduce that

$$\lim_{\mathrm{size}(\mathcal{T})\to 0}\left(\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\overline{w}_{K,\sigma}(\overline{u}_\sigma-u_K)\varphi(x_K) - \sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}(\overline{u}_\sigma-u_K)\int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\varphi\,d\gamma\right)=0. \tag{4.4.38}$$

We have, gathering by edges and since $\varphi=0$ on σ whenever $\sigma\in\mathcal{E}_{\mathrm{ext}}$,

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\overline{u}_\sigma\int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\varphi\,d\gamma = \sum_{\sigma\in\mathcal{E}_{\mathrm{int}}}\overline{u}_\sigma\left(\int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\varphi\,d\gamma+\int_\sigma \mathbf{w}\cdot\mathbf{n}_{L,\sigma}\varphi\,d\gamma\right)=0$$

since $\mathbf{n}_{K,\sigma}=-\mathbf{n}_{L,\sigma}$ if $\sigma=K|L\in\mathcal{E}_{\mathrm{int}}$. Moreover,

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}u_K\int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\varphi\,d\gamma = \sum_{K\in\mathcal{T}}u_K\int_K \mathrm{div}(\varphi\mathbf{w})\longrightarrow\int_\Omega u\mathrm{div}(\varphi\mathbf{w})$$

as $\mathrm{size}(\mathcal{T})\to 0$. Thus, (4.4.38) implies

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\overline{w}_{K,\sigma}(\overline{u}_\sigma-u_K)\varphi(x_K)\longrightarrow-\int_\Omega u\mathrm{div}(\varphi\mathbf{w})$$

as $\mathrm{size}(\mathcal{T})\to 0$. Together with (4.4.34), (4.4.35), (4.4.36) and (4.4.37), this gives

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}_{\mathrm{int}}}m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}w_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}w_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L)) \\ \longrightarrow\ -\int_\Omega u\varphi\mathrm{div}(\mathbf{w})+\int_\Omega u\mathrm{div}(\varphi\mathbf{w})=\int_\Omega u\mathbf{w}\cdot\nabla\varphi\end{aligned} \tag{4.4.39}$$

as $\mathrm{size}(\mathcal{T})\to 0$.
By noticing that

$$\left|\int_\Omega u\mathbf{w}\cdot\nabla\varphi-\int_\Omega u\mathbf{v}\cdot\nabla\varphi\right|\leq||u||_{L^2(\Omega)}||\,|\mathbf{v}-\mathbf{w}|\,||_{L^2(\Omega)}||\,|\nabla\varphi|\,||_{L^\infty(\Omega)}\leq C_7\varepsilon$$

where C_7 only depends on u and φ, (4.4.33) and (4.4.39) allow to write

$$\begin{aligned}\limsup_{\mathrm{size}(\mathcal{T})\to 0}\left|\sum_{\sigma\in\mathcal{E}_{\mathrm{int}}}m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+}-\frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L))-\int_\Omega u\mathbf{v}\cdot\nabla\varphi\right| \\ \leq\ (C_1\sqrt{C_4 d}+C_7)\varepsilon\end{aligned}$$

for all $\varepsilon > 0$ and with C_1, C_4 and C_7 not depending on ε, that is to say

$$\sum_{\sigma\in\mathcal{E}_{\rm int}} m(\sigma)\left(\frac{d_{L,\sigma}}{d_\sigma}v_{L,\sigma}u_{L,\sigma,+} - \frac{d_{K,\sigma}}{d_\sigma}v_{K,\sigma}u_{K,\sigma,+}\right)(\varphi(x_K)-\varphi(x_L)) \longrightarrow \int_\Omega u\mathbf{v}\cdot\nabla\varphi \tag{4.4.40}$$

as size$(\mathcal{T})\to 0$.

Gathering (4.4.27), (4.4.32) and (4.4.40) in (4.4.26), we see that u satisfies the equation of (4.1.3). ■

4.5 Another scheme

The scheme of Section 4.2 is based on a discretization of (4.1.1) that brings in approximate values of $\int_\sigma \mathbf{v}\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ and $\int_\sigma G\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ based on the values of $\mathbf{v}$ and G on a subset of K (the "half-diamond"). The choice of such approximate values seems to be quite adapted when there is a link between the mesh and $\mathbf{v}$ or G: for example, if $\mathbf{v}$ or G is constant on each side of an hyperplane and if we take meshes such that each control volume is on one side of this hyperplane.
But when there is no relation between $\mathbf{v}$ or G and the mesh, the reasons for using the values of $\mathbf{v}$ or G only on K to approximate $\int_\sigma \mathbf{v}\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ or $\int_\sigma G\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ are not so clear: we could approximate $\mathbf{v}$ or G on σ by some quantity $\mathbf{v}_\sigma$ or G_σ, and then consider $m(\sigma)\mathbf{v}_\sigma\cdot\mathbf{n}_{K,\sigma}$ or $m(\sigma)G_\sigma\cdot\mathbf{n}_{K,\sigma}$ as a coherent approximate value of $\int_\sigma \mathbf{v}\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ or $\int_\sigma G\cdot\mathbf{n}_{K,\sigma}\,d\gamma$. This is what the following scheme does.

Let $\mathcal{T}$ be an admissible mesh. If $\sigma = K|L\in\mathcal{E}_{\rm int}$, we define the "full-diamond" around σ by $\triangle_\sigma = \triangle_{K,\sigma}\cup\triangle_{L,\sigma}$; if $\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K$, the "full-diamond" around σ is simply $\triangle_\sigma = \triangle_{K,\sigma}$. We let then, for $K\in\mathcal{T}$ and $\sigma\in\mathcal{E}$,

$$\mathbf{v}_\sigma = \frac{1}{\mathrm{meas}(\triangle_\sigma)}\int_{\triangle_\sigma}\mathbf{v}\,,\quad b_K = \frac{1}{\mathrm{meas}(K)}\int_K b\,,$$
$$f_K = \frac{1}{\mathrm{meas}(K)}\int_K f \quad\text{and}\quad G_\sigma = \frac{1}{\mathrm{meas}(\triangle_\sigma)}\int_{\triangle_\sigma} G.$$

The new scheme for (4.1.1) is

$$\begin{aligned}\forall K\in\mathcal{T}\,,\ &\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)\mathbf{v}_\sigma\cdot\mathbf{n}_{K,\sigma}u_{\sigma,+} + \mathrm{meas}(K)b_Ku_K\\ &\qquad = \mathrm{meas}(K)f_K + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)G_\sigma\cdot\mathbf{n}_{K,\sigma}\,,\end{aligned} \tag{4.5.1}$$

$$\begin{aligned}&\forall K\in\mathcal{T}\,,\forall\sigma = K|L\in\mathcal{E}_K\cap\mathcal{E}_{\rm int}\,,\quad F_{K,\sigma} = \tfrac{m(\sigma)}{d_\sigma}(u_K-u_L)\,,\\ &\forall K\in\mathcal{T}\,,\forall\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm ext}\,,\qquad\quad F_{K,\sigma} = \tfrac{m(\sigma)}{d_\sigma}u_K\,,\end{aligned} \tag{4.5.2}$$

$$\begin{aligned}&\forall\sigma = K|L\in\mathcal{E}_{\rm int}\,,\quad u_{\sigma,+} = u_K \ \text{ if }\ \mathbf{v}_\sigma\cdot\mathbf{n}_{K,\sigma}\ge 0\,,\quad u_{\sigma,+} = u_L\ \text{ otherwise,}\\ &\forall\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K\,,\quad u_{\sigma,+} = u_K \ \text{ if }\ \mathbf{v}_\sigma\cdot\mathbf{n}_{K,\sigma}\ge 0\,,\quad u_{\sigma,+} = 0\ \text{ otherwise.}\end{aligned} \tag{4.5.3}$$

In fact, we can remark that (4.5.1)—(4.5.3) is exactly (4.2.2)—(4.2.5), provided that we define $v_{K,\sigma} = \mathbf{v}_\sigma\cdot\mathbf{n}_{K,\sigma}$, $G_{K,\sigma} = G_\sigma\cdot\mathbf{n}_{K,\sigma}$ and let $u_{K,\sigma,+} = u_{\sigma,+}$ (for all $K\in\mathcal{T}$ and all $\sigma\in\mathcal{E}_K$). Indeed, in this case, if $\sigma = K|L\in\mathcal{E}_{\rm int}$, we have $v_{K,\sigma} = -v_{L,\sigma}$, so that (4.5.3) is equivalent to (4.2.5) (with the notation $u_{K,\sigma,+} = u_{\sigma,+}$), and $G_{K,\sigma} = -G_{L,\sigma}$, so that (4.2.4) comes down to $F_{K,\sigma} = -F_{L,\sigma}$ (or $u_\sigma = 0$ if $\sigma\in\mathcal{E}_{\rm ext}$) which, associated to (4.2.3), is equivalent to (4.5.2).
Thus, we easily see that the preceding techniques to obtain *a priori* estimates on the solutions to (4.2.2)—(4.2.5) give us estimates on the solutions to (4.5.1)—(4.5.3), which proves the existence and uniqueness of the solution to this problem. The convergence proof also works as before, and we deduce that, if $\alpha > 0$ is fixed and $u_\mathcal{T}$ denotes the solution to (4.5.1)—(4.5.3), then $u_\mathcal{T}$ converges in $L^q(\Omega)$, for all $q < \frac{2d}{2-d}$, to the unique solution of (4.1.3), as size$(\mathcal{T})\to 0$ with reg$(\mathcal{T})\ge\alpha$.

Chapitre 5

Error estimates for the convergence of a finite volume discretization of convection-diffusion equations

Reference: J. Droniou. *J. Numer. Math.* **11** (2003), no. 1, 1-32.

Abstract: we study error estimates for a finite volume discretization of an elliptic equation. We prove that, for $s \in [0,1]$, if the exact solution belongs to H^{1+s} and the right-hand side is $f + \mathrm{div}(G)$ with $f \in L^2$ and $G \in (H^s)^N$, then the solution of the finite volume scheme converges in discrete H^1-norm to the exact solution, with a rate of convergence of order h^s (where h is the size of the mesh).

5.1 Introduction

5.1.1 The problem

Let Ω be a polygonal open subset of $\mathbb{R}^N$ ($N = 2$ or 3). We study a finite volume discretization of

$$\begin{cases} -\Delta \overline{u} + \mathrm{div}(\mathbf{v}\overline{u}) + b\overline{u} = f + \mathrm{div}(G) & \text{in } \Omega, \\ \overline{u} = 0 & \text{on } \partial\Omega \end{cases} \tag{5.1.1}$$

where $\mathbf{v} \in (C(\overline{\Omega}))^N$, $b \in L^\infty(\Omega)$ is nonnegative, $f \in L^2(\Omega)$ and $G \in (H^s(\Omega))^N$ with $s \in [0,1]$ (if $s = 0$, $H^0(\Omega)$ is to be understood as $L^2(\Omega)$).
The solution to (5.1.1) is taken in a weak sense as in [31], that is to say $\overline{u} \in H^1_0(\Omega)$ and the partial differential equation is satisfied in the distributional sense.

Finite volume methods have been widely used to approximate the solutions of convection-diffusion equations (see e.g. [38], [48], [33]...). The convergence of the approximations is well-known (see e.g. [38] (Theorem 9.1)) and some error estimates in the H^2 framework have been obtained in [48] (Theorem 3.2). These schemes have mainly been considered when the right-hand side of the elliptic equation belongs to $L^2(\Omega)$ (i.e. $G = 0$ in (5.1.1)); but, recently, [33] has presented a finite volume scheme capable of handling (5.1.1) with any $G \in (L^2(\Omega))^N$. The case $G \in (H^1(\Omega))^N$ being (roughly speaking) the H^2 framework studied in [48], it seems natural to hope, via interpolation techniques, for error estimates when G belongs to intermediate spaces between $L^2(\Omega)$ and $H^1(\Omega)$ (these estimates are well-known for the finite element methods, see for example [17] (Theorem 5.1)). We prove here such error estimates, thus filling a gap between finite volume methods and finite element methods.

It can be interesting to notice that, in the case $N = 2$ and for "finite volume element" schemes (which are somewhat a mixing between the finite element methods and the finite volume methods, and are different from the ones we present here), some error estimates in the H^{1+s} framework have been obtained in [21] (Theorem 4.1, p. 176), when $s \in]1/2, 1]$.
We also emphasize on a noticeable feature of (5.1.1): its non-coercivity. It is not supposed that $\frac{1}{2}\mathrm{div}(\mathbf{v}) + b \geq 0$, so that the bilinear form associated to (5.1.1) may be non-coercive. However, under the sole hypotheses stated after (5.1.1), existence and uniqueness of a solution to this equation is known (see [31] (Theorem 2.1)). The *a priori* estimates on this equation and its discretization are harder to obtain than in the coercive case, but the techniques that give such estimates are now well-known (see [31], [33]).

In the following subsection, we present the finite volume scheme used to discretize (5.1.1); this scheme is in fact a simplified version of the one presented in [33] (simplified because we take into account the additional regularity we have on $\mathbf{v}$ with respect to [33]). In Section 2, we state the main result concerning estimates on the difference between the approximate solution and the exact solution; notice that we have to use a different discretization of the exact solution than in [48], because we do not only intend to study the case when this solution belongs to $H^2(\Omega)$, but also when it belongs to $H^{1+s}(\Omega)$ for $s \in [0, 1]$; hence we cannot, in contrary to [48], discretize the solution by taking its values on points. In Section 3, we study the case when the solution belongs to H^1; in this case, the "error estimate" reduces to a bound in $\mathcal{O}(1)$. In Section 4, we prove a $\mathcal{O}(h)$ convergence when the exact solution belongs to $H^2(\Omega)$ and G belongs to $(H^1(\Omega))^N$; note that, since our discretization of the solution is not the same as in [48], we cannot directly refer to this paper and we must re-make the whole work (moreover, our scheme is different to the one presented in [48], because of the presence of G); in particular, it appears through this study that the way we discretize the solution is crucial to obtain good error estimates. In Section 5, we use the results of Sections 3 and Sections 4 to prove, via interpolation results, the theorem stated in Section 2. Section 6 presents some numerical results, and Section 7 is an appendix which gathers some technical lemmas useful throughout the paper.

5.1.2 Definition of the scheme

We use meshes of Ω similar to the ones presented in [38].

Definition 5.1.1 *An admissible mesh $\mathcal{T}$ of Ω is a finite family of polygonal open convex subsets of Ω (the "control volumes"), together with a finite family $\mathcal{E}$ of disjoint subsets of $\overline{\Omega}$ consisting in non-empty open convex subsets of affine hyperplanes (the "edges") and a family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points in Ω such that*

i) $\overline{\Omega} = \bigcup_{K\in\mathcal{T}} \overline{K}$,

ii) each $\sigma \in \mathcal{E}$ is contained in ∂K for some $K \in \mathcal{T}$,

iii) by denoting $\mathcal{E}_K = \{\sigma \in \mathcal{E} \mid \sigma \subset \partial K\}$, $\partial K = \cup_{\sigma\in\mathcal{E}_K} \overline{\sigma}$ for all $K \in \mathcal{T}$,

iv) for all $K \neq L$ in $\mathcal{T}$, either the $(N-1)$-dimensional measure of $\overline{K} \cap \overline{L}$ is null, or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, that we denote then $\sigma = K|L$,

v) for all $K \in \mathcal{T}$, $x_K \in K$,

vi) for all $\sigma = K|L \in \mathcal{E}$, the line (x_K, x_L) intersects and is orthogonal to σ,

vii) for all $\sigma \in \mathcal{E}$, $\sigma \subset \partial\Omega \cap \partial K$, the line which is orthogonal to σ and going through x_K intersects σ.

If $K \in \mathcal{T}$, h_K denotes the diameter of K; the size of $\mathcal{T}$ is $h_\mathcal{T} = \sup_{K\in\mathcal{T}} h_K$. The unit normal to $\sigma \in \mathcal{E}_K$ outward to K is denoted by $\mathbf{n}_{K,\sigma}$.
We define $\mathcal{E}_{\mathrm{int}} = \{\sigma \in \mathcal{E} \mid \sigma \notin \partial\Omega\}$ (interior edges) and $\mathcal{E}_{\mathrm{ext}} = \mathcal{E}\backslash\mathcal{E}_{\mathrm{int}}$. We denote by m the $(N-1)$-dimensional measure on the edges of the mesh so that, if $\sigma \in \mathcal{E}$, $m(\sigma)$ is the $(N-1)$-dimensional measure

of σ. If $\sigma = K|L \in \mathcal{E}_{\rm int}$, d_σ is the Euclidean distance between the points (x_K, x_L) and $d_{K,\sigma}$ denotes the distance between x_K and σ; if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, $d_\sigma = d_{K,\sigma}$ is the distance between x_K and σ.
If $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, the "half-diamond" $\triangle_{K,\sigma}$ is defined by $\triangle_{K,\sigma} = \{tx_K + (1-t)x\,,\ t \in]0,1[\,,\ x \in \sigma\}$ (the convex hull of $\{x_K\} \cup \sigma$). We notice that $|\triangle_{K,\sigma}| = m(\sigma)d_{K,\sigma}/N$ (where $|\cdot|$ denotes the Lebesgue measure in $\mathbb{R}^N$).

We also make the following hypotheses on the meshes:

$$\exists \zeta > 0 \text{ such that } \forall K \in \mathcal{T}\,,\ d_{K,\sigma} \geq \zeta d_\sigma\,, \tag{5.1.2}$$

$$\exists \alpha > 0 \text{ such that } \forall K \in \mathcal{T}\,,\ B(x_K, \alpha h_K) \subset K\,, \tag{5.1.3}$$

$$\exists M > 0 \text{ such that } \forall K \in \mathcal{T}\,,\ \mathrm{card}(\mathcal{E}_K) \leq M \tag{5.1.4}$$

($B(x,\eta)$ denotes the ball of center x and radius η).
Hypothesis (5.1.2) is classical when discrete Sobolev inequalities are needed. These inequalities are useful in *a priori* estimates on the scheme (to control the convective term of the equation), which appear in [33] or [32]. Here, we will directly use the results of [32], hence the need for (5.1.2) will not be glaring. Notice that if $\mathbf{v} = 0$ (or $\mathbf{v}$ is regular and $\mathrm{div}(\mathbf{v}) \geq 0$), then Hypothesis (5.1.2) can be dropped.
(5.1.3) and (5.1.4) appear in [38] for the same kind of results that we present here (see Remarks 5.3.1 and 5.4.3).

Remark 5.1.1 *As an example of admissible meshes, we can take regular uniform meshes as in Section 5.6, but also triangular meshes (provided that all angles of the triangles are less than $\pi/2$ — this can be relaxed, see Example 9.1 in [38]) and most of Voronoï meshes (see Example 9.2 in [38]).*

The finite volume discretization is obtained by an integration of $-\Delta\overline{u} + \mathrm{div}(\mathbf{v}\overline{u}) + b\overline{u} = f + \mathrm{div}(G)$ on a control volume K: with some integrates by parts, we formally obtain

$$\sum_{\sigma\in\mathcal{E}_K} -\int_\sigma \nabla\overline{u}\cdot\mathbf{n}_{K,\sigma}\,dm + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma \overline{u}\mathbf{v}\cdot\mathbf{n}_{K,\sigma}\,dm + \int_K b\overline{u} = \int_K f + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma G\cdot\mathbf{n}_{K,\sigma}\,dm.$$

To discretize this equation, we define, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$v_{K,\sigma} = \left(\frac{1}{m(\sigma)}\int_\sigma \mathbf{v}(\xi)\,dm(\xi)\right)\cdot\mathbf{n}_{K,\sigma}\,,\quad b_K = \frac{1}{|K|}\int_K b(x)\,dx\,,$$
$$f_K = \frac{1}{|K|}\int_K f(x)\,dx \quad\text{and}\quad G_{K,\sigma} = \left(\frac{1}{|\triangle_{K,\sigma}|}\int_{\triangle_{K,\sigma}} G(x)\,dx\right)\cdot\mathbf{n}_{K,\sigma}\,,$$

which are approximate values of $\mathbf{v}\cdot\mathbf{n}_{K,\sigma}$ and $G\cdot\mathbf{n}_{K,\sigma}$ on σ, and of b and f on K.
Then, letting u_K and u_σ be approximate values of $\overline{u}$ on K and σ, the finite volume discretization of (5.1.1) is written

$$\forall K \in \mathcal{T}\,,\ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)v_{K,\sigma}u_{\sigma,+} + |K|b_Ku_K = |K|f_K + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)G_{K,\sigma}\,, \tag{5.1.5}$$

$$\forall K \in \mathcal{T}\,, \forall \sigma \in \mathcal{E}_K\,,\quad F_{K,\sigma} = -\tfrac{m(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)\,, \tag{5.1.6}$$

$$\begin{aligned} &\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,,\quad F_{K,\sigma} - m(\sigma)G_{K,\sigma} = -(F_{L,\sigma} - m(\sigma)G_{L,\sigma})\,,\\ &\forall \sigma \in \mathcal{E}_{\rm ext}\,,\qquad\qquad u_\sigma = 0\,, \end{aligned} \tag{5.1.7}$$

$$\begin{aligned} &\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,,\quad u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0\,,\quad u_{\sigma,+} = u_L \text{ otherwise,}\\ &\forall \sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K\,,\quad u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0\,,\quad u_{\sigma,+} = 0 \text{ otherwise} \end{aligned} \tag{5.1.8}$$

($F_{K,\sigma}$ is of course a discretization of $-\int_\sigma \nabla u\cdot\mathbf{n}_{K,\sigma}\,dm$).

Remark 5.1.2 *(5.1.8) is a classical upwind choice of the discretization of the convection term* $\mathrm{div}(\mathbf{v}\overline{u})$ *(see [38], p. 766). This choice brings stability to the scheme and allows unconditional* a priori *estimates (see [32] (Proposition 3.2, p. 72)); notice however that, since our problem is non-coercive (i.e.* $\mathrm{div}(\mathbf{v})$ *is not supposed nonnegative), the maximum principle on the scheme (5.1.5)—(5.1.8) is not known.*

The scheme (5.1.5)—(5.1.8) is not exactly the same as in [33], because $\mathbf{v}$ has not been discretized the same way. In fact, in [33], $\mathbf{v}$ is less regular, so it must be discretized using mean values on $\triangle_{K,\sigma}$, not on σ; to obtain error estimates, we must assume here that $\mathbf{v}$ is more regular than in [33], so it seems more natural (and easier!), since the regularity of $\mathbf{v}$ allows it, to consider mean values of $\mathbf{v}$ on σ rather than $\triangle_{K,\sigma}$.
In fact, the unknowns $(u_\sigma)_{\sigma\in\mathcal{E}}$ in (5.1.5)—(5.1.8) can be immediately eliminated thanks to (5.1.7), and (5.1.5)—(5.1.8) reduces thus to a system with unknowns $(u_K)_{K\in\mathcal{T}}$, which reads

$$\begin{aligned}\forall K\in\mathcal{T}\,,\ &\sum_{\sigma\in\mathcal{E}_K}\frac{m(\sigma)}{d_\sigma}(u_K-u_L)+\sum_{\sigma\in\mathcal{E}_K}m(\sigma)v_{K,\sigma}u_{\sigma,+}+|K|b_Ku_K\\ &\qquad=|K|f_K+\sum_{\sigma\in\mathcal{E}_K}m(\sigma)\left(\frac{d_{K,\sigma}}{d_\sigma}G_{K,\sigma}-\frac{d_{L,\sigma}}{d_\sigma}G_{L,\sigma}\right)\\ &\text{(where }\sigma=K|L\text{ if }\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm int}\text{ and }u_L=d_{L,\sigma}=G_{L,\sigma}=0\text{ if }\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm ext})\,,\end{aligned}\tag{5.1.9}$$

$$\begin{aligned}&\forall\sigma=K|L\in\mathcal{E}_{\rm int}\,,\quad u_{\sigma,+}=u_K\ \text{ if }\ v_{K,\sigma}\geq 0\,,\quad u_{\sigma,+}=u_L\ \text{ otherwise,}\\ &\forall\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K\,,\quad u_{\sigma,+}=u_K\ \text{ if }\ v_{K,\sigma}\geq 0\,,\quad u_{\sigma,+}=0\ \text{ otherwise.}\end{aligned}\tag{5.1.10}$$

A priori estimates on (5.1.9)—(5.1.10) are then direct consequences of Proposition 3.2, p.72, in [32] (see (5.3.6) in the proof of Corollary 5.3.1 below). These estimates show that the linear system (5.1.9)—(5.1.10) is invertible and thus that there exists a unique solution $(u_K)_{K\in\mathcal{T}}$ to this system.

Remark 5.1.3 *In [70], some non-coercive problems are also handled from a numerical point of view. However, the scheme used (namely a Finite Volume Element scheme) is not a Finite Volume scheme as the one we present here, and the regularity on the datas (*$\mathbf{v}$ *and the mesh) are quite stronger. Moreover, in this reference, the existence of a solution to the Finite Volume Element scheme is obtained only for* $h_\mathcal{T}$ *small enough.*

5.2 Statement of the main result

The discretization $\overline{u}_\mathcal{T}=(\overline{u}_K)_{K\in\mathcal{T}}$ of the exact solution $\overline{u}$ on an admissible mesh $\mathcal{T}$ is defined by

$$\forall K\in\mathcal{T}\,,\ \overline{u}_K=\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)}\overline{u}(x)\,dx.\tag{5.2.1}$$

A more natural way to discretize the exact solution would perhaps be to take the mean value of $\overline{u}$ on each cell K. This is not a problem when handling the H^1 case (see Remark 5.3.1), but such a discretization would lead to bad constitency errors in the H^2 case (see Remark 5.4.1).

If $\mathcal{T}$ is an admissible mesh, we identify the elements $(v_K)_{K\in\mathcal{T}}\in\mathbb{R}^{\mathrm{Card}(\mathcal{T})}$ to functions $v_\mathcal{T}$ defined a.e. on Ω and constant on each cell $K\in\mathcal{T}$. We denote by $X(\mathcal{T})$ this space of functions, and it is endowed with the discrete H^1-norm (which is a natural norm when considering Finite Volume discretizations of elliptic equations):

$$||v_\mathcal{T}||_{1,\mathcal{T}}=\left(\sum_{\sigma\in\mathcal{E}}\frac{m(\sigma)}{d_\sigma}(D_\sigma v_\mathcal{T})^2\right)^{1/2}$$

where $D_\sigma v_\mathcal{T}=|v_K-v_L|$ if $\sigma=K|L\in\mathcal{E}_{\rm int}$ and $D_\sigma v_\mathcal{T}=|v_K|$ if $\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\rm ext}$.

In the sequel, all the estimates will be made through this norm. Notice that $||\cdot||_{L^2(\Omega)} \leq \mathrm{diam}(\Omega)||\cdot||_{1,\mathcal{T}}$ on $X(\mathcal{T})$ (see [38] (Lemma 9.1) for a proof of this), so that an estimate in $X(\mathcal{T})$ gives a similar estimate in $L^2(\Omega)$.

The main result of this paper is the following theorem.

Theorem 5.2.1 *We suppose that $N = 2$ or that Ω is convex; we also assume that $\mathrm{div}(\mathbf{v}) \in L^2(\Omega)$. Let $s \in [0,1]$ and $\mathcal{T}$ be an admissible mesh which satisfies Hypotheses (5.1.2), (5.1.3) and (5.1.4). Then there exists C only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, ||b||_{L^\infty(\Omega)}, \zeta, \alpha, M)$ such that, if $G \in (H^s(\Omega))^N$ and if the solution $\overline{u}$ to (5.1.1) belongs to $H^{1+s}(\Omega)$, we have*

$$||\overline{u}_\mathcal{T} - u_\mathcal{T}||_{1,\mathcal{T}} \leq C \left(||\overline{u}||_{H^{1+s}(\Omega)} + ||G||_{(H^s(\Omega))^N} + ||f||_{L^2(\Omega)} \right) {h_\mathcal{T}}^s$$

where $\overline{u}_\mathcal{T} = (\overline{u}_K)_{K\in\mathcal{T}}$ is defined by (5.2.1) and $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ is the solution to (5.1.9)—(5.1.10).

Remark 5.2.1 *The hypotheses "$N = 2$ or Ω is convex" and "$\mathrm{div}(\mathbf{v}) \in L^2(\Omega)$" are technical hypotheses useful to identify interpolation spaces (see the proof of Theorem 5.2.1 and Subsection 5.7.2). In fact, we believe that these hypotheses are not necessary to compute the interpolation spaces that appear in our work, but we have found no result in interpolation literature that allows to get rid of them (for example, to be able to handle non-convex polygonal open sets in dimension 2, we use [2] whose generalization to $N = 3$ does not seem easy at all).*
It is also to be noticed that these hypotheses are useless if $s = 1$ (see Theorem 5.4.1).

Remark 5.2.2 *Notice that, with our hypotheses, Ω can be a non-convex polygonal open set of $\mathbb{R}^2$, so that assuming $G \in (H^s(\Omega))^N$ does not necessarily implies $\overline{u} \in H^{1+s}(\Omega)$.*

5.3 The H^1 framework

Proposition 5.3.1 *If $\mathcal{T}$ is an admissible mesh which satisfies Hypotheses (5.1.3) and (5.1.4), there exists C only depending on (N, α, M) such that $||\overline{u}_\mathcal{T}||_{1,\mathcal{T}} \leq C||\overline{u}||_{H_0^1(\Omega)}$.*

Remark 5.3.1 *In [38] (Lemma 9.4), a similar result is proved (also using Hypotheses (5.1.3) and (5.1.4)) with $\overline{u}_K$ replaced by the mean value of $\overline{u}$ on K.*

Proof of Proposition 5.3.1
Let $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$. We have

$$\begin{aligned}
&\left| \frac{1}{|B(x_K, \alpha h_K)|} \int_{B(x_K,\alpha h_K)} \overline{u}(x)\, dx - \frac{1}{m(\sigma)} \int_\sigma \overline{u}(\xi)\, dm(\xi) \right| \\
&\leq \left| \frac{1}{|B(x_K, \alpha h_K)|} \int_{B(x_K,\alpha h_K)} \overline{u}(x)\, dx - \frac{1}{|\triangle_{K,\sigma}|} \int_{\triangle_{K,\sigma}} \overline{u}(y)\, dy \right| \\
&\quad + \left| \frac{1}{|\triangle_{K,\sigma}|} \int_{\triangle_{K,\sigma}} \overline{u}(y)\, dy - \frac{1}{m(\sigma)} \int_\sigma \overline{u}(\xi)\, dm(\xi) \right| \qquad (5.3.2)
\end{aligned}$$

Since $B(x_K, \alpha h_K)$ and $\triangle_{K,\sigma}$ are both contained in K which is convex and has diameter h_K, Lemma 5.7.1 in the Appendix allows to write

$$\begin{aligned}
&\left| \frac{1}{|B(x_K, \alpha h_K)|} \int_{B(x_K,\alpha h_K)} \overline{u}(x)\, dx - \frac{1}{|\triangle_{K,\sigma}|} \int_{\triangle_{K,\sigma}} \overline{u}(y)\, dy \right|^2 \\
&\leq \frac{C_1 h_K^{N+2}}{|B(x_K, \alpha h_K)|\,|\triangle_{K,\sigma}|} \int_{\mathrm{co}(B(x_K,\alpha h_K)\cup\triangle_{K,\sigma})} |\nabla\overline{u}(z)|^2\, dz \\
&\leq \frac{C_2 h_K^2}{|\triangle_{K,\sigma}|} \int_K |\nabla\overline{u}(z)|^2\, dz
\end{aligned}$$

with C_1 and C_2 only depending on (N,α). Using this inequality and Lemma 5.7.2 (from the Appendix) in (5.3.2), we obtain

$$\left| \frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} \overline{u}(x)\,dx - \frac{1}{m(\sigma)}\int_\sigma \overline{u}(\xi)\,dm(\xi)\right|^2 \le \frac{C_3 h_K^2}{|\triangle_{K,\sigma}|}\int_K |\nabla\overline{u}(z)|^2\,dz \tag{5.3.3}$$

where C_3 only depends on (N,α).

Let $\sigma\in\mathcal{E}_{\rm ext}$. Since $\overline{u}\in H_0^1(\Omega)$, (5.3.3) shows that, denoting by K the cell such that $\sigma\in\mathcal{E}_K$, $(D_\sigma\overline{u}_\mathcal{T})^2 \le \frac{C_3h_K^2}{|\triangle_{K,\sigma}|}\int_K|\nabla\overline{u}(x)|^2\,dx$. We have $|\triangle_{K,\sigma}| = m(\sigma)d_\sigma/N$ (recall that $d_{K,\sigma}=d_\sigma$ since $\sigma\in\mathcal{E}_{\rm ext}$); moreover, by Hypothesis (5.1.3), $\alpha h_K \le d_{K,\sigma}=d_\sigma$; thus,

$$\frac{m(\sigma)}{d_\sigma}(D_\sigma\overline{u}_\mathcal{T})^2 \le \frac{NC_3h_K^2}{d_\sigma^2}\int_K|\nabla\overline{u}(x)|^2\,dx \le \frac{NC_3}{\alpha^2}\int_K|\nabla\overline{u}(x)|^2\,dx. \tag{5.3.4}$$

Let $\sigma=K|L\in\mathcal{E}_{\rm int}$. We have

$$\begin{aligned} D_\sigma\overline{u}_\mathcal{T} &\le \left|\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)}\overline{u}(x)\,dx - \frac{1}{m(\sigma)}\int_\sigma\overline{u}(\xi)\,dm(\xi)\right| \\ &+ \left|\frac{1}{m(\sigma)}\int_\sigma\overline{u}(\xi)\,dm(\xi) - \frac{1}{|B(x_L,\alpha h_L)|}\int_{B(x_L,\alpha h_L)}\overline{u}(x)\,dx\right| \end{aligned}$$

and (5.3.3) gives thus

$$(D_\sigma\overline{u}_\mathcal{T})^2 \le \frac{2C_3h_K^2}{|\triangle_{K,\sigma}|}\int_K|\nabla\overline{u}(x)|^2\,dx + \frac{2C_3h_L^2}{|\triangle_{L,\sigma}|}\int_L|\nabla\overline{u}(x)|^2\,dx.$$

We have $|\triangle_{K,\sigma}| = m(\sigma)d_{K,\sigma}/N$ and $d_\sigma\ge d_{K,\sigma}\ge\alpha h_K$ (and the same properties with K replaced by L), so that

$$\frac{m(\sigma)}{d_\sigma}(D_\sigma\overline{u}_\mathcal{T})^2 \le \frac{2C_3N}{\alpha^2}\left(\int_K|\nabla\overline{u}(x)|^2\,dx + \int_L|\nabla\overline{u}(x)|^2\,dx\right) \tag{5.3.5}$$

(5.3.4) and (5.3.5) show that, for all $\sigma\in\mathcal{E}$,

$$\frac{m(\sigma)}{d_\sigma}(D_\sigma\overline{u}_\mathcal{T})^2 \le C_4\sum_{K\in\mathcal{T}\ |\ \sigma\in\mathcal{E}_K}\int_K|\nabla\overline{u}(x)|^2\,dx$$

with C_4 only depending on (N,α). Summing these inequalities on $\sigma\in\mathcal{E}$, we find

$$||\overline{u}_\mathcal{T}||_{1,\mathcal{T}}^2 \le C_4\sum_{\sigma\in\mathcal{E}}\sum_{K\in\mathcal{T}\ |\ \sigma\in\mathcal{E}_K}\int_K|\nabla\overline{u}(x)|^2\,dx \le C_4\sum_{K\in\mathcal{T}}\int_K|\nabla\overline{u}(x)|^2\,dx\ {\rm card}(\mathcal{E}_K)$$

and (5.1.4) concludes the proof of the proposition. ■

Corollary 5.3.1 *Assume that $\mathcal{T}$ is an admissible mesh which satisfies (5.1.2), (5.1.3) and (5.1.4). There exists C only depending on $(\Omega,||\mathbf{v}||_{(L^\infty(\Omega))^N},\zeta,\alpha,M)$ such that, if $\overline{u}$ is the variational solution to (5.1.1), $\overline{u}_\mathcal{T}$ is defined by (5.2.1) and $u_\mathcal{T}=(u_K)_{K\in\mathcal{T}}$ is the solution to (5.1.9)—(5.1.10), then*

$$||\overline{u}_\mathcal{T}-u_\mathcal{T}||_{1,\mathcal{T}} \le C(||\overline{u}||_{H_0^1(\Omega)} + ||f||_{L^2(\Omega)} + ||\,|G|\,||_{L^2(\Omega)}).$$

Remark 5.3.2 *In fact, by [31] (Theorem 2.1),* $||\overline{u}||_{H_0^1(\Omega)}$ *is controlled by* $||f||_{L^2(\Omega)} + ||\,|G|\,||_{L^2(\Omega)}$ *and we could thus drop it in the preceding inequality.*

Proof of Corollary 5.3.1
The right-hand side of (5.1.9) is written $|K|f_K + \sum_{\sigma\in\mathcal{E}_K} m(\sigma)Q_{K,\sigma}$ with $(Q_{K,\sigma})_{K\in\mathcal{T},\,\sigma\in\mathcal{E}_K}$ which satisfies, for all $\sigma = K|L \in \mathcal{E}_{\rm int}$, $Q_{K,\sigma} = -Q_{L,\sigma}$ (conservativity); thus, by Proposition 3.2, p.72, in [32], there exists C_1 only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, \zeta)$ such that

$$||u_\mathcal{T}||_{1,\mathcal{T}} \le C_1 \left(\sum_{K\in\mathcal{T}} |K| f_K^2\right)^{1/2} + C_1 \left(\sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma Q_\sigma^2\right)^{1/2} \tag{5.3.6}$$

where $Q_\sigma = |Q_{K,\sigma}|$ for some $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$ (by conservativity, this definition of Q_σ does not depend on the choice of such a K).
Since $|\,\triangle_{K,\sigma}\,| = m(\sigma)d_{K,\sigma}/N$ for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we have, by convexity of $X \to X^2$ and for all $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$\begin{aligned} Q_\sigma^2 &\le \frac{d_{K,\sigma}}{d_\sigma} \frac{1}{|\,\triangle_{K,\sigma}\,|} \int_{\triangle_{K,\sigma}} |G(x)|^2\,dx + \frac{d_{L,\sigma}}{d_\sigma} \frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} |G(x)|^2\,dx \\ &\le \frac{N}{m(\sigma)d_\sigma} \left(\int_{\triangle_{K,\sigma}} |G(x)|^2\,dx + \int_{\triangle_{L,\sigma}} |G(x)|^2\,dx \right) \end{aligned}$$

(notice that, suppressing the term involving L, this estimate is still true if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext}$). Thus,

$$\begin{aligned} \sum_{\sigma\in\mathcal{E}} m(\sigma) d_\sigma Q_\sigma^2 &\le N \sum_{\sigma\in\mathcal{E}} \sum_{K\in\mathcal{T}\,|\,\sigma\in\mathcal{E}_K} \int_{\triangle_{K,\sigma}} |G(x)|^2\,dx \\ &\le N \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \int_{\triangle_{K,\sigma}} |G(x)|^2\,dx \\ &= N \sum_{K\in\mathcal{T}} \int_K |G(x)|^2\,dx = N \int_\Omega |G(x)|^2\,dx. \end{aligned}$$

Moreover, $f_K^2 \le \frac{1}{|K|}\int_K |f(x)|^2\,dx$ and (5.3.6) gives then $||u_\mathcal{T}||_{1,\mathcal{T}} \le C_1||f||_{L^2(\Omega)} + C_1\sqrt{N}||\,|G|\,||_{L^2(\Omega)}$. Combined with Proposition 5.3.1, this concludes the proof. ■

5.4 The H^2 framework

Proposition 5.4.1 *Assume that* $\mathcal{T}$ *is an admissible mesh which satisfies Hypothesis (5.1.3) and that* $\overline{u} \in H^2(\Omega) \cap H_0^1(\Omega)$. *Define, for* $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$R_{K,\sigma} = \frac{\overline{u}_K - \overline{u}_L}{d_\sigma} + \frac{1}{m(\sigma)} \int_\sigma \nabla\overline{u}(\xi)\cdot\mathbf{n}_{K,\sigma}\,dm(\xi).$$

and, for $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$,

$$R_{K,\sigma} = \frac{\overline{u}_K}{d_\sigma} + \frac{1}{m(\sigma)} \int_\sigma \nabla\overline{u}(\xi)\cdot\mathbf{n}_{K,\sigma}\,dm(\xi).$$

Then there exists C *only depending on* (N,α) *such that*

$$m(\sigma) d_\sigma R_{K,\sigma}^2 \le C \sum_{L\in\mathcal{T}\,|\,\sigma\in\mathcal{E}_L} h_L^2 \int_L |D^2\overline{u}(x)|^2\,dx.$$

Remark 5.4.1 *Notice that this result is false in general if we replace $\overline{u}_K$ by the mean value of $\overline{u}$ on K and if x_K is not the gravity center of K (consider $K =]0,h[^2$, $L =]-h,0[\times]0,h[$, $x_K = (2h/3, h/2)$, $x_L = (-2h/3, h/2)$ and $\overline{u}(x,y) = x$; we have then $R_{K,\sigma} = -1/4$ but $D^2\overline{u} = 0$).*

Remark 5.4.2 *It would be tempting to try to use the Bramble-Hilbert result (see [18] (Theorem 2)) to obtain the estimate of Proposition 5.4.1 (and also in Proposition 5.3.1 and Lemmas 5.7.1, 5.7.2). This theorem can be used in Finite Element methods thanks to a "reference finite element": for example, in triangular meshes, each finite element can be transformed, by some simple linear application, into some reference triangle; this allows to easily obtain estimates only depending on the size of the element, not its geometry (because they all have the same geometry: that of a triangle).*
We do not have such reference control volume in our meshes (our control volumes can have very different geometries); so, in order to prove that the estimates on $R_{K,\sigma}$ only depend on the size of K and not on its geometry, Bramble-Hilbert's result is useless and we have to make the whole proof.

Proof of Proposition 5.4.1
Due to technical reasons, we must first replace the mean value of $\overline{u}$ on $B(x_K, \alpha h_K)$ by the mean value on $B(x_K, \frac{\alpha h_K}{2})$; step 1 is the study of the consistency error for $\sigma \in \mathcal{E}_{\rm int}$ with these new mean values (and if $\overline{u}$ is regular). In step 2, we prove that the error introduced by the use of the mean values on $B(x_K, \frac{\alpha h_K}{2})$ can be controlled, and we conclude the proof for interior edges. In step 3, the case of boundary edges is handled thanks to a symmetry trick which brings us back to the case of interior edges.

Step 1: we suppose that $\overline{u}$ is regular and we take $\sigma = K|L \in \mathcal{E}_{\rm int}$.
We define

$$\begin{aligned} R'_{K,\sigma} &= \frac{1}{d_\sigma}\left(\frac{1}{|B(x_K,\frac{\alpha h_K}{2})|}\int_{B(x_K,\frac{\alpha h_K}{2})}\overline{u}(x)\,dx - \frac{1}{|B(x_L,\frac{\alpha h_L}{2})|}\int_{B(x_L,\frac{\alpha h_L}{2})}\overline{u}(y)\,dy\right)\\ &\quad + \frac{1}{m(\sigma)}\int_\sigma \nabla\overline{u}(\xi)\cdot\mathbf{n}_{K,\sigma}\,dm(\xi). \end{aligned}$$

Let $\xi \in \sigma$. By Taylor's expansions, we have, for all $x \in B(x_K, \alpha h_K/2)$ and all $y \in B(x_L, \alpha h_L/2)$,

$$\overline{u}(x) = \overline{u}(\xi) + \nabla\overline{u}(\xi)\cdot(x-\xi) + \int_0^1 (1-t)D^2\overline{u}(\xi + t(x-\xi))(x-\xi)\cdot(x-\xi)\,dt$$
$$\overline{u}(y) = \overline{u}(\xi) + \nabla\overline{u}(\xi)\cdot(y-\xi) + \int_0^1 (1-t)D^2\overline{u}(\xi + t(y-\xi))(y-\xi)\cdot(y-\xi)\,dt.$$

Subtracting these equations and taking the mean value on $\xi \in \sigma$, we find

$$\begin{aligned} \overline{u}(x) - \overline{u}(y) &= \frac{1}{m(\sigma)}\int_\sigma \nabla\overline{u}(\xi)\,dm(\xi)\cdot(x-y)\\ &\quad + \frac{1}{m(\sigma)}\int_\sigma\int_0^1 (1-t)D^2\overline{u}(\xi + t(x-\xi))(x-\xi)\cdot(x-\xi)\,dt dm(\xi)\\ &\quad - \frac{1}{m(\sigma)}\int_\sigma\int_0^1 (1-t)D^2\overline{u}(\xi + t(y-\xi))(y-\xi)\cdot(y-\xi)\,dt dm(\xi). \end{aligned}$$

We now take the mean values on $x \in B(x_K, \alpha h_K/2)$ and $y \in B(x_L, \alpha h_L/2)$; since the mean values of $x \to x$ and $y \to y$ on these sets are respectively x_K and x_L and since $x_L - x_K = d_\sigma \mathbf{n}_{K,\sigma}$, dividing by d_σ, we obtain

$$\begin{aligned} R'_{K,\sigma} &= \frac{1}{d_\sigma m(\sigma)|B(x_K,\frac{\alpha h_K}{2})|}\int_{B(x_K,\frac{\alpha h_K}{2})}\int_\sigma\int_0^1 (1-t)D^2\overline{u}(\xi + t(x-\xi))(x-\xi)\cdot(x-\xi)\,dt dm(\xi)dx\\ &\quad - \frac{1}{d_\sigma m(\sigma)|B(x_L,\frac{\alpha h_L}{2})|}\int_{B(x_L,\frac{\alpha h_L}{2})}\int_\sigma\int_0^1 (1-t)D^2\overline{u}(\xi + t(y-\xi))(y-\xi)\cdot(y-\xi)\,dt dm(\xi)dy. \end{aligned}$$

By Jensen's inequality, and since $|x-\xi| \le h_K$ if $x \in K$ and $\xi \in \sigma$, we obtain

$$(R'_{K,\sigma})^2 \le \frac{2h_K^4}{d_\sigma^2 m(\sigma)|B(x_K, \frac{\alpha h_K}{2})|} \int_{B(x_K,\frac{\alpha h_K}{2})} \int_\sigma \int_0^1 (1-t)^2 |D^2\overline{u}(\xi + t(x-\xi))|^2 \, dt dm(\xi) dx$$
$$+ \frac{2h_L^4}{d_\sigma^2 m(\sigma)|B(x_L, \frac{\alpha h_L}{2})|} \int_{B(x_L,\frac{\alpha h_L}{2})} \int_\sigma \int_0^1 (1-t)^2 |D^2\overline{u}(\xi + t(y-\xi))|^2 \, dt dm(\xi) dy. \quad (5.4.7)$$

By translation, we can suppose that $\sigma = \{0\} \times \widetilde{\sigma} \subset \{0\} \times \mathbb{R}^{N-1}$. Let $x \in B(x_K, \alpha h_K/2)$; we use the change of variable $(t, \xi') \in]0,1[\times\widetilde{\sigma} \to z = (0,\xi') + t(x - (0,\xi')) \in V_x \subset K$ (K is convex), whose jacobian determinant is $|x_1|(1-t)^{N-1}$ (where x_1 is the first component of x — we will see below that $x_1 \neq 0$). Since $N \le 3$, we have, if $t \in]0,1[$, $(1-t)^2 \le (1-t)^{N-1}$ and we can thus write

$$\begin{aligned} \int_\sigma \int_0^1 (1-t)^2 |D^2\overline{u}(\xi + t(x-\xi))|^2 \, dt dm(\xi) &\le \int_\sigma \int_0^1 (1-t)^{N-1} |D^2\overline{u}(\xi + t(x-\xi))|^2 \, dt dm(\xi) \\ &\le |x_1|^{-1} \int_K |D^2\overline{u}(z)|^2 \, dz. \end{aligned} \quad (5.4.8)$$

Write $x_K = (a,b)$ with $a \in \mathbb{R}$ and $b \in \mathbb{R}^{N-1}$. The straight line going through x_K and orthogonal to $\sigma \subset \{0\} \times \mathbb{R}^{N-1}$, (i.e. the line $\mathbb{R} \times \{b\}$) intersects σ (i.e. $(0,b) \in \sigma$). Thus, $|a| = |(a,b) - (0,b)| \ge \mathrm{dist}(x_K, \sigma) \ge \mathrm{dist}(x_K, \partial K) \ge \alpha h_K$ (recall that $B(x_K, \alpha h_K) \in K$). Thus, if $x \in B(x_K, \alpha h_K/2)$, we have $|x_1| \ge |a| - |x_1 - a| \ge \alpha h_K - |x - x_K| \ge \frac{\alpha h_K}{2}$ and (5.4.8) gives then

$$\int_\sigma \int_0^1 (1-t)^2 |D^2\overline{u}(\xi + t(x-\xi))|^2 \, dt dm(\xi) \le \frac{2}{\alpha h_K} \int_K |D^2\overline{u}(z)|^2 \, dz.$$

Therefore,

$$\frac{1}{|B(x_K, \frac{\alpha h_K}{2})|} \int_{B(x_K,\frac{\alpha h_K}{2})} \int_\sigma \int_0^1 (1-t)^2 |D^2\overline{u}(\xi + t(x-\xi))|^2 \, dt dm(\xi) dx \le \frac{2}{\alpha h_K} \int_K |D^2\overline{u}(z)|^2 \, dz.$$

Coming back to (5.4.7) (and using the preceding inequality also with L instead of K), we obtain

$$(R'_{K,\sigma})^2 \le \frac{4h_K^3}{\alpha d_\sigma^2 m(\sigma)} \int_K |D^2\overline{u}(z)|^2 \, dz + \frac{4h_L^3}{\alpha d_\sigma^2 m(\sigma)} \int_L |D^2\overline{u}(z)|^2 \, dz.$$

Since $d_\sigma \ge d_{K,\sigma} \ge \alpha h_K$, we deduce that

$$m(\sigma) d_\sigma (R'_{K,\sigma})^2 \le \frac{4h_K^2}{\alpha^2} \int_K |D^2\overline{u}(z)|^2 \, dz + \frac{4h_L^2}{\alpha^2} \int_L |D^2\overline{u}(z)|^2 \, dz \quad (5.4.9)$$

Step 2: we now estimate the difference between the mean values of $\overline{u}$ on $B(x_K, \alpha h_K)$ and on $B(x_K, \frac{\alpha h_K}{2})$, and we conclude for $\sigma \in \mathcal{E}_{\mathrm{int}}$.
Let $v(x) = \overline{u}(x) - \left(\frac{1}{|B(x_K,\alpha h_K)|} \int_{B(x_K,\alpha h_K)} \nabla\overline{u}(y) \, dy\right) \cdot (x - x_K)$. We have, for $x \in B(0, \alpha h_K)$,

$$v(x_K + x) - v\left(x_K + \frac{x}{2}\right) = \int_0^1 \nabla v \left(x_K + \frac{1+t}{2} x\right) \cdot \frac{x}{2} \, dt.$$

Integrating on $x \in B(0, \alpha h_K)$, dividing by $|B(x_K, \alpha h_K)|$ and thanks to the change of variable $y = x_K + \frac{x}{2}$ in the second integral, we find

$$\begin{aligned} &\frac{1}{|B(x_K,\alpha h_K)|} \int_{B(x_K,\alpha h_K)} v(x) \, dx - \frac{1}{|B(x_K, \frac{\alpha h_K}{2})|} \int_{B(x_K,\frac{\alpha h_K}{2})} v(y) \, dy \\ &= \frac{1}{|B(x_K,\alpha h_K)|} \int_{B(0,\alpha h_K)} \int_0^1 \nabla v \left(x_K + \frac{1+t}{2} x\right) \cdot \frac{x}{2} \, dt dx. \end{aligned}$$

Since the mean values on $B(x_K, \alpha h_K)$ and on $B(x_K, \frac{\alpha h_K}{2})$ of $x \to x - x_K$ are null, the mean values of v on these sets are equal to the mean values of $\overline{u}$ on the same sets. Thus, denoting

$$I_{K,\sigma} = \frac{1}{d_\sigma}\left(\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} \overline{u}(x)\,dx - \frac{1}{|B(x_K,\frac{\alpha h_K}{2})|}\int_{B(x_K,\frac{\alpha h_K}{2})} \overline{u}(y)\,dy\right),$$

we have just proved that

$$|I_{K,\sigma}| \le \frac{\alpha h_K}{2d_\sigma |B(x_K,\alpha h_K)|}\int_{B(0,\alpha h_K)}\int_0^1 \left|\nabla v\left(x_K + \frac{1+t}{2}x\right)\right| dtdx.$$

Using the change of variable $x \to z = x_K + \frac{1+t}{2}x$ (which sends $B(0,\alpha h_K)$ into (but not onto) $B(x_K,\alpha h_K)$), we deduce

$$\begin{aligned}(I_{K,\sigma})^2 &\le \frac{\alpha^2 h_K^2}{4d_\sigma^2}\left(\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)}\int_0^1 |\nabla v(z)|\left(\frac{2}{1+t}\right)^N dtdz\right)^2 \\ &\le \frac{2^{2N-2}\alpha^2 h_K^2}{d_\sigma^2}\left(\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |\nabla v(z)|\,dz\right)^2. \qquad (5.4.10)\end{aligned}$$

We have $\nabla v(z) = \nabla \overline{u}(z) - \frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} \nabla \overline{u}(y)\,dy$. Thus, thanks to Lemma 5.7.1,

$$\begin{aligned}&\left(\frac{1}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |\nabla v(z)|\,dz\right)^2 \\ &\le \left(\frac{1}{|B(x_K,\alpha h_K)|\,|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)}\int_{B(x_K,\alpha h_K)} |\nabla \overline{u}(z) - \nabla \overline{u}(y)|\,dzdy\right)^2 \\ &\le \frac{C_1 h_K^{N+2}}{|B(x_K,\alpha h_K)|\,|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |D^2\overline{u}(x)|^2\,dx \\ &\le \frac{C_2 h_K^2}{|B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |D^2\overline{u}(x)|^2\,dx\end{aligned}$$

where C_1 and C_2 only depend on (N,α).
Coming back to (5.4.10), we obtain C_3 only depending on (N,α) such that

$$(I_{K,\sigma})^2 \le \frac{C_3 h_K^4}{d_\sigma^2 |B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |D^2\overline{u}(x)|^2\,dx.$$

σ is of diameter less than h_K, so that $m(\sigma) \le C_4 h_K^{N-1}$ with C_4 only depending on N. Moreover, $d_\sigma \ge d_{K,\sigma} \ge \alpha h_K$; hence,

$$m(\sigma) d_\sigma (I_{K,\sigma})^2 \le \frac{C_5 h_K^{N+3}}{\alpha h_K |B(x_K,\alpha h_K)|}\int_{B(x_K,\alpha h_K)} |D^2\overline{u}(x)|^2\,dx \le C_6 h_K^2 \int_K |D^2\overline{u}(x)|^2\,dx \qquad (5.4.11)$$

with C_5 and C_6 only depending on (N,α).

We have $R_{K,\sigma} = I_{K,\sigma} + R'_{K,\sigma} - I_{L,\sigma}$. Thanks to (5.4.9) and (5.4.11), we deduce

$$m(\sigma) d_\sigma R_{K,\sigma}^2 \le C_7 h_K^2 \int_K |D^2\overline{u}(x)|^2\,dx + C_7 h_L^2 \int_L |D^2\overline{u}(x)|^2\,dx \qquad (5.4.12)$$

with C_7 only depending on (N, α). This estimate has been obtained for $\overline{u}$ regular, but, by the density result of Lemma 5.7.3 (found in the Appendix), it is also satisfied by functions in $H^2(K \cup \sigma \cup L)$ (thus by functions in $H^2(\Omega)$). This concludes the proof if $\sigma \in \mathcal{E}_{\rm int}$.

Step 3: suppose now that $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.
Since $\overline{u} \in H^2(\Omega) \cap H^1_0(\Omega)$, we have $\overline{u} = 0$ on σ. Denoting by S the orthogonal symmetry with respect to the hyperplane generated by σ, it is then well known that the function $\mathcal{U} : K \cup \sigma \cup S(K) \to \mathbb{R}$ which is equal to $\overline{u}$ on K and to $-\overline{u} \circ S$ on $S(K)$ belongs to $H^2(K \cup \sigma \cup S(K))$.
We notice that all the hypotheses on (K, x_K, σ, L, x_L) used in Steps 1 and 2 (and in the proof of Lemma 5.7.3) are satisfied here by $(K, x_K, \sigma, S(K), S(x_K))$ (and with $\mathrm{dist}(x_K, S(x_K)) = 2\mathrm{dist}(x_K, \sigma) = 2d_\sigma$ instead of d_σ).
The result (5.4.12) of Step 2 hence applies with $\mathcal{U}$ instead of $\overline{u}$ and we can write, defining $\mathcal{U}_K$ and $\mathcal{U}_{S(K)}$ as the mean values of $\mathcal{U}$ on $B(x_K, \alpha h_K)$ and $B(x_{S(K)}, \alpha h_{S(K)})$ respectively,

$$\begin{aligned} & 2m(\sigma)d_\sigma \left(\frac{\mathcal{U}_K - \mathcal{U}_{S(K)}}{2d_\sigma} + \frac{1}{m(\sigma)} \int_\sigma \nabla \mathcal{U}(\xi) \cdot \mathbf{n}_{K,\sigma} \, dm(\xi) \right)^2 \\ \leq \quad & C_7 h_K^2 \int_K |D^2 \mathcal{U}(x)|^2 \, dx + C_7 h_{S(K)}^2 \int_{S(K)} |D^2 \mathcal{U}(x)|^2 \, dx. \end{aligned}$$

Since $\mathcal{U} = \overline{u}$ on K, we have $\mathcal{U}_K = \overline{u}_K$ and, in the sense of the traces, $\nabla \mathcal{U} = \nabla \overline{u}$ on σ. Moreover, $\mathcal{U} = -\overline{u} \circ S$ on $S(K)$, so that $\mathcal{U}_{S(K)} = -\overline{u}_K$ (notice that $h_{S(K)} = h_K$, which implies $B(S(x_K), \alpha h_{S(K)}) = S(B(x_K, \alpha h_K))$) and $|D^2 \mathcal{U}(x)| = |D^2 \overline{u}(S(x))|$ for $x \in S(K)$. Thus, the preceding estimate yields

$$2m(\sigma)d_\sigma \left(\frac{\overline{u}_K}{d_\sigma} + \frac{1}{m(\sigma)} \int_\sigma \nabla \overline{u}(\xi) \cdot \mathbf{n}_{K,\sigma} \, dm(\xi) \right)^2 \leq 2C_7 h_K^2 \int_K |D^2 \overline{u}(x)|^2 \, dx,$$

which is exactly the desired result for $\sigma \in \mathcal{E}_{\rm ext}$. ■

Theorem 5.4.1 *Assume that $G \in (H^1(\Omega))^N$ and that the variational solution $\overline{u}$ to (5.1.1) belongs to $H^2(\Omega) \cap H^1_0(\Omega)$. If $\mathcal{T}$ is an admissible mesh which satisfies (5.1.2), (5.1.3) and (5.1.4), there exists C only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, ||b||_{L^\infty(\Omega)}, \zeta, \alpha, M)$ such that, $\overline{u}_\mathcal{T}$ being defined by (5.2.1) and $u_\mathcal{T} = (u_K)_{K \in \mathcal{T}}$ being the solution to (5.1.9)—(5.1.10),*

$$||\overline{u}_\mathcal{T} - u_\mathcal{T}||_{1,\mathcal{T}} \leq C(||\overline{u}||_{H^2(\Omega)} + ||\, |\nabla G|\, ||_{L^2(\Omega)}) h_\mathcal{T}.$$

Remark 5.4.3 *A similar result, with $\overline{u}_K$ replaced by $\overline{u}(x_K)$ and $G = 0$, is proved (using (5.1.3)) in [38] and [48] (Theorem 3.2). Here, we also need (5.1.4) because, $\overline{u}_K$ taking into account all the values of $\overline{u}$ on a ball around x_K, we cannot control $R_{K,\sigma}$ in Proposition 5.4.1 (for example) only by means of $\int_{\triangle_{K,\sigma}} |D^2 \overline{u}|^2$, as it is done in [48].*

Proof of Theorem 5.4.1
By (5.1.1), $\mathrm{div}(\mathbf{v}\overline{u}) \in L^2(\Omega) \subset L^1(\Omega)$ and, $\mathbf{v}$ and $\overline{u}$ being continuous on $\overline{\Omega}$ (because $\overline{u} \in H^2(\Omega)$ and $N \leq 3$), we can apply Lemma 5.7.4 (see the Appendix) to compute the integral of $\mathrm{div}(\mathbf{v}\overline{u})$ on a convex open subset of Ω. Thus, integrating (5.1.1) on a control volume $K \in \mathcal{T}$, we obtain

$$\begin{aligned} & -\sum_{\sigma \in \mathcal{E}_K} \int_\sigma \nabla \overline{u}(\xi) \cdot \mathbf{n}_{K,\sigma} \, dm(\xi) + \sum_{\sigma \in \mathcal{E}_K} \int_\sigma \mathbf{v}(\xi) \cdot \mathbf{n}_{K,\sigma} \overline{u}(\xi) \, dm(\xi) + \int_K b(x) \overline{u}(x) \, dx \\ & \qquad = \quad |K| f_K + \sum_{\sigma \in \mathcal{E}_K} \int_\sigma G(\xi) \cdot \mathbf{n}_{K,\sigma} \, dm(\xi). \end{aligned} \tag{5.4.13}$$

Denote, as in Proposition 5.4.1,

$$R_{K,\sigma} = \frac{\overline{u}_K - \overline{u}_L}{d_\sigma} - \left(-\frac{1}{m(\sigma)} \int_\sigma \nabla \overline{u}(\xi) \cdot \mathbf{n}_{K,\sigma}\, dm(\xi) \right)$$

with $\sigma = K|L$ if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm int}$ and $\overline{u}_L = 0$ if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext}$.
Let

$$r_{K,\sigma} = v_{K,\sigma} \overline{u}_{\sigma,+} - \frac{1}{m(\sigma)} \int_\sigma \mathbf{v}(\xi) \cdot \mathbf{n}_{K,\sigma} \overline{u}(\xi)\, dm(\xi)$$

where $\overline{u}_{\sigma,+} = \overline{u}_K$ if $v_{K,\sigma} \geq 0$, $\overline{u}_{\sigma,+} = \overline{u}_L$ if $v_{K,\sigma} < 0$ and $\sigma = K|L \in \mathcal{E}_K \cap \mathcal{E}_{\rm int}$, and $\overline{u}_{\sigma,+} = 0$ if $v_{K,\sigma} < 0$ and $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext}$.
Finally, define

$$\rho_K = b_K \overline{u}_K - \frac{1}{|K|} \int_K b(x) \overline{u}(x)\, dx$$

and

$$M_{K,\sigma} = \frac{1}{m(\sigma)} \int_\sigma G(\xi) \cdot \mathbf{n}_{K,\sigma}\, dm(\xi) - \left(\frac{d_{K,\sigma}}{d_\sigma} G_{K,\sigma} - \frac{d_{L,\sigma}}{d_\sigma} G_{L,\sigma} \right)$$

(with the convention that $\sigma = K|L$ if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm int}$ and that $d_{L,\sigma} = G_{L,\sigma} = 0$ if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext}$).

(5.4.13) shows that $(\overline{u}_K)_{K \in \mathcal{T}}$ satisfies (5.1.9)—(5.1.10), provided that we add

$$|K| \rho_K + \sum_{\sigma \in \mathcal{E}_K} m(\sigma)(R_{K,\sigma} + r_{K,\sigma} + M_{K,\sigma})$$

to the right-hand side of (5.1.9). Therefore, subtracting the equations satisfied by $(u_K)_{K \in \mathcal{T}}$ to the equations satisfied by $(\overline{u}_K)_{K \in \mathcal{T}}$, we see that $(e_K)_{K \in \mathcal{T}} = (\overline{u}_K - u_K)_{K \in \mathcal{T}}$ satisfies

$$\begin{aligned} \forall K \in \mathcal{T},\ & \sum_{\sigma \in \mathcal{E}_K} \frac{m(\sigma)}{d_\sigma} (e_K - e_L) + \sum_{\sigma \in \mathcal{E}_K} m(\sigma) v_{K,\sigma} e_{\sigma,+} + |K| b_K e_K \\ & \qquad = |K| \rho_K + \sum_{\sigma \in \mathcal{E}_K} m(\sigma)(R_{K,\sigma} + r_{K,\sigma} + M_{K,\sigma}) \\ & \text{(where } \sigma = K|L \text{ if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm int} \text{ and } e_L = d_{L,\sigma} = G_{L,\sigma} = 0 \text{ if } \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext}), \end{aligned}$$

$$\begin{aligned} & \forall \sigma = K|L \in \mathcal{E}_{\rm int}, \quad e_{\sigma,+} = e_K \ \text{ if } \ v_{K,\sigma} \geq 0, \quad e_{\sigma,+} = e_L \ \text{ otherwise}, \\ & \forall \sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K, \quad e_{\sigma,+} = e_K \ \text{ if } \ v_{K,\sigma} \geq 0, \quad e_{\sigma,+} = 0 \ \text{ otherwise}. \end{aligned}$$

By definition, $(R_{K,\sigma})_{K \in \mathcal{T},\, \sigma \in \mathcal{E}_K}$, $(r_{K,\sigma})_{K \in \mathcal{T},\, \sigma \in \mathcal{E}_K}$ and $(M_{K,\sigma})_{K \in \mathcal{T},\, \sigma \in \mathcal{E}_K}$ are conservative: for all $\sigma = K|L \in \mathcal{E}_{\rm int}$, we have $R_{K,\sigma} = -R_{L,\sigma}$, $r_{K,\sigma} = -r_{L,\sigma}$ and $M_{K,\sigma} = -M_{L,\sigma}$ (notice that $\mathbf{n}_{K,\sigma} = -\mathbf{n}_{L,\sigma}$). Hence, by Proposition 3.2 in [32], we deduce that there exists C_0 only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, \zeta)$ such that

$$||e_{\mathcal{T}}||_{1,\mathcal{T}} \leq C_0 \left(\sum_{K \in \mathcal{T}} |K| \rho_K^2 \right)^{1/2} + C_0 \left(\sum_{\sigma \in \mathcal{E}} m(\sigma) d_\sigma A_\sigma^2 \right)^{1/2}, \tag{5.4.14}$$

where we have denoted $A_\sigma = |R_{K,\sigma} + r_{K,\sigma} + M_{K,\sigma}|$ for some $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$ (by conservativity of these quantities, this definition does not depend on the choice of such a K).

We have

$$|\rho_K| = \left| \frac{1}{|K|} \int_K b(x)(\overline{u}_K - \overline{u}(x))\, dx \right| \leq \frac{||b||_{L^\infty(\Omega)}}{|K|\, |B(x_K, \alpha h_K)|} \int_K \int_{B(x_K, \alpha h_K)} |\overline{u}(y) - \overline{u}(x)|\, dy dx.$$

Therefore, by Lemma 5.7.1, there exists C_1 only depending on $(N, ||b||_{L^\infty(\Omega)}, \alpha)$ such that

$$|K| \rho_K^2 \leq C_1 h_K^2 \int_K |\nabla \overline{u}(x)|^2\, dx. \tag{5.4.15}$$

By definition,

$$|r_{K,\sigma}| = \left| \frac{1}{m(\sigma)} \int_\sigma \mathbf{v}(\xi) \cdot \mathbf{n}_{K,\sigma} (\overline{u}_{\sigma,+} - \overline{u}(\xi))\, dm(\xi) \right| \leq \frac{||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}}{m(\sigma)} \int_\sigma |\overline{u}_{\sigma,+} - \overline{u}(\xi)|\, dm(\xi) \qquad (5.4.16)$$

and either $\overline{u}_{\sigma,+} = \overline{u}_L$ for some $L \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_L$, or $\overline{u}_{\sigma,+} = 0$ and $\sigma \in \mathcal{E}_{\rm ext}$. In the second case, since $u \in H^1_0(\Omega)$, we obtain $|r_{K,\sigma}| \leq 0$ (i.e. $r_{K,\sigma} = 0$). In the first case, we define $v = |\overline{u}_L - \overline{u}| \in H^1(\Omega)$ and, using Lemma 5.7.2, we see that

$$\begin{aligned}
&\left(\frac{1}{m(\sigma)} \int_\sigma |\overline{u}_L - \overline{u}(\xi)|\, dm(\xi) \right)^2 \\
&= \left(\frac{1}{m(\sigma)} \int_\sigma v(\xi)\, dm(\xi) \right)^2 \\
&\leq 2 \left(\frac{1}{m(\sigma)} \int_\sigma v(\xi)\, dm(\xi) - \frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} v(x)\, dx \right)^2 + 2 \left(\frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} v(x)\, dx \right)^2 \\
&\leq \frac{C_2 h_L^2}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} |\nabla v(x)|^2\, dx + 2 \left(\frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} v(x)\, dx \right)^2
\end{aligned}$$

with C_2 only depending on (N, α). But $\nabla v = \mathrm{sgn}(\overline{u}_L - \overline{u}) \nabla \overline{u}$ and

$$v(x) = \left| \frac{1}{|B(x_L, \alpha h_L)|} \int_{B(x_L, \alpha h_L)} \overline{u}(y)\, dy - \overline{u}(x) \right| \leq \frac{1}{|B(x_L, \alpha h_L)|} \int_{B(x_L, \alpha h_L)} |\overline{u}(y) - \overline{u}(x)|\, dy,$$

so that, by Lemma 5.7.1,

$$\begin{aligned}
&\left(\frac{1}{m(\sigma)} \int_\sigma |\overline{u}_L - \overline{u}(\xi)|\, dm(\xi) \right)^2 \\
&\leq \frac{C_2 h_L^2}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} |\nabla \overline{u}(x)|^2\, dx + 2 \left(\frac{1}{|\,\triangle_{L,\sigma}\,|\,|B(x_L, \alpha h_L)|} \int_{\triangle_{L,\sigma}} \int_{B(x_L, \alpha h_L)} |\overline{u}(y) - \overline{u}(x)|\, dy dx \right)^2 \\
&\leq \frac{C_3 h_L^2}{|\,\triangle_{L,\sigma}\,|} \int_L |\nabla \overline{u}(x)|^2\, dx
\end{aligned}$$

where C_3 only depends on (N, α). Using this in (5.4.16), and since $|\,\triangle_{L,\sigma}\,| = m(\sigma) d_{L,\sigma}/N \geq \alpha m(\sigma) h_L/N$ (because $d_{L,\sigma} = \mathrm{dist}(x_L, \sigma) \geq \mathrm{dist}(x_L, \partial L)$ and $B(x_L, \alpha h_L) \subset L$), we deduce that there exists C_4 only depending on $(N, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, \alpha)$ such that

$$m(\sigma) d_\sigma r_{K,\sigma}^2 \leq C_3 ||\mathbf{v}||^2_{(L^\infty(\Omega))^N} \frac{m(\sigma) d_\sigma h_L^2}{|\,\triangle_{L,\sigma}\,|} \int_L |\nabla u(x)|^2\, dx \leq C_4 h_{\mathcal{T}} h_L \int_L |\nabla u(x)|^2\, dx \qquad (5.4.17)$$

for some $L \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_L$ (we have used the fact that $d_\sigma \leq 2 h_{\mathcal{T}}$). Notice that this estimate is also true (for any $L \in \mathcal{T}$...) in the case where $u_{\sigma,+} = 0$ with $\sigma \in \mathcal{E}_{\rm ext}$, since $r_{K,\sigma}$ is then null.
We have, for $\sigma = K|L \in \mathcal{E}_{\rm int}$, since $d_{K,\sigma} + d_{L,\sigma} = d_\sigma$,

$$\begin{aligned}
M_{K,\sigma} &= \frac{d_{K,\sigma}}{d_\sigma} \left(\frac{1}{m(\sigma)} \int_\sigma G(\xi) \cdot \mathbf{n}_{K,\sigma}\, dm(\xi) - \frac{1}{|\,\triangle_{K,\sigma}\,|} \int_{\triangle_{K,\sigma}} G(x) \cdot \mathbf{n}_{K,\sigma}\, dx \right) \\
&\quad + \frac{d_{L,\sigma}}{d_\sigma} \left(\frac{1}{m(\sigma)} \int_\sigma G(\xi) \cdot \mathbf{n}_{K,\sigma}\, dm(\xi) + \frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} G(x) \cdot \mathbf{n}_{L,\sigma}\, dx \right) \\
&= \frac{d_{K,\sigma}}{d_\sigma} \left(\frac{1}{m(\sigma)} \int_\sigma G(\xi)\, dm(\xi) - \frac{1}{|\,\triangle_{K,\sigma}\,|} \int_{\triangle_{K,\sigma}} G(x)\, dx \right) \cdot \mathbf{n}_{K,\sigma} \\
&\quad + \frac{d_{L,\sigma}}{d_\sigma} \left(\frac{1}{m(\sigma)} \int_\sigma G(\xi)\, dm(\xi) - \frac{1}{|\,\triangle_{L,\sigma}\,|} \int_{\triangle_{L,\sigma}} G(x)\, dx \right) \cdot \mathbf{n}_{K,\sigma}.
\end{aligned}$$

Hence, $X \to X^2$ being convex, Lemma 5.7.2 gives C_5 only depending on (N, α) such that

$$M_{K,\sigma}^2 \leq C_5 \frac{d_{K,\sigma}}{d_\sigma} \frac{h_K^2}{|\triangle_{K,\sigma}|} \int_{\triangle_{K,\sigma}} |\nabla G(x)|^2 \, dx + C_5 \frac{d_{L,\sigma}}{d_\sigma} \frac{h_L^2}{|\triangle_{L,\sigma}|} \int_{\triangle_{L,\sigma}} |\nabla G(x)|^2 \, dx.$$

But $|\triangle_{K,\sigma}| = m(\sigma) d_{K,\sigma}/N$ and $|\triangle_{L,\sigma}| = m(\sigma) d_{L,\sigma}/N$, so that

$$m(\sigma) d_\sigma M_{K,\sigma}^2 \leq C_5 N \left(h_K^2 \int_{\triangle_{K,\sigma}} |\nabla G(x)|^2 \, dx + h_L^2 \int_{\triangle_{L,\sigma}} |\nabla G(x)|^2 \, dx \right). \tag{5.4.18}$$

Notice that, suppressing the term involving L, this estimate is still true if $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\text{ext}}$.

We now gather (5.4.15), (5.4.17), (5.4.18) and the estimates of Proposition 5.4.1 in (5.4.14); using Hypothesis (5.1.4), we find thus C_6 only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, ||b||_{L^\infty(\Omega)}, \zeta, \alpha)$ such that

$$\begin{aligned}
||e_{\mathcal{T}}||_{1,\mathcal{T}} &\leq C_6 h_{\mathcal{T}} \left(\sum_{K \in \mathcal{T}} \int_K |\nabla \overline{u}(x)|^2 \, dx \right)^{1/2} \\
&\quad + C_6 h_{\mathcal{T}} \left(\sum_{\sigma \in \mathcal{E}} \sum_{L \in \mathcal{T} \mid \sigma \in \mathcal{E}_L} \int_L |D^2 \overline{u}(x)|^2 \, dx + \int_L |\nabla \overline{u}(x)|^2 \, dx + \int_L |\nabla G(x)|^2 \, dx \right)^{1/2} \\
&\leq C_6 h_{\mathcal{T}} || \, |\nabla \overline{u}| \, ||_{L^2(\Omega)} \\
&\quad + C_6 h_{\mathcal{T}} \left(M \sum_{L \in \mathcal{T}} \int_L |D^2 \overline{u}(x)|^2 \, dx + \int_L |\nabla \overline{u}(x)|^2 \, dx + \int_L |\nabla G(x)|^2 \, dx \right)^{1/2} \\
&\leq C_6 h_{\mathcal{T}} || \, |\nabla \overline{u}| \, ||_{L^2(\Omega)} + \sqrt{M} C_6 h_{\mathcal{T}} \left(\int_\Omega |D^2 \overline{u}(x)|^2 + |\nabla \overline{u}(x)|^2 + |\nabla G(x)|^2 \, dx \right)^{1/2},
\end{aligned}$$

which concludes the proof. ■

5.5 Proof of the main result

We can now prove, using interpolation techniques, Theorem 5.2.1.

Proof of Theorem 5.2.1
Let

$$B = \{(\overline{u}, G, f) \in H_0^1(\Omega) \times (L^2(\Omega))^N \times L^2(\Omega) \mid \Delta \overline{u} - \operatorname{div}(\mathbf{v}\overline{u}) - b\overline{u} + \operatorname{div}(G) + f = 0\}$$

(endowed with the norm of $H_0^1(\Omega) \times (L^2(\Omega))^N \times L^2(\Omega)$) and define $T : B \to X(\mathcal{T})$ by $T(\overline{u}, G, f) = (\overline{u}_K - u_K)_{K \in \mathcal{T}}$, where $(u_K)_{K \in \mathcal{T}}$ is the solution to (5.1.9)—(5.1.10).
Corollary 5.3.1 shows that, if $X(\mathcal{T})$ is endowed with the discrete H^1-norm, T is linear and continuous with a norm bounded by C_0 only depending on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, \alpha, M)$.
If we define

$$A = \{(\overline{u}, G, f) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times (H^1(\Omega))^N \times L^2(\Omega) \mid \Delta \overline{u} - \operatorname{div}(\mathbf{v}\overline{u}) - b\overline{u} + \operatorname{div}(G) + f = 0\}$$

(endowed with the norm of $(H^2(\Omega) \cap H_0^1(\Omega)) \times (H^1(\Omega))^N \times L^2(\Omega)$), Theorem 5.4.1 shows that, $X(\mathcal{T})$ still being endowed with the discrete H^1-norm, $T : A \to X(\mathcal{T})$ is continuous with a norm bounded by $C_1 h_{\mathcal{T}}$, where C_1 only depends on $(\Omega, ||\mathbf{v}||_{(L^\infty(\Omega))^N}, ||b||_{L^\infty(\Omega)}, \zeta, \alpha, M)$.
Thus, by classical interpolation results (see e.g. [6], Theorem 4.1.2, p.88), T is linear continuous $[A, B]_{1-s} \to X(\mathcal{T})$ with a norm bounded by $C_0^{1-s}(C_1 h_{\mathcal{T}})^s \leq \max(C_0, 1) \max(C_1, 1) h_{\mathcal{T}}^s$.

Subsection 5.7.2 in the Appendix shows that

$$\begin{aligned}[A,B]_{1-s} &= \{(\overline{u},G,f) \in [H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s} \times (H^s(\Omega))^N \times L^2(\Omega) \mid \\ &\quad \Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} + \mathrm{div}(G) + f = 0\}\end{aligned}$$

(with equivalent norms). To conclude the proof of the theorem, it remains therefore to see that $[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s} = H^{1+s}(\Omega)\cap H_0^1(\Omega)$.

In the case $N=2$, i.e. if Ω is a polygonal open subset of $\mathbb{R}^2$, this result is proved in [2] (Theorem 3.1). If Ω is convex, we propose the following simple proof (which do not uses the fact that Ω is polygonal). First of all, notice that we have, by definition, $[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s} \hookrightarrow H_0^1(\Omega)$ and, since the inclusions

$$H^2(\Omega)\cap H_0^1(\Omega) \hookrightarrow H^2(\Omega) \quad \text{and} \quad H_0^1(\Omega) \hookrightarrow H^1(\Omega)$$

are continuous, by interpolation, there is a continuous inclusion

$$[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s} \hookrightarrow [H^2(\Omega), H^1(\Omega)]_{1-s} = H^{1+s}(\Omega).$$

This shows that $[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s} \hookrightarrow H^{1+s}(\Omega)\cap H_0^1(\Omega)$.
To prove the reverse inclusion, denote $S=\Delta^{-1}$ with Dirichlet boundary conditions. Since Ω is convex, S is linear continuous

$$H^{-1}(\Omega) \to H_0^1(\Omega) \quad \text{and} \quad L^2(\Omega) \to H^2(\Omega)\cap H_0^1(\Omega).$$

Δ being linear continuous

$$H^1(\Omega) \to H^{-1}(\Omega) \quad \text{and} \quad H^2(\Omega) \to L^2(\Omega),$$

we deduce that $S\circ\Delta$ is linear continuous

$$H^1(\Omega) \to H_0^1(\Omega) \quad \text{and} \quad H^2(\Omega) \to H^2(\Omega)\cap H_0^1(\Omega).$$

By interpolation, $S\circ\Delta$ is thus linear continuous

$$[H^2(\Omega), H^1(\Omega)]_{1-s} = H^{1+s}(\Omega) \to [H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s}.$$

But $S\circ\Delta = Id$ on $H_0^1(\Omega)$, and this shows therefore that $H^{1+s}(\Omega)\cap H_0^1(\Omega)$ is continuously imbedded in $[H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_{1-s}$, which concludes the proof of the theorem. ■

5.6 Numerical results

We present here a few numerical results which illustrate the convergence results we have just proved.
In all these tests, the open set is $\Omega =]-1,1[^2$ and we have taken no lower order term, i.e. $\mathbf{v}=0$ and $b=0$ in (5.1.1); as a right-hand side, we have let $f=0$ and $G=-\nabla u$ (some tests have also been made with $G=-\nabla u+\mathbf{w}$ where $\mathbf{w}$ is divergence free, and the results are similar, provided that $\mathbf{w}$ has the required regularity). The meshes used are regular cartesian grids, and we analyse the rate of convergence by showing, in each case, the discrete H^1 norm of the error versus the size of the mesh, in log-log scale.

Our first test function is a pyramid, based on the function $(x,y)\to(1-|x|)(1-|y|)$ that we have twisted in order that the peak be at $(1/\sqrt{2}, 1/\sqrt{2})$ instead of $(0,0)$ (this has been done to avoid too good convergence results due to symetries between the function and the mesh). The results are shown in Figure 5.1. The dots on this figure indicates a reference slope; as we can see, the rate of convergence is roughly 0.5, which is the expected result since the function is here in $H^{3/2-\epsilon}$ for all $\epsilon>0$.
Then, we have taken $(x,y)\to(1-x^2)(1-y^2)|(x,y)|^s$, which belongs (if $s\in]0,1[$) to $H^{1+s-\epsilon}$ for all $\epsilon>0$. Figure 5.2 shows different cases for s which confirm the result of Theorem 5.2.1.

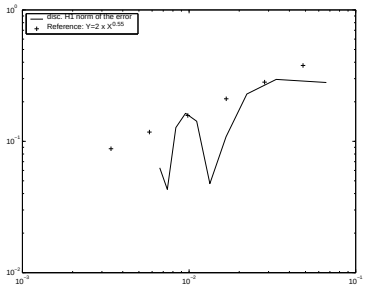


Figure 5.1: Reference slope: 0.55

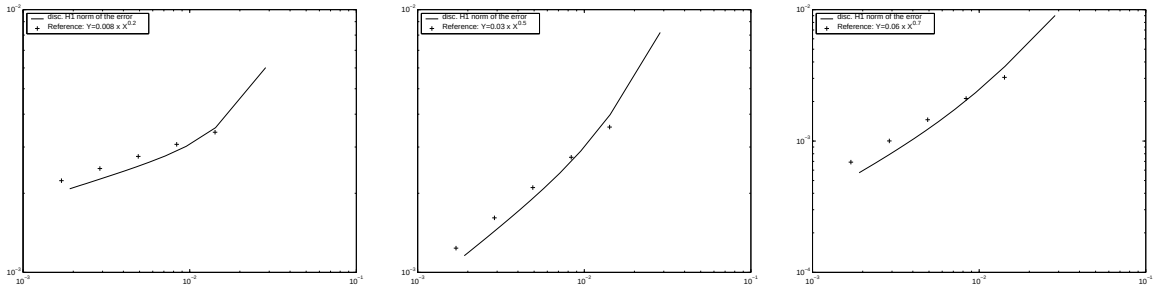


Figure 5.2: Cases $s = 0.2$ (reference slope: 0.2), $s = 0.5$ (reference slope: 0.5), $s = 0.7$ (reference slope: 0.7)

If we add a convection term, with negative divergence, which provokes the loss of coercivity in (5.1.1), the results are similar to the preceding ones, the only difference being that the constant "C" appearing in Theorem 5.2.1 is much bigger. For example, coming back to the first test function, the constant in the reference slope of figure 5.1 is 2 whereas, if we add a convection term with $\mathbf{v} = -10(x, y)$, it becomes 80 (and the slope does not change).

5.7 Appendix

5.7.1 Technical lemmas

Lemma 5.7.1 *There exists $C > 0$ only depending on N such that, if U and V are non-empty open subsets of $\mathbb{R}^N$ contained in a same ball of radius R, we have, for all $v \in H^1(\mathrm{co}(U \cup V))$,*

$$\left| \frac{1}{|U|} \int_U v(x)\,dx - \frac{1}{|V|} \int_V v(y)\,dy \right|^2$$
$$\leq \left(\frac{1}{|U|\,|V|} \int_U \int_V |v(x) - v(y)|\,dxdy \right)^2 \leq \frac{CR^{N+2}}{|U|\,|V|} \int_{\mathrm{co}(U \cup V)} |\nabla v(z)|^2\,dz.$$

Proof of Lemma 5.7.1
$\mathrm{co}(U \cup V)$ is a convex open subset of $\mathbb{R}^N$. Thus, its boundary is Lipschitz-continuous and the regular functions are dense in $H^1(\mathrm{co}(U \cup V))$. We therefore just have to prove the lemma for regular functions. The first inequality is obvious. Let us prove the second. If v is regular, we have, for all $x \in U$ and all

$y \in V$, $v(x) - v(y) = \int_0^1 \nabla v(tx + (1-t)y) \cdot (x-y)\, dt$. This implies

$$\frac{1}{|U|\,|V|} \int_U \int_V |v(x) - v(y)|\, dxdy \le \frac{1}{|U|\,|V|} \int_U \int_V \int_0^1 |\nabla v(tx + (1-t)y)|\, |x-y|\, dtdydx$$

and, since $|x-y| \le 2R$ for all $x \in U$ and all $y \in V$ (U and V are contained in a same ball of radius R), Jensen's inequality gives

$$\left(\frac{1}{|U|\,|V|} \int_U \int_V |v(x) - v(y)|\, dxdy\right)^2 \le \frac{4R^2}{|U|\,|V|} \int_U \int_V \int_0^1 |\nabla v(tx + (1-t)y)|^2\, dtdydx. \tag{5.7.1}$$

Let $y \in V$. Using the change of variable $x \in U \to z = tx + (1-t)y \in tU + (1-t)y \subset \mathrm{co}(U \cup V)$ and Fubini's theorem, we find

$$\int_U \int_V \int_0^1 |\nabla v(tx + (1-t)y)|^2\, dtdxdy \le \int_{\mathrm{co}(U\cup V)} |\nabla v(z)|^2 \int_V \int_{I(z,y)} t^{-N}\, dtdydz \tag{5.7.2}$$

where $I(z,y) = \{t \in [0,1] \mid \exists x \in U\,,\ tx + (1-t)y = z\}$. If $t \in I(z,y)$, then $t(x-y) = z - y$ for some $x \in U$; since U and V are contained in a same ball of radius R, we have then $2Rt \ge t|x-y| = |z-y|$ and thus $I(z,y) \subset [\frac{|z-y|}{2R}, 1]$. We deduce that

$$\int_{I(z,y)} t^{-N}\, dt \le \int_{\frac{|z-y|}{2R}}^1 t^{-N}\, dt \le \frac{1}{N-1} \frac{(2R)^{N-1}}{|z-y|^{N-1}}.$$

Thus, there exists C_0 only depending on N such that, for all $z \in \mathrm{co}(U \cup V)$,

$$\int_V \int_{I(z,y)} t^{-N}\, dtdy \le C_0 R^{N-1} \int_V \frac{1}{|z-y|^{N-1}}\, dy = C_0 R^{N-1} \int_{z-V} \frac{1}{|\xi|^{N-1}}\, d\xi.$$

U and V are included in a same ball of radius R; thus, $\mathrm{co}(U \cup V)$ is also included in this ball and, for all $z \in \mathrm{co}(U \cup V)$, $z - V$ is therefore contained in $B(0, 2R)$, which allows to write, using polar coordinates,

$$\int_V \int_{I(z,y)} t^{-N}\, dtdy \le C_0 R^{N-1} \int_{B(0,2R)} \frac{1}{|\xi|^{N-1}}\, d\xi = C_0 R^{N-1} C_1 \int_0^{2R} \frac{1}{\rho^{N-1}} \rho^{N-1}\, d\rho = 2C_0 C_1 R^N$$

where C_1 is the $(N-1)$-dimensional measure of $\partial B(0,1)$ (C_1 only depends on N).
Gathering this last inequality, (5.7.2) and (5.7.1), we conclude the proof of the lemma. ■

Lemma 5.7.2 *If $\mathcal{T}$ is an admissible mesh which satisfies Hypothesis (5.1.3), there exists C only depending on (N, α) such that, if $v \in H^1(\Omega)$, then, for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$,*

$$\left| \frac{1}{|\,\triangle_{K,\sigma}\,|} \int_{\triangle_{K,\sigma}} v(x)\, dx - \frac{1}{m(\sigma)} \int_\sigma v(\xi)\, dm(\xi) \right|^2 \le \frac{C d_{K,\sigma}^2}{|\,\triangle_{K,\sigma}\,|} \int_{\triangle_{K,\sigma}} |\nabla v(x)|^2\, dx.$$

Proof of Lemma 5.7.2
The regular functions being dense in $H^1(\Omega)$, it is sufficient to prove the lemma for $v \in C^1(\mathbb{R}^N)$.
By translation and rotation, we can suppose that $\sigma = \{0\} \times \widetilde{\sigma}$ with $\widetilde{\sigma} \subset \mathbb{R}^{N-1}$ and that $x_K = (d_{K,\sigma}, 0)$.
For $a \in [0, d_{K,\sigma}]$, we denote $\widetilde{\sigma}_a = \{y \in \mathbb{R}^{N-1} \mid (a,y) \in \triangle_{K,\sigma}\}$. By definition, $(a,y) \in \triangle_{K,\sigma}$ if and only if there exists $t \in [0,1]$ and $z \in \widetilde{\sigma}$ such that $t(d_{K,\sigma}, 0) + (1-t)(0,z) = (a,y)$; this is equivalent to $t = \frac{a}{d_{K,\sigma}}$ and $y = (1-t)z = \left(1 - \frac{a}{d_{K,\sigma}}\right) z$. Thus, $\widetilde{\sigma}_a = \left(1 - \frac{a}{d_{K,\sigma}}\right) \widetilde{\sigma}$.

For all $y \in \widetilde{\sigma}$ and all $a \in [0, d_{K,\sigma}]$, we have

$$v(0,y) - v\left(a, \left(1 - \frac{a}{d_{K,\sigma}}\right) y\right) = \int_0^1 \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \cdot \left(-a, \frac{a}{d_{K,\sigma}} y\right) dt.$$

Integrating on $y \in \widetilde{\sigma}$ and using the change of variable $z = \left(1 - \frac{a}{d_{K,\sigma}}\right) y$, we find

$$\int_\sigma v(\xi)\, dm(\xi) - \frac{1}{\left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1}} \int_{\widetilde{\sigma}_a} v(a,z)\, dz = \int_{\widetilde{\sigma}} \int_0^1 \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \cdot \left(-a, \frac{a}{d_{K,\sigma}} y\right) dt dy.$$

Multiplying by $\left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1}$ and integrating on $a \in [0, d_{K,\sigma}]$, we obtain

$$\begin{aligned} &\int_\sigma v(\xi)\, dm(\xi) \int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} da - \int_0^{d_{K,\sigma}} \int_{\widetilde{\sigma}_a} v(a,z)\, dz da \\ &= \int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \cdot \left(-a, \frac{a}{d_{K,\sigma}} y\right) dt dy da. \end{aligned} \tag{5.7.3}$$

But $\int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} da = \frac{d_{K,\sigma}}{N}$ and $| \triangle_{K,\sigma} | = \frac{m(\sigma) d_{K,\sigma}}{N}$; therefore, (5.7.3) gives, thanks to Fubini's theorem,

$$\begin{aligned} &\frac{1}{m(\sigma)} \int_\sigma v(\xi)\, dm(\xi) - \frac{1}{| \triangle_{K,\sigma} |} \int_{\triangle_{K,\sigma}} v(x)\, dx \\ &= \frac{1}{| \triangle_{K,\sigma} |} \int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \cdot \left(-a, \frac{a}{d_{K,\sigma}} y\right) dt dy da \end{aligned} \tag{5.7.4}$$

By definition of an admissible mesh, the straight line going through $x_K = (d_{K,\sigma}, 0)$ and orthogonal to $\sigma \subset \{0\} \times \mathbb{R}^{N-1}$ intersects σ; this means that $0 \in \widetilde{\sigma}$. Moreover, σ is contained in $\overline{K}$ which has diameter h_K; thus, $\widetilde{\sigma}$ has diameter less than or equal to h_K. By Hypothesis (5.1.3), $d_{K,\sigma} = \mathrm{dist}(x_K, \sigma) \geq \mathrm{dist}(x_K, \partial K) \geq \alpha h_K$.
We deduce that, for all $y \in \widetilde{\sigma}$, $|y| = |y - 0| \leq \mathrm{diam}(\widetilde{\sigma}) \leq h_K \leq \frac{1}{\alpha} d_{K,\sigma}$. Thus,

$$\begin{aligned} &\left| \int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \cdot \left(-a, \frac{a}{d_{K,\sigma}} y\right) dt dy da \right| \\ &\leq C_0 \int_0^{d_{K,\sigma}} \left(1 - \frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \left| \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \right| a\, dt dy da \end{aligned} \tag{5.7.5}$$

where C_0 only depends on α.
Let $a \in]0, d_{K,\sigma}[$. By the change of variable $\varphi_a : (t,y) \in]0,1[\times \widetilde{\sigma} \to z = \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \in \varphi_a(]0,1[\times \widetilde{\sigma})$ (whose Jacobian determinant is $a \left(1 - t\frac{a}{d_{K,\sigma}}\right)^{N-1} = a \left(1 - \frac{z_1}{d_{K,\sigma}}\right)^{N-1}$ since $z_1 = ta$), we have

$$\int_{\widetilde{\sigma}} \int_0^1 \left| \nabla v \left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) \right| dt dy = \int_{\varphi_a(]0,1[\times \widetilde{\sigma})} |\nabla v(z)| a^{-1} \left(1 - \frac{z_1}{d_{K,\sigma}}\right)^{-N+1} dz.$$

But $\varphi_a(]0,1[\times \widetilde{\sigma}) \subset \triangle_{K,\sigma}$ (because $\left(ta, \left(1 - t\frac{a}{d_{K,\sigma}}\right) y\right) = \frac{ta}{d_{K,\sigma}} (d_{K,\sigma}, 0) + \left(1 - \frac{ta}{d_{K,\sigma}}\right) (0,y)$) and Fubini's

theorem allows thus to write

$$\begin{aligned}
&\int_0^{d_{K,\sigma}} \left(1-\frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \left|\nabla v\left(ta, \left(1-t\frac{a}{d_{K,\sigma}}\right) y\right)\right| a\, dt dy da \\
&= \int_0^{d_{K,\sigma}} \int_{\varphi_a(]0,1[\times\widetilde{\sigma})} |\nabla v(z)| \left(1-\frac{z_1}{d_{K,\sigma}}\right)^{-N+1} \left(1-\frac{a}{d_{K,\sigma}}\right)^{N-1} dz da \\
&\leq \int_{\triangle_{K,\sigma}} |\nabla v(z)| \int_{a\in[0,d_{K,\sigma}] \mid z\in\varphi_a(]0,1[\times\widetilde{\sigma})} \left(\frac{1-\frac{a}{d_{K,\sigma}}}{1-\frac{z_1}{d_{K,\sigma}}}\right)^{N-1} da dz.
\end{aligned}$$

If $z \in \varphi_a(]0,1[\times\widetilde{\sigma})$, we have $z_1 = ta$ for some $t \in]0,1[$, i.e. $0 \leq z_1 \leq a$. Therefore, $1-\frac{a}{d_{K,\sigma}} \leq 1-\frac{z_1}{d_{K,\sigma}}$ and

$$\int_{a\in[0,d_{K,\sigma}] \mid z\in\varphi_a(]0,1[\times\widetilde{\sigma})} \left(\frac{1-\frac{a}{d_{K,\sigma}}}{1-\frac{z_1}{d_{K,\sigma}}}\right)^{N-1} da \leq d_{K,\sigma}.$$

We deduce that

$$\int_0^{d_{K,\sigma}} \left(1-\frac{a}{d_{K,\sigma}}\right)^{N-1} \int_{\widetilde{\sigma}} \int_0^1 \left|\nabla v\left(ta, \left(1-t\frac{a}{d_{K,\sigma}}\right) y\right)\right| a\, dt dy da \leq d_{K,\sigma} \int_{\triangle_{K,\sigma}} |\nabla v(z)|\, dz.$$

We now use this inequality in (5.7.5) and introduce the resulting estimate in (5.7.4) to obtain

$$\left|\frac{1}{|\triangle_{K,\sigma}|}\int_{\triangle_{K,\sigma}} v(x)\, dx - \frac{1}{m(\sigma)}\int_\sigma v(\xi)\, dm(\xi)\right| \leq \frac{C_0 d_{K,\sigma}}{|\triangle_{K,\sigma}|}\int_{\triangle_{K,\sigma}} |\nabla v(x)|\, dx.$$

Jensen's inequality concludes then the proof of the lemma. ■

Lemma 5.7.3 *Let $\mathcal{T}$ be an admissible mesh, $\sigma = K|L \in \mathcal{E}_{\rm int}$ and $U = K \cup \sigma \cup L$. Then U is an open subset of $\mathbb{R}^N$ and $C^\infty(\overline{U})$ is dense in $H^2(U)$.*

Proof of Lemma 5.7.3
Step 1: we prove that U is open.
By translation and rotation, we can suppose that $\sigma = \{0\} \times \widetilde{\sigma}$ with $\widetilde{\sigma} \subset \mathbb{R}^{N-1}$ and, since the line going through (x_K, x_L) is orthogonal to σ, that $0 \in \sigma$, $x_K = (b,0) \in]0,\infty[\times\{0\}$ and $x_L = (a,0) \in]-\infty,0[\times\{0\}$. Define then $\varphi : (t,y) \in]-1,1[\times\mathbb{R}^{N-1} \to (t^-a + t^+b, (1-|t|)y) \in]a,b[\times\mathbb{R}^{N-1}$ (where $t^+ = \max(0,t)$ and $t^- = \max(0,-t)$); φ is an homeomorphism (the inverse mapping is $\psi(z_1,z') = (\frac{z_1^+}{b} + \frac{z_1^-}{a}, (1-\frac{z_1^+}{b} - \frac{z_1^-}{a})^{-1} z')$). It is easy to see that $\varphi(]-1,1[\times\widetilde{\sigma}) = \triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma}$; indeed, $\varphi(0,\widetilde{\sigma}) = \sigma$, $\varphi(]0,1[\times\widetilde{\sigma}) = \triangle_{K,\sigma}$ (recall that $x_K = (b,0)$) and $\varphi(]-1,0[,\widetilde{\sigma}) = \triangle_{L,\sigma}$ (recall that $x_L = (a,0)$); by hypothesis on the egdes, $\widetilde{\sigma}$ is open in $\mathbb{R}^{N-1}$ and $\varphi(]-1,1[\times\widetilde{\sigma}) = \triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma}$ is thus open in $]a,b[\times\mathbb{R}^{N-1}$, i.e. in $\mathbb{R}^N$.
We have $U = K \cup \sigma \cup L = K \cup \triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma} \cup L$ (because $\triangle_{K,\sigma} \subset K$ and $\triangle_{L,\sigma} \subset L$ by convexity of K and L), and this proves that U is open in $\mathbb{R}^N$ (K and L are open in $\mathbb{R}^N$).

Step 2: we prove that, for all $\lambda > 1$, $\overline{U} \subset \lambda U$ (recall that $0 \in \sigma \subset U$).
To see this, it is sufficient to show that, for all $z \in \overline{U}$, $]0,z[\subset U$; indeed, once we have obtained this result, we write, for $z \in \overline{U}\backslash\{0\}$ (the case $z = 0$ is obvious) and $\lambda > 1$, $z = \lambda(\frac{1}{\lambda}z)$ and, since $\frac{1}{\lambda}z = \frac{1}{\lambda}z + (1-\frac{1}{\lambda}) \times 0 \in]0,z[\subset U$ (because $\frac{1}{\lambda} \in]0,1[$), we deduce that $z \in \lambda U$.
Let us take $z \in \overline{U} = \overline{L} \cup \overline{\sigma} \cup \overline{K}$.

Assume first that $z \in \overline{\sigma}$. Then, since σ is an open convex subset of $\{0\} \times \mathbb{R}^{N-1}$ and $0 \in \sigma$, a classical convexity lemma tells us that $]0,z[\subset \sigma \subset U$, which concludes this case.

Assume now that $z \in \overline{K}\backslash\overline{\sigma}$ (the case $z \in \overline{L}\backslash\overline{\sigma}$ being treated the same way).
If $z \in K$, then by the same convexity lemma as before, since $0 \in \overline{K}$ and K is convex and open, we have $]0,z[\subset K \subset U$. We can thus suppose that $z \in \partial K\backslash\overline{\sigma}$.

Let us stop a moment to prove the following geometrical fact: $\partial K \cap (\{0\} \times \mathbb{R}^{N-1}) = \overline{\sigma}$.
We first notice that $K \cap (\{0\} \times \mathbb{R}^{N-1}) = \emptyset$: indeed, if it is not the case, then, taking $a \in K \cap (\{0\} \times \mathbb{R}^{N-1})$, since $0 \in \overline{K}$, we have $]0,a] \subset K$; but $0 \in \sigma$ which is open in $\{0\} \times \mathbb{R}^{N-1}$ and, since $]0,a] \subset \{0\} \times \mathbb{R}^{N-1}$ with $a \neq 0$ (because $0 \in \partial K$, which does not intersects K), we can find $b \in]0,a] \cap \sigma$; this means that $b \in \partial K \cap K$, which is not possible.
Thus, $A := \partial K \cap (\{0\} \times \mathbb{R}^{N-1})$ is equal to $\overline{K} \cap (\{0\} \times \mathbb{R}^{N-1})$ and is therefore convex (because $\overline{K}$ and $\{0\} \times \mathbb{R}^{N-1}$ are convex).
We have $\sigma \subset A$ (and thus $\overline{\sigma} \subset A$, A being closed); take $c \in A \subset \{0\} \times \mathbb{R}^{N-1}$. Since A is convex, the set $O = \cup_{0 \leq t < 1}(tc + (1-t)\sigma)$ is contained in A; we want to show that $O\backslash\overline{\sigma} = \emptyset$.
σ being open in $\{0\} \times \mathbb{R}^{N-1}$, O and thus $O\backslash\overline{\sigma}$ are also open in $\{0\} \times \mathbb{R}^{N-1}$. We have $O\backslash\overline{\sigma} \subset \partial K$, which implies $O\backslash\overline{\sigma} \subset \cup_{\{\sigma' \in \mathcal{E}_K\,,\ \sigma' \neq \sigma\}} \overline{\sigma'}$. Suppose that $O\backslash\overline{\sigma}$ is not empty; then, since it is an open subset of $\{0\} \times \mathbb{R}^{N-1}$ and $\{\sigma' \in \mathcal{E}_K,\ \sigma' \neq \sigma\}$ is finite, there exists $\sigma' \in \mathcal{E}_K\backslash\{\sigma\}$ whose adherence contains N points of $O\backslash\overline{\sigma} \subset \{0\} \times \mathbb{R}^{N-1}$ in general position. Thus, the hyperplane containing σ' is the affine space generated by these points, that is to say $\{0\} \times \mathbb{R}^{N-1}$. By hypothesis on the edges, the intersection of $\{0\} \times \mathbb{R}^{N-1}$ and of the line going through x_K and orthogonal to $\{0\} \times \mathbb{R}^{N-1}$ belongs to both σ and σ' (since σ and σ' both generate $\{0\} \times \mathbb{R}^{N-1}$), which is a contradiction with the fact that σ and σ' are disjoint.
Thus, $O\backslash\overline{\sigma}$ is empty, and $O \subset \overline{\sigma}$. Since $c \in \overline{O}$, we deduce that $c \in \overline{\sigma}$, and this concludes the proof that $A \subset \overline{\sigma}$, i.e. that $\partial K \cap (\{0\} \times \mathbb{R}^{N-1}) = \overline{\sigma}$.
We also notice that, since $K \cap (\{0\} \times \mathbb{R}^{N-1}) = \emptyset$ (see above) and $x_K \in]0,\infty[\times \mathbb{R}^{N-1}$, by connexity of K, we have $K \subset]0,\infty[\times \mathbb{R}^{N-1}$. We prove the same way that $L \subset]-\infty,0[\times \mathbb{R}^{N-1}$.

Let us now return to the proof that, if $z \in \partial K\backslash\overline{\sigma}$, then $]0,z[\subset U$.
As we have seen before, $\triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma}$ is an open set and, since 0 belongs to this open set (and $z \neq 0$), $]0,z[\cap (\triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma})$ is not empty. We have $z \in \overline{K} \subset [0,\infty[\times \mathbb{R}^{N-1}$; since $z \in \partial K\backslash\overline{\sigma}$ and $\partial K \cap (\{0\} \times \mathbb{R}^{N-1}) = \overline{\sigma}$, this implies $z \notin \{0\} \times \mathbb{R}^{N-1}$ and thus $z \in]0,\infty[\times \mathbb{R}^{N-1}$. Hence, $]0,z[\subset]0,\infty[\times \mathbb{R}^{N-1}$.
But $\triangle_{L,\sigma} \subset L \subset]-\infty,0[\times \mathbb{R}^{N-1}$ and $\sigma \subset \{0\} \times \mathbb{R}^{N-1}$; therefore, $]0,z[\cap (\triangle_{K,\sigma} \cup \sigma \cup \triangle_{L,\sigma}) =]0,z[\cap \triangle_{K,\sigma}$.
This set being non-empty, we can take $c \in]0,z[\cap \triangle_{K,\sigma} \subset K$. Since 0 and z belong to $\overline{K}$, $]0,c]$ and $[c,z[$ are contained in K, which implies that $]0,z[=]0,c] \cup [c,z[\subset K$ and concludes this step.

Step 3: we prove the density result.
Take $v \in H^2(U)$ and define, for $\lambda > 1$, $v_\lambda(x) = v(x/\lambda)$; v_λ belongs to $H^2(\lambda U)$ and the restriction of v_λ to $U \subset \lambda U$ converges, as $\lambda \to 1$, to v in $H^2(U)$.
Indeed, to see the convergence in $L^2(U)$, we take $\varepsilon > 0$ and $w \in C_c(U)$ such that $||v - w||_{L^2(U)} < \varepsilon$; we then write, with $w_\lambda(x) = w(x/\lambda)$, $||v_\lambda - v||_{L^2(U)} \leq ||v_\lambda - w_\lambda||_{L^2(U)} + ||w_\lambda - w||_{L^2(U)} + ||w - v||_{L^2(U)}$. By a change of variable, we have $||v_\lambda - w_\lambda||_{L^2(U)} \leq ||v_\lambda - w_\lambda||_{L^2(\lambda U)} = \lambda^{N/2}||v - w||_{L^2(U)} \leq \lambda^{N/2}\varepsilon$, so that $||v_\lambda - v||_{L^2(U)} \leq (\lambda^{N/2} + 1)\varepsilon + ||w_\lambda - w||_{L^2(U)}$. Since $w \in C_c(U)$, the dominated convergence theorem (for example) gives $||w_\lambda - w||_{L^2(U)} \to 0$ as $\lambda \to 1$ and this concludes the proof of the L^2 convergence.
The first and second derivatives of v_λ being (with evident notations) $\lambda^{-1}(\nabla v)_\lambda$ and $\lambda^{-2}(D^2 v)_\lambda$, the H^2 convergence is an immediate consequence of the L^2 convergence showed above.
To approximate v in $H^2(U)$ by regular function, we thus just need to approximate v_λ in this space. We extend v_λ to $\mathbb{R}^N$ by 0 outside λU and take $(\rho_n)_{n \geq 1}$ a smoothing kernel; $v_\lambda * \rho_n \in C_c^\infty(\mathbb{R}^N)$ and, since U is relatively compact in λU and $v_\lambda \in H^2(\lambda U)$, we have $v_\lambda * \rho_n \to v_\lambda$ in $H^2(U)$ as $n \to \infty$ (because, for n large enough — such that $U + \mathrm{supp}(\rho_n) \subset \lambda U$ —, we have $\nabla(v_\lambda * \rho_n) = (\nabla v_\lambda) * \rho_n$ and $D^2(v_\lambda * \rho_n) = (D^2 v_\lambda) * \rho_n$ on U). This concludes the proof of the lemma. ■

Lemma 5.7.4 *If U is a convex open bounded set in $\mathbb{R}^N$ and $\mathbf{w} \in (C(\overline{U}))^N$ is such that* $\mathrm{div}(\mathbf{w}) \in L^1(U)$,

then

$$\int_U \mathrm{div}(\mathbf{w})(x)\,dx = \int_{\partial U} \mathbf{w}(\xi)\cdot\mathbf{n}(\xi)\,dm(\xi)$$

(m denotes here the $(N-1)$-dimensional measure on ∂U and $\mathbf{n}$ is the unit normal to ∂U outward to U — notice that, since U is convex, it has a Lipschitz-continuous boundary).

Proof of Lemma 5.7.4
By translation, we can suppose that $0 \in U$. Then, for all $z \in \overline{U}$, since U is convex, $[0, z[\in U$; as we have seen in step 2 of the proof of Lemma 5.7.3, this implies that, for all $\lambda > 1$, $\overline{U} \subset \lambda U$.
Let $\mathbf{w}_\lambda(x) = \mathbf{w}(x/\lambda)$; we have $\mathbf{w}_\lambda \in (C(\lambda\overline{U}))^N$ and $\mathbf{w}_\lambda \to \mathbf{w}$ uniformly on $\overline{U}$ (and thus on ∂U) as $\lambda \to 1$ (this is due to the uniform continuity of $\mathbf{w}$ on this set). Moreover, $\mathrm{div}(W_\lambda) = \lambda^{-1}(\mathrm{div}(\mathbf{w}))_\lambda \in L^1(\lambda U)$ and, as in step 3 of the proof of Lemma 5.7.3, we deduce that $\mathrm{div}(\mathbf{w}_\lambda) \to \mathrm{div}(\mathbf{w})$ in $L^1(U)$ as $\lambda \to 1$. It is thus sufficient to prove that, for all $\lambda > 1$, $\mathbf{w}_\lambda$ satisfies the result of the Lemma.
Let $(\rho_n)_{n\geq 1}$ be a smoothing kernel. Extend $\mathbf{w}_\lambda$ to $\mathbb{R}^N$ by 0 outside λU and define $\mathbf{w}_{n,\lambda} = \mathbf{w}_\lambda * \rho_n \in C_c^\infty(\mathbb{R}^N)$. By regularity of $W_{n,\lambda}$, we have

$$\int_U \mathrm{div}(\mathbf{w}_{n,\lambda})(x)\,dx = \int_{\partial U} \mathbf{w}_{n,\lambda}(\xi)\cdot\mathbf{n}(\xi)\,dm(\xi). \tag{5.7.6}$$

But $\mathbf{w}_\lambda$ is uniformly continuous on the open set λU, which contains the compact set ∂U; thus, $\mathbf{w}_{n,\lambda} \to \mathbf{w}_\lambda$ uniformly on ∂U.
We have, in the sense of the distributions on $\mathbb{R}^N$, $\mathrm{div}(\mathbf{w}_{n,\lambda}) = \mathrm{div}(\mathbf{w}_\lambda) * \rho_n$. Since $\mathrm{div}(\mathbf{w}_\lambda) \in L^1(\lambda U)$ and U is relatively compact in λU, we deduce that $\mathrm{div}(\mathbf{w}_{n,\lambda}) \to \mathrm{div}(\mathbf{w}_\lambda)$ in $L^1(U)$ as $n \to \infty$.
These convergences allow to pass to the limit in (5.7.6) to see that $\mathbf{w}_\lambda$ satisfies the result of the lemma, which concludes the proof. ■

5.7.2 Interpolation

We prove in this subsection that, if Ω is a bounded open subset of $\mathbb{R}^N$ ($N = 2$ or 3) with a Lipschitz-continous boundary and $\mathbf{v} \in (C(\overline{\Omega}))^N$ satisfies $\mathrm{div}(\mathbf{v}) \in L^2(\Omega)$, then, for all $\theta \in]0,1[$, the interpolate space of order θ between

$$A = \{(\overline{u}, G, f) \in (H^2(\Omega)\cap H_0^1(\Omega)) \times (H^1(\Omega))^N \times L^2(\Omega) \mid \Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} + \mathrm{div}(G) + f = 0\}$$

and

$$B = \{(\overline{u}, G, f) \in H_0^1(\Omega) \times (L^2(\Omega))^N \times L^2(\Omega) \mid \Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} + \mathrm{div}(G) + f = 0\}$$

is (with equivalent norms)

$$\begin{aligned} C \quad = \quad & \{(\overline{u}, G, f) \in [H^2(\Omega)\cap H_0^1(\Omega), H_0^1(\Omega)]_\theta \times (H^{1-\theta}(\Omega))^N \times L^2(\Omega) \mid \\ & \Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} + \mathrm{div}(G) + f = 0\}, \end{aligned}$$

each of these spaces being endowed by its natural norm (notice that $[H^1(\Omega), L^2(\Omega)]_\theta = H^{1-\theta}(\Omega)$).
This result is quite natural, but not so easy to prove.

To simplify the notations, we let $V = H_0^1(\Omega)$ and $W = H^2(\Omega)\cap H_0^1(\Omega)$. The proof relies on a result in [66]. Define the linear application

$$T \in \mathcal{L}(V \times (L^2(\Omega))^N \times L^2(\Omega); H^{-1}(\Omega)) \cap \mathcal{L}(W \times (H^1(\Omega))^N \times L^2(\Omega); L^2(\Omega))$$

by $T(\overline{u}, G, f) = \Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} + \mathrm{div}(G) + f$; this application is continuous $W \times (H^1(\Omega))^N \times L^2(\Omega) \to L^2(\Omega)$ because, since $\mathrm{div}(\mathbf{v}) \in L^2(\Omega)$ — this is the only place where we need this hypothesis — and $W \subset C(\overline{\Omega})$ — recall that $N \leq 3$ —, we have $\mathrm{div}(\mathbf{v}\overline{u}) = \mathrm{div}(\mathbf{v})\overline{u} + \mathbf{v}\cdot\nabla\overline{u} \in L^2(\Omega)$ when $\overline{u} \in W$. Then

$$A = \{x \in W \times (H^1(\Omega))^N \times L^2(\Omega) \mid T(x) = 0\}, \quad B = \{x \in V \times (L^2(\Omega))^N \times L^2(\Omega) \mid T(x) = 0\}$$

and Theorem 14.3 in [66] ([1]) allows to see that

$$[A,B]_\theta = \{x \in [W \times (H^1(\Omega))^N \times L^2(\Omega), V \times (L^2(\Omega))^N \times L^2(\Omega)]_\theta \mid T(x) = 0\} = C$$

with equivalent norms (notice that this last space is equal to C because the interpolate space of a product of spaces is the product of the corresponding interpolate spaces), provided that we can construct an application

$$R \in \mathcal{L}(H^{-1}(\Omega)); V \times (L^2(\Omega))^N \times L^2(\Omega)) \cap \mathcal{L}(L^2(\Omega); W \times (H^1(\Omega))^N \times L^2(\Omega))$$

such that $T \circ R = Id$ on $H^{-1}(\Omega)$.
The rest of this subsection is devoted to the construction of such a R.

The main difficulty in constructing this application is the lack of regularity of $\partial\Omega$. If Ω is a regular (or convex) open set and $\mathbf{v} \in (C^1(\overline{\Omega}))^N$ (for example), then R is quite easy to build: take, for $L \in H^{-1}(\Omega)$, $R(L) = (\overline{u}, 0, 0)$ where $\overline{u}$ is the variational solution of $\Delta\overline{u} - \mathrm{div}(\mathbf{v}\overline{u}) - b\overline{u} = L$ with Dirichlet boundary conditions; the regularity of $\partial\Omega$ ensures then that R is continuous $L^2(\Omega) \to W \times (H^1(\Omega))^N \times L^2(\Omega)$.
If Ω is a polygonal non-convex open set, we must find another way to construct R. The main idea is to get rid of Ω and to bring ourselves back to $\mathbb{R}^N$.

Following an idea of [17], we first build $r \in \mathcal{L}(L^2(\mathbb{R}^N); L^2(\Omega)) \cap \mathcal{L}(H^1(\mathbb{R}^N); H^1_0(\Omega))$ such that $r(\varphi) = \varphi$ for all $\varphi \in \mathcal{D}(\Omega)$.
To do so, we notice that, since Ω has a bounded Lipschitz boundary, so does $\mathbb{R}^N \backslash \Omega$; there exists thus an extension operator E which is continuous $H^1(\mathbb{R}^N \backslash \Omega) \to H^1(\mathbb{R}^N)$ and $L^2(\mathbb{R}^N \backslash \Omega) \to L^2(\mathbb{R}^N)$ (the classical extension operators constructed via symetries satisfy these continuities). We define r by $r(\varphi) = (\varphi - E(\varphi_{|\mathbb{R}^N \backslash \Omega}))_{|\Omega}$. Since $\varphi - E(\varphi_{|\mathbb{R}^N \backslash \Omega}) = 0$ on $\mathbb{R}^N \backslash \Omega$, we clearly have $r(\varphi) \in H^1_0(\Omega)$ if $\varphi \in H^1(\mathbb{R}^N)$; moreover, if $\varphi \in \mathcal{D}(\Omega)$, $\varphi_{|\mathbb{R}^N \backslash \Omega} = 0$ so that $r(\varphi) = \varphi$, and r has thus the desired properties.

r allows us to extend elements of $H^{-1}(\Omega)$ into elements of $H^{-1}(\mathbb{R}^N)$, in such a way that elements of $L^2(\Omega) \subset H^{-1}(\Omega)$ are extented into elements of $L^2(\mathbb{R}^N)$.
Indeed, $r^* : H^{-1}(\Omega) \to H^{-1}(\mathbb{R}^N)$ is linear continuous and, since $r : L^2(\mathbb{R}^N) \to L^2(\Omega)$ is continuous, r^* is also continuous $L^2(\Omega) \to L^2(\mathbb{R}^N)$ (we have identified, as usual, the dual space of L^2 to L^2 itself). Moreover, if $L \in H^{-1}(\Omega)$, one has $L = r^*(L)$ in $\mathcal{D}'(\Omega)$: indeed, for all $\varphi \in \mathcal{D}(\Omega)$, $r^*(L)(\varphi) = L(r(\varphi)) = L(\varphi)$.

Let $L \in H^{-1}(\Omega)$; since $r^*(L) \in H^{-1}(\mathbb{R}^N)$, we can define $w^L \in H^1(\mathbb{R}^N)$ as the variational solution of $-\Delta w^L + w^L = r^*(L)$ on $\mathbb{R}^N$, and, since $r^* : H^{-1}(\Omega) \to H^{-1}(\mathbb{R}^N)$ is linear continuous, the application $L \in H^{-1}(\Omega) \to w^L \in H^1(\mathbb{R}^N)$ is linear continuous. Moreover, r^* is also linear continuous $L^2(\Omega) \to L^2(\mathbb{R}^N)$ so that, by the regularity properties of $-\Delta + Id$ on $\mathbb{R}^N$, $L \to w^L$ is linear continuous $L^2(\Omega) \to H^2(\mathbb{R}^N)$. Define now $R(L) = (0, -\nabla(w^L_{|\Omega}), w^L_{\Omega})$. Since the restriction to Ω is linear continuous $L^2(\mathbb{R}^N) \to L^2(\Omega)$, $H^1(\mathbb{R}^N) \to H^1(\Omega)$ and $H^2(\mathbb{R}^N) \to H^2(\Omega)$, R is linear continuous $H^{-1}(\Omega) \to V \times (L^2(\Omega))^N \times L^2(\Omega)$ and $L^2(\Omega) \to W \times (H^1(\Omega))^N \times L^2(\Omega)$.
Moreover, for all $L \in H^{-1}(\Omega)$, $T \circ R(L) = \mathrm{div}(-\nabla(w^L_{|\Omega})) + w^L_{|\Omega} = -\Delta(w^L_{|\Omega}) + w^L_{|\Omega} = r^*(L)_{|\Omega} = L$ in $\mathcal{D}'(\Omega)$ (by properties of r^*), thus also in $H^{-1}(\Omega)$. This concludes the construction of R and this appendix.

[1] In fact, this theorem concerns the interpolation of complex banach spaces, and we consider here spaces of real-valued functions; but it is not very difficult to see, since we handle spaces of functions, that the result of this theorem is also valid in our case of real interpolation.

Chapitre 6

A finite volume scheme for a noncoercive elliptic equation with measure data

Reference: J. Droniou, T. Gallouët and R. Herbin. *SIAM J. Numer. Anal.* **41** (2003), no. 6, 1997-2031.

Abstract We show here the convergence of the finite volume approximate solutions of a convection-diffusion equation to a weak solution, without the usual coercitivity assumption on the elliptic operator and with weak regularity assumptions on the data. Numerical experiments are performed to obtain some rates of convergence in two and three space dimensions.

6.1 Introduction

The scope of this work is the discretization by the cell-centered finite volume method of convection-diffusion problems on general structured or non structured grids. Let Ω be a polygonal (or polyhedral) open subset of $\mathbb{R}^d$ ($d = 2$ or 3); the problem under study writes:

$$\begin{cases} -\Delta u + \mathrm{div}(\mathbf{v}u) + bu = \mu & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{6.1.1}$$

with the following hypotheses on the data:

$$\begin{gathered} \mathbf{v} \in (C(\overline{\Omega}))^d\,, \\ b \in L^2(\Omega)\,,\ b \geq 0 \text{ a.e. on } \Omega\,, \\ \mu \in M(\overline{\Omega})\,, \end{gathered} \tag{6.1.2}$$

where $M(\overline{\Omega}) = (C(\overline{\Omega}))'$ is the dual space of $C(\overline{\Omega})$, which may also be identified to the set of bounded measures on $\overline{\Omega}$. In the sequel, we shall consider the usual infinity norm on $C(\overline{\Omega})$), and we shall denote by $||\cdot||_{M(\overline{\Omega})}$ its dual norm on $M(\overline{\Omega})$.
Our purpose is to prove the convergence of the cell-centered finite volume scheme for the discretization of Problem (6.1.1). Cell-centered schemes for convection-diffusion equations using rectangular, triangular or Voronoï grids were analysed in a number of papers including [68],[45], [55], [63], [71], [27]. The analysis which we develop here uses some of the tools which were developped in [38], [48], [39] and [46]. In [39], a convergence result without any assumption of regularity of the solution is proved. An approximate gradient was constructed in [40]. Noncoercive elliptic equations with a regular H^{-1} right-hand-side were also recently studied [33]. Finally, a thorough study of finite volume schemes for linear or nonlinear

elliptic, parabolic and hyperbolic equations may be found in [38], which we refer to for further details. The discretization grids which are considered here and in these latter works consist of polygonal (or polyhedral) control volumes satisfying adequate geometrical conditions (which are stated in the sequel) and not necessarily ordered in a cartesian grid.

Let us remark that the analysis which is developed here still holds for equations of the type

$$-\mathrm{div}\Big(k(x)\nabla u(x)\Big) + \mathrm{div}\Big(\mathbf{v}(x)\,u(x)\Big) + b(x)u(x) = f(x), \quad x \in \Omega, \tag{6.1.3}$$

with the following hypotheses on k:

$$\begin{array}{l} k \text{ is a piecewise } C^1 \text{ function from } \overline{\Omega} \text{ to } \mathbb{R}; \\ \text{there exists } k_0 \in \mathbb{R}_+^* \text{ such that } k(x) \geq k_0 \text{ for a.e. } x \in \Omega. \end{array} \tag{6.1.4}$$

For the sake of the simplicity of notations we prefer to deal with the Laplace operator here but we shall point out the modifications which take place if the operator $\mathrm{div}(k\nabla.)$ is considered instead: see remarks 6.2.2, 6.2.4 and 6.2.6. If now k is a tensor satisfying the following hypotheses:

$$\begin{array}{l} k \text{ is a piecewise } C^1 \text{ function from } \overline{\Omega} \text{ to } \mathbb{R}^{d\times d}, \\ \text{for all } x \in \overline{\Omega}, k(x) \text{ is a symmetric matrix}, \\ \text{there exists } k_0 \in \mathbb{R}_+^* \text{ such that } k(x)\xi\cdot\xi \geq k_0 \text{ for a.e. } x \in \Omega \text{ and for all } \xi \in \mathbb{R}^d, \end{array} \tag{6.1.5}$$

then one may still write the finite volume scheme and obtain some error estimates in the regular case, but the assumptions on the mesh have to be modified see [48], [56] and [27]. However if the mesh is Cartesian and if for all $x \in \overline{\Omega}$ the matrix $k(x)$ is diagonal then it is "aligned" with the grid and the analysis is similar to the (non constant) scalar case of Equation (6.1.3).
The originality of the present work with respect to the above cited works is threefold: first, the elliptic operator associated to the convection-diffusion equation is not assumed to be coercive; second, the convection velocity $\mathbf{v}$ is only assumed to be continuous (it was assumed C^1 in previous works); third, the right hand side μ is only supposed to be a Radon measure.

In the next section, the finite volume scheme for the discretization of (6.1.1) is presented, along with the admissible meshes. We then state the main convergence theorem of this paper (Theorem 6.2.1), along with some preliminary technical results similar to those used in [38], [48], [39], and the proof of which is given in an appendix. Section 3 is devoted to *a priori* estimates on the approximate solutions (existence is not proven at this stage), which will be needed in order to obtain compactness results, and which also yield the existence and uniqueness of the approximate solution. The proof of Theorem 6.2.1, that is the proof of the convergence of the approximate solutions to the weak solution of (6.1.1), is then given in Section 4. Section 5 presents a modified finite volume scheme where the measure data whose support is on the edges of the mesh are taken into account through a jump of the flux between two neighboring cells; comparing this scheme to the scheme of Section 2, the convergence result is easy to obtain. Finally, we present in Section 6 some numerical results in two and three space dimensions, using Cartesian or unstructured triangular meshes (in 2D), as well as for a spherical geometry. These results allow to derive some rates of convergence of the method, even though no error estimate is known theoretically.

6.2 Conservative finite volume discretization and convergence result

Definition 6.2.1 *An admissible mesh of Ω, denoted by $\mathcal{M}$, is given by a finite partition $\mathcal{T}$ of Ω in polygonal (or polyhedral) convex sets (the "control volumes"), by a finite family $\mathcal{E}$ of disjoint subsets of $\overline{\Omega}$ contained in affine hyperplanes (the "edges") and by a family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points in Ω such that*

i) each $\sigma \in \mathcal{E}$ is a non-empty open subset of ∂K for some $K \in \mathcal{T}$,

ii) by denoting $\mathcal{E}_K = \{\sigma \in \mathcal{E} \mid \sigma \subset \partial K\}$, *one has* $\partial K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$ *for all* $K \in \mathcal{T}$,

iii) for all $K \neq L$ *in* $\mathcal{T}$, *either the* $(d-1)$*-dimensional measure of* $\overline{K} \cap \overline{L}$ *is null, or* $\overline{K} \cap \overline{L} = \overline{\sigma}$ *for some* $\sigma \in \mathcal{E}$, *that we denote then* $\sigma = K|L$,

iv) for all $K \in \mathcal{T}$, x_K *is in the interior of* K,

v) for all $\sigma = K|L \in \mathcal{E}$, *the line* (x_K, x_L) *intersects and is orthogonal to* σ,

vi) for all $\sigma \in \mathcal{E}$, $\sigma \subset \partial\Omega \cap \partial K$, *the line which is orthogonal to* σ *and going through* x_K *intersects* σ.

Remark 6.2.1 (Other admissible meshes) *Note that Property v) in the above definition is required so as to obtain a consistent discretization of the normal fluxes over the boundary of the control domains when using the two points finite difference scheme to discretize the normal flux. In fact, the above definition of an admissible mesh may be extended to other geometries of* Ω *than a polygone or a polyhedron. For instance, if* $\Omega = \{x \in \mathbb{R}^d; |x| \leq r\}$ *is a spherical ball of radius* r, *then a natural mesh is defined by the control volumes* $K_0 = \{x \in \mathbb{R}^d; |x| \leq r_{1/2}\}$ *and, for* $i = 1, N$, $K_i = \{x \in \mathbb{R}^d; r_{i-1/2} \leq |x| \leq r_{i+1/2}\}$ *where* $(r_{i+1/2})_{i=1,N} \subset (0, r]$ *is a given increasing sequence such that* $r_{N+1/2} = r$. *Let* $x_0 = 0$ *and, for* $i = 1, \ldots, N$, $r_i \in (r_{i-1/2}, r_{i+1/2})$, *then a discretization of the normal diffusive flux* $\nabla u \cdot \mathbf{n}$ *(where* $\mathbf{n}$ *is the outward normal unit vector) over the sphere* $\{x \in \mathbb{R}^d; |x| = r_{i+1/2}\}$ *by the two points scheme* $\frac{u_{i+1} - u_i}{r_{i+1} - r_i}$ *is clearly consistent if the solution* u *to (6.1.1) only depends on* r. *Moreover, if* $r_{i+1/2} = \frac{1}{2}(r_{i+1} - r_i)$, *it is consistent of order 2. Hence this class of spherical discretizations is clearly admissible for the analysis which will be derived in the sequel.*

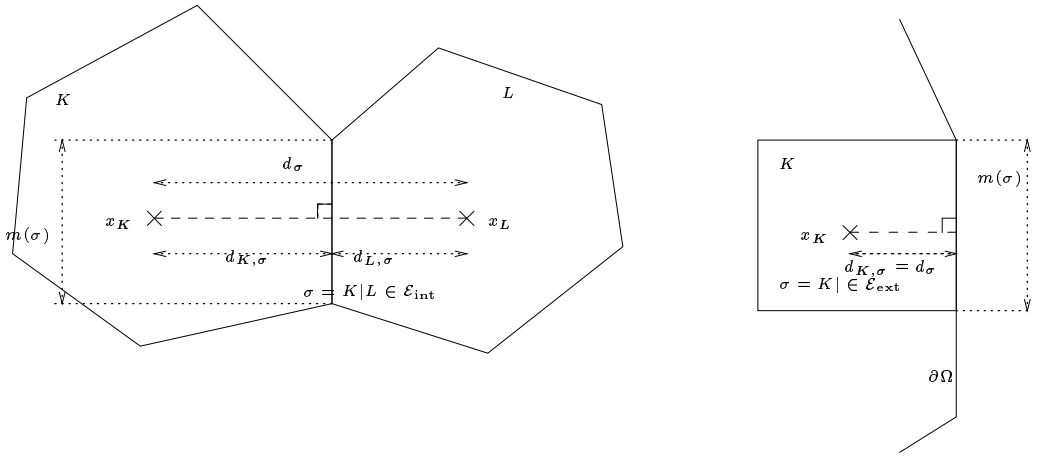


Figure 6.1: Notations for an admissible mesh

The size of the mesh is then defined by $\text{size}(\mathcal{M}) = \sup_{K \in \mathcal{T}} \text{diam}(K)$. We denote by $\text{meas}(K)$ the Lebesgue measure of $K \in \mathcal{T}$. The unit normal to $\sigma \in \mathcal{E}_K$ outward to K is denoted by $\mathbf{n}_{K,\sigma}$.
We define $\mathcal{E}_{\rm int} = \{\sigma \in \mathcal{E} \mid \sigma \not\subset \partial\Omega\}$ and $\mathcal{E}_{\rm ext} = \mathcal{E} \backslash \mathcal{E}_{\rm int}$. If $\sigma \in \mathcal{E}$, $\text{meas}(\sigma)$ is the $(d-1)$-dimensional measure of σ; if $\sigma = K|L \in \mathcal{E}_{\rm int}$, d_σ is the distance between the points (x_K, x_L) and $d_{K,\sigma}$ denotes the distance between x_K and σ; if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, $d_\sigma = d_{K,\sigma}$ is the distance between x_K and σ. The transmissivity through an edge σ is

$$\tau_\sigma = \frac{\text{meas}(\sigma)}{d_\sigma}.$$

Within the integrals, the letter λ (resp. γ) stands for the d (resp.$(d-1)$)-dimensional measure on the domain Ω (resp. on the edges of the mesh). Note that moth measures are denoted by "meas" when applied to a control volume or an edge.

We shall naturally identify the set $\mathbb{R}^{\mathrm{Card}(\mathcal{T})}$ to the set $X(\mathcal{T})$ of functions defined a.e. on Ω and constant on each control volume $K \in \mathcal{T}$.

Remark 6.2.2 *In the case of the operator* $\mathrm{div}(k\nabla .)$ *which is considered in Equation (6.1.3) where* k *is a function from* $\overline{\Omega}$ *to* $\mathbb{R}$ *or* $\mathbb{R}^{d\times d}$ *which satisfies (6.1.4) or (6.1.5), admissible meshes must satisfy the following additional condition:*

(vi) For any $K \in \mathcal{T}$, *the restriction* $k|_K$ *of the function* k *to any given control volume* K *belongs to* $C^1(\overline{K})$.

Furthermore if k *is a piecewise* C^1 *function from* $\overline{\Omega}$ *to* $\mathbb{R}^{d\times d}$, *the orthogonality conditions* (iv) *and* (v) *have to be modified into:*

(iv)' For any $K \in \mathcal{T}$, *let* k_K *denote the mean value of* k *on* K, *that is*

$$k_K = \frac{1}{meas(K)} \int_K k d\lambda. \tag{6.2.1}$$

The set $\mathcal{T}$ *is such that there exists a family of points*

$$\mathcal{P} = (x_K)_{K\in\mathcal{T}} \text{ such that } x_K = \cap_{\sigma\in\mathcal{E}_K} \mathcal{D}_{K,\sigma,k} \in \overline{K},$$

where $\mathcal{D}_{K,\sigma,k}$ *is a straight line perpendicular to* σ *with respect to the scalar product induced by* k_K^{-1} *such that* $\mathcal{D}_{K,\sigma,k} \cap \sigma = \mathcal{D}_{L,\sigma,k} \cap \sigma \neq \emptyset$ *if* $\sigma = K|L$. *Furthermore, if* $\sigma = K|L$, *let* $y_\sigma = \mathcal{D}_{K,\sigma,k} \cap \sigma (= \mathcal{D}_{L,\sigma,k} \cap \sigma)$ *and assume that* $x_K \neq x_L$.

(v)' For any $\sigma \in \mathcal{E}_{\mathrm{ext}}$, *let* K *be the control volume such that* $\sigma \in \mathcal{E}_K$ *and let* $\mathcal{D}_{K,\sigma,k}$ *be the straight line going through* x_K *and orthogonal to* σ *with respect to the scalar product induced by* k_K^{-1}; *then, there exists* $y_\sigma \in \sigma \cap \mathcal{D}_{K,\sigma,k}$.

If $\mathcal{M}$ is an admissible mesh, and under Hypothesis (6.1.2), we can define the finite volume discretization of (6.1.1).
By denoting, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$b_K = \frac{1}{\mathrm{meas}(K)} \int_K b d\lambda \quad \text{and} \quad v_{K,\sigma} = \int_\sigma \mathbf{v}\cdot\mathbf{n}_{K,\sigma}\, d\gamma \tag{6.2.2}$$

the scheme is defined by

$$\forall K \in \mathcal{T}\,, \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+} + \mathrm{meas}(K) b_K u_K = \mu(K)\,, \tag{6.2.3}$$

$$\begin{array}{l} \forall \sigma = K|L \in \mathcal{E}_{\mathrm{int}}\,, \quad F_{K,\sigma} = -\tau_\sigma(u_L - u_K)\,, \\ \forall \sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K\,, \quad F_{K,\sigma} = \tau_\sigma u_K\,, \end{array} \tag{6.2.4}$$

$$\begin{array}{l} \forall \sigma = K|L \in \mathcal{E}_{\mathrm{int}}\,, \quad u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = u_L \text{ otherwise,} \\ \forall \sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K\,, \quad u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = 0 \text{ otherwise.} \end{array} \tag{6.2.5}$$

Equations (6.2.3)—(6.2.5) form a linear system in $(u_K)_{K\in\mathcal{T}}$ of size $\mathrm{Card}(\mathcal{T})$. Notice that this scheme is conservative in the sense that if $\sigma = K|L$, then $F_{K,\sigma} = -F_{L,\sigma}$ and $v_{K,\sigma} = -v_{L,\sigma}$.

Remark 6.2.3 *The approximation (6.2.5) of the convective flux is the classical upwind scheme, which we choose here because it ensures both the existence of a solution to the scheme (and the maximum principle) without any condition on the size of the mesh. If instead of the upwind scheme, we used the central difference scheme, then we would need a condition on the size of the mesh in order to have existence of a solution to the scheme, and in order for the maximum principle to hold. However, when*

the size of the mesh tends to 0, the centered scheme may also be shown to converge. The upwind scheme is often preferred in applications because of its robustness on coarse meshes.
Also note that if $v_{K,\sigma} = 0$, for some $\sigma = K|L$ for example, then (6.2.5) does not determine $u_{\sigma,+}$ uniquely since one may take either $u_{\sigma,+} = u_K$ (since $v_{K,\sigma} \geq 0$) or $u_{\sigma,+} = u_L$ (since $v_{L,\sigma} = -v_{K,\sigma} = 0 \geq 0$). However, this is no real problem since $u_{\sigma,+}$ always appears multiplied by $v_{K,\sigma}$ or $v_{L,\sigma}$ and thus, if $v_{K,\sigma} = 0$, the value of $u_{\sigma,+}$ does not matter (one can, for example, reduce the second sum of (6.2.3) to the $\sigma \in \mathcal{E}_K$ such that $v_{K,\sigma} \neq 0$).

Remark 6.2.4 *In the case of a non constant diffusion coefficient as in Equation (6.1.3) where k is a function from Ω to $\mathbb{R}$ satisfying (6.1.4) or from Ω to $\mathbb{R}^{d\times d}$ satisfying (6.1.5), one considers admissible meshes satisfying (vi) of Remark 6.2.2 and in the tensor case also $(iv)'$ and $(v)'$ instead of (iv) and (v). For $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, let*

$$k_{K,\sigma} = \left| \frac{1}{\mathrm{meas}(K)} \int_K k \, d\lambda \, \mathbf{n}_{K,\sigma} , \right| \tag{6.2.6}$$

(where $|\cdot|$ denotes the Euclidean norm). Note that in the scalar case, this yields in fact $k_{K,\sigma} = \frac{1}{\mathrm{meas}(K)} \int_K k d\lambda$. The exact diffusion fluxes $k(x)\nabla u \cdot \mathbf{n}_{K,\sigma}$ on an edge σ of the mesh may then be approximated in a consistent way (see [38] and [56]) by replacing the formulae in (6.2.4) by:

- *internal edges:*

$$F_{K,\sigma} = -\tau_\sigma(u_L - u_K), \ \textit{if } \sigma \in \mathcal{E}_{\mathrm{int}}, \, \sigma = K|L, \tag{6.2.7}$$

where

$$\tau_\sigma = \mathrm{meas}(\sigma) \frac{k_{K,\sigma} k_{L,\sigma}}{k_{K,\sigma} d_{L,\sigma} + k_{L,\sigma} d_{K,\sigma}};$$

- *boundary edges:*

$$F_{K,\sigma} = -\tau_\sigma(u_\sigma - u_K), \ \textit{if } \sigma \in \mathcal{E}_{\mathrm{ext}} \textit{ and } x_K \notin \sigma, \tag{6.2.8}$$

where

$$\tau_\sigma = \mathrm{meas}(\sigma) \frac{k_{K,\sigma}}{d_{K,\sigma}}.$$

Let us now state our main result, which we shall prove in the following sections.

Theorem 6.2.1 *If $\mathcal{M}$ is an admissible mesh, then there exists a unique solution to (6.2.3)—(6.2.5). Moreover, if $(\mathcal{M}_n)_{n\geq 1}$ is a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying*

$$\textit{for all } n \geq 1, \textit{ for all } K \in \mathcal{T}_n, \textit{ for all } \sigma \in \mathcal{E}_K, \ d_{K,\sigma} \geq \zeta d_\sigma ,$$

and such that $\mathrm{size}(\mathcal{M}_n) \to 0$, *then, by denoting $u_n \in X(\mathcal{T}_n)$ the solution of (6.2.3)—(6.2.5) with $\mathcal{M} = \mathcal{M}_n$, $(u_n)_{n\geq 1}$ converges to u in $L^p(\Omega)$ for all $p \in [1, \frac{d}{d-2})$, where u is the unique solution to (6.1.1) in the sense*

$$\begin{cases} u \in \bigcap\limits_{q<\frac{d}{d-1}} W_0^{1,q}(\Omega) , \\ \displaystyle\int_\Omega \nabla u \cdot \nabla \varphi d\lambda - \int_\Omega u\mathbf{v} \cdot \nabla \varphi d\lambda + \int_\Omega bu\varphi d\lambda = \int_\Omega \varphi \, d\mu \, , \ \forall \varphi \in \bigcup\limits_{s>d} W_0^{1,s}(\Omega), \end{cases} \tag{6.2.9}$$

where $\int_\Omega \varphi \, d\mu =< \mu, \varphi >_{(C(\overline{\Omega}))', C(\overline{\Omega})}$. (We recall that $W^{1,q}(\Omega)$ is the set of functions which belong to $L^q(\Omega)$ and such that their derivatives are also in $L^q(\Omega)$, and $W_0^{1,q}(\Omega) = \overline{C_c^\infty(\Omega)}^{W^{1,q}(\Omega)}$. We also recall that $W_0^{1,s}(\Omega) \subset C_0(\overline{\Omega})$ for $s > d$.)

Remark 6.2.5 *Notice that we do not suppose the existence and uniqueness of a solution to (6.2.9); we will prove both.*

Remark 6.2.6 *A convergence result still holds if a non constant piecewise C^1 diffusion scalar coefficient is considered i.e. if k satisfies (6.1.4) and if Equation (6.1.3) is discretized by the scheme (6.2.2),(6.2.3),(6.2.6)—(6.2.8). In fact, in the two-dimensional case, the proof follows the one given below in the case $k = Id$. In the three-dimensional case however, the regularity of the solution to the dual problem (6.4.1), which is used in the proof of the uniqueness of a solution to (6.2.9) (see section 6.4) is not so clear. Hence in the 3D case, uniqueness of a solution to (6.2.9) is not known, and the convergence result of Theorem (6.2.1) still holds, but only up to a subsequence.*
If one now considers the general tensor case, then some more restrictive assumptions are needed on the mesh in order to obtain consistency of the fluxes, see [38] and [56].

The proof of existence and uniqueness of a solution to (6.2.3)—(6.2.5) is based on *a priori* estimates on the solutions to this problem, which are obtained with the following discrete $W_0^{1,q}$ norm, defined as follows.

Definition 6.2.2 (Discrete $W^{1,q}$ norm) *If $\mathcal{M}$ is an admissible mesh, $v_\mathcal{T} = (v_K)_{K\in\mathcal{T}} \in \mathbb{R}^{\mathrm{Card}(\mathcal{T})}$ and $1 \le q < \infty$, we define*

$$||v_\mathcal{T}||_{1,q,\mathcal{M}} = \left(\sum_{\sigma\in\mathcal{E}} meas(\sigma) d_\sigma \left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q\right)^{1/q},$$

where $D_\sigma v_\mathcal{T} = |v_K - v_L|$ if $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$ and $D_\sigma v_\mathcal{T} = |v_K|$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$.

Let us now state the main *a priori* estimate, which will be proven in Section 6.3. This estimate is crucial to prove the existence of a solution to (6.2.3)—(6.2.5), and also to obtain the compactness properties on approximate solutions which will eventually yield the convergence result.

Theorem 6.2.2 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying*

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}_K,\ d_{K,\sigma} \ge \zeta d_\sigma. \tag{6.2.10}$$

Then, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on $(\Omega, \mathbf{v}, q, \zeta)$ such that, if $u_\mathcal{T} \in X(\mathcal{T})$ is a solution to (6.2.3)—(6.2.5), then $||u_\mathcal{T}||_{1,q,\mathcal{M}} \le C||\mu||_{M(\overline{\Omega})}$.

In the sequel, we shall use the following properties of the discrete $W_0^{1,q}$ norm:

Proposition 6.2.1 (Discrete Poincaré inequality) *If $1 \le q \le 2$, $\mathcal{M}$ is an admissible mesh and $v_\mathcal{T} \in X(\mathcal{T})$, then*

$$\|v_\mathcal{T}\|_{L^q(\Omega)} \le \mathrm{diam}(\Omega)||v_\mathcal{T}||_{1,q,\mathcal{M}}. \tag{6.2.11}$$

Proposition 6.2.2 (Discrete Sobolev Inequality) *Let $1 \le q \le 2$, $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (6.2.10). Then, with $q^* = \frac{dq}{d-q}$ if $q < d$ and $q^* < \infty$ if $q = d = 2$, there exists $C > 0$ only depending on (Ω, q, q^*, ζ) such that, for all $v_\mathcal{T} \in X(\mathcal{T})$,*

$$||v_\mathcal{T}||_{L^{q*}(\Omega)} \le C||v_\mathcal{T}||_{1,q,\mathcal{M}}.$$

In fact, it is easily seen that the above inequality also holds for any $r \le q^*$, that is:

$$||v_\mathcal{T}||_{L^r(\Omega)} \le C||v_\mathcal{T}||_{1,q,\mathcal{M}}, \text{ for any } r \le q^*.$$

Proposition 6.2.3 (Discrete Rellich Theorem) *Let $1 \leq q \leq 2$ and $\mathcal{M}$ be an admissible mesh. Then there exists $C > 0$ only depending on (Ω, q) such that, for all $h \in \mathbb{R}^d$ and all $v_\mathcal{T} \in X(\mathcal{T})$, denoting $w_\mathcal{T}$ the extension of $v_\mathcal{T}$ to $\mathbb{R}^d$ by 0 outside Ω, we have*

$$\int_{\mathbb{R}^d} |w_\mathcal{T}(x+h) - w_\mathcal{T}(x)|^q \, d\lambda(x) \leq |h|(|h| + C\text{size}(\mathcal{M}))^{q-1} ||v_\mathcal{T}||^q_{1,q,\mathcal{M}}. \tag{6.2.12}$$

In particular, if $(\mathcal{M}_n)_{n\geq 1}$ is a sequence of admissible meshes and $v_n \in X(\mathcal{T}_n)$ is such that $(||v_n||_{1,q,\mathcal{M}_n})_{n\geq 1}$ is bounded, then $(v_n)_{n\geq 1}$ is relatively compact in $L^q(\Omega)$.

Proposition 6.2.4 (Regularity of the limit) *Let $q \in (1,2]$ and $(\mathcal{M}_n)_{n\geq 1}$ be a sequence of admissible meshes such that $\text{size}(\mathcal{M}_n) \to 0$. If $v_n \in X(\mathcal{T}_n)$, $(||v_n||_{1,q,\mathcal{M}_n})_{n\geq 1}$ is bounded and $v_n \to v$ in $L^q(\Omega)$, then $v \in W^{1,q}_0(\Omega)$.*

These propositions are easy adaptations of similar results in [38] for the case $q = 2$ (see also [26] for Proposition 6.2.2 and [46] for Proposition 6.2.3). We sketch the proofs of these propositions in the appendix for the sake of completeness.

6.3 A Priori Estimates

The aim of this section is to prove the discrete $W^{1,q}$ *a priori* estimate of Theorem 6.2.2, which is crucial in the proof of existence of the scheme, and also in the obtention of a compactness result which will allow to prove the convergence of a sequence of approximate solutions (Theorem 6.2.1 and its proof in Section 4).
Such *a priori* estimates were already used for the study of the finite volume approximation of nonlinear elliptic or parabolic equations, see e.g. [39], [41]. But in these previous works, the estimates were obtained in a discrete H^1 norm, accordingly with the regularity of the solution of the continuous problem.
We prove here some *a priori* estimates on the solution to (6.2.3)—(6.2.5) in a discrete $W^{1,q}$ norm, since the solution to the continuous problem is in $W^{1,q}$. As in the continuous case, it is difficult to obtain an estimate on $u_\mathcal{T}$ itself (note that in the continuous case, u is not allowed as a test function in (6.2.9)). Hence, as in [46], we shall obtain estimates on truncations of the approximate solutions, that is the functions $T_k(u_\mathcal{T})$, where T_k is defined in Figure 6.2. However, in [46], we only dealt with the Laplace operator, whereas here we allow non-coercive convection-diffusion operators. Because of this non-coercivity, we shall need to start with some weaker estimates, namely an estimate on $\ln(1 + |u_\mathcal{T}|)$, as was done in [31] in the continuous case. In order to obtain this estimate, we shall obtain some estimate on $S_k(u_\mathcal{T})$, where $S_k = Id - T_k$ is also defined in Figure 6.2 and section 6.3.2. Note that in the diffusion dominated case, the operator is becomes coercive and the discrete $W^{1,q}$ estimate may be directly obtained from the estimates on $T_k(u_\mathcal{T})$ as in [46].
Since the function T_k is bounded, the estimate on $T_k(u_\mathcal{T})$ is easy to obtain. The estimate on $S_k(u_\mathcal{T})$ is more tricky. The convective term is controlled through a bound of $\text{meas}(E_k)$ where $E_k = \{|u_\mathcal{T}| > k\}$ (see Corollary 6.3.1), which is a consequence of an estimate on $\ln(1 + |u_\mathcal{T}|)$ (see Proposition 6.3.1).

Each of the estimates we present here has a continuous counterpart; see for example [11], [14] for estimates on nonlinear elliptic equations with measure data and [31], [32] for estimates on linear and nonlinear noncoercive variational elliptic problems. Mixing the techniques of [14] and [31] (or [32]), we can prove estimates (and an existence result) on solutions to linear or nonlinear noncoercive elliptic equations with measure data.
To obtain the estimates on the solutions to (6.2.3)—(6.2.5), we adapt to the discrete setting this mix of techniques of [14] and [31]. Thus, to make the following proofs easier to understand, we sketch, for each of the discrete estimate, the proof of the corresponding continuous estimate.

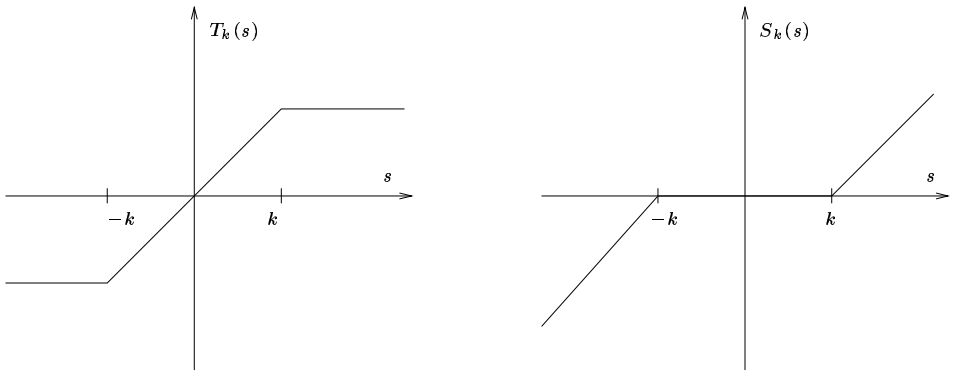


Figure 6.2: The functions T_k and S_k

6.3.1 Estimate on $\ln(1+|u_\mathcal{T}|)$

Proposition 6.3.1 *Let $\mathcal{M}$ be an admissible mesh. If $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ is a solution to (6.2.3)—(6.2.5), then*

$$|| \ln(1+|u_\mathcal{T}|)||^2_{1,2,\mathcal{T}} \leq 2||\mu||_{M(\overline{\Omega})} + d\text{meas}(\Omega)\, ||\, |\mathbf{v}|\, ||^2_{L^\infty(\Omega)} \tag{6.3.1}$$

(where $|\mathbf{v}|$ denotes the euclidean norm of $\mathbf{v}$ in $\mathbb{R}^d$).

Before we prove Proposition 6.3.1, let us state an easy corollary, which is used in the proof of the estimate of Proposition 6.3.2.

Corollary 6.3.1 *Let $\mathcal{M}$ be an admissible mesh. If $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ is a solution to (6.2.3)—(6.2.5) and, for $k>0$, $E_k = \{|u_\mathcal{T}| > k\}$, then there exists $C \in \mathbb{R}^*_+$ only depending on $(\Omega, \mathbf{v})$ such that*

$$\text{meas}(E_k) \leq \frac{C(1+||\mu||_{M(\overline{\Omega})})}{(\ln(1+k))^2}.$$

Proof of Corollary 6.3.1
By Proposition 6.3.1, we get that

$$|| \ln(1+|u_\mathcal{T}|)||^2_{1,2,\mathcal{T}} \leq (2 + d\text{meas}(\Omega)\, ||\, |\mathbf{v}|\, ||^2_{L^\infty(\Omega)})(1+||\mu||_{M(\overline{\Omega})}).$$

Therefore, using the discrete Poincaré inequality (Proposition 6.2.1), we get that there exists $C \in \mathbb{R}^*_+$ only depending on $(\Omega, \mathbf{v})$ such that:

$$|| \ln(1+|u_\mathcal{T}|)||^2_{L^2(\Omega)} \leq C(1+||\mu||_{M(\overline{\Omega})}).$$

Finally, since $\text{meas}(E_k) = \text{meas}(\{\ln(1+|u_\mathcal{T}|) \geq \ln(1+k)\})$, the Chebyshev inequality yields that $\text{meas}(E_k) \leq \frac{C(1+||\mu||_{M(\overline{\Omega})})}{(\ln(1+k))^2}$. ■

Proof of Proposition 6.3.1
Step 0: sketch of the proof in the continuous case.
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. Suppose that $\mu \in H^{-1}(\Omega) \cap L^1(\Omega)$ and let $u \in H^1_0(\Omega)$ be a variational solution of (6.1.1). Using $\varphi(u)$ as a test function in the equation satisfied by u, and since φ is bounded by 1, we find:

$$\int_\Omega \nabla u \cdot \frac{\nabla u}{(1+|u|)^2}\, d\lambda + \int_\Omega bu\varphi(u)\, d\lambda \leq ||\mu||_{L^1(\Omega)} + ||\, |\mathbf{v}|\, ||_{L^\infty(\Omega)} \int_\Omega |u| \frac{|\nabla u|}{(1+|u|)^2}\, d\lambda \leq C + C\int_\Omega \frac{|\nabla u|}{(1+|u|)}\, d\lambda,$$

where C only depends on $||\mu||_{L^1(\Omega)}$ and $\mathbf{v}$. Since $\nabla(\ln(1+|u|)) = \mathrm{sgn}(u)\frac{\nabla u}{(1+|u|)}$ and $bu\varphi(u) \geq 0$ (b is nonnegative and $\varphi(s)$ has the same sign as s), we deduce that

$$|| \, |\nabla(\ln(1+|u|))| \, ||^2_{L^2(\Omega)} \leq C + C\mathrm{meas}(\Omega)^{1/2} || \, |\nabla(\ln(1+|u|))| \, ||_{L^2(\Omega)},$$

which gives an estimate on $|| \, |\nabla(\ln(1+|u|))| \, ||_{L^2(\Omega)}$ (and thus, by the Poincaré inequality, also on $||\ln(1+|u|)||_{L^2(\Omega)}$).

Step 1: proof of a first discrete estimate.
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. Multiplying each equality of (6.2.3) by $\varphi(u_K)$ and summing on $K \in \mathcal{T}$, we have

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}\varphi(u_K) + \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+}\varphi(u_K) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(u_K) = \sum_{K\in\mathcal{T}} \mu(K)\varphi(u_K). \tag{6.3.2}$$

Gathering by edges and using (6.2.4), we can write

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \tag{6.3.3}$$

where we let $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\mathrm{int}}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$.
By the conservativity of the fluxes, still gathering by edges, we find

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} u_{\sigma,+} v_{K,\sigma}(\varphi(u_K) - \varphi(u_L))$$

(recall that $u_L = 0$ — so that $\varphi(u_L) = 0$ — if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$). If $\sigma \in \mathcal{E}$, we denote $v_\sigma = |v_{K,\sigma}|$ for a $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$ (the definition of v_σ does not depend on the choice of such a K) and $u_{\sigma,-}$ the downstream choice of u, i.e. $u_{\sigma,-}$ is such that $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$ (where $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\mathrm{int}}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$).
Let $\sigma \in \mathcal{E}$; if $v_{K,\sigma} \geq 0$, then $u_{\sigma,+} = u_K$ and $u_{\sigma,-} = u_L$ so that $v_{K,\sigma}(\varphi(u_K) - \varphi(u_L)) = v_\sigma(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-}))$; if $v_{K,\sigma} < 0$, then $u_{\sigma,+} = u_L$ and $u_{\sigma,-} = u_K$, which gives $v_{K,\sigma}(\varphi(u_K) - \varphi(u_L)) = -v_\sigma(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) = v_\sigma(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-}))$. Thus,

$$\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-})). \tag{6.3.4}$$

b being nonnegative and $\varphi(s)$ having the same sign as s,

$$\sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(u_K) \geq 0. \tag{6.3.5}$$

Since φ is bounded by 1 and $\mathcal{T}$ is a partition of Ω,

$$\left| \sum_{K\in\mathcal{T}} \mu(K)\varphi(u_K) \right| \leq \sum_{K\in\mathcal{T}} |\mu(K)| \leq |\mu|(\Omega) = ||\mu||_{M(\overline{\Omega})}. \tag{6.3.6}$$

Using (6.3.3), (6.3.4), (6.3.5) and (6.3.6) in (6.3.2), we get

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \leq ||\mu||_{M(\overline{\Omega})} + \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})). \tag{6.3.7}$$

We now study each term of the last sum a little more precisely. We use the fact that φ is nondecreasing.

- If $u_{\sigma,+} \geq u_{\sigma,-}$ and $u_{\sigma,+} \geq 0$, then $\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+}) \leq 0$ and $u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$.
- If $u_{\sigma,+} \geq u_{\sigma,-}$ and $u_{\sigma,+} < 0$, then $0 > u_{\sigma,+} \geq u_{\sigma,-}$, so that $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$.
- If $u_{\sigma,+} < u_{\sigma,-}$ and $u_{\sigma,+} \geq 0$, then $0 \leq u_{\sigma,+} < u_{\sigma,-}$, so that $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$.
- If $u_{\sigma,+} < u_{\sigma,-}$ and $u_{\sigma,+} < 0$, then $\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+}) \geq 0$ and $u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$.

By denoting $\mathcal{A} = \{\sigma \in \mathcal{E} \mid u_{\sigma,+} \geq u_{\sigma,-},\ u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E} \mid u_{\sigma,+} < u_{\sigma,-},\ u_{\sigma,+} \geq 0\}$, we notice thus that, for all $\sigma \in \mathcal{E} \backslash \mathcal{A}$, $v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$. This gives

$$\sum_{\sigma \in \mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq \sum_{\sigma \in \mathcal{A}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})).$$

As $v_\sigma \leq \mathrm{meas}(\sigma)|| \, |\mathbf{v}| \, ||_{L^\infty(\Omega)}$, we deduce, using the Cauchy-Schwarz inequality, that

$$\begin{aligned}
&\sum_{\sigma \in \mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \\
&\quad \leq \quad || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \sum_{\sigma \in \mathcal{A}} \mathrm{meas}(\sigma) |u_{\sigma,+}| |\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})| \\
&\quad \leq \quad || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \left(\sum_{\sigma \in \mathcal{A}} \mathrm{meas}(\sigma) d_\sigma \right)^{\frac{1}{2}} \left(\sum_{\sigma \in \mathcal{A}} \tau_\sigma u_{\sigma,+}^2 (\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+}))^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

But $\sum_{\sigma \in \mathcal{A}} \mathrm{meas}(\sigma) d_\sigma \leq \sum_{\sigma \in \mathcal{E}} \mathrm{meas}(\sigma) d_\sigma = d\mathrm{meas}(\Omega)$ and, if $\sigma \in \mathcal{A}$, $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$, thus, by Lemma 6.3.1 below and Young's inequality,

$$\begin{aligned}
&\sum_{\sigma \in \mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \\
&\quad \leq \quad (d\mathrm{meas}(\Omega))^{1/2} || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \left(\sum_{\sigma \in \mathcal{A}} \tau_\sigma (u_{\sigma,-} - u_{\sigma,+})(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \right)^{\frac{1}{2}} \\
&\quad \leq \quad \frac{1}{2} d\mathrm{meas}(\Omega) \, || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)}^2 + \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \tau_\sigma (u_{\sigma,-} - u_{\sigma,+})(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})).
\end{aligned}$$

For all $\sigma \in \mathcal{E}$, we have $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$, so that $(u_{\sigma,-} - u_{\sigma,+})(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) = (u_K - u_L)(\varphi(u_K) - \varphi(u_L))$. Coming back to (6.3.7), we obtain

$$\sum_{\sigma \in \mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \leq 2||\mu||_{M(\overline{\Omega})} + d\mathrm{meas}(\Omega) \, || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)}^2, \tag{6.3.8}$$

which concludes this step.

Step 2: Estimate on $\ln(1 + |u_\mathcal{T}|)$.
We notice that, for all $s \in \mathbb{R}$, $\ln(1 + |s|) = \int_0^s \frac{\mathrm{sgn}(t)\, dt}{1+|t|}$. Thus, for all $(x, y) \in \mathbb{R}^2$, by the Cauchy-Schwarz inequality and since φ is nondecreasing,

$$\begin{aligned}
(\ln(1 + |x|) - \ln(1 + |y|))^2 \quad &= \quad \left(\int_y^x \frac{\mathrm{sgn}(t)\, dt}{1 + |t|} \right)^2 \\
&\leq \quad |x - y| \left| \int_y^x \frac{dt}{(1 + |t|)^2} \right| = |x - y||\varphi(x) - \varphi(y)| = (x - y)(\varphi(x) - \varphi(y)).
\end{aligned}$$

Using this upper bound and (6.3.8), we deduce the result of the proposition. ■

Let us now state and prove the technical result which was used in Step 1 of the above proof.

Lemma 6.3.1 *Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. If $(x,y) \in \mathbb{R}^2$ have the same sign and $|x| \leq |y|$, then*

$$x^2(\varphi(y) - \varphi(x))^2 \leq (y-x)(\varphi(y) - \varphi(x)). \tag{6.3.9}$$

Proof of Lemma 6.3.1
Since φ is C^1-continuous on $\mathbb{R}$, there exists $\theta \in [x,y]$ such that $\varphi(y) - \varphi(x) = \varphi'(\theta)(y-x)$, so that, since φ is nondecreasing,

$$\begin{aligned} x^2(\varphi(y) - \varphi(x))^2 &\leq \frac{x^2}{(1+|\theta|)^2}|y-x|\,|\varphi(y) - \varphi(x)| \\ &\leq \frac{x^2}{(1+|\theta|)^2}(y-x)(\varphi(y) - \varphi(x)). \end{aligned}$$

But $|x| \leq |y|$ and x and y have the same sign, so that, since $\theta \in [x,y]$, we have $|\theta| \geq |x|$, and (6.3.9) is thus a consequence of the previous inequality. ■

6.3.2 Estimate on $||u_{\mathcal{T}}||_{1,q,\mathcal{M}}$

We denote, for $k > 0$, $T_k(s) = \max(-k, \min(s,k))$ and $S_k(s) = s - T_k(s)$ (see Figure 6.2).

Proposition 6.3.2 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (6.2.10). We suppose that μ satisfies $||\mu||_{M(\overline{\Omega})} \leq 1$. Then there exists $k_0 > 0$ only depending on $(\Omega, \mathbf{v}, \zeta)$ and, for all $m \in (1,2)$, $C > 0$ only depending on $(\Omega, \mathbf{v}, m, \zeta)$ such that, if $u_{\mathcal{T}}$ is a solution to (6.2.3)—(6.2.5) and $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$, we have*

$$\sum_{\sigma \in \mathcal{E}} \tau_\sigma (S_{k_0}(u_K) - S_{k_0}(u_L))(\varphi_m(S_{k_0}(u_K)) - \varphi_m(S_{k_0}(u_L))) \leq C \tag{6.3.10}$$

and

$$\sum_{\sigma \in \mathcal{E}} \tau_\sigma (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \leq C, \tag{6.3.11}$$

where we let $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.

Remark 6.3.1 *Problem (6.2.3)—(6.2.5) being linear, there is no loss of generality in the estimate if we consider measures of norm less than 1, as we will see in Theorem 6.2.2*

Proof of Proposition 6.3.2
Step 0: sketch of the estimate in the continuous case.
Suppose that $u \in H_0^1(\Omega)$ is a variational solution of (6.1.1) with $\mu \in H^{-1}(\Omega) \cap L^1(\Omega)$ satisfying $||\mu||_{L^1(\Omega)} \leq 1$, and take $\varphi_m(S_k(u))$ as a test function in (6.2.9). Using the fact that $bu\varphi_m(S_k(u)) \geq 0$ (b is nonnegative and $\varphi_m(s)$ and $S_k(s)$ have the same sign as s), that $\nabla(S_k(u)) = \nabla u$ where $\nabla(S_k(u)) \neq 0$ and that φ_m is bounded by $1/(m-1)$, we have

$$\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda \leq \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \int_\Omega |u| \frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^m}\,d\lambda.$$

But $|u| \leq k + |S_k(u)|$ and $(1+|S_k(u)|)^{2m} \geq (1+|S_k(u)|)^m$, so that, by the Cauchy-Schwarz inequality,

$$\begin{aligned} &\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda \\ &\quad\leq \frac{1}{m-1} + C_1 k \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^{2m}}\,d\lambda\right)^{\frac{1}{2}} + C_1 \int_\Omega \frac{|S_k(u)|}{(1+|S_k(u)|)^{\frac{m}{2}}} \frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^{\frac{m}{2}}}\,d\lambda \\ &\quad\leq \frac{1}{m-1} + C_1 k \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda\right)^{\frac{1}{2}} + C_1 ||\psi(S_k(u))||_{L^2(\Omega)} \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda\right)^{\frac{1}{2}}, \end{aligned} \tag{6.3.12}$$

where C_1 only depends on $(\Omega, \mathbf{v})$ and $\psi(s) = \frac{|s|}{(1+|s|)^{\frac{m}{2}}}$.
Now, by the Hölder inequality and the Sobolev injection, and since $\psi(S_k(u)) = 0$ outside $E_k = \{|u| > k\}$, there exists $r > 2$ only depending on d, and C_2 only depending on (Ω, r) (notice that a dependence on Ω takes into account a dependence on d), such that

$$||\psi(S_k(u))||_{L^2(\Omega)} \leq \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} ||\psi(S_k(u))||_{L^r(\Omega)} \leq C_2 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} || \, |\nabla(\psi(S_k(u)))| \, ||_{L^2(\Omega)}. \quad (6.3.13)$$

Since $|\psi'(s)| \leq \frac{1+\frac{m}{2}}{(1+|s|)^{\frac{m}{2}}} \leq \frac{2}{(1+|s|)^{\frac{m}{2}}}$, one has

$$|| \, |\nabla(\psi(S_k(u)))| \, ||_{L^2(\Omega)} \leq 2 \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m} \, d\lambda \right)^{\frac{1}{2}}. \quad (6.3.14)$$

Gathering (6.3.12), (6.3.13) and (6.3.14), we find C_3 only depending on $(\Omega, \mathbf{v})$ such that

$$\begin{aligned} &\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m} \, d\lambda \\ &\quad \leq \frac{C_1}{m-1} + C_1 k \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m} \, d\lambda \right)^{\frac{1}{2}} + C_3 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} \int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m} \, d\lambda. \end{aligned} \quad (6.3.15)$$

Thanks to a continuous equivalent of Corollary 6.3.1, there exists C_4 only depending on $(\Omega, \mathbf{v})$ such that $\mathrm{meas}(E_k) \leq \frac{C_4}{(\ln(1+k))^2}$. Thus, there exists $k_0 > 0$ only depending on (C_4, C_3, r) (i.e. on $(\Omega, \mathbf{v})$) such that $C_3 \mathrm{meas}(E_{k_0})^{\frac{1}{2}-\frac{1}{r}} \leq \frac{1}{2}$. Applying (6.3.15) to this k_0 gives

$$\int_\Omega \frac{|\nabla(S_{k_0}(u))|^2}{(1+|S_{k_0}(u)|)^m} \, d\lambda \leq C_5$$

where C_5 only depends on $(\Omega, \mathbf{v}, m)$, which is the continuous equivalent of (6.3.10).

The estimate on $T_{k_0}(u)$ is quite simple and well known (see [11]). Take $\varphi_m(T_{k_0}(u))$ as a test function in the equation satisfied by u; since $\nabla(T_{k_0}(u)) = 0$ outside $\{|u| \leq k_0\}$ and $(1+|T_{k_0}(u)|)^{2m} \geq (1+|T_{k_0}(u)|)^m$, we find

$$\begin{aligned} \int_\Omega \frac{|\nabla(T_{k_0}(u))|^2}{(1+|T_{k_0}(u)|)^m} \, d\lambda &\leq \frac{1}{m-1} + || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \int_{\{|u| \leq k_0\}} |u| \frac{|\nabla(T_{k_0}(u))|}{(1+|T_{k_0}(u)|)^m} \, d\lambda \\ &\leq \frac{1}{m-1} + || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} k_0 \mathrm{meas}(\Omega)^{1/2} \left(\int_\Omega \frac{|\nabla(T_{k_0}(u))|^2}{(1+|T_{k_0}(u)|)^m} \, d\lambda \right)^{\frac{1}{2}}. \end{aligned}$$

This gives an estimate on $T_{k_0}(u)$ which is the continuous equivalent of (6.3.11).

Step 1: estimate on $S_k(u_\mathcal{T})$.
Let $\mathcal{M}$ be an admissible mesh and take $u_\mathcal{T}$ a solution of (6.2.3)—(6.2.5). Multiplying each equation of (6.2.3) by $\varphi_m(S_k(u_K))$, summing on $K \in \mathcal{T}$ and gathering by edges, we find

$$\begin{aligned} &\sum_{\sigma \in \mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) + \sum_{K \in \mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi_m(S_k(u_K)) \\ &\quad = \sum_{K \in \mathcal{T}} \mu(K) \varphi_m(S_k(u_K)) - \sum_{\sigma \in \mathcal{E}} v_{K,\sigma} u_{\sigma,+} (\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \end{aligned} \quad (6.3.16)$$

(recall that, if $\sigma \in \mathcal{E}_{\mathrm{int}}$, we use the notation $\sigma = K|L$ and, if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$, we set $u_L = 0$).
The function φ_m is bounded by $\frac{1}{m-1}$ and $\mathcal{T}$ is a partition of Ω, so that

$$\left| \sum_{K \in \mathcal{T}} \mu(K) \varphi_m(S_k(u_K)) \right| \leq \frac{1}{m-1} \sum_{K \in \mathcal{T}} |\mu(K)| \leq \frac{||\mu||_{M(\overline{\Omega})}}{m-1} \leq \frac{1}{m-1} \quad (6.3.17)$$

We again denote $u_{\sigma,-}$ the downstream choice of u_σ (i.e. $u_{\sigma,-} = u_L$ if $v_{K,\sigma} \geq 0$ and $u_{\sigma,-} = u_K$ otherwise) and $v_\sigma = |v_{K,\sigma}|$ (for a $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$); we have then:

$$-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}\, u_{\sigma,+}(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) = \sum_{\sigma\in\mathcal{E}} v_\sigma\, u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))).$$

But, as in the proof of Proposition 6.3.1 (because $\varphi_m \circ S_k$ is nondecreasing), we have $u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))) \leq 0$ if $\sigma \notin \mathcal{A}$, where $\mathcal{A} = \{\sigma \in \mathcal{E} \mid u_{\sigma,+} \geq u_{\sigma,-}\ ,\ u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E} \mid u_{\sigma,+} < u_{\sigma,-}\ ,\ u_{\sigma,+} \geq 0\}$. Thus,

$$\begin{aligned}
&-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}\, u_{\sigma,+}(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \\
&\leq \sum_{\sigma\in\mathcal{A}} v_\sigma\, u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))) \\
&\leq ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))|. \qquad (6.3.18)
\end{aligned}$$

Let $a_{k,\sigma} = \int_0^1 \varphi_m'(S_k(u_{\sigma,+}) + t(S_k(u_{\sigma,-}) - S_k(u_{\sigma,+})))\, dt \geq 0$, so that

$$\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+})) = a_{k,\sigma}(S_k(u_{\sigma,-}) - S_k(u_{\sigma,+})). \qquad (6.3.19)$$

We can write

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))| \\
&= \sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma) a_{k,\sigma}^{1/2}|u_{\sigma,+}| a_{k,\sigma}^{1/2}|S_k(u_{\sigma,-}) - S_k(u_{\sigma,+})| \\
&\leq \left(\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma) d_\sigma a_{k,\sigma} u_{\sigma,+}^2\right)^{\frac{1}{2}} \left(\sum_{\sigma\in\mathcal{A}} \tau_\sigma a_{k,\sigma}(S_k(u_{\sigma,-}) - S_k(u_{\sigma,+}))^2\right)^{\frac{1}{2}}.
\end{aligned}$$

But, by (6.3.19), $a_{k,\sigma}(S_k(u_{\sigma,-}) - S_k(u_{\sigma,+}))^2 = (S_k(u_{\sigma,-}) - S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+})))$, so that

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))| \\
&\leq \left(\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma) d_\sigma a_{k,\sigma} u_{\sigma,+}^2\right)^{\frac{1}{2}} \\
&\quad \times \left(\sum_{\sigma\in\mathcal{A}} \tau_\sigma (S_k(u_{\sigma,-}) - S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+})))\right)^{\frac{1}{2}}. \qquad (6.3.20)
\end{aligned}$$

Moreover, for all $\sigma \in \mathcal{A}$, $u_{\sigma,+}$ and $u_{\sigma,-}$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$. Thus, for such σ, $(S_k(u_{\sigma,+}), S_k(u_{\sigma,-}))$ have the same sign and $|S_k(u_{\sigma,+})| \leq |S_k(u_{\sigma,-})|$ and, by Lemma 6.3.2 stated after this proof, we deduce that

$$a_{k,\sigma} \leq \frac{1}{(1 + |S_k(u_{\sigma,+})|)^m} \leq 1.$$

Since $|u_{\sigma,+}| \leq k + |S_k(u_{\sigma,+})|$, we deduce that

$$a_{k,\sigma} u_{\sigma,+}^2 \leq 2k^2 + 2\frac{|S_k(u_{\sigma,+})|^2}{(1 + |S_k(u_{\sigma,+})|)^m},$$

which gives, in (6.3.20), using $\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma \le \sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma = d\mathrm{meas}(\Omega)$ and $(\alpha+\beta)^{1/2}\le \alpha^{1/2}+\beta^{1/2}$ for all nonnegative (α,β),

$$\begin{aligned}&\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|\\ &\quad\le\ \sqrt{2d\mathrm{meas}(\Omega)}kA_k+\sqrt{2}A_k\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\psi(S_k(u_{\sigma,+}))^2\right)^{\frac12},\end{aligned}\tag{6.3.21}$$

where $\psi(s)=\frac{|s|}{(1+|s|)^{\frac{m}{2}}}$ and

$$\begin{aligned}A_k &= \left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(S_k(u_{\sigma,-})-S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+})))\right)^{\frac12}\\ &= \left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(S_k(u_K)-S_k(u_L))(\varphi_m(S_k(u_K))-\varphi_m(S_k(u_L)))\right)^{\frac12}\end{aligned}$$

(recall that $\sigma=K|L$ if $\sigma\in\mathcal{E}_{\rm int}$, that $u_L=0$ if $\sigma\in\mathcal{E}_{\rm ext}\cap\mathcal{E}_K$ and that $\{u_{\sigma,+},u_{\sigma,-}\}=\{u_K,u_L\}$ for all $\sigma\in\mathcal{E}$).
We have, since $d_{K,\sigma}\ge\zeta d_\sigma$ for all $K\in\mathcal{T}$ and all $\sigma\in\mathcal{E}_K$,

$$\begin{aligned}\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\psi(S_k(u_{\sigma,+}))^2 &\le \sum_{K\in\mathcal{T}}\psi(S_k(u_K))^2\left(\sum_{\sigma\in\mathcal{A}\cap\mathcal{E}_K\ |\ v_{K,\sigma}\ge0}\mathrm{meas}(\sigma)d_\sigma\right)\\ &\le \frac{1}{\zeta}\sum_{K\in\mathcal{T}}\psi(S_k(u_K))^2\left(\sum_{\sigma\in\mathcal{E}_K}\mathrm{meas}(\sigma)d_{K,\sigma}\right)\\ &= \frac{1}{\zeta}\sum_{K\in\mathcal{T}}\psi(S_k(u_K))^2\times d\mathrm{meas}(K)=\frac{d}{\zeta}||\psi(S_k(u_\mathcal{T}))||^2_{L^2(\Omega)}.\end{aligned}$$

By Proposition 6.2.2, and since $\psi(S_k(u_\mathcal{T}))=0$ outside $E_k=\{|u_\mathcal{T}|>k\}$, we can thus find $r>2$ and $C_1>0$ only depending on (Ω,ζ) such that

$$\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\psi(S_k(u_{\sigma,+}))^2\right)^{\frac12}\le C_1\mathrm{meas}(E_k)^{\frac12-\frac1r}||\psi(S_k(u_\mathcal{T}))||_{1,2,\mathcal{M}}.$$

But, by Lemma 6.3.3 below and the definition of A_k,

$$||\psi(S_k(u_\mathcal{T}))||^2_{1,2,\mathcal{M}}=\sum_{\sigma\in\mathcal{E}}\tau_\sigma(\psi(S_k(u_K))-\psi(S_k(u_L))^2\le 4A_k^2,$$

so that

$$\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\psi(S_k(u_{\sigma,+}))^2\right)^{\frac12}\le 2C_1A_k\mathrm{meas}(E_k)^{\frac12-\frac1r}.$$

Returning to (6.3.21), we thus find

$$\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|\le\sqrt{2d\mathrm{meas}(\Omega)}kA_k+2\sqrt{2}C_1\mathrm{meas}(E_k)^{\frac12-\frac1r}A_k^2.\tag{6.3.22}$$

(6.3.16), (6.3.17), (6.3.18), (6.3.22) and the fact that $b_K u_K \varphi_m(S_k(u_K)) \geq 0$ then give

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \\
&\quad\leq \quad \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \sqrt{2d\mathrm{meas}(\Omega)} k A_k + 2\sqrt{2}||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} C_1 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} A_k^2 \\
&\quad\leq \quad \frac{1}{m-1} + C_2 k^2 + \frac{1}{2} A_k^2 + C_2 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} A_k^2, \qquad (6.3.23)
\end{aligned}$$

where C_2 only depends on $(\Omega, \mathbf{v}, \zeta)$. But φ_m and S_k are nondecreasing and S_k is Lipschitz-continuous with Lipschitz constant 1 so that

$$(S_k(u_K) - S_k(u_L))(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \leq (u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L)))$$

and (6.3.23) gives

$$\begin{aligned}
\sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_k(u_K) - S_k(u_L))(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) &\leq \frac{2}{m-1} + 2C_2 k^2 \\
&+ 2C_2 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} \sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_k(u_K) - S_k(u_L))(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \qquad (6.3.24)
\end{aligned}$$

By Corollary 6.3.1, there exists $k_0 > 0$ only depending on $(\Omega, \mathbf{v}, C_2, r)$ (i.e. only depending on $(\Omega, \mathbf{v}, \zeta)$) such that $2C_2 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} \leq \frac{1}{2}$. We deduce from (6.3.24) that

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_{k_0}(u_K) - S_{k_0}(u_L))(\varphi_m(S_{k_0}(u_K)) - \varphi_m(S_{k_0}(u_L))) \leq \frac{4}{m-1} + 4C_2 k_0^2,$$

which gives (6.3.10).

Step 2: Estimate on $T_{k_0}(u_\mathcal{T})$.
Multiplying each equation of (6.2.3) by $\varphi_m(T_{k_0}(u_K))$, summing on $K \in \mathcal{T}$ and re-ordering the sums on the edges, we find

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi_m(T_{k_0}(u_K)) \\
&\quad= \quad \sum_{K\in\mathcal{T}} \mu(K) \varphi_m(T_{k_0}(u_K)) - \sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+} (\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))). \qquad (6.3.25)
\end{aligned}$$

As before, we have

$$\left| \sum_{K\in\mathcal{T}} \mu(K) \varphi_m(T_{k_0}(u_K)) \right| \leq \frac{1}{m-1} \qquad (6.3.26)$$

and, with the previous notations,

$$\begin{aligned}
-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+} (\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) &= \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+} (\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+}))) \\
&\leq \sum_{\sigma\in\mathcal{A}} v_\sigma u_{\sigma,+} (\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+}))).
\end{aligned}$$

If $\sigma \in \mathcal{A}$, then $0 \leq u_{\sigma,+} \leq u_{\sigma,-}$ or $u_{\sigma,-} \leq u_{\sigma,+} \leq 0$. In either case, if $|u_{\sigma,+}| \geq k_0$, then $T_{k_0}(u_{\sigma,+}) = T_{k_0}(u_{\sigma,-})$, so that $\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+})) = 0$. Thus, in the previous sum, we can suppress the

terms $\sigma \in \mathcal{A}$ such that $|u_{\sigma,+}| \geq k_0$ and we have

$$\begin{aligned}
&-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+}(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \\
&\leq\ k_0 \sum_{\sigma\in\mathcal{A}} v_\sigma |\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+}))| \\
&\leq\ k_0 ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma\right)^{\frac{1}{2}} \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+})))^2\right)^{\frac{1}{2}}.
\end{aligned}$$

φ_m and T_{k_0} are nondecreasing and φ_m is Lipschitz-continuous with Lipschitz constant 1, thus, for all $\sigma \in \mathcal{E}$,

$$\begin{aligned}
(\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+})))^2 &\leq (T_{k_0}(u_{\sigma,-}) - T_{k_0}(u_{\sigma,+}))(\varphi_m(T_{k_0}(u_{\sigma,-})) - \varphi_m(T_{k_0}(u_{\sigma,+}))) \\
&= (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))).
\end{aligned}$$

Using this inequality and the fact that $\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma = d\mathrm{meas}(\Omega)$, we find

$$\begin{aligned}
&-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+}(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \\
&\leq\ k_0 ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \sqrt{d\mathrm{meas}(\Omega)} \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L)))\right)^{\frac{1}{2}} \quad (6.3.27)
\end{aligned}$$

Since φ_m and T_{k_0} are nondecreasing and T_{k_0} is Lipschitz-continuous with Lipschitz constant 1, we have

$$(T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \leq (u_K - u_L)(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))).$$

Combined with (6.3.25), (6.3.26), (6.3.27) and the fact that $b_K u_K \varphi_m(T_{k_0}(u_K)) \geq 0$, this inequality gives

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} \tau_\sigma (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \\
&\leq\ \frac{1}{m-1} + k_0 ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \sqrt{d\mathrm{meas}(\Omega)} \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L)))\right)^{\frac{1}{2}}
\end{aligned}$$

from which we deduce (6.3.11). ■

There remains to state and prove the two technical lemmas which were used in Step 1 of the above proof.

Lemma 6.3.2 *Let $m \in (1,2)$ and $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$. If (x,y) have the same sign and $|x| \leq |y|$, then*

$$\int_0^1 \varphi_m'(x + t(y-x))\, dt \leq \frac{1}{(1+|x|)^m}.$$

Proof of Lemma 6.3.2
Suppose that $0 \leq x \leq y$. Then, for all $t \in [0,1]$, $0 \leq x \leq x + t(y-x)$, so that $\varphi_m'(x + t(y-x)) = \frac{1}{(1+(x+t(y-x)))^m} \leq \frac{1}{(1+|x|)^m}$. Integrating this relation on $[0,1]$ gives the desired inequality. If $y \leq x \leq 0$, we use the fact that φ_m' is even and apply the previous result to $(-x,-y)$. ■

Lemma 6.3.3 *Let $m \in (1,2)$, $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$ and $\psi(s) = \frac{|s|}{(1+|s|)^{\frac{m}{2}}}$. Then for all $(x,y) \in \mathbb{R}^2$, one has*

$$(\psi(x) - \psi(y))^2 \leq 4(x-y)(\varphi_m(x) - \varphi_m(y)).$$

Proof of Lemma 6.3.3
The function ψ is Lipschitz-continuous and, for all $s \in \mathbb{R}$,

$$|\psi'(s)| = \left| \frac{\mathrm{sgn}(s)}{(1+|s|)^{\frac{m}{2}}} - \frac{\frac{m}{2}\mathrm{sgn}(s)|s|}{(1+|s|)^{1+\frac{m}{2}}} \right| \leq \frac{1+\frac{m}{2}}{(1+|s|)^{\frac{m}{2}}} \leq \frac{2}{(1+|s|)^{\frac{m}{2}}},$$

so that, for all $(x,y) \in \mathbb{R}^2$, by the Cauchy-Schwarz inequality,

$$|\psi(x)-\psi(y)| = \left| \int_y^x \psi'(s)\,ds \right| \leq \left| \int_y^x \frac{4\,ds}{(1+|s|)^m} \right|^{1/2} |x-y|^{1/2} \leq 2|\varphi_m(x)-\varphi_m(y)|^{1/2}|x-y|^{1/2}.$$

Taking the power 2 of this inequality and using the fact that φ_m is nondecreasing, we deduce the desired inequality. ■

We shall now deduce the key estimate on $u_{\mathcal{T}}$ (Theorem 6.2.2) from Proposition 6.3.2 and the following lemma.

Lemma 6.3.4 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (6.2.10). Let $F : (1,2) \to \mathbb{R}^+$ be a function. For $m \in (1,2)$, we denote $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$. If $v_{\mathcal{T}} = (v_K)_{K\in\mathcal{T}} \in X(\mathcal{T})$ satisfies, for all $m \in (1,2)$,*

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (v_K - v_L)(\varphi_m(v_K) - \varphi_m(v_L)) \leq F(m)$$

(where we have denoted, as usual, $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$), then, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on (Ω, ζ, F, q) such that $||v_{\mathcal{T}}||_{1,q,\mathcal{M}} \leq C$.

Proof of Lemma 6.3.4
Let $q \in [1, \frac{d}{d-1})$.
Take $m \in (1,2)$ (fixed later on as a function of d and q) and denote $a_{m,\sigma} = \int_0^1 \varphi_m'(v_K + t(v_L - v_K))\,dt$. We have $\varphi_m(v_K) - \varphi_m(v_L) = (v_K - v_L)a_{m,\sigma}$, so that

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma a_{m,\sigma} (D_\sigma v_{\mathcal{T}})^2 \leq F(m).$$

By Hölder's inequality, we have, since $1 \leq q < 2$,

$$\begin{aligned}
\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma}\right)^q &\leq \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma}\right)^2 \right)^{\frac{q}{2}} \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma}^{-\frac{q}{2-q}} \right)^{\frac{2-q}{2}} \\
&\leq F(m)^{\frac{q}{2}} \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma}^{-\frac{q}{2-q}} \right)^{\frac{2-q}{2}}. \qquad (6.3.28)
\end{aligned}$$

For all $(x,y) \in \mathbb{R}^2$ and all $t \in [0,1]$, one has $|x + t(y-x)| \leq \sup(|x|,|y|)$, so that

$$\varphi_m'(x+t(y-x)) = \frac{1}{(1+|x+t(y-x)|)^m} \geq \frac{1}{(1+\sup(|x|,|y|))^m} \geq \inf\left(\frac{1}{(1+|x|)^m}, \frac{1}{(1+|y|)^m} \right).$$

Taking $x = v_K$, $y = v_L$ and integrating the previous inequality on $t \in [0,1]$, we find

$$a_{m,\sigma} \geq \inf\left(\frac{1}{(1+|v_K|)^m}, \frac{1}{(1+|v_L|)^m} \right),$$

which implies

$$a_{m,\sigma}^{-\frac{q}{2-q}} \leq \sup\left((1+|v_K|)^{\frac{mq}{2-q}}, (1+|v_L|)^{\frac{mq}{2-q}} \right) \leq 2^{\frac{mq}{2-q}} (1 + |v_K|^{\frac{mq}{2-q}} + |v_L|^{\frac{mq}{2-q}}).$$

We deduce from (6.3.28), using the fact that $\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma = d\mathrm{meas}(\Omega)$ and re-ordering the sum on the control volumes,

$$||v_\mathcal{T}||^q_{1,q,\mathcal{M}} \le C_1 \left(1 + \sum_{K\in\mathcal{T}} |v_K|^{\frac{mq}{2-q}} \left(\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_\sigma\right)\right)^{\frac{2-q}{2}},$$

where C_1 only depends on (F, m, q, Ω). But since $d_{K,\sigma} \ge \zeta d_\sigma$ for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$, we have $\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_\sigma \le \frac{1}{\zeta}\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_{K,\sigma} = \frac{d}{\zeta}\mathrm{meas}(K)$ and we obtain thus

$$||v_\mathcal{T}||^q_{1,q,\mathcal{M}} \le C_2 \left(1 + ||v_\mathcal{T}||^{\frac{mq}{2}}_{L^{\frac{mq}{2-q}}(\Omega)}\right), \tag{6.3.29}$$

where C_2 only depends on (F, m, q, Ω) (notice that, since $m > 1$, we always have $\frac{mq}{2-q} \ge 1$).
By Proposition 6.2.2, there exists C_3 only depending on (Ω, q, ζ) such that, if $q^* = \frac{dq}{d-q}$ (note that $q < \frac{d}{d-1} \le d$),

$$||v_\mathcal{T}||_{L^{q^*}(\Omega)} \le C_3||v_\mathcal{T}||_{1,q,\mathcal{M}}.$$

Using this in (6.3.29), we obtain

$$||v_\mathcal{T}||^q_{L^{q^*}(\Omega)} \le C_3^q C_2 \left(1 + ||v_\mathcal{T}||^{\frac{mq}{2}}_{L^{\frac{mq}{2-q}}(\Omega)}\right).$$

If $q < \frac{d}{d-1}$, one has $\frac{q}{2-q} < q^*$, so that we can choose $m \in (1, 2)$ (only depending on (q, d)) such that $\frac{mq}{2-q} \le q^*$. We obtain thus, with such a choice of m and Hölder's inequality,

$$||v_\mathcal{T}||^q_{L^{q^*}(\Omega)} \le C_4 \left(1 + ||v_\mathcal{T}||^{\frac{mq}{2}}_{L^{q^*}(\Omega)}\right),$$

where C_4 only depends on (Ω, ζ, q, F). Since $\frac{mq}{2} < q$ (recall that $m < 2$), this inequality gives us C_5 only depending on (Ω, ζ, q, F) such that $||v_\mathcal{T}||_{L^{q^*}(\Omega)} \le C_5$ and, returning to (6.3.29), we deduce the desired estimate on $||v_\mathcal{T}||_{1,q,\mathcal{M}}$. ■

6.3.3 Proof of Theorem 6.2.2

We give here the proof of the key estimate on $||u_\mathcal{T}||_{1,q,\mathcal{M}}$ which was stated in Theorem 6.2.2, and which is crucial to show existence and convergence of the solution to the finite volume scheme.
Proof of Theorem 6.2.2
Let $\Lambda > ||\mu||_{M(\overline{\Omega})}$ (to avoid dividing by 0). (6.2.3)—(6.2.5) being a linear problem, we see that $u_\mathcal{T}/\Lambda$ is a solution to (6.2.3)—(6.2.5) with μ/Λ instead of μ.
Since $||\mu/\Lambda||_{M(\overline{\Omega})} \le 1$, we can apply Proposition 6.3.2 to $u_\mathcal{T}/\Lambda$; let $k_0 > 0$ only depending on $(\Omega, \mathbf{v}, \zeta)$ given by this proposition. $S_{k_0}(u_\mathcal{T}/\Lambda)$ and $T_{k_0}(u_\mathcal{T}/\Lambda)$ satisfy then the hypotheses of Lemma 6.3.4 with a function F only depending on $(\Omega, \mathbf{v}, \zeta)$. We deduce from this lemma that, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on $(\Omega, \mathbf{v}, \zeta, q)$ such that

$$||S_{k_0}(u_\mathcal{T}/\Lambda)||_{1,q,\mathcal{M}} \le C \quad \text{and} \quad ||T_{k_0}(u_\mathcal{T}/\Lambda)||_{1,q,\mathcal{M}} \le C.$$

Since $u_\mathcal{T}/\Lambda = S_{k_0}(u_\mathcal{T}/\Lambda) + T_{k_0}(u_\mathcal{T}/\Lambda)$ and $||\cdot||_{1,q,\mathcal{M}}$ is a norm, this gives $||u_\mathcal{T}/\Lambda||_{1,q,\mathcal{M}} \le C$, that is to say $||u_\mathcal{T}||_{1,q,\mathcal{M}} \le C\Lambda$. Letting then Λ tend to $||\mu||_{M(\overline{\Omega})}$, we obtain the desired estimate on $u_\mathcal{T}$. ■

6.4 Proof of Theorem 6.2.1

We first prove the uniqueness of the solution to (6.2.9), which does not involve numerical analysis methods, and then the existence and convergence of the approximate solutions (which yields the existence of a solution to (6.2.9)).

Proof of the uniqueness of the solution to (6.2.9)
This proof uses the regularity results of [52] on the variational solution to $-\Delta v = f \in L^2(\Omega)$, $v_{|\partial\Omega} = 0$, for Ω polygonal (or polyhedral) open set in $\mathbb{R}^d$, $d = 2$ or 3.
Problem (6.2.9) being linear, it is sufficient to prove that, if u is a solution to (6.2.9) with $\mu = 0$, then $u = 0$.
Let $\theta \in L^\infty(\Omega)$ and take $\varphi \in H_0^1(\Omega) \cap L^\infty(\Omega)$ the solution to

$$\int_\Omega \nabla\varphi \cdot \nabla\psi \, d\lambda - \int_\Omega \psi \mathbf{v} \cdot \nabla\varphi \, d\lambda + \int_\Omega b\varphi\psi \, d\lambda = \int_\Omega \theta\psi \, d\lambda \,, \ \forall \psi \in H_0^1(\Omega). \tag{6.4.1}$$

The existence of such a φ is ensured by the results of [31]. Letting $\Theta = \theta + \mathbf{v} \cdot \nabla\varphi - b\varphi \in L^2(\Omega)$, we see that $\varphi \in H_0^1(\Omega)$ satisfies $-\Delta\varphi = \Theta$ on Ω.
Since Ω is a polygonal (or polyhedral) open set in $\mathbb{R}^2$ or $\mathbb{R}^3$, the results of [52] give us $\eta > 0$ such that $\varphi \in H^{\frac{3}{2}+\eta}(\Omega)$. Thus, by the Sobolev injections (see [1]), there exists $s > d$ such that $\varphi \in W_0^{1,s}(\Omega)$ (in the case $d = 2$, to obtain such a $s > 2$, we could also have used the result of [69] — which is stated for regular open sets but is also true for open sets with Lipschitz-continuous boundary, see [49]).
Thanks to this additional regularity, a density argument allows to see that (6.4.1) is also true for $\psi \in W_0^{1,s'}(\Omega)$, where s' is the conjugate exponent to s, that is, such that $\frac{1}{s} + \frac{1}{s'} = 1$.
We can thus use φ in the equation satisfied by u and u in the equation satisfied by φ to obtain

$$0 = \int_\Omega \nabla u \cdot \nabla\varphi \, d\lambda - \int_\Omega u\mathbf{v} \cdot \nabla\varphi \, d\lambda + \int_\Omega bu\varphi \, d\lambda = \int_\Omega \theta u \, d\lambda.$$

We deduce from this equality, satisfied for all $\theta \in L^\infty(\Omega)$, that $u = 0$, i.e. the uniqueness of the solution to (6.2.9). ■

Proof of the existence and convergence results
The existence of a unique solution to (6.2.3)—(6.2.5) is an immediate consequence of the estimate of Theorem 6.2.2: indeed, if $\mu = 0$, then this theorem shows that any solution to (6.2.3)—(6.2.5) is null, that is to say that the square matrix defining this linear system is invertible.

Let us now prove the convergence result. The techniques used here are easy adaptations of the convergence proof of [38].
Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of functions of $L^2(\Omega)$ such that u_n is solution to (6.2.3)—(6.2.5) for $\mathcal{M} = \mathcal{M}_n$, where $(\mathcal{M}_n)_{n\in\mathbb{N}}$ is a family of admissible meshes $\mathcal{M}$ satisfying (6.2.10) (for some fixed $\zeta > 0$), and such that size$(\mathcal{M}_n)$ tends to 0 as n tends to $+\infty$.
We first prove (steps 0 to 5), that if $(u_n)_{n\in\mathbb{N}}$ tends to u in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$, as n tends to $+\infty$ (and size$(\mathcal{M}_n) \to 0$), with $u \in \cap_{q<\frac{d}{d-1}} W_0^{1,q}(\Omega)$, then u is a solution to (6.2.9).
We then prove (step 6), thanks to the a priori estimates of Section 3, the compactness of the sequence $(u_n)_{n\in\mathbb{N}}$ and conclude, thanks to the uniqueness result which was proved above, to the convergence of $(u_n)_{n\in\mathbb{N}}$ to the solution u to (6.2.9).

Step 0: Density argument
By density of $C_c^\infty(\Omega)$ in $W_0^{1,s}(\Omega)$ for all $s \in (d, \infty)$ and by the regularity results on u, it is clearly sufficient to prove that u satisfies the equation of (6.2.9) for all $\varphi \in C_c^\infty(\Omega)$. Take such a φ. Multiplying (6.2.3) by $\varphi(x_K)$ and summing over $K \in \mathcal{T}$ we have, by conservativity of the fluxes and by dropping the index n:

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(x_K) - \varphi(x_L)) + \sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+}(\varphi(x_K) - \varphi(x_L)) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(x_K) = \sum_{K\in\mathcal{T}} \varphi(x_K)\mu(K). \tag{6.4.2}$$

We shall now pass to the limit as size($\mathcal{M}$) tends to 0 in (6.4.2), and prove the convergence of each of the terms to the corresponding term in (6.2.9). In fact, the proof of convergence of the first and third terms of the left hand side can be found in [38] or [39] and so can the proof of the second term under a stronger regularity condition. The proof of convergence of the right hand side may be found in [46], so that the only new part in this proof is Step 4 which shows the convergence of the convective term with a continuous convection velocity (rather than C^1 in previous works). However, we give a quick proof for all terms for the sake of completeness.

Step 1: convergence of the lower order terms.
Denote $\varphi_{\mathcal{T}} \in X(\mathcal{T})$ the function defined by $\varphi_K = \varphi(x_K)$ for all $K \in \mathcal{T}$. By regularity of φ, we have $\varphi_{\mathcal{T}} \to \varphi$ uniformly on Ω as size($\mathcal{M}$) $\to 0$, thus

$$\sum_{K\in\mathcal{T}} \varphi(x_K)\mu(K) = \int_\Omega \varphi_{\mathcal{T}}\, d\mu \to \int_\Omega \varphi\, d\mu \tag{6.4.3}$$

as size($\mathcal{M}$) $\to 0$ (notice that $\varphi_{\mathcal{T}} = 0$ near $\partial\Omega$ for size($\mathcal{M}$) small enough).
By regularity of b, $b_{\mathcal{T}} = (b_K)_{K\in\mathcal{T}}$ tends to b in $L^2(\Omega)$ as size($\mathcal{M}$) $\to 0$; thus, since $\varphi_{\mathcal{T}} \to \varphi$ in $L^\infty(\Omega)$ and $u_{\mathcal{T}} \to u$ in $L^2(\Omega)$ (because $2 < d/(d-2)$) as size($\mathcal{M}$) $\to 0$, we have

$$\sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(x_K) = \int_\Omega b_{\mathcal{T}} u_{\mathcal{T}} \varphi_{\mathcal{T}}\, d\lambda \to \int_\Omega bu\varphi\, d\lambda \tag{6.4.4}$$

as size($\mathcal{M}$) $\to 0$.

Step 2: convergence of the diffusion term.
Gathering by control volumes, we have

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) = \sum_{K\in\mathcal{T}} u_K \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\varphi(x_K) - \varphi(x_L)).$$

But, by regularity of φ,

$$\tau_\sigma(\varphi(x_K) - \varphi(x_L)) = -\int_\sigma \nabla\varphi\cdot\mathbf{n}_{K,\sigma}\, d\gamma + \mathrm{meas}(\sigma) R_{K,\sigma},$$

where $|R_{K,\sigma}| \le C_1 \mathrm{size}(\mathcal{M})$ (C_1 does not depend on the mesh) and $R_{K,\sigma} = -R_{L,\sigma}$ whenever $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$. Thus, gathering by edges,

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) + \sum_{K\in\mathcal{T}} u_K \int_{\partial K} \nabla\varphi\cdot\mathbf{n}_K\, d\gamma &= \sum_{K\in\mathcal{T}} u_K \sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma) R_{K,\sigma} \\ &= \sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) R_{K,\sigma}(u_K - u_L).\end{aligned}$$

But

$$\left|\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) R_{K,\sigma}(u_K - u_L)\right| \le C_1 \mathrm{size}(\mathcal{M}) \sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma \frac{D_\sigma u_{\mathcal{T}}}{d_\sigma} = C_1 \mathrm{size}(\mathcal{M}) ||u_{\mathcal{T}}||_{1,1,\mathcal{M}},$$

and this last quantity tends to 0 as size($\mathcal{M}$) $\to 0$ (because, by Theorem 6.2.2, $||u_{\mathcal{T}}||_{1,1,\mathcal{M}}$ stays bounded).
By noticing that

$$\sum_{K\in\mathcal{T}} u_K \int_{\partial K} \nabla\varphi\cdot\mathbf{n}_K\, d\gamma = \sum_{K\in\mathcal{T}} u_K \int_K \Delta\varphi\, d\lambda = \int_\Omega u_{\mathcal{T}} \Delta\varphi\, d\lambda,$$

and since $u_{\mathcal{T}} \to u$ in $L^1(\Omega)$ as size$(\mathcal{M}) \to 0$, we deduce that

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) \to -\int_\Omega u\Delta\varphi \, d\lambda = \int_\Omega \nabla u \cdot \nabla\varphi \, d\lambda \tag{6.4.5}$$

as size$(\mathcal{M}) \to 0$.

Step 3: Preliminary to the convergence of the convection term (in fact, we prove here the convergence of the convection term if $\mathbf{v}$ is regular).
Let $\mathbf{w} \in (C^1(\overline{\Omega}))^d$ and define, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, $w_{K,\sigma} = \int_\sigma \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma$ (notice that, if $\sigma = K|L \in \mathcal{E}_{\rm int}$, then $w_{K,\sigma} = -w_{L,\sigma}$). We want to study the limit, as size$(\mathcal{M}) \to 0$, of $\sum_{\sigma\in\mathcal{E}} w_{K,\sigma} u_{\sigma,+}(\varphi(x_K) - \varphi(x_L))$ (that is to say the convection term of (6.4.2) with $\mathbf{w}$ instead of $\mathbf{v}$).
We have

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} w_{K,\sigma} u_{\sigma,+}(\varphi(x_K) - \varphi(x_L)) \\
&\quad = \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} u_{\sigma,+} \varphi(x_K) \\
&\quad = \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+} - u_K)\varphi(x_K) + \sum_{K\in\mathcal{T}} \varphi(x_K) u_K \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}.
\end{aligned} \tag{6.4.6}$$

Since $\sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} = \int_{\partial K} \mathbf{w} \cdot \mathbf{n}_K \, d\gamma = \int_K \mathrm{div}(\mathbf{w}) \, d\lambda$, we have

$$\sum_{K\in\mathcal{T}} \varphi(x_K) u_K \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} = \int_\Omega u_{\mathcal{T}} \varphi_{\mathcal{T}} \mathrm{div}(\mathbf{w}) \, d\lambda \to \int_\Omega u\varphi \mathrm{div}(\mathbf{w}) \, d\lambda \tag{6.4.7}$$

as size$(\mathcal{M}) \to 0$.
Moreover,

$$\begin{aligned}
\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+} - u_K)\varphi(x_K) &= \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} (u_{\sigma,+} - u_K) \int_\sigma \varphi \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma \\
&\quad + \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} (u_{\sigma,+} - u_K) \int_\sigma (\varphi(x_K) - \varphi) \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma.
\end{aligned}$$

Since, for size$(\mathcal{M})$ small enough, the support of φ does not intersect the cells K such that $\partial K \cap \partial\Omega \neq \emptyset$, we have

$$\sum_{K\in\mathcal{T}} \sum_{\sigma=K|L\in\mathcal{E}_K} u_{\sigma,+} \int_\sigma \varphi \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma = \sum_{\sigma\in\mathcal{E}_{\rm int}} u_{\sigma,+} \left(\int_\sigma \varphi \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma + \int_\sigma \varphi \mathbf{w} \cdot \mathbf{n}_{L,\sigma} \, d\gamma \right) = 0,$$

because $\mathbf{n}_{K,\sigma} = -\mathbf{n}_{L,\sigma}$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$. On the other hand,

$$-\sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} u_K \int_\sigma \varphi \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma = -\sum_{K\in\mathcal{T}} u_K \int_K \mathrm{div}(\varphi\mathbf{w}) \, d\lambda = -\int_\Omega u_{\mathcal{T}} \mathrm{div}(\varphi\mathbf{w}) \, d\lambda \to -\int_\Omega u \mathrm{div}(\varphi\mathbf{w}) \, d\lambda$$

as size$(\mathcal{M}) \to 0$. By regularity of φ, we have C_5 only depending on φ such that

$$\begin{aligned}
&\left| \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} (u_{\sigma,+} - u_K) \int_\sigma (\varphi(x_K) - \varphi) \mathbf{w} \cdot \mathbf{n}_{K,\sigma} \, d\gamma \right| \\
&\quad \leq C_5 || \, |\mathbf{w}| \, ||_{C(\overline{\Omega})} \mathrm{size}(\mathcal{M}) \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma) |u_{\sigma,+} - u_K| \\
&\quad \leq C_5 || \, |\mathbf{w}| \, ||_{C(\overline{\Omega})} \mathrm{size}(\mathcal{M}) \sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) D_\sigma u_{\mathcal{T}} \\
&\quad = C_5 || \, |\mathbf{w}| \, ||_{C(\overline{\Omega})} \mathrm{size}(\mathcal{M}) ||u_{\mathcal{T}}||_{1,1,\mathcal{M}}.
\end{aligned}$$

The last quantity tending to 0 as size($\mathcal{M}$) $\to 0$, we deduce from what preceeds that

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+}-u_K)\varphi(x_K) \to -\int_\Omega u\mathrm{div}(\varphi\mathbf{w})\,d\lambda \tag{6.4.8}$$

as size($\mathcal{M}$) $\to 0$.
Using (6.4.7) and (6.4.8) in (6.4.6), we obtain

$$\sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) \to \int_\Omega u\varphi\mathrm{div}(\mathbf{w})\,d\lambda - \int_\Omega u\mathrm{div}(\varphi\mathbf{w})\,d\lambda = -\int_\Omega u\mathbf{w}\cdot\nabla\varphi\,d\lambda \tag{6.4.9}$$

as size($\mathcal{M}$) $\to 0$.

Step 4: convergence of the convection term.
Let $\varepsilon>0$ and take $\mathbf{w}\in(C^1(\overline{\Omega}))^d$ such that $||\,|\mathbf{v}-\mathbf{w}|\,||_{C(\overline{\Omega})}\le\varepsilon$. By regularity of φ,

$$\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) - \sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right| \le C_2\varepsilon\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma|u_{\sigma,+}|$$

where C_2 only depends on φ. Gathering by control volumes, we deduce that

$$\begin{aligned}\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) - \sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right| \\ \le\ C_2\varepsilon\sum_{K\in\mathcal{T}}|u_K|\sum_{\sigma\in\mathcal{E}_K\,|\,v_{K,\sigma}\ge 0}\mathrm{meas}(\sigma)d_\sigma.\end{aligned}$$

But, by hypothesis on the mesh, $\sum_{\sigma\in\mathcal{E}_K\,|\,v_{K,\sigma}\ge0}\mathrm{meas}(\sigma)d_\sigma \le \zeta^{-1}\sum_{\sigma\in\mathcal{E}_K}\mathrm{meas}(\sigma)d_{K,\sigma} = \zeta^{-1}d\mathrm{meas}(K)$, so that

$$\begin{aligned}\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) - \sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right| &\le C_3\varepsilon\sum_{K\in\mathcal{T}}\mathrm{meas}(K)|u_K| \\ &\le C_4\varepsilon,\end{aligned} \tag{6.4.10}$$

where C_3 and C_4 do not depend on the mesh $\mathcal{M}$ nor on ε ($\sum_{K\in\mathcal{T}}\mathrm{meas}(K)|u_K| = ||u_\mathcal{T}||_{L^1(\Omega)}$ is bounded).
We also notice that

$$\left|\int_\Omega u\mathbf{v}\cdot\nabla\varphi\,d\lambda - \int_\Omega u\mathbf{w}\cdot\nabla\varphi\,d\lambda\right| \le C_6\varepsilon \tag{6.4.11}$$

where C_6 does not depend on ε.
Using then (6.4.9) and (6.4.11) in (6.4.10), we obtain

$$\limsup_{\mathrm{size}(\mathcal{M})\to 0}\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) + \int_\Omega u\mathbf{v}\cdot\nabla\varphi\,d\lambda\right| \le C_7\varepsilon,$$

where C_7 does not depend on ε. This being true for any $\varepsilon>0$, we deduce that

$$\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) \to -\int_\Omega u\mathbf{v}\cdot\nabla\varphi\,d\lambda \tag{6.4.12}$$

as size($\mathcal{M}$) $\to 0$.

Step 5: Passage to the limit in the scheme.

Using (6.4.3), (6.4.4), (6.4.5) and (6.4.12), we may pass to the limit in (6.4.2) to obtain:

$$\int_\Omega \nabla u \cdot \nabla \varphi \, d\lambda - \int_\Omega u\mathbf{v} \cdot \nabla \varphi \, d\lambda + \int_\Omega bu\varphi \, d\lambda = \int_\Omega \varphi \, d\mu,$$

which proves that u is a solution to (6.2.9).

Step 6: proof of the convergence of $(u_n)_{n\in\mathbb{N}}$.
Thanks to Theorem 6.2.2 and to Propositions 6.2.3 and 6.2.4, we see that $(u_n)_{n\geq 1}$ is relatively compact in $L^q(\Omega)$ for all $q \in [1, \frac{d}{d-1})$ and that the adherence values (in $L^q(\Omega)$) of this sequence are in $W_0^{1,q}(\Omega)$ (for $q \in (1, \frac{d}{d-1})$). Up to a subsequence, we can thus suppose that $u_n \to u$ in $L^q(\Omega)$ for all $q \in [1, \frac{d}{d-1})$, with $u \in \cap_{q<\frac{d}{d-1}} W_0^{1,q}(\Omega)$; by Proposition 6.2.2 and Theorem 6.2.2, $(u_n)_{n\geq 1}$ is also bounded in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$ so that, by an easy consequence of the Vitali convergence theorem, $u_n \to u$ in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$.
By what we have just proved, we see that u is then a solution to (6.2.9); since this solution is unique, this proves that the whole sequence $(u_n)_{n\geq 1}$ converges to u.
As a by-product, this convergence entails the existence of a solution to (6.2.9) (which can be deduced from previous works, [11] and [31] for instance). ■

6.5 A scheme with jump of the fluxes

Until now, we considered, in the definition of "admissible mesh", a partition of Ω into convex polygonal (or polyhedral) sets. We then defined a finite volume scheme where the conservativity of the numerical fluxes writes : $F_{K,\sigma} = -F_{L,\sigma}$ for all $\sigma = K|L \in \mathcal{E}_{\rm int}$.
There is, however, another manner to deal with the discretization of a right-hand side measure, which was implemented, for instance, in [44] for the numerical simulation of fuel cells. In this formulation, we write that if the support of the measure intersects a given edge, then there is a jump of the flux on this edge. This leads to the following scheme.

The mesh $\mathcal{M}$ we consider now is defined by a finite family $\mathcal{T}$ of polygonal (or polyhedral) open disjoint subsets of Ω, by a finite family $\mathcal{E}$ of subsets of $\overline{\Omega}$ contained in affine hyperplanes and by a finite family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points of Ω such that

a) $\mathcal{T} \cup \mathcal{E}$ is a partition of $\overline{\Omega}$,

b) for each $\sigma \in \mathcal{E}$, there exists $K \in \mathcal{T}$ and a non-empty open subset O of ∂K such that $O \subset \sigma \subset \overline{O}$,

c) items iii)—vi) of Definition 6.2.1 hold.

The notations concerning the mesh are the same as before, and the reader can easily verify that Propositions 6.2.1 — 6.2.4 are still true for such meshes.
Still defining $(b_K)_{K\in\mathcal{T}}$ and $(v_{K,\sigma})_{K\in\mathcal{T},\,\sigma\in\mathcal{E}_K}$ by (6.2.2), the new scheme is

$$\forall K \in \mathcal{T}\,, \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+} + \mathrm{meas}(K) b_K u_K = \mu(K)\,, \tag{6.5.1}$$

$$\begin{aligned} &\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,, \quad F_{K,\sigma} = -\tfrac{\mathrm{meas}(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)\,, \\ &\forall \sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K\,, \quad F_{K,\sigma} = \tau_\sigma u_K\,, \end{aligned} \tag{6.5.2}$$

$$\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,, \quad F_{K,\sigma} + F_{L,\sigma} = -\mu(\sigma)\,, \tag{6.5.3}$$

$$\begin{aligned} &\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,, \quad u_{\sigma,+} = u_K \ \text{ if } v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = u_L \ \text{ otherwise,} \\ &\forall \sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K\,, \quad u_{\sigma,+} = u_K \ \text{ if } v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = 0 \ \text{ otherwise.} \end{aligned} \tag{6.5.4}$$

Notice that the unknowns of this scheme are $(u_K)_{K\in\mathcal{T}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$ (which represent approximate values on the edges), but that Relation (6.5.3) allows to eliminate the $(u_\sigma)_{\sigma\in\mathcal{E}}$; this scheme can thus be considered as a linear system on $(u_K)_{K\in\mathcal{T}}$.

In fact, the elimination of u_σ thanks to (6.5.3) gives, for $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$F_{K,\sigma} = \frac{\text{meas}(\sigma)}{d_\sigma}(u_K - u_L) - \frac{d_{L,\sigma}}{d_\sigma}\mu(\sigma).$$

Thus, this new scheme is in fact the scheme (6.2.3)—(6.2.5) where we have changed, for all $K \in \mathcal{T}$, $\mu(K)$ by $\widetilde{\mu}_K = \mu(K) + \sum_{\sigma\in\mathcal{E}_K} \frac{d_{L,\sigma}}{d_\sigma}\mu(\sigma)$ (with $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $d_{L,\sigma} = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$), which is just another way to discretize the measure μ (forgetting the values of μ on the boundary of the domain, which does not modify the problem since we consider Dirichlet boundary conditions).
The matrix of (6.5.1)—(6.5.4) is thus the same as the matrix of (6.2.3)—(6.2.5) and, since $(\widetilde{\mu}_K)_{K\in\mathcal{T}}$ satisfies

$$\sum_{K\in\mathcal{T}} |\widetilde{\mu}_K| \le \sum_{K\in\mathcal{T}} |\mu(K)| + \sum_{\sigma\in\mathcal{E}} \left(\frac{d_{K,\sigma}}{d_\sigma} + \frac{d_{L,\sigma}}{d_\sigma}\right) |\mu(\sigma)| = \sum_{K\in\mathcal{T}} |\mu(K)| + \sum_{\sigma\in\mathcal{E}} |\mu(\sigma)| \le ||\mu||_{M(\overline{\Omega})}$$

(because $\mathcal{T}\cup\mathcal{E}$ is a partition of $\overline{\Omega}$), the *a priori* estimates on the solutions to (6.5.1)—(6.5.4) are obtained exactly the same way as the estimates on the solutions to (6.2.3)—(6.2.5).
We also have, for $\varphi \in C_c(\Omega)$, for $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$\left| \frac{d_{L,\sigma}}{d_\sigma}\varphi(x_K)\mu(\sigma) + \frac{d_{K,\sigma}}{d_\sigma}\varphi(x_L)\mu(\sigma) - \int_\sigma \varphi\, d\mu \right| \le \omega(\varphi, \text{size}(\mathcal{M}))|\mu(\sigma)|,$$

where $\omega(\varphi, h)$ is the modulus of continuity of φ; thus,

$$\sum_{K\in\mathcal{T}} \varphi(x_K)\widetilde{\mu}_K \to \int_\Omega \varphi\, d\mu$$

as $\text{size}(\mathcal{M}) \to 0$ and the convergence of the solution of (6.5.1)—(6.5.4) as $\text{size}(\mathcal{M}) \to 0$ is obtained by the same technique as in the proof of Theorem 6.2.1.

6.6 Numerical results

We performed a few simple numerical experiments on problems to which the exact solution is known, in order to try and obtain some rates of convergence of the finite volume scheme in presence of a non regular right hand side. Numerical results were also shown in [34] in the non coercive case with right hand side in H^{-1}, so we shall concentrate here on tests in the irregular data case.

6.6.1 Comparison of the two finite volume schemes

The first numerical experiment is concerned with the comparison of the treatment of the singularity in the one-dimensional case. In this case, the Dirac is not a very "mean" measure, in the sense that the solution of the problem is continous, the jump is only on the derivative. In the first version of the FV scheme (scheme (6.2.3)-(6.2.5), which we shall call Scheme 1 in the sequel), the Dirac measure is taken in its integral form in the right hand side while in the second version (scheme (6.5.1)-(6.5.4), which we shall call Scheme 2), the mesh is adapted so as to be able to write the numerical jump of the flux on a cell interface. We solve $-u'' = \delta_{1/2}$, $u(0) = 0$, $u(1) = 0$, on the interval $(0,1)$; the exact solution is $u(x) = \frac{x}{2}$ for $x < .5$, $u(x) = \frac{(1-x)}{2}$ for $x \ge .5$. We use a uniform mesh, and ensure that the number of cells is even, so that in the second scheme, the flux jump is located on a cell interface. The error function

e is defined by $e(x) = u(x_K) - u_K$ for any $x \in K$, where $u(x_K)$ denotes the value of the exact solution of the continuous problem at point x_K and $(u_K)_{K \in \mathcal{T}}$ the solution to the finite volume scheme.
We analyse the rate of convergence by showing the L^1, L^2 and L^∞ norms of the error e versus the number of cells with a log-log scale in Figure 6.3.

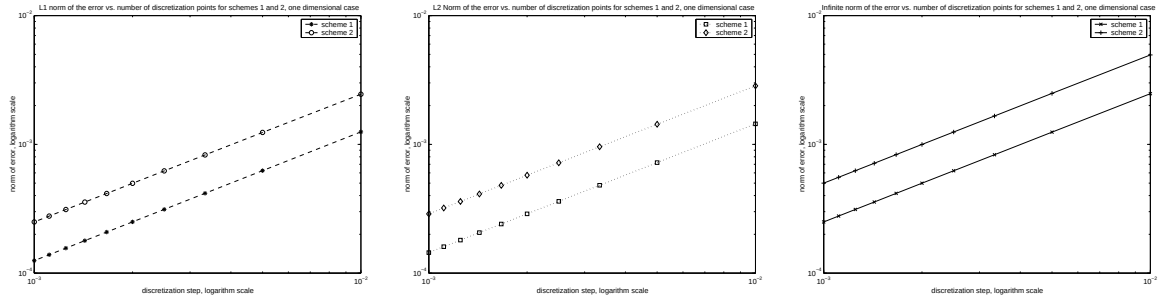


Figure 6.3: Convergence rate in the one-dimensional case.

The results show straight lines for all three norms, so that it is natural to try and evaluate the norms of the error as $\|e\| \equiv Ch^\alpha$. The computation of the coefficients C and α from the numerical results are given in Table 6.1. These coefficients are computed using the two finest meshes.

α	L^1 norm	L^2 norm	L^∞ norm
Scheme 1	1.0000	1.0000	0.9961
Scheme 2	0.9923	0.9941	0.9961

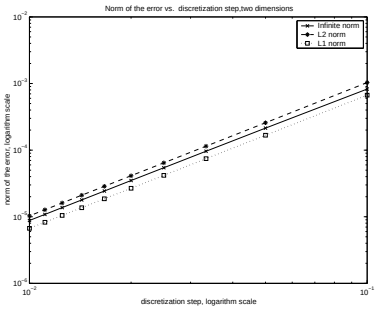
C	L^1 norm	L^2 norm	L^∞ norm
Scheme 1	0.1250	0.1443	0.2431
Scheme 2	0.2365	0.2768	0.4861

Table 6.1: Values of (C, α) for schemes 1 and 2, one dimensional case.

These results show that the two schemes have a rate of convergence which is roughly the same (close to one) and that the constant C is about twice as large for Scheme 2 (jump of flux) than for Scheme 1 (Dirac in one cell). This is quite in accordance with what can be seen from the implementation the scheme, because Scheme 2 amounts to spreading the Dirac measure over two cells, instead of one in Scheme 1.

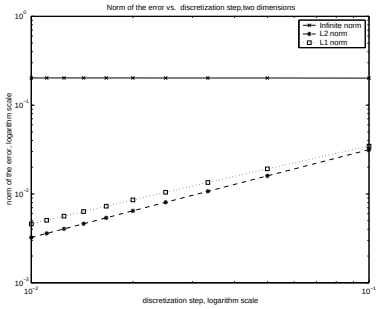
6.6.2 Two and three-dimensional tests on a Cartesian mesh

We also implemented the finite volume scheme on the square (resp. cubic) domain $\Omega = (-1, 1)^2$ (resp. $\Omega = (-1, 1)^3$). The domain is discretized with a uniform mesh, and the L^p norm of the error is computed for an increasing number of cells, so as to evaluate the rate of convergence.
We first tested the two-dimensional code for a regular data, yielding the exact solution $u(x, y) = \sin x \sin y$. In this case, since the mesh is rectangular and the exact solution regular, the consistency error on the flux is of order 2 and the rate of convergence in the L^2 norm can be theoretically shown to be of order 2 ([38], [48], see also [22] for a related co-volume scheme). The rate of convergence was computed for the piecewise constant error function defined by $e_K = u(x_K) - u_K$ for $K \in \mathcal{T}$, where u is the exact solution and $(u_K)_{K \in \mathcal{T}}$ is the solution to the finite volume scheme.
We then performed some tests with a right hand side given by a Dirac measure at 0. The boundary conditions were taken such that the exact solution be the restriction of the solution of $-\Delta u = \delta_0$ in the whole set $\mathbb{R}^2$ (resp. $\mathbb{R}^3$). It is well-known that this function lies in $L^p(\mathbb{R}^2)$ for $p \in [1, +\infty)$ (resp. $L^p(\mathbb{R}^3)$ for $p \in [1, 3)$).
We obtain the results (in log-log scale) given in Figure 6.5. The coefficients C and α such that $\|e\| = Ch^\alpha$ are again evaluated for the norms $L^1(\Omega)$ and $L^2(\Omega)$, and are also given in Figure 6.5.
In these tests, the mesh is such that the point $(0, 0)$ is located at the corner of the cell $[0, h] \times [0, h]$, where h is the discretization step of the mesh. Hence the radial symmetry of the solution is broken by the



	α	C
L^1 norm	2.0000	.1031
L^2 norm	2.0000	.0428
L^∞ norm	1.7931	.0690

Figure 6.4: Convergence rate, two dimensional case, regular right hand side.



	α	C
L^1 norm	.9047	.2421
L^2 norm	.9965	.3181

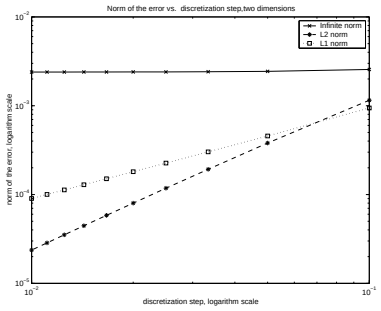
Figure 6.5: Convergence rate, two dimensional case, right hand side Dirac at zero, non symmetric discrete problem.

mesh. If we restore it by allocating one fourth of the Dirac measure to each of the four cells $[0,h]\times[0,h]$, $[0,h]\times[0,-h]$, $[-h,0]\times[0,h]$ and $[-h,0]\times[-h,0]$, we gain in the order of convergence, as can be seen in Figure 6.6. Hence the order of convergence depends on the singularity of the data, but also on the preservation of the symmetry of the solution.

A question of interest is to know whether the singular data influences the rate of convergence outside of the region of singularity. In order to check this point, we compute the norm of the error between the exact and approximate solutions on the region $\{x \leq -.5\}\times\{y \leq -.5\}$. We find that in this case, we recover an order of convergence close to one in all norms if the Dirac is located at the corner of a cell, in which case the symmetry of the solution is not preserved by the discretization (see Figure 6.7). In this case, the rate of convergence in the regular zone is perturbed by the singularity outside this zone (recall that the theoretical rate of convergence for regular solutions on rectangular meshes is 2 [48], [38]). However, if we restore the symmetry of the problem as described above, then the rate of convergence is close to two (see Figure 6.8).

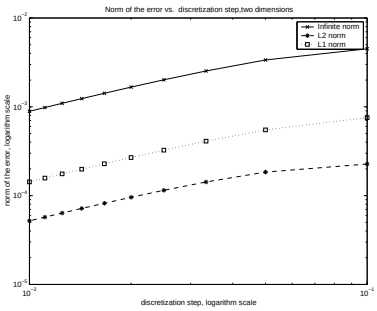
We then implemented a three dimensional cartesian mesh and found, for the non-symmetric discrete problem (Dirac located at a corner of the cell $[0,h]^3$) a rate of convergence close to 1 in norm L^1 and .5 in norm L^2, as shown in Figure 6.9. Recall that in this case the exact solution is in L^p for $1 \leq p < 3$.

If the Dirac is distributed on the eight cells neighbouring the origin, in order to symmetrize the discrete problem, as was done in the two-dimensional case, then one obtains a rate of convergence of 1.631 in the L^1 norm and .504 in the L^2 norm. This seems to indicate a super-convergence in the L^1 norm, although not to the second order (see also Remark 6.6.1).



	α	C
L^1 norm	1.7740	.0073
L^2 norm	1.0010	.0837

Figure 6.6: Convergence rate, two dimensional case, right hand side Dirac at zero, symmetric discrete problem.



	α	C
L^1 norm	0.9131	.0035
L^2 norm	0.9350	.0086
L^∞ norm	0.9360	.0586

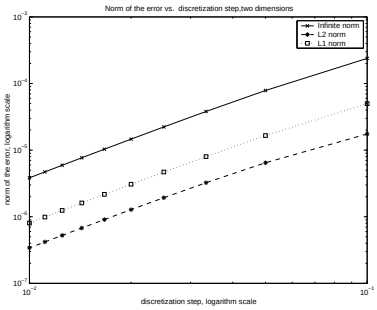
Figure 6.7: Convergence rate, two dimensional case, right hand side Dirac at zero, non symmetric discrete problem, norm computed on a "regular zone".

6.6.3 Two-dimensional tests on an unstructured mesh

We also tested our algorithm on an unstructured triangular mesh. Numerical experiments for the cell centered scheme on triangular meshes were performed in [15] and in [40] in the case of coercive convection diffusion equations and regular data. These experiments show a convergence rate of order two, as in the finite element case, although this superconvergence is still, to our knowledge, an open problem in the finite volume case. We show in figure 6.10 the rate of convergence which we obtain for the Poisson equation where the right hand side is a Dirac at 0 and the boundary conditions are such that the exact solution is $u(x_1, x_2) = \ln(x_1^2 + x_2^2)$. The refined meshes are not imbedded, so that the convergence lines are not straight, but one can figure out that the L^1 and L^2 norms of the error between the exact and approximate solutions are bounded by 0.1 size $(\mathcal{M})^{0.7}$.

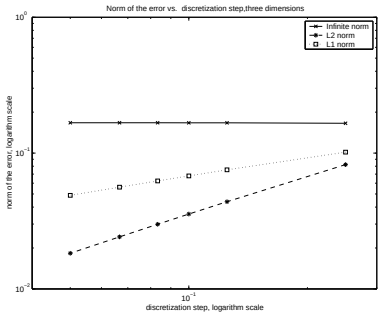
6.6.4 Spherical domain and mesh

We also made some experiments for a three dimensional spherical problem : we search for the solution of $-\Delta u = \delta_0$ on the Euclidean unit ball $B(0,1)$ of $\mathbb{R}^3$, with boundary conditions such that the exact solution be the restriction of the solution of $-\Delta u = \delta_0$ in the whole set $\mathbb{R}^3$. The control volumes are defined by $K_i = \{x \in B(0,1); ih \leq |x| \leq (i+1)h\}$, for $i = 0, \ldots, N$, where $h = \frac{1}{N+1/2}$. As we noted in Remark 6.2.1, such domain and mesh are not strictly contained in Definition 6.2.1 of an admissible mesh,



	α	C
L^1 norm	1.9486	.0247
L^2 norm	1.9571	.0042
L^∞ norm	1.9305	.0025

Figure 6.8: Convergence rate, two dimensional case, right hand side Dirac at zero, located at the center of the center cell, norm computed on a "regular zone".



	α	C
L^1 norm	0.9670	.3314
L^2 norm	0.4809	.2062

Figure 6.9: Convergence rate, three dimensional case, right hand side Dirac at zero, nonsymmetric discrete problem.

since a sphere is hardly a polyhedral domain, but in fact, the discretization of the normal flux on the boundaries of such a spherical mesh is clearly consistent when looking at spherical solutions of Problem 6.1.1. Indeed, the numerical flux at interface $i+1/2$ is taken as $F_{i+1/2} = \frac{4\pi i^2 h^2}{h}(u_{i+1} - u_i)$, where the $(u_i)_{i=0,\dots,N}$ denote the discrete unknowns. In this case, the rate of convergence of the method was found to be 2 in norm L^1 and .5 in norm L^2 : see Figure 6.11.
Hence the symmetry of the problem seems to improve the performance of the method, at least on the L^1 norm.

Remark 6.6.1 *We recall that in the three-dimensional case, the exact solution $-\Delta u = \delta_0$ is in $L^{3-\varepsilon}$ for any $\varepsilon > 0$, hence we can expect a convergence in L^p for $1 \leq p < 3$. From a convergence in $L^{3-\varepsilon}$ for any $\varepsilon > 0$, and a convergence with a rate h^α in L^1, one may deduce (from Hölder's inequality) a convergence in the L^2 norm with a rate of at least $h^{\frac{\alpha}{4}-\varepsilon}$ for any $\varepsilon > 0$. The above numerical results are in accordance with this estimate, both in the spherical case and in the Cartesian case of section 6.6.2.*

We also give in Table 6.2 below the rate of convergence obtained when computing the norm of the error on a zone where the solution is regular, i.e. on the set $\{x \in \mathbb{R}^3, |x| > 1/2\}$. Again, we find in this case a rate of convergence of 2 (even a little more than 2) for all norms.
If we now search for the solution of $-\Delta u = \mu$ on the three-dimensional unit ball, with μ the two dimensional Lebesgue measure supported on the sphere of radius .5, then the obtained convergence rate

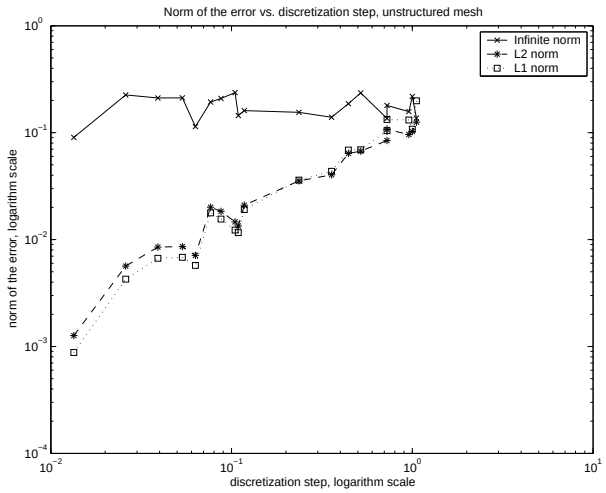
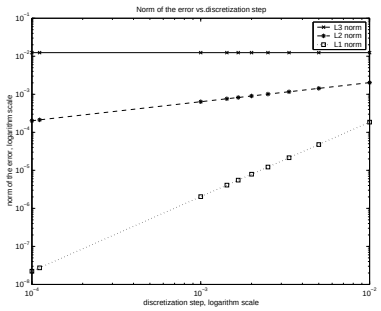


Figure 6.10: Convergence rate, two dimensional case, right hand side Dirac at zero, triangular mesh.



	α	C
L^1 norm	1.9288	.4993
L^2 norm	0.5000	.1879

Figure 6.11: Convergence rate, three dimensional case, right hand side Dirac at zero, spherical case.

is again 1, even though the exact solution is more regular than the solution to the Dirac problem, see Figure 8. Note that in this case, the exact solution is in L^∞ (and even in H^1).

6.7 Appendix

Throughout this section, for any $q \in (1, +\infty)$, we denote by q' its conjugate exponent, that is, $q' \in (1, +\infty)$ such that $\frac{1}{q} + \frac{1}{q'} = 1$.

Proof of Proposition 6.2.1

The case $q = 2$ is done in [38]. We use the same method for $q \in [1, 2)$.

Define, for $\sigma \in \mathcal{E}$ and $(x, y) \in \mathbb{R}^d$, $\chi_\sigma(x, y) = 1$ if $\sigma \cap [x, y] \neq \emptyset$ and $\chi_\sigma(x, y) = 0$ otherwise. Let $\mathbf{d}$ be a unit vector and define, for $x \in \Omega$, $y(x)$ as the point on the semi-line, with origin x and direction $\mathbf{d}$, such that $y(x) \in \partial\Omega$ and $[x, y(x)] \subset \overline{\Omega}$. If $\sigma \in \mathcal{E}$, we let $c_\sigma = |\mathbf{n}_\sigma \cdot \mathbf{d}|$, where $\mathbf{n}_\sigma$ is a unit normal to σ.

For all $x \in \Omega$ such that x does not belong to an affine hyperplane generated by some $\sigma \in \mathcal{E}$, i.e. for a.e. $x \in \Omega$, we have

$$|v_\mathcal{T}(x)| \leq \sum_{\sigma \in \mathcal{E}} \chi_\sigma(x, y(x)) D_\sigma v_\mathcal{T}$$

	α	C
L^1 norm	2.0506	.3411
L^2 norm	2.1164	.1720
L^∞ norm	2.1295	.2331

	α	C
L^1 norm	1.0506	.1874
L^2 norm	0.9993	.1787
L^∞ norm	0.9983	.2006

Table 6.2: Convergence rate, three dimensional case. left: the right hand side is a Dirac measure at zero, spherical case, norm computed on a "regular zone", right: the right hand side is a two dimensional Lebesgue measure supported on the sphere of radius 1/2. The norm is computed on the whole set Ω.

(recall that $v_{\mathcal{T}}(x) = v_K$ for the $K \in \mathcal{T}$ such that $x \in K$). Take such an x and suppose that, for some $\sigma \in \mathcal{E}$, $c_\sigma = 0$; we have then $\chi_\sigma(x, y(x)) = 0$ (indeed, otherwise x would belong to the affine hyperplane generated by σ). Thus, the preceding sum can be reduced to the $\sigma \in \mathcal{E}$ such that $c_\sigma \neq 0$ and we can write, thanks to Hölder's inequality, for a.e. $x \in \Omega$,

$$|v_{\mathcal{T}}(x)|^q \leq \left(\sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \chi_\sigma(x, y(x)) d_\sigma c_\sigma^{-\frac{q}{q'}} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \right)^q \right) \left(\sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \chi_\sigma(x, y(x)) d_\sigma c_\sigma \right)^{\frac{q}{q'}}. \tag{6.7.1}$$

Since

$$\sum_{\sigma \in \mathcal{E}} \chi_\sigma(x, y(x)) d_\sigma c_\sigma \leq \mathrm{diam}(\Omega) \quad \text{for all} \quad x \in \Omega$$

(see [38]) and $\int_\Omega \chi_\sigma(x, y(x))\, d\lambda(x) \leq \mathrm{diam}(\Omega)\mathrm{meas}(\sigma)c_\sigma$, we obtain, integrating (6.7.1) on Ω,

$$\int_\Omega |v_{\mathcal{T}}|^q \, d\lambda \leq \mathrm{diam}(\Omega)^{\frac{q}{q'}} \sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \mathrm{diam}(\Omega)\mathrm{meas}(\sigma) d_\sigma c_\sigma^{1-\frac{q}{q'}} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \right)^q.$$

But $q \leq 2$, so that $1 - \frac{q}{q'} = 2 - q \geq 0$ and $c_\sigma^{2-q} \leq 1$, which concludes this proof. ■

Proof of Proposition 6.2.2
The case $d = 2$ has already been done in the course of the proof of the discrete Sobolev inequalities in [38] (inequality (9.73), page 791).
For $d = 3$, the case $q = 2$ may be found in [27]. The case of a general q is similar; we use the following inequality (inequality (9.75) page 793 of [38]) : for any $w_{\mathcal{T}} \in X(\mathcal{T})$,

$$\int_\Omega |w_{\mathcal{T}}|^{\frac{3}{2}} \, d\lambda \leq \left(\sum_{\sigma \in \mathcal{E}} \mathrm{meas}(\sigma) D_\sigma w_{\mathcal{T}} \right)^{3/2}.$$

Applying this to $w_K = |v_K|^{\frac{2q}{3-q}} \mathrm{sgn}(v_K)$, and since $D_\sigma w_{\mathcal{T}} \leq \frac{2q}{3-q}(|v_K|^{\frac{3(q-1)}{3-q}} + |v_L|^{\frac{3(q-1)}{3-q}}) D_\sigma v_{\mathcal{T}}$ (with $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$ or $v_L = 0$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$), we deduce, by the Hölder inequality,

$$\begin{aligned}
&\left(\int_\Omega |v_{\mathcal{T}}|^{\frac{3q}{3-q}} \, d\lambda \right)^{2/3} \\
&\quad \leq \frac{2q}{3-q} \sum_{\sigma \in \mathcal{E}} \mathrm{meas}(\sigma) d_\sigma (|v_K|^{\frac{3(q-1)}{3-q}} + |v_L|^{\frac{3(q-1)}{3-q}}) \frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \\
&\quad \leq \frac{2q}{3-q} \left(\sum_{\sigma \in \mathcal{E}} \mathrm{meas}(\sigma) d_\sigma \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \right)^q \right)^{1/q} \left(\sum_{\sigma \in \mathcal{E}} \mathrm{meas}(\sigma) d_\sigma (2^{q'-1} |v_K|^{\frac{3q}{3-q}} + 2^{q'-1} |v_L|^{\frac{3q}{3-q}}) \right)^{1/q'}.
\end{aligned}$$

But, by hypothesis on ζ,

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma|v_K|^{\frac{3q}{3-q}} &= \sum_{K\in\mathcal{T}}|v_K|^{\frac{3q}{3-q}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{meas}(\sigma)d_\sigma\\ &\le \frac{1}{\zeta}\sum_{K\in\mathcal{T}}|v_K|^{\frac{3q}{3-q}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{meas}(\sigma)d_{K,\sigma}\\ &= \frac{3}{\zeta}\sum_{K\in\mathcal{T}}\mathrm{meas}(K)|v_K|^{\frac{3q}{3-q}}\\ &= \frac{3}{\zeta}||v_\mathcal{T}||^{\frac{3q}{3-q}}_{L^{\frac{3q}{3-q}}(\Omega)}.\end{aligned}$$

Thus, we finally have

$$\left(\int_\Omega|v_\mathcal{T}|^{\frac{3q}{3-q}}\,d\lambda\right)^{2/3} \le C||v_\mathcal{T}||_{1,q,\mathcal{M}}||v_\mathcal{T}||^{\frac{3(q-1)}{3-q}}_{L^{\frac{3q}{3-q}}(\Omega)}$$

where C only depends on (q,ζ), and this gives the desired estimate. ■

Proof of Proposition 6.2.3
Define $\chi_\sigma(x,y)$ as at the beginning of the proof of Proposition 6.2.1.
Suppose first that $q>1$ and take $h\in\mathbb{R}^d\backslash\{0\}$. Denote, for $\sigma\in\mathcal{E}$, $c_\sigma=|\mathbf{n}_\sigma\cdot\frac{h}{|h|}|$ (where $\mathbf{n}_\sigma$ is a unit normal to σ).
We have, for a.e. $x\in\Omega$ (in fact for all x which does not belong to an affine hyperplane generated by some $\sigma\in\mathcal{E}$),

$$|w_\mathcal{T}(x+h)-w_\mathcal{T}(x)|\le\sum_{\sigma\in\mathcal{E}}\chi_\sigma(x+h,x)D_\sigma v_\mathcal{T}. \tag{6.7.2}$$

As in the proof of Proposition 6.2.1, this sum can be limited to those $\sigma\in\mathcal{E}$ such that $c_\sigma\neq0$, and we have then, by Hölder, for a.e. $x\in\Omega$,

$$|w_\mathcal{T}(x+h)-w_\mathcal{T}(x)|\le\left(\sum_{\sigma\in\mathcal{E}\ |\ c_\sigma\neq0}\frac{\chi_\sigma(x+h,x)d_\sigma}{c_\sigma}\left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q\right)^{1/q}\left(\sum_{\sigma\in\mathcal{E}}\chi_\sigma(x+h,x)d_\sigma c_\sigma^{q'/q}\right)^{1/q'}.$$

Since $q\le2$ (and hence $q'/q\ge1$) and $c_\sigma\in[0,1]$, we have $c_\sigma^{q'/q}\le c_\sigma$; but (see [38]) $\sum_{\sigma\in\mathcal{E}}\chi_\sigma(x+h,x)d_\sigma c_\sigma\le|h|+C\mathrm{size}(\mathcal{M})$, where C only depends on Ω. Thus,

$$|w_\mathcal{T}(x+h)-w_\mathcal{T}(x)|^q\le(|h|+C\mathrm{size}(\mathcal{M}))^{q-1}\sum_{\sigma\in\mathcal{E}\ |\ c_\sigma\neq0}\frac{\chi_\sigma(x+h,x)d_\sigma}{c_\sigma}\left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q.$$

Since $\int_{\mathbb{R}^d}\chi_\sigma(x+h,x)\,d\lambda(x)\le\mathrm{meas}(\sigma)c_\sigma|h|$, we deduce, after integrating, the desired estimate (6.2.12).
If $q=1$, we simply integrate (6.7.2) and this directly gives (bounding $\int_{\mathbb{R}^d}\chi_\sigma(x+h,x)\,d\lambda(x)$ by $\mathrm{meas}(\sigma)|h|$) the estimate.
The compactness result is then an immediate application of Kolmogorov's Theorem, with the use of Proposition 6.2.1 to obtain a bound in $L^q(\Omega)$. ■

Proof of Proposition 6.2.4
Applying (6.2.12) to v_n and passing to the limit $n\to\infty$, we get, for $h\in\mathbb{R}^d\backslash\{0\}$,

$$\int_{\mathbb{R}^d}\frac{|w(x+h)-w(x)|^q}{h^q}\,d\lambda(x)\le C,$$

where C does not depend on h and w is the extension of v to $\mathbb{R}^d$ by 0 outside Ω.
Since $q>1$, this estimate classically gives $w\in W^{1,q}(\mathbb{R}^d)$ and, by the regularity of Ω, since w is the extension of v by 0 outside Ω, $v\in W_0^{1,q}(\Omega)$. ■

Partie IV

Approximation parabolique d'une équation hyperbolique en domaine borné

Chapitre 7

An error estimate for the parabolic approximation of multidimensional scalar conservation laws with boundary conditions

Reference: J. Droniou, C. Imbert and J. Vovelle, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** (2004), no. 5, 689-714.

Abstract We study the parabolic approximation of a multidimensional scalar conservation law with initial and boundary conditions. We prove that the rate of convergence of the viscous approximation to the weak entropy solution is of order $\eta^{1/3}$, where η is the size of the artificial viscosity. We use a kinetic formulation and kinetic techniques for initial-boundary value problems developed by the last two authors in a previous work.

keywords: conservation law, initial-boundary value problem, error estimates, parabolic approximation, kinetic techniques.
AMS classification: 35L65, 35D99, 35F25, 35F30, 35A35

7.1 Introduction

Let Ω be a bounded open subset of $\mathbb{R}^d$ with Lipschitz continuous boundary. Let $n(\overline{x})$ be the outward unit normal to Ω at a point $\overline{x} \in \partial\Omega$, $Q = (0,+\infty) \times \Omega$ and $\Sigma = (0,+\infty) \times \partial\Omega$. We consider the following multidimensional scalar conservation law

$$\partial_t u + \operatorname{div} A(u) = 0 \text{ in } Q, \tag{7.1.1a}$$

with the initial condition

$$u(0,x) = u_0(x), \forall x \in \Omega, \tag{7.1.1b}$$

and the boundary condition

$$u(s,y) = u_b(s,y), \forall (s,y) \in \Sigma. \tag{7.1.1c}$$

It is known that entropy solutions must be considered if one wants to solve scalar conservation laws (Equation (7.1.1a) is replaced by a family of inequalities — see [60] for the Cauchy problem) and that the Dirichlet boundary conditions are to be understood in a generalized sense (see [3] for regular initial and boundary conditions and [73] for merely bounded data).

In this paper, we estimate the difference between the weak entropy solution of (7.1.1) and the smooth solution of the regularized parabolic equation

$$\partial_t v + \mathrm{div} A(v) = \eta \Delta v \text{ in } Q, \tag{7.1.2}$$

satisfying the same initial and boundary conditions. Throughout the paper, we make the following hypotheses on the data: the flux function A belongs to $\mathsf{C}^2(\mathbb{R})$, the initial condition u_0 is in $\mathsf{C}^2(\overline{\Omega})$, the boundary $\partial\Omega$ of the domain Ω is C^2, the boundary condition u_b belongs to $\mathsf{C}^2(\overline{\Sigma})$. In that case, there exists a unique solution v^η (regular outside $\{0\} \times \partial\Omega$) to the problem (7.1.2)-(7.1.1b)-(7.1.1c).

The aim of this paper is to prove the following error estimate.

Theorem 7.1.1 *Suppose that Ω is C^2, $A \in \mathsf{C}^2(\mathbb{R})$, $u_0 \in \mathsf{C}^2(\overline{\Omega})$ and $u_b \in \mathsf{C}^2(\overline{\Sigma})$. Let u be the weak entropy solution of (7.1.1) and let v^η be the solution of the approximate parabolic problem (7.1.2)-(7.1.1b)-(7.1.1c). Let $T_0 > 0$; there exists a constant C only depending on $(\Omega, u_b, u_0, A, T_0)$ such that, for all $t \in [0, T_0]$,*

$$||u(t) - v^\eta(t)||_{\mathsf{L}^1(\Omega)} \leqslant C\eta^{1/3}\,. \tag{7.1.3}$$

We now recall what is known about error estimates for approximations of conservation laws.
In the case where the function u is smooth (a feature which, we recall, requires the data to be smooth, compatible and the time T_0 to be small enough), error estimates of order

$$\begin{cases} \eta^{1/2} & \text{if the boundary is characteristic} \\ \eta & \text{if the boundary is not characteristic} \end{cases} \tag{7.1.4}$$

in $\mathsf{L}^\infty(0,T;\mathsf{L}^1(\Omega))$ have been given (see Gues [53], Gisclon and Serre [50], Grenier and Gues [51], Joseph and LeFloch [59], Chainais-Hillairet and Grenier [20] and references therein). The technique of *boundary layer analysis* developed in those articles is devoted to the investigation of the initial-boundary value problem for *systems* of conservation laws (and not only for a single equation). Roughly speaking, the viscous approximation v^η is decomposed as $v^\eta = u + c^\eta +$ (remainder) where c^η characterizes the boundary layer which appear in the vicinity of $\partial\Omega$. Estimates on $v^\eta - u$ are then consequences of estimates on $c^\eta +$ (remainder) (see Appendix 7.8.1).
To our knowledge, there does not exist other techniques of analysis which would confirm the error estimate (7.1.4). On the contrary, many techniques have been set and improved to analyse the error of approximation for the *Cauchy Problem* ($\Omega = \mathbb{R}^d$) for conservation laws (and results of sharpness of error estimates have also been delivered). The first error estimate for the Cauchy problem is given by Kuznetčov in 1976 [61]: an adaptation of the proof of the result of comparison between two weak entropy solutions given by Kružkov [60] yields an error estimate of order 1/2 in the L^1-norm. The reader interested in more precise, more general and more recent results is invited to consult the compilation made by Tang [79], the introduction of [78], and references therein.
We establish here Estimate (7.1.3) for arbitrary times T_0; in particular, the possible occurence of shocks is taken into account: u is the *weak* entropy solution to Problem (7.1.1) and has no more regularity, in general, than the ones stated in Proposition 7.2.1. As a consequence, u may be irregular in the vicinity of $\partial\Omega$ and this constitutes an obstacle to the analysis of the rate of convergence of v^η. To circumvent this obstacle, we use the kinetic formulation of [58] (an adaptation to boundary problems of the kinetic formulation introduced in [67]) and adapt the technique of error estimate developed by Perthame for the analysis of the Cauchy Problem [74]. We then obtain a rate of convergence of 1/3. The accuracy or non-sharpeness of this order (compare to (7.1.4)) remains an open problem for us.

The paper is dedicated to the proof of Theorem 7.1.1. It is organized as follows. We begin with some preliminaries, mainly to state (or recall) the kinetic formulations of both hyperbolic and parabolic equations.

In order to enlight the key ideas of this rather technical proof, we present its skeleton in Subsection 7.2.4. In Section 7.3, we obtain a first estimate in the interior of the domain; then, in Sections 7.4 and 7.5, we transport the equations so that Ω becomes a half space and we regularize them in order to use the solution of one of them as a test function in the other. Eventually, in Section 7.6, we conclude the proof of Theorem 7.1.1 by getting an estimate of the boundary term which appears at the end of Section 7.5.

7.2 Preliminaries

In order to clarify computations, we drop the superscript η in v^η and simply write v for the approximate solution. We prove Theorem 7.1.1 in several steps.

7.2.1 Known estimates on u and v

We gather in the following proposition the estimates we will need to prove Theorem 7.1.1. We refer to [3] for a proof of these results.

Proposition 7.2.1 *Assume that Ω is C^2, $A \in \mathsf{C}^2(\mathbb{R})$, $u_0 \in \mathsf{C}^2(\overline{\Omega})$ and $u_b \in \mathsf{C}^2(\overline{\Sigma})$. There exists $C > 0$ only depending on $(\Omega, u_b, u_0, A, T_0)$ such that*

1. *the functions $u, v : [0,T_0] \to \mathsf{L}^1(\Omega)$ are C-Lipschitz continuous*
2. *for all $t \in (0,T_0)$, $\int_\Omega |\partial_t v(t,\cdot)| \leqslant C$*
3. *for all $t \in [0,T_0]$, $|u(t,\cdot)|_{\mathsf{BV}(\Omega)} \leqslant C$ and $|v(t,\cdot)|_{\mathsf{BV}(\Omega)} \leqslant C$.*

7.2.2 Notations

Let us introduce some local charts of Ω. Since Ω is C^2 and bounded, we can find a finite cover $\{O_i\}_{i\in\{0,\dots,n\}}$ of $\overline{\Omega}$ by open sets of $\mathbb{R}^d$ such that $\overline{O_0} \subset \Omega$ and that, for all $i \in \{1,\dots,n\}$, there exists a C^2-diffeomorphism $h_i : O_i \to B^d$ (the unit ball in $\mathbb{R}^d$) satisfying

- $\partial\Omega \subset \cup_{i=1}^n O_i$;
- $h_i(O_i \cap \partial\Omega) = B^{d-1} := B^d \cap (\mathbb{R}^{d-1} \times \{0\})$;
- $h_i(O_i \cap \Omega) = B^d_+ := B^d \cap (\mathbb{R}^{d-1} \times (0,+\infty))$.

Let $(\lambda_i)_{i\in\{0,\dots,n\}}$ be a partition of the unity on $\overline{\Omega}$, subordinate to the cover $\{O_i\}_{i\in\{0,\dots,n\}}$.

In the following, when a quantity appears with a bar above, it denotes something related to the boundary of Ω (possibly transported on B^{d-1} by a chart): either a variable on $\partial\Omega$ or the value of a function on this boundary. The values of a function ϕ at $t = 0$ are denoted by $\phi^{(t=0)}$.
Here are other general notations, related to the regularization of the equations. Let $\theta \in \mathsf{C}^\infty_\mathsf{c}(]1/2,1[;\mathbb{R}^+)$ be such that $\int_\mathbb{R} \theta = 1$ and define, for $\tau > 0$, $\theta_\tau(\cdot) = \frac{1}{\tau}\theta(\frac{\cdot}{\tau})$ (right-decentred regularizing kernel). When necessary, we define regularizing kernels ρ_μ in space (either the whole space or on the (transported) boundary of Ω) or space-time variables; when such a kernel on $\mathbb{R}^N$ ($N = d-1$, $N = d$ or $N = d+1$) is given and f is a function defined and locally integrable on a set $S \subset \mathbb{R}^N$, we denote, for $z \in \mathbb{R}^N$,

$$f^\mu(z) = \int_S f(r)\rho_\mu(z-r)\,dr\,,$$

i.e. f^μ is the convolution of ρ_μ by the extension of f by 0 outside S. We have then, for all $\phi \in \mathsf{L}^1(\mathbb{R}^N)$ with compact support,

$$\int_S f(\phi \star \check{\rho_\mu}) = \int_{\mathbb{R}^N} f^\mu \phi$$

(where $\check{\rho_\mu}(z) = \rho_\mu(-z)$).

7.2.3 Kinetic formulations of (7.1.1) and (7.1.2)

The function sgn_+ is defined by $\mathrm{sgn}_+(s) = 0$ if $s \leqslant 0$ and $\mathrm{sgn}_+(s) = 1$ if $s > 0$; similarly, $\mathrm{sgn}_-(s) = -1$ if $s < 0$ and $\mathrm{sgn}_-(s) = 0$ if $s \geqslant 0$. Let $D = \sup(||u_b||_\infty, ||u_0||_\infty)$.
Let us recall the kinetic formulation of (7.1.1) obtained in [58]: there exists a bounded nonnegative measure $m \in \mathcal{M}^+(Q \times \mathbb{R}_\xi)$, which has a compact support with respect to ξ, and two nonnegative measurable functions $m_+^b, m_-^b \in \mathsf{L}^\infty_{\mathrm{loc}}(\Sigma \times \mathbb{R}_\xi)$ such that the function m_+^b vanishes for $\xi \gg 1$ (resp. the function m_-^b vanishes for $\xi \ll -1$) and such that the functions $f_\pm(t,x,\xi) = \mathrm{sgn}_\pm(u(t,x) - \xi)$ associated with u satisfy, for any $\phi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R}^{d+2})$

$$\int_{Q\times\mathbb{R}_\xi} f_\pm(\partial_t + a\cdot\nabla)\phi + \int_{\Omega\times\mathbb{R}_\xi} f_\pm^0\phi^{(t=0)} + \int_{\Sigma\times\mathbb{R}_\xi} (-a\cdot n) f_\pm^\tau \overline{\phi} = \int_{Q\times\mathbb{R}_\xi} \partial_\xi\phi dm \tag{7.2.1}$$

where $a = A'$, $f_\pm^0(x,\xi) = \mathrm{sgn}_\pm(u_0(x) - \xi)$ and $f_\pm^\tau(t,\overline{x},\xi) = \mathrm{sgn}_\pm(\overline{u}(t,\overline{x}) - \xi)$ satisfies

$$(-a\cdot n) f_\pm^\tau = M f_\pm^b + \partial_\xi m_\pm^b \tag{7.2.2}$$

with $f_\pm^b(t,\overline{x},\xi) = \mathrm{sgn}_\pm(u_b(t,\overline{x}) - \xi)$ and M the Lipschitz constant of the flux function A on $[-D,D]$. This formula is the kinetic formulation of the BLN condition (see [3]).
We next give a kinetic formulation for the approximate solution. Consider two test functions $\varphi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R}_t \times \mathbb{R}^d)$, $\psi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R}_\xi)$ and define $E(\alpha) = \int \psi(\xi)\mathrm{sgn}_\pm(\alpha - \xi)d\xi$ and $H(\alpha) = \int a(\xi)\psi(\xi)\mathrm{sgn}_\pm(\alpha - \xi)d\xi$. Note that $E' = \psi$ and $H' = E'a$. Now multiply the equation $\partial_t v + \mathrm{div}A(v) = \eta\Delta v$ by $\varphi(t,x)\psi(v(t,x)) = \varphi(t,x)E'(v(t,x))$, integrate over Q and integrate by part (using the fact that v is C^2 outside $\{0\}\times\partial\Omega$)

$$\int_Q E(v)\partial_t\varphi + H(v)\cdot\nabla\varphi + \int_\Omega E(u_0)\varphi^{(t=0)} - \int_\Sigma H(u_b)\cdot n\overline{\varphi}$$
$$= \int_Q \eta E'(v)\nabla v\cdot\nabla\varphi - \int_\Sigma \eta E'(u_b)\overline{\nabla v}\cdot n\overline{\varphi} + \int_Q \eta E''(v)|\nabla v|^2\,\varphi.$$

Using the definition of E and H, we obtain, denoting $g_\pm(t,x,\xi) = \mathrm{sgn}_\pm(v(t,x) - \xi)$,

$$\int_{Q\times\mathbb{R}_\xi} g_\pm(\partial_t + a\cdot\nabla)\phi - \int_{Q\times\mathbb{R}_\xi} \eta\delta_v\nabla v\cdot\nabla\phi + \int_{\Omega\times\mathbb{R}_\xi} f_\pm^0\phi^{(t=0)} + \int_{\Sigma\times\mathbb{R}_\xi} \overline{G_\pm}\ \overline{\phi} = \int_{Q\times\mathbb{R}_\xi} \partial_\xi\phi dq \tag{7.2.3}$$

where $\phi(t,x,\xi) = \varphi(t,x)\psi(\xi)$ and

$$\overline{G_\pm} = (-a\cdot n) f_\pm^b + \eta\delta_{u_b}\overline{\nabla v}\cdot n$$
$$q = \eta\delta_v|\nabla v|^2 \geqslant 0$$

(notice that the support of q is compact with respect to ξ). Using a classical argument relying on convolution techniques, we claim that (7.2.3) holds true for any test function $\phi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R}^{d+2})$.

Remark 7.2.1 *i) Since f_+, f_+^0, f_+^τ and m vanish for $\xi \gg 1$, Equation (7.2.1) with f_+ holds true when the support of the test function ϕ is merely lower bounded (and not necesseraly compact) with respect to ξ. Similarly, we can apply (7.2.3) with g_- to test functions the support of which is only upper bounded with respect to ξ. Notice also that, in all the following, though we write integrals in ξ on the whole of $\mathbb{R}_\xi$, the integrands we consider are null outside a fixed compact (namely $[-D,D]$) of $\mathbb{R}_\xi$; we use this in some estimates, without recalling it.*

ii) Equations (7.2.1) and (7.2.3) can be applied to certain test functions which are not fully regular but have some monotony properties with respect to ξ, provided we replace the equality by an inequality (the sign of which is given by the monotony of the test function). More precisely, we consider, in the following, test functions of the kind $\phi(t,x,\xi) = \int_0^\infty\int_\Omega \varphi(t,x,s,y)\mathrm{sgn}_\pm(W(s,y) - \xi)\,dy\,ds$, where W is bounded and φ is regular and has a fixed sign; we can approximate $\mathrm{sgn}_\pm$ by some non-decreasing and regular functions $\mathrm{sgn}_{\pm,\delta}$; then, applying (7.2.1) or (7.2.3) to $\phi_\delta(t,x,\xi) = \int_0^\infty\int_\Omega \varphi(t,x,s,y)\mathrm{sgn}_{\pm,\delta}(W(s,y) - \xi)\,dy\,ds$, which is regular and has the same monotony properties as ϕ (with respect to ξ), we notice that the right-hand side has a fixed sign; then, passing to the limit $\delta \to 0$, we see that these inequalities are satisfied with ϕ.

7.2.4 Main ideas of the proof

We present here formal manipulations which enable to understand the key steps of the proof. Let $(t,x) \mapsto \varphi(t,x)$ be a non-negative regular function. Plugging $\phi = \varphi g_-$ in (7.2.1) and $\phi = \varphi f_+$ in (7.2.3), we obtain

$$\int_{Q\times\mathbb{R}_\xi} f_+(\partial_t + a\cdot\nabla)(\varphi g_-) + \int_{\Sigma\times\mathbb{R}_\xi} (-a\cdot n) f_+^\tau f_-^b \overline{\varphi} \leqslant 0$$

and

$$\int_{Q\times\mathbb{R}_\xi} g_-(\partial_t + a\cdot\nabla)(\varphi f_+) + \int_{\Sigma\times\mathbb{R}_\xi} (-a\cdot n) f_-^b f_+^\tau \overline{\varphi} - \eta \int_{Q\times\mathbb{R}_\xi} \delta_v \nabla v\cdot\nabla(f_+\varphi) + \eta \int_{\Sigma\times\mathbb{R}_\xi} \delta_{u_b} \overline{\nabla v}\cdot n \overline{f_+}\overline{\varphi} \leqslant 0$$

(since $f_+^0 f_-^0 = 0$). Summing these inequalities and integrating by parts, it comes

$$\int_{Q\times\mathbb{R}_\xi} f_+ g_-(\partial_t + a\cdot\nabla)\varphi \leqslant -\int_{\Sigma\times\mathbb{R}_\xi} (-a\cdot n) f_+^\tau f_-^b \overline{\varphi} - \eta \int_{\Sigma\times\mathbb{R}_\xi} \delta_{u_b} \overline{\nabla v}\cdot n\overline{f_+}\overline{\varphi} + \eta \int_{Q\times\mathbb{R}_\xi} \delta_v \nabla v\cdot\nabla(f_+\varphi).$$

Taking $\varphi(t,x) = \omega_\zeta(t)$ with $(\omega_\zeta)_{\zeta>0}$ which converges to the characteristic function of $[0,T]$ and $\omega_\zeta' \to -\delta_T$, this gives

$$\begin{aligned}\int_\Omega (u-v)^+(T,x)\,dx &= \int_{\Omega\times\mathbb{R}_\xi} (-f_+ g_-)^{(t=T)} \\ &\leqslant -\int_{[0,T]\times\partial\Omega\times\mathbb{R}_\xi} (-a\cdot n) f_+^\tau f_-^b - \eta \int_{[0,T]\times\partial\Omega\times\mathbb{R}_\xi} \delta_{u_b} \overline{\nabla v}\cdot n\overline{f_+} + \eta \int_{[0,T]\times\Omega\times\mathbb{R}_\xi} \delta_v \nabla v\cdot\nabla f_+. \quad (7.2.4)\end{aligned}$$

The functions f_+ and g_- are not regular enough to justify such manipulations, which are therefore performed with f_+^ε and g_-^ν, regularized versions of these applications. The smoothing of g_- is purely technical and we immediately let $\nu \to 0$; at the contrary, the way we define f_+^ε is crucial for the proof. A decentralizing regularization allows to get rid of the second term of the right-hand side of (7.2.4); the size of the regularization being ε, $||\nabla f_+^\varepsilon||_\infty$ is bounded by C/ε and the last term of (7.2.4) is of order η/ε. There remains to estimate the first term of the right-hand side of (7.2.4), which is the aim of a whole section (Section 7.6); the idea is to re-use the kinetic equation satisfied by v.

7.3 Estimate in the interior of the domain

In this section, we let $\lambda = \lambda_0$ (we drop the subscript 0) and $K := \mathrm{supp}(\lambda_0)$. In order to obtain an estimate on the interior of the domain, we need to localize using λ, regularize both kinetic equations in order to combine them, proceeding as we did when proving the Comparison Theorem in [58]. This step is more or less classical.

Let $\alpha > 0$ and $0 < \varepsilon < \mathrm{dist}(K, \partial\Omega)$; denote $\gamma_\varepsilon(x) = \prod_{i=1}^d \theta_\varepsilon(x_i)$. Taking $\phi \in \mathsf{C}_\mathsf{c}^\infty(\mathbb{R}^{d+2})$ with support in $\mathbb{R}\times K\times\mathbb{R}_\xi$ and using $\phi \star (\check{\gamma_\varepsilon}\otimes\check{\theta_\alpha})$ [1] — notice that this function is null on the boundary of Ω — as a test function in (7.2.1) with f_+, we find

$$\int_{\mathbb{R}^{d+2}} f_+^{\alpha,\varepsilon}(\partial_t + a\cdot\nabla)\phi + \int_{\mathbb{R}^{d+2}} f_+^{0\,\varepsilon}\otimes\theta_\alpha\phi = \int_{\mathbb{R}^{d+2}} \partial_\xi\phi\, dm^{\alpha,\varepsilon} \quad (7.3.1)$$

(where $m^{\alpha,\varepsilon}$ is the convolution in (t,x) of $\gamma_\varepsilon\otimes\theta_\alpha$ by the extention of m by 0 outside $Q\times\mathbb{R}_\xi$). We next regularize the equation satisfied by g_-, using the same method but different parameters $\beta > 0$ and $0 < \nu < \mathrm{dist}(K,\partial\Omega)$: we obtain for the same ϕ's

$$\int_{\mathbb{R}^{d+2}} g_-^{\beta,\nu}(\partial_t + a\cdot\nabla)\phi + \int_{\mathbb{R}^{d+2}} f_-^{0\,\nu}\otimes\theta_\beta\phi - \eta\int_{Q\times\mathbb{R}_\xi} \delta_v\nabla v\cdot(\nabla\phi)\star(\check{\gamma_\nu}\otimes\check{\theta_\beta}) = \int_{\mathbb{R}^{d+2}} \partial_\xi\phi\, dq^{\beta,\nu}. \quad (7.3.2)$$

[1] Here and after, the tensorial product is used to recall that $\check{\gamma_\varepsilon}$ and $\check{\theta_\alpha}$ use different variables (for example, $\check{\gamma_\varepsilon}\otimes\check{\theta_\alpha}(t,x) = \check{\gamma_\varepsilon}(x)\check{\theta_\alpha}(t)$) and the convolution product $\star$ never involves the kinetic variable ξ.

Suppose that $\phi \in \mathsf{C}_{\mathsf{c}}^{\infty}(\mathbb{R}^{d+1})$ is non-negative with support in $\mathbb{R}\times K$ and apply (7.3.1) to the test function $-g_-^{\beta,\nu}(t,x,\xi)\phi(t,x)$ and (7.3.2) to $-f_+^{\alpha,\varepsilon}(t,x,\xi)\phi(t,x)$, and sum the two equations; using the fact that $-f_+^{\alpha,\varepsilon}$ and $-g_-^{\beta,\nu}$ are non-decreasing with respect to ξ, we find, after some integrate by parts,

$$-\int_{\mathbb{R}^{d+2}} f_+^{\alpha,\varepsilon} g_-^{\beta,\nu}(\partial_t + a\cdot\nabla)\phi + \eta\int_{Q\times\mathbb{R}_\xi} \delta_v \nabla v\cdot(\nabla(f_+^{\alpha,\varepsilon}\phi)) \star (\check{\gamma_\nu}\otimes\check{\theta_\beta}) - \int_{\mathbb{R}^{d+2}} \left[f_+^{0\,\varepsilon}\otimes\theta_\alpha g_-^{\beta,\nu} + f_-^{0\,\nu}\otimes\theta_\beta f_+^{\alpha,\varepsilon}\right]\phi \geqslant 0. \quad (7.3.3)$$

Thanks to the decentred regularization, $f_+^{\alpha,\varepsilon}(t,x,\xi)$ is null if $t\leqslant\alpha/2$; hence, for $\beta\leqslant\alpha/2$, $f_-^{0\,\nu}\otimes\theta_\beta f_+^{\alpha,\varepsilon}\equiv 0$. Moreover, the function which associates t with

$$\begin{aligned}\int_{\Omega\times\mathbb{R}_\xi} f_+^{0\,\varepsilon}(x,\xi) g_-(t,x,\xi)\phi(t,x)\,dx\,d\xi &= \int_\Omega\int_\Omega\int_{\mathbb{R}_\xi} f_+^0(y,\xi) g_-(t,x,\xi)\gamma_\varepsilon(x-y)\phi(t,x)\,d\xi\,dy\,dx \\ &= -\int_\Omega\int_\Omega (u_0(y)-v(t,x))^+\gamma_\varepsilon(x-y)\phi(t,x)\,d\xi\,dy\,dx \quad (7.3.4)\end{aligned}$$

is continuous (because $v\in\mathsf{C}([0,T_0];L^1(\Omega))$ and we have $g_-(0,\cdot,\cdot)=f_-^0$. Therefore, letting β, ν and α successively tend to zero in (7.3.3), we have

$$\int_{Q\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)(\partial_t+a\cdot\nabla)\phi + \eta\int_{Q\times\mathbb{R}_\xi}\delta_v\nabla v\cdot\nabla(f_+^\varepsilon\phi) - \int_{\Omega\times\mathbb{R}_\xi} f_+^{0\,\varepsilon} f_-^0\phi^{(t=0)} \geqslant 0.$$

Choose $T\in[0,T_0]$ and let $\phi(t,x)=\lambda(x)w_\beta(t)$ where $w_\beta(t)=\int_{t-T}^{+\infty}\theta_\beta(r)dr$; we obtain

$$\int_{Q\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)[-\theta_\beta(t-T)\lambda + w_\beta a\cdot\nabla\lambda] + \eta\int_{Q\times\mathbb{R}_\xi}\delta_v\nabla v\cdot\nabla(f_+^\varepsilon\lambda)w_\beta - \int_{\Omega\times\mathbb{R}_\xi} f_+^{0\,\varepsilon} f_-^0\lambda \geqslant 0.$$

The function $t\to\int_{\Omega\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)(t,x,\xi)\lambda(x)\,dx$ is continuous (it is similar to (7.3.4)); thus, letting $\beta\to 0$,

$$-\int_{\Omega\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)^{(t=T)}\lambda + \int_{Q^T\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)a\cdot\nabla\lambda + \int_{Q^T\times\mathbb{R}_\xi}\eta\delta_v\nabla v\cdot\nabla(f_+^\varepsilon\lambda) - \int_{\Omega\times\mathbb{R}_\xi} f_+^{0\,\varepsilon} f_-^0\lambda \geqslant 0$$

where $Q^T=(0,T)\times\Omega$. We therefore obtain

$$T_1 \leqslant T_2 + T_{IC} + T_D \quad (7.3.5)$$

where

$$\begin{aligned} T_1 &= \int_{\Omega\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)^{(t=T)}\lambda\,, \\ T_2 &= \int_{Q^T\times\mathbb{R}_\xi}(-f_+^\varepsilon g_-)a\cdot\nabla\lambda\,, \\ T_D &= \int_{Q^T\times\mathbb{R}_\xi}\eta\delta_v\nabla v\cdot\nabla(f_+^\varepsilon\lambda)\,, \\ T_{IC} &= -\int_{\Omega\times\mathbb{R}_\xi} f_+^{0\,\varepsilon} f_-^0\lambda. \end{aligned}$$

We now estimate these terms. We have

$$\begin{aligned} T_1 &= \int_K\int_\Omega\left[\int_{\mathbb{R}_\xi}(-f_+(T,y,\xi)g_-(T,x,\xi))\,d\xi\right]\lambda(x)\gamma_\varepsilon(x-y)\,dy\,dx \\ &= \int_K\int_\Omega (u(T,y)-v(T,x))^+\lambda(x)\gamma_\varepsilon(x-y)\,dy\,dx \\ &\geqslant \int_K\int_\Omega (u(T,x)-v(T,x))^+\lambda(x)\gamma_\varepsilon(x-y)\,dy\,dx - \int_K\int_\Omega (u(T,x)-u(T,y))^+\lambda(x)\gamma_\varepsilon(x-y)\,dy\,dx. \end{aligned}$$

But, if $x \in K$, we have, by choice of ε, $x - \Omega \supset \mathrm{supp}(\gamma_\varepsilon)$, hence $\int_\Omega \gamma_\varepsilon(x-y)\,dy = 1$. Moreover, since $u(T,\cdot) \in \mathsf{BV}(\Omega)$, by Lemma 7.8.1 (see the appendix),

$$\int_K \int_\Omega (u(T,x) - u(T,y))^+ \lambda(x) \gamma_\varepsilon(x-y)\,dy\,dx \leqslant \int_\Omega \int_\Omega |u(T,x) - u(T,y)| \gamma_\varepsilon(x-y)\,dy\,dx \leqslant C\varepsilon.$$

Hence,

$$T_1 \geqslant \int_\Omega (u(T,x) - v(T,x))^+ \lambda(x)\,dx - C\varepsilon.$$

Next, reasoning as for T_1,

$$\begin{aligned} T_2 &= \int_0^T \int_K \int_\Omega \int_{\mathbb{R}_\xi} (-f_+(t,y,\xi) g_-(t,x,\xi)) \gamma_\varepsilon(x-y) a(\xi)\cdot\nabla\lambda(x)\,dy\,dx\,dt \\ &\leqslant C \int_0^T \int_\Omega \int_\Omega (u(t,y) - v(t,x))^+ \gamma_\varepsilon(x-y)\,dy\,dx\,dt \\ &\leqslant C\varepsilon + C \int_0^T \int_\Omega (u(t,x) - v(t,x))^+\,dx\,dt. \end{aligned}$$

Let us estimate the diffusion term T_D. First, we write: $T_D = T_D^1 + T_D^2$ with

$$T_D^1 = \int_{Q^T \times \mathbb{R}_\xi} \eta \delta_v \nabla v \cdot f_+^\varepsilon \nabla\lambda = \int_{Q^T \times \mathbb{R}_\xi} \eta \nabla v \cdot \left(\int_{\mathbb{R}_\xi} \delta_v f_+^\varepsilon \right) \nabla\lambda \leqslant \eta \int_{Q^T} |\nabla v| |\nabla\lambda| \leqslant C\eta$$

and

$$T_D^2 = \int_{Q^T \times \mathbb{R}_\xi} \eta \delta_v \nabla v \cdot \lambda \nabla f_+^\varepsilon \leqslant C\eta \int_{Q^T} |\nabla v|(t,x) \sup_\xi |\nabla f_+^\varepsilon|(t,x,\xi)\,dtdx.$$

But $\nabla f_+^\varepsilon(t,x,\xi) = \int_\Omega f_+(t,y,\xi) \nabla\gamma_\varepsilon(x-y)\,dy$, so that $|\nabla f_+^\varepsilon|(t,x,\xi) \leqslant ||\nabla\gamma_\varepsilon||_{L^1(\mathbb{R}^d)} \leqslant C/\varepsilon$. Hence,

$$T_D^2 \leqslant \frac{C\eta}{\varepsilon} \int_{Q^T} |\nabla v| \leqslant \frac{C\eta}{\varepsilon}.$$

Using Lemma 7.8.1, a straightforward computation gives $T_{IC} \leqslant C\varepsilon$. We finally gather the different estimates in (7.3.5) and get, for all ε,

$$\int_\Omega (u(T,x) - v(T,x))^+ \lambda(x) dx \leqslant C\left(\varepsilon + \frac{\eta}{\varepsilon}\right) + C \int_0^T \int_\Omega (u(t,x) - v(t,x))^+\,dx\,dt.$$

Minimizing on ε, we obtain (recall that $\lambda = \lambda_0$ here)

$$\int_\Omega (u(T,x) - v(T,x))^+ \lambda_0(x) dx \leqslant C\sqrt{\eta} + C \int_0^T \int_\Omega (u(t,x) - v(t,x))^+\,dx\,dt. \tag{7.3.6}$$

7.4 Transport and regularization of the kinetic equations

In order to estimate $(u(T,\cdot) - v(T,\cdot))^+$ near the boundary of Ω, we choose a chart (O_i, h_i, λ_i) and we transport the equations to B_+^d. In the following, we drop the subscript i.

7.4.1 Transport of the kinetic equations

We now write the kinetic equations satisfied by u and v once they have been transported on B_+^d. Consider a test function $\Psi \in \mathsf{C}_\mathsf{c}^\infty(\mathbb{R} \times B^d \times \mathbb{R}_\xi)$ and set $\phi(t,x,\xi) = \Psi(t,h(x),\xi) \in \mathsf{C}_\mathsf{c}^2(\mathbb{R}_t \times O \times \mathbb{R}_\xi)$. Next, extend

ϕ by 0 to get a function $\phi \in \mathsf{C}^2_{\mathsf{c}}(\mathbb{R}^{d+2})$ and plug it into $(7.2.1)_+$ (ϕ is not C^∞ but is regular enough to be taken as a test function in this equation). This gives

$$\int_{\mathbb{R}\times O\times\mathbb{R}_\xi} f_+\left[(\partial_t\Psi)\circ h + a\cdot h'^T(\nabla\Psi)\circ h\right] + \int_{O\times\mathbb{R}_\xi} f_+^0\Psi^{(t=0)}\circ h + \int_{\mathbb{R}\times(\partial\Omega\cap O)\times\mathbb{R}_\xi}(-a\cdot n)f_+^\tau\overline{\Psi}\circ h$$
$$= \int_{\mathbb{R}\times O\times\mathbb{R}_\xi}(\partial_\xi\Psi)\circ h\,dm.$$

Through the change of variables $y = h(x)$, and by definition of the measure on Σ, we obtain

$$\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}|Jh^{-1}|f_+\circ h^{-1}(\partial_t\Psi + h'\circ h^{-1}a\cdot\nabla\Psi) + \int_{B_+^d}\int_{\mathbb{R}_\xi}|Jh^{-1}|f_+^0\circ h^{-1}\Psi^{(t=0)}$$
$$+\int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi}(-a\cdot nf_+^\tau)\circ h^{-1}\overline{\Psi}\left|\frac{\partial h^{-1}}{\partial x_1}\wedge\cdots\wedge\frac{\partial h^{-1}}{\partial x_{d-1}}\right|dx_1\ldots dx_{d-1}$$
$$= \int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(\partial_\xi\Psi)\,d(h_*m).$$

In the following, we adopt the notations

$$j(x) = |Jh^{-1}(x)| \quad\text{and}\quad H(x) = h'\circ h^{-1}(x) \quad\text{and}\quad l(x) = \left|\frac{\partial h^{-1}}{\partial x_1}\wedge\cdots\wedge\frac{\partial h^{-1}}{\partial x_{d-1}}\right|(x).$$

Moreover, for any function $r(t,x,\xi)$, we write $\tilde{r}(t,x,\xi)$ for $r(t,h^{-1}(x),\xi)$. Therefore, the previous equality reads

$$\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi} j\tilde{f}_+(\partial_t\Psi + Ha\cdot\nabla\Psi) + \int_{B_+^d}\int_{\mathbb{R}_\xi} j\tilde{f}_+^0\Psi^{(t=0)} + \int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi} l(-a\cdot\tilde{n})\tilde{f}_+^\tau\overline{\Psi}$$
$$= \int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\partial_\xi\Psi\,d(h_*m). \quad (7.4.1)$$

Similar computations are achieved on the kinetic equation satisfied by v. We obtain

$$\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi} j\tilde{g_-}(\partial_t\Psi + Ha\cdot\nabla\Psi) + \int_{B_+^d}\int_{\mathbb{R}_\xi} j\tilde{f}_-^0\Psi^{(t=0)}$$
$$+\int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi} l(-a\cdot\tilde{n})\tilde{f}_-^b\overline{\Psi} + \int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi} l\tilde{D}\delta_{\tilde{u_b}}\overline{\Psi}$$
$$+\eta\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\tilde{Z}\delta_{\tilde{v}}\cdot\nabla\Psi = \int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\partial_\xi\Psi\,d(h_*q) \quad (7.4.2)$$

where $D(t,\overline{x}) = \eta\nabla v(t,\overline{x})\cdot n(\overline{x})$ and $Z(t,x) = -h'(x)\nabla v(t,x)$. Notice that

$$Z \text{ is bounded in } L^1((0,T)\times\Omega) \text{ for all } T \geqslant 0. \quad (7.4.3)$$

This property, as well as the Lipschitz continuity $[0,\infty)\to\mathsf{L}^1$ of u and v (with a Lipschitz constant for v independent of η) and the bounds on $|u(t,\cdot)|_{\mathsf{BV}}$ and $|v(t,\cdot)|_{\mathsf{BV}}$, are conserved by the transport by h.

7.4.2 Transport of the BLN condition

We state here the only consequence of (7.2.2) that we use in the following.

Let $\Psi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R} \times B^{d-1} \times \mathbb{R}_\xi)$ be non-negative and non-decreasing with respect to ξ. The function $\phi(t, \overline{x}, \xi) = \Psi(t, h(\overline{x}), \xi)(1 - f^b_+(t, \overline{x}, \xi))$ is non-decreasing with respect to ξ (since Ψ and $1 - f^b_+$ are non-negative and non-decreasing with respect to ξ). Hence, (7.2.2) implies

$$\int_0^\infty \int_{\partial\Omega \cap O} \int_{\mathbb{R}_\xi} (-a \cdot n) f^\tau_+ \Psi \circ h (1 - f^b_+) \leqslant \int_0^\infty \int_{\partial\Omega \cap O} \int_{\mathbb{R}_\xi} M f^b_+ (1 - f^b_+) \Psi \circ h.$$

But $f^b_+(1 - f^b_+) = 0$ so that, transporting this equation with h^{-1} on B^{d-1}, we deduce that, for all $\Psi \in \mathsf{C}^\infty_{\mathsf{c}}(\mathbb{R} \times B^{d-1} \times \mathbb{R}_\xi)$ which is non-negative and non-decreasing with respect to ξ,

$$\int_0^\infty \int_{B^{d-1}} \int_{\mathbb{R}_\xi} l(-a \cdot \tilde{n}) \tilde{f}^\tau_+ \Psi \leqslant \int_0^\infty \int_{B^{d-1}} \int_{\mathbb{R}_\xi} l(-a \cdot \tilde{n}) \tilde{f}^\tau_+ \tilde{f}^b_+ \Psi. \tag{7.4.4}$$

We also need to understand how the unit normal is transported by the chart (O, h).

Lemma 7.4.1 *For all $\overline{y} \in B^{d-1}$ and all $X \in \mathbb{R}^d$, we have $l(\overline{y}) X \cdot \tilde{n}(\overline{y}) = -j(\overline{y})(H(\overline{y})X)_d$, where $(H(\overline{y})X)_d$ is the d-th coordinate of $H(\overline{y})X$.*

Proof of Lemma 7.4.1
Let $\psi \in \mathsf{C}^\infty_{\mathsf{c}}(B^d)$ and $\phi = \psi \circ h \in \mathsf{C}^2_{\mathsf{c}}(O)$ (extented by 0 outside O). Integrating by parts, we have

$$\int_\Omega X \cdot \nabla\phi(x)\, dx = \int_{\partial\Omega} \phi(\overline{x}) X \cdot n(\overline{x})\, d\sigma(\overline{x}).$$

Since $\nabla\phi(x) = h'(x)^T \nabla\psi(h(x))$, transporting these integrals by h (all the integrands are null outside O), we find

$$\int_{B^d_+} j(x) H(x) X \cdot \nabla\psi(x)\, dx = \int_{B^d_+} X \cdot (h'(h^{-1}(x)))^T \nabla\psi(x) |J h^{-1}(x)|\, dx = \int_{B^{d-1}} \psi(\overline{x}) X \cdot n(h^{-1}(\overline{x})) l(\overline{x})\, d\overline{x}.$$

Another integrate by parts then yields

$$\int_{B^{d-1}} \psi(\overline{x}) X \cdot n(h^{-1}(\overline{x})) l(\overline{x})\, d\overline{x} = \int_{B^{d-1}} (-j(\overline{x})(H(\overline{x})X)_d) \psi(\overline{x})\, d\overline{x} - \int_{B^d_+} \mathrm{div}(jHX)(x)\psi(x)\, dx.$$

(the unit normal to B^d_+ on B^{d-1} is $(0, \ldots, 0, -1)$). Taking first $\psi \in \mathsf{C}^\infty_{\mathsf{c}}(B^d_+)$, we see that $\mathrm{div}(jHX) = 0$ on B^d_+; thus, for all $\psi \in \mathsf{C}^\infty_{\mathsf{c}}(B^d)$, $\int_{B^{d-1}} \psi(\overline{x}) X \cdot n(h^{-1}(\overline{x})) l(\overline{x})\, d\overline{x} = \int_{B^{d-1}} (-j(\overline{x})(H(\overline{x})X)_d) \psi(\overline{x})\, d\overline{x}$, which concludes the proof. ■

7.4.3 Regularization of the transported equations

From now on, we work on B^d and we thus simply write r for $\tilde{r}$. Let $K := \mathrm{supp}(\lambda)$ (compact subset of B^d). We now regularize equations (7.4.1) and (7.4.2).

For $\overline{\varepsilon} > 0$, we denote $\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}) = \prod_{i=1}^{d-1} \theta_{\overline{\varepsilon}}(x_i)$; we take $\varepsilon_d > 0$ and we denote $\varepsilon = (\overline{\varepsilon}, \varepsilon_d)$, $\gamma_\varepsilon(x) = \overline{\gamma}_{\overline{\varepsilon}}(\overline{x}) \theta_{\varepsilon_d}(x_d)$. We choose $\overline{\varepsilon} + \varepsilon_d < \mathrm{dist}(K, \partial B^d)$. Let $\Psi \in \mathsf{C}^2_{\mathsf{c}}(\mathbb{R} \times B^d \times \mathbb{R}_\xi)$ with support in $\mathbb{R} \times K \times \mathbb{R}_\xi$; then, $\Psi \star (\check{\gamma}_\varepsilon \otimes \check{\theta}_\alpha)$ is compactly supported in $\mathbb{R} \times B^d \times \mathbb{R}_\xi$. Using $\Psi \star (\check{\gamma}_\varepsilon \otimes \check{\theta}_\alpha)$ in (7.4.1), we get

$$\int_{\mathbb{R}^{d+2}} (j f_+)^{\alpha,\varepsilon} \partial_t \Psi + (j f_+ H)^{\alpha,\varepsilon} a \cdot \nabla\Psi + \int_{\mathbb{R}^{d+2}} (j f^0_+)^\varepsilon \otimes \theta_\alpha \Psi$$
$$+ \int_{\mathbb{R}^{d+2}} (l(-a \cdot n) f^\tau_+)^{\alpha,\overline{\varepsilon}} \otimes \theta_{\varepsilon_d} \Psi = \int_{\mathbb{R}^{d+2}} \partial_\xi \Psi\, d(h_* m)^{\alpha,\varepsilon}. \tag{7.4.5}$$

The same test function with parameters β and ν in (7.4.2) gives

$$\int_{\mathbb{R}^{d+2}} (jg_-)^{\beta,\nu}\partial_t\Psi + (jg_-H)^{\beta,\nu}a\cdot\nabla\Psi + \int_{\mathbb{R}^{d+2}} (jf_-^0)^\nu\otimes\theta_\beta\Psi$$
$$+\int_{\mathbb{R}^{d+2}} (l(-a\cdot n)f_-^b)^{\beta,\overline{\nu}}\otimes\theta_{\nu_d}\Psi + \int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi} lD\delta_{u_b}\overline{\Psi\star(\check{\gamma_\nu}\otimes\check{\theta_\beta})}$$
$$+\eta\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi} Z\delta_v\cdot\nabla\Psi\star(\check{\gamma_\nu}\otimes\check{\theta_\beta}) = \int_{\mathbb{R}^{d+2}}\partial_\xi\Psi\, d(h_*q)^{\beta,\nu}. \quad (7.4.6)$$

7.5 Combination of the equations and new estimates

The next step consists in combining the two preceding kinetic equations. Choose a non-negative regular function $\phi(t,x)$, with support in $]-\infty,T_0]\times K$, and apply $(-jf_+)^{\alpha,\varepsilon}(t,x,\xi)\phi(t,x)$ as a test function in (7.4.6) and $(-jg_-)^{\beta,\nu}(t,x,\xi)\phi(t,x)$ as a test function in (7.4.5). These two test functions are non-decreasing with respect to ξ so that, summing the results, we get $U_1^{\beta,\nu}+U_2^{\beta,\nu}+U_3^{\beta,\nu}+U_4^{\beta,\nu}+U_5^{\beta,\nu}+U_6^{\beta,\nu}\geqslant 0$, where

$$\begin{aligned}
U_1^{\beta,\nu} &= \int_{\mathbb{R}^{d+2}} (jf_+)^{\alpha,\varepsilon}(\partial_t(-jg_-)^{\beta,\nu}\phi + (-jg_-)^{\beta,\nu}\partial_t\phi) + (-jg_-)^{\beta,\nu}(\partial_t(jf_+)^{\alpha,\varepsilon}\phi + (jf_+)^{\alpha,\varepsilon}\partial_t\phi)\\
U_2^{\beta,\nu} &= \int_{\mathbb{R}^{d+2}} (jf_+H)^{\alpha,\varepsilon}a\cdot\big(\nabla(-jg_-)^{\beta,\nu}\phi + (-jg_-)^{\beta,\nu}\nabla\phi\big)\\
&\qquad +(-jg_-H)^{\beta,\nu}a\cdot(\nabla(jf_+)^{\alpha,\varepsilon}\phi + (jf_+)^{\alpha,\varepsilon}\nabla\phi)\\
U_3^{\beta,\nu} &= \int_{\mathbb{R}^{d+2}} (jf_+^0)^\varepsilon\otimes\theta_\alpha(-jg_-)^{\beta,\nu}\phi + (-jf_-^0)^\nu\otimes\theta_\beta(jf_+)^{\alpha,\varepsilon}\phi\\
U_4^{\beta,\nu} &= -\int_0^\infty\int_{B^{d-1}}\int_{\mathbb{R}_\xi} lD\delta_{u_b}\overline{((jf_+)^{\alpha,\varepsilon}\phi)\star(\check{\gamma_\nu}\otimes\check{\theta_\beta})}\\
U_5^{\beta,\nu} &= \int_{\mathbb{R}^{d+2}} (l(-a\cdot n)f_+^\tau)^{\alpha,\overline{\varepsilon}}\otimes\theta_{\varepsilon_d}(-jg_-)^{\beta,\nu}\phi + (-l(-a\cdot n)f_-^b)^{\beta,\overline{\nu}}\otimes\theta_{\nu_d}(jf_+)^{\alpha,\varepsilon}\phi\\
U_6^{\beta,\nu} &= \eta\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\delta_v Z\cdot\nabla((jf_+)^{\alpha,\varepsilon}\phi)\star(\check{\gamma_\nu}\otimes\check{\theta_\beta}).
\end{aligned}$$

7.5.1 Passing to the limit in β and ν

We study the limits of $U_1^{\beta,\nu},\dots,U_6^{\beta,\nu}$ as β and ν tend to 0.

The first term $U_1^{\beta,\nu}$
Integrating by parts, we have

$$U_1^{\beta,\nu} = \int_{\mathbb{R}^{d+2}} (jf_+)^{\alpha,\varepsilon}(-jg_-)^{\beta,\nu}\partial_t\phi$$

and thus, as $\beta\to 0$ and $\nu\to 0$,

$$U_1^{\beta,\nu}\to\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(-jg_-)\partial_t\phi. \quad (7.5.1)$$

The second term $U_2^{\beta,\nu}$

The first step, here, is to get H out of the regularizations $(jf_+H)^{\alpha,\varepsilon}$ and $(jg_-H)^{\beta,\nu}$. To do this, we notice that, for all $(t,x,\xi)\in\mathbb{R}\times B^d\times\mathbb{R}$, since H is C^1,

$$\begin{aligned}
&|(jf_+H)^{\alpha,\varepsilon}(t,x,\xi)-H(x)(jf_+)^{\alpha,\varepsilon}(t,x,\xi)|\\
&\qquad = \left|\int_0^\infty\int_{B_+^d} j(y)f_+(s,y,\xi)H(y)\theta_\alpha(t-s)\gamma_\varepsilon(x-y)\,ds\,dy\right.\\
&\qquad\quad \left.-H(x)\int_0^\infty\int_{B_+^d} j(y)f_+(s,y,\xi)\theta_\alpha(t-s)\gamma_\varepsilon(x-y)\,ds\,dy\right|\\
&\qquad \leqslant \int_0^\infty\int_{B_+^d} j(y)f_+(s,y,\xi)|H(y)-H(x)|\theta_\alpha(t-s)\gamma_\varepsilon(x-y)\,ds\,dy\\
&\qquad \leqslant C(\overline{\varepsilon}+\varepsilon_d)\int_0^\infty\int_{B_+^d}\theta_\alpha(t-s)\gamma_\varepsilon(x-y)\,ds\,dy\leqslant C(\overline{\varepsilon}+\varepsilon_d)
\end{aligned}$$

(here, "$|\cdot|$" is a matrice norm). Hence,

$$\begin{aligned}
&\left|\int_{\mathbb{R}\times B^d\times\mathbb{R}_\xi}(jf_+H)^{\alpha,\varepsilon}a\cdot\left(\nabla(-jg_-)^{\beta,\nu}\phi+(-jg_-)^{\beta,\nu}\nabla\phi\right)\right.\\
&\qquad \left.-\int_{\mathbb{R}\times B^d\times\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}Ha\cdot\left(\nabla(-jg_-)^{\beta,\nu}\phi+(-jg_-)^{\beta,\nu}\nabla\phi\right)\right|\\
&\leqslant C(\overline{\varepsilon}+\varepsilon_d)\left(||\nabla(-jg_-)^{\beta,\nu}||_{\mathsf{L}^1(]-\infty,T_0]\times K\times[-D,D])}||\phi||_\infty+||(-jg_-)^{\beta,\nu}||_{\mathsf{L}^\infty(\mathbb{R}^{d+2})}||\nabla\phi||_\infty\right) \quad (7.5.2)
\end{aligned}$$

(recall that $\mathrm{supp}(\phi)\subset]-\infty,T_0]\times K$). But, by Lemma 7.8.2 (see the appendix), for all $s\in\mathbb{R}^+$,

$$\int_K\int_{-D}^D|\nabla(j\mathrm{sgn}_-(v(s,\cdot)-\xi))^\nu|\leqslant C(1+|v(s,\cdot)|_{\mathsf{BV}(B_+^d)})\leqslant C.$$

Therefore,

$$\begin{aligned}
||\nabla(-jg_-)^{\beta,\nu}||_{\mathsf{L}^1(]-\infty,T_0]\times K\times[-D,D])} &\leqslant \int_{-\infty}^{T_0}\int_0^\infty\int_K\int_{-D}^D|\nabla(j\mathrm{sgn}_-(v(s,\cdot)-\xi))^\nu|\theta_\beta(t-s)\,ds\,dt\\
&\leqslant C\int_0^{T_0}\int_0^\infty\theta_\beta(t-s)\,ds\,dt\leqslant C. \qquad (7.5.3)
\end{aligned}$$

Noticing that, thanks to ϕ, the integrals in $U_2^{\beta,\nu}$ are in fact on $\mathbb{R}\times B^d\times\mathbb{R}_\xi$, we deduce from (7.5.2), (7.5.3) and similar estimates for the second part of $U_2^{\beta,\nu}$ that $U_2^{\beta,\nu}$ is equal to

$$\begin{aligned}
&\int_{\mathbb{R}\times B^d\times\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}Ha\cdot(\nabla(-jg_-)^{\beta,\nu}\phi+(-jg_-)^{\beta,\nu}\nabla\phi)+(-jg_-)^{\beta,\nu}Ha\cdot(\nabla(jf_+)^{\alpha,\varepsilon}\phi+(jf_+)^{\alpha,\varepsilon}\nabla\phi)\\
&\qquad\qquad +\mathcal{O}((\overline{\varepsilon}+\varepsilon_d+\overline{\nu}+\nu_d)(||\phi||_\infty+||\nabla\phi||_\infty))\\
&\qquad =\int_{\mathbb{R}\times B^d\times\mathbb{R}_\xi}\phi Ha\cdot\nabla((jf_+)^{\alpha,\varepsilon}(-jg_-)^{\beta,\nu})+2(jf_+)^{\alpha,\varepsilon}(-jg_-)^{\beta,\nu}Ha\cdot\nabla\phi\\
&\qquad\qquad +\mathcal{O}((\overline{\varepsilon}+\varepsilon_d+\overline{\nu}+\nu_d)(||\phi||_\infty+||\nabla\phi||_\infty))\\
&=\int_{\mathbb{R}\times B^d\times\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(-jg_-)^{\beta,\nu}(2Ha\cdot\nabla\phi-\mathrm{div}(\phi Ha))+\mathcal{O}((\overline{\varepsilon}+\varepsilon_d+\overline{\nu}+\nu_d)(||\phi||_\infty+||\nabla\phi||_\infty))
\end{aligned}$$

(we used the fact that ϕ has a compact support in $\mathbb{R}\times B^d$). Letting β and ν tend to 0, this gives

$$\limsup_{\beta,\nu\to0}U_2^{\beta,\nu}\leqslant\mathcal{O}((\overline{\varepsilon}+\varepsilon_d)(||\phi||_\infty+||\nabla\phi||_\infty))+\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(-jg_-)(2Ha\cdot\nabla\phi-\mathrm{div}(\phi Ha)). \quad (7.5.4)$$

The third, fourth and fifth terms
By the choice of a decentred convolution kernel, we have for β and ν_d small enough

$$\theta_\beta(\cdot)(jf_+)^{\alpha,\varepsilon}(\cdot,x,\xi)\equiv 0,\qquad \overline{((jf_+)^{\alpha,\varepsilon}\phi)\star(\check{\gamma}_\nu\otimes\check{\theta}_\beta)}\equiv 0,\qquad \theta_{\nu_d}(\cdot)(jf_+)^{\alpha,\varepsilon}(t,\overline{x},\cdot,\xi)\equiv 0.$$

Therefore, as β and ν go to 0,

$$U_3^{\beta,\nu} \to \int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+^0)^\varepsilon\otimes\theta_\alpha(-jg_-)\phi \tag{7.5.5}$$

$$U_4^{\beta,\nu} \to 0 \tag{7.5.6}$$

$$U_5^{\beta,\nu} \to \int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(l(-a\cdot n)f_+^\tau)^{\alpha,\overline{\varepsilon}}\otimes\theta_{\varepsilon_d}(-jg_-)\phi. \tag{7.5.7}$$

The sixth term $U_6^{\beta,\nu}$
Since $(jf_+)^{\alpha,\varepsilon}\phi$ is regular, we have

$$U_6^{\beta,\nu}\to\eta\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\delta_v Z\cdot\nabla((jf_+)^{\alpha,\varepsilon}\phi). \tag{7.5.8}$$

as β and ν tend to 0.

Using (7.5.1), (7.5.4), (7.5.5), (7.5.6), (7.5.7) and (7.5.8) in $U_1^{\beta,\nu}+U_2^{\beta,\nu}+U_3^{\beta,\nu}+U_4^{\beta,\nu}+U_5^{\beta,\nu}+U_6^{\beta,\nu}\geqslant 0$, we obtain as β and ν go to 0

$$\begin{aligned}
&-\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(-jg_-)\partial_t\phi\\
&\leqslant C((\overline{\varepsilon}+\varepsilon_d)(||\phi||_\infty+||\nabla\phi||_\infty))+\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(-jg_-)(2Ha\cdot\nabla\phi-\operatorname{div}(\phi Ha))\\
&\quad+\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+^0)^\varepsilon\otimes\theta_\alpha(-jg_-)\phi+\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}(l(-a\cdot n)f_+^\tau)^{\alpha,\overline{\varepsilon}}\otimes\theta_{\varepsilon_d}(-jg_-)\phi\\
&\quad+\eta\int_0^\infty\int_{B_+^d}\int_{\mathbb{R}_\xi}\delta_v Z\cdot\nabla((jf_+)^{\alpha,\varepsilon}\phi).
\end{aligned}\tag{7.5.9}$$

7.5.2 Choice of ϕ and continuation of the estimates

We now take $T\in[0,T_0]$ and $\phi(t,x)=\lambda(x)w_\beta(t)$, where $w_\beta(t)=\int_{t-T}^\infty\check{\theta}_\beta(s)\,ds$ (notice that w_β has its support in $]-\infty,T_0]$). The function w_β converges, as $\beta\to 0$, to the characteristic function of $]-\infty,T]$ and $w'_\beta(t)=-\check{\theta}_\beta(t-T)$ converges to $-\delta_T$. Since $t\to\int_{B_+^d}\int_{\mathbb{R}_\xi}(jf_+)^{\alpha,\varepsilon}(t,x,\xi)(-jg_-)(t,x,\xi)\lambda(x)\,dx\,d\xi$ is continuous (it is similar to (7.3.4)), we deduce from (7.5.9) that

$$T_1^{\alpha,\varepsilon}\leq T_2^{\alpha,\varepsilon}+T_3^{\alpha,\varepsilon}+T_4^{\alpha,\varepsilon}+T_5^{\alpha,\varepsilon}+C(\overline{\varepsilon}+\varepsilon_d) \tag{7.5.10}$$

where

$$\begin{aligned}
T_1^{\alpha,\varepsilon} &= \int_{B_+^d}\int_{\mathbb{R}_\xi} ((jf_+)^{\alpha,\varepsilon}(-jg_-))^{(t=T)}\lambda \\
T_2^{\alpha,\varepsilon} &= \int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} (jf_+)^{\alpha,\varepsilon}(-jg_-)Y \\
T_3^{\alpha,\varepsilon} &= \int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} (jf_+^0)^{\varepsilon}\otimes\theta_\alpha(-jg_-)\lambda \\
T_4^{\alpha,\varepsilon} &= \eta\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} \delta_v Z\cdot\nabla((jf_+)^{\alpha,\varepsilon}\lambda) \\
T_5^{\alpha,\varepsilon} &= \int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} (l(-a\cdot n)f_+^\tau)^{\alpha,\overline{\varepsilon}}\otimes\theta_{\varepsilon_d}(-jg_-)\lambda
\end{aligned}$$

and $Y = 2Ha\cdot\nabla\lambda - \mathrm{div}(\lambda Ha) \in \mathsf{L}^\infty((0,T_0)\times B_+^d\times\mathbb{R}_\xi)$. Our aim is to obtain an inequality of the kind of (7.3.6); we now estimate each term $T_i^{\alpha,\varepsilon}$.

The first term $T_1^{\alpha,\varepsilon}$
We have

$$\begin{aligned}
T_1^{\alpha,\varepsilon} &= \int_{B_+^d}\int_0^\infty\int_{B_+^d} j(y)j(x)(u(s,y)-v(T,x))^+\theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x)\,dy\,ds\,dx \\
&\geqslant \int_{B_+^d}\int_0^\infty\int_{B_+^d} j(y)j(x)(u(T,x)-v(T,x))^+\theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x)\,dy\,ds\,dx \\
&\quad - \int_{B_+^d}\int_0^\infty\int_{B_+^d} j(y)j(x)(u(T,x)-u(s,y))^+\theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x)\,dy\,ds\,dx. \quad (7.5.11)
\end{aligned}$$

Lemma 7.8.1 and Proposition 7.2.1 give

$$\int_{B_+^d}\int_0^\infty\int_{B_+^d} |u(T,x)-u(s,y)|\theta_\alpha(T-s)\gamma_\varepsilon(x-y)\,dy\,ds\,dx \leqslant C(\overline{\varepsilon}+\varepsilon_d+\alpha). \quad (7.5.12)$$

Since j is bounded from below by $\underline{j}>0$, we have

$$\begin{aligned}
&\int_{B_+^d}\int_0^\infty\int_{B_+^d} j(y)j(x)(u(T,x)-v(T,x))^+\theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x)\,dy\,ds\,dx \\
&\geqslant \underline{j}^2\int_{B_+^d}\lambda(x)(u(T,x)-v(T,x))^+\left(\int_0^\infty\theta_\alpha(T-s)\,ds\right)\left(\int_{B_+^d}\gamma_\varepsilon(x-y)\,dy\right)dx \\
&\geqslant \underline{j}^2\int_{K\cap B_+^d\cap\{x_d>\varepsilon_d\}}\lambda(x)(u(T,x)-v(T,x))^+\left(\int_0^\infty\theta_\alpha(T-s)\,ds\right)\left(\int_{B_+^d}\gamma_\varepsilon(x-y)\,dy\right)dx
\end{aligned}$$

(recall that K is the support of λ).

If $T\geqslant\alpha$, then $\int_0^\infty\theta_\alpha(T-s)\,ds = \int_{-\infty}^T\theta_\alpha = 1$. Moreover, if $x\in K$ and $x_d>\varepsilon_d$, we have $]0,\overline{\varepsilon}[^{d-1}\times]0,\varepsilon_d[\subset x-B_+^d$ (indeed, if $z\in]0,\overline{\varepsilon}[^{d-1}\times]0,\varepsilon_d[$ then, since $x\in K$, we have $x-z\in B^d$ and, since $x_d>\varepsilon_d>z_d$, $x-z\in B_+^d$); hence, for those x's, $\int_{B_+^d}\gamma_\varepsilon(x-y)\,dy = 1$ since the support of γ_ε is contained in $]0,\overline{\varepsilon}[^{d-1}\times]0,\varepsilon_d[$.

Thus, for $T \geqslant \alpha$,

$$\underline{j}^2 \int_{B_+^d \cap \{x_d > \varepsilon_d\}} \lambda(x)(u(T,x) - v(T,x))^+ \, dx \\ \leqslant \int_{B_+^d} \int_0^\infty \int_{B_+^d} j(y)j(x)(u(T,x) - v(T,x))^+ \theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x) \, dy \, ds \, dx.$$

Since u and v are bounded,

$$\int_{B_+^d \cap \{x_d \leqslant \varepsilon_d\}} \lambda(x)(u(T,x) - v(T,x))^+ \, dx \leqslant \int_{B_+^d \cap \{x_d \leqslant \varepsilon_d\}} C \, dx \leqslant C\varepsilon_d.$$

Hence, if $T \geqslant \alpha$,

$$\underline{j}^2 \int_{B_+^d} \lambda(x)(u(T,x) - v(T,x))^+ \, dx \\ \leqslant \int_{B_+^d} \int_0^\infty \int_{B_+^d} j(y)j(x)(u(T,x) - v(T,x))^+ \theta_\alpha(T-s)\gamma_\varepsilon(x-y)\lambda(x) \, dy \, ds \, dx + C\varepsilon_d. \quad (7.5.13)$$

Equations (7.5.11), (7.5.12) and (7.5.13) give, if $T \geqslant \alpha$,

$$T_1^{\alpha,\varepsilon} \geqslant -C(\overline{\varepsilon} + \varepsilon_d + \alpha) + \underline{j}^2 \int_{B_+^d} \lambda(x)(u(T,x) - v(T,x))^+ \, dx.$$

But u and v are Lipschitz continuous $[0,T_0] \to \mathsf{L}^1(B_+^d)$ (with a Lipschitz constant not depending on η) and equal to u_0 at $t = 0$; hence, for $T \leqslant \alpha$,

$$\int_{B_+^d} \lambda(x)(u(T,x) - v(T,x))^+ \, dx \leqslant C \int_{B_+^d} |u(T,x) - u_0(x)| \, dx + C \int_{B_+^d} |v(T,x) - u_0(x)| \, dx \leqslant C\alpha.$$

Therefore, $T_1^{\alpha,\varepsilon}$ being non-negative, we have, for all $T \in [0,T_0]$,

$$T_1^{\alpha,\varepsilon} \geqslant -C(\overline{\varepsilon} + \varepsilon_d + \alpha) + \underline{j}^2 \int_{B_+^d} \lambda(x)(u(T,x) - v(T,x))^+ \, dx. \quad (7.5.14)$$

The second term $T_2^{\alpha,\varepsilon}$
We have

$$\begin{aligned} T_2^{\alpha,\varepsilon} &= \int_0^T \int_{B_+^d} \int_0^\infty \int_{B_+^d} \int_{\mathbb{R}_\xi} j(y)j(x) f_+(s,y,\xi)(-g_-(t,x,\xi))Y(t,x)\theta_\alpha(t-s)\gamma_\varepsilon(x-y) \, d\xi \, dy \, ds \, dx \, dt \\ &\leqslant C \int_0^T \int_{B_+^d} \int_0^\infty \int_{B_+^d} (u(s,y) - v(t,x))^+ \theta_\alpha(t-s)\gamma_\varepsilon(x-y) \, dy \, ds \, dx \, dt \\ &\leqslant C \int_0^T \int_{B_+^d} \int_0^\infty \int_{B_+^d} (u(s,y) - u(t,x))^+ \theta_\alpha(t-s)\gamma_\varepsilon(x-y) \, dy \, ds \, dx \, dt \\ &\quad + C \int_0^T \int_{B_+^d} \int_0^\infty \int_{B_+^d} (u(t,x) - v(t,x))^+ \theta_\alpha(t-s)\gamma_\varepsilon(x-y) \, dy \, ds \, dx \, dt \\ &\leqslant C(\overline{\varepsilon} + \varepsilon_d + \alpha) + C \int_0^T \int_{B_+^d} (u(t,x) - v(t,x))^+ \, dx \, dt \end{aligned}$$

(we used (7.5.12) with $T = t$).

Therefore,

$$T_2^{\alpha,\varepsilon} \leqslant C(\overline{\varepsilon} + \varepsilon_d + \alpha) + C\int_0^T \int_{B_+^d} (u(t,x) - v(t,x))^+ \, dx \, dt. \tag{7.5.15}$$

The third term $T_3^{\alpha,\varepsilon}$
We write

$$\begin{aligned} T_3^{\alpha,\varepsilon} &= \int_0^T \int_{B_+^d} \int_{B_+^d} \int_{\mathbb{R}_\xi} j(y) j(x) f_+^0(y,\xi)(-g_-(t,x,\xi))\theta_\alpha(t)\gamma_\varepsilon(x-y)\lambda(x) \, d\xi \, dy \, dx \, dt \\ &\leqslant C\int_0^T \int_{B_+^d} \int_{B_+^d} (u_0(y) - v(t,x))^+ \theta_\alpha(t)\gamma_\varepsilon(x-y) \, dy \, dx \, dt. \end{aligned}$$

But $v(0,x) = u_0(x)$ so that, v being Lipschitz continuous $[0,T_0] \to \mathsf{L}^1(B_+^d)$ (with a Lipschitz constant not depending on η) and u_0 being in $\mathsf{BV}(B_+^d)$, by Lemma 7.8.1,

$$\begin{aligned} T_3^{\alpha,\varepsilon} &\leqslant C\int_0^T \int_{B_+^d} \int_{B_+^d} |u_0(y) - u_0(x)|\theta_\alpha(t)\gamma_\varepsilon(x-y) \, dy \, dx \, dt \\ &\quad + C\int_0^T \int_{B_+^d} \int_{B_+^d} |v(0,x) - v(t,x)|\theta_\alpha(t)\gamma_\varepsilon(x-y) \, dy \, dx \, dt \\ &\leqslant C(\overline{\varepsilon} + \varepsilon_d + \alpha). \end{aligned} \tag{7.5.16}$$

The fourth term $T_4^{\alpha,\varepsilon}$
We have, for all (t,x,ξ),

$$\begin{aligned} |\nabla((jf_+)^{\alpha,\varepsilon}\lambda)(t,x,\xi)| &= \left| \int_0^\infty \int_{B_+^d} j(y) f_+(s,y,\xi)\theta_\alpha(t-s)(\nabla\gamma_\varepsilon(x-y)\lambda(x) + \gamma_\varepsilon(x-y)\nabla\lambda(x)) \, dy \, ds \right| \\ &\leqslant C||\nabla\gamma_\varepsilon||_{L^1(\mathbb{R}^d)} + C||\nabla\lambda||_{L^\infty(\mathbb{R}^d)} \leqslant \frac{C}{\overline{\varepsilon}} + \frac{C}{\varepsilon_d} + C \leqslant \frac{C}{\overline{\varepsilon}} + \frac{C}{\varepsilon_d} \end{aligned}$$

(recall that $\varepsilon_d \leqslant 1$). Hence, by (7.4.3),

$$T_4^{\alpha,\varepsilon} \leqslant \eta \int_0^T \int_{B_+^d} |Z|(t,x) \left(\sup_\xi |\nabla((jf_+)^{\alpha,\varepsilon}\lambda)|(t,x,\xi) \right) dt dx \leqslant \frac{C\eta}{\overline{\varepsilon}} + \frac{C\eta}{\varepsilon_d}. \tag{7.5.17}$$

To sum up, gathering (7.5.10), (7.5.14), (7.5.15), (7.5.16) and (7.5.17), we have proved so far that

$$\begin{aligned} &\int_{B_+^d} \lambda(x)(u(T,x) - v(T,x))^+ \, dx \\ &\qquad \leqslant C\left(\overline{\varepsilon} + \varepsilon_d + \alpha + \frac{\eta}{\overline{\varepsilon}} + \frac{\eta}{\varepsilon_d}\right) + C\int_0^T \int_{B_+^d} (u(t,x) - v(t,x))^+ \, dx \, dt + T_5^{\alpha,\varepsilon}. \end{aligned} \tag{7.5.18}$$

The aim of the following section is to estimate $T_5^{\alpha,\varepsilon}$. Using boundary layers arguments (see the introduction), we give in subsection 7.8.1 of the appendix an insight of the reason why this term can be bounded. However, this is only an insight: since we also want to consider irregular solutions to (7.1.1), we cannot in general estimate $T_5^{\alpha,\varepsilon}$ using boundary layers analysis.

7.6 Estimate for the boundary term

This estimate is made in several steps. First, using the BLN condition, we introduce f_+^b and give an upper bound $\overline{T_5}^{\alpha,\varepsilon}$ to $T_5^{\alpha,\varepsilon}$. Then, we want to see $(Ha)_d$ in $\overline{T_5}^{\alpha,\varepsilon}$, in order to express $\overline{T_5}^{\alpha,\varepsilon}$ as a part of the interior term in (7.4.2); to this end, we use Lemma 7.4.1. Finally, we must regularize the function f_+^b introduced above in order that $\overline{T_5}^{\alpha,\varepsilon}$ appears in (7.4.2) for some *regular* Ψ. The resulting term $S^{\alpha,\varepsilon}$ is then estimated.

7.6.1 Introduction of f_+^b

We have

$$T_5^{\alpha,\varepsilon} = \int_0^\infty \int_{B^{d-1}} \int_{\mathbb{R}_\xi} l(\overline{y})(-a(\xi)\cdot n(\overline{y}))f_+^\tau(s,\overline{y},\xi)\Psi(s,\overline{y},\xi)\,d\xi\,d\overline{y}\,ds$$

where

$$\Psi(s,\overline{y},\xi) = \int_0^T \int_{B_+^d} \theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\theta_{\varepsilon_d}(x_d)(-j(x)g_-(t,x,\xi))\lambda(x)\,dx\,dt.$$

As $(-g_-)$, Ψ is non-negative and non-decreasing with respect to ξ. Thus, (7.4.4) implies

$$\begin{aligned} T_5^{\alpha,\varepsilon} &\leqslant \overline{T}_5^{\alpha,\varepsilon} := \int_0^\infty \int_{B^{d-1}} \int_{\mathbb{R}_\xi} l(\overline{y})(-a(\xi)\cdot n(\overline{y}))f_+^\tau(s,\overline{y},\xi)f_+^b(s,\overline{y},\xi)\Psi(s,\overline{y},\xi)\,d\xi\,d\overline{y}\,ds \quad (7.6.1)\\ &= \int_0^T \int_{B_+^d} \int_{\mathbb{R}_\xi} j(x)g_-(t,x,\xi)\Phi_0(t,x,\xi)\,d\xi\,dt\,dx \end{aligned}$$

with

$$\Phi_0(t,x,\xi) = \theta_{\varepsilon_d}(x_d)\int_0^\infty \int_{B^{d-1}} \lambda(x)l(\overline{y})(a(\xi)\cdot n(\overline{y}))f_+^\tau(s,\overline{y},\xi)f_+^b(s,\overline{y},\xi)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds.$$

7.6.2 Apparition of Ha

By Lemma 7.4.1, we have $l(\overline{y})(a(\xi)\cdot n(\overline{y})) = -j(\overline{y})(H(\overline{y})a(\xi))_d$. Thus, if we define

$$\Phi(t,x,\xi) = \theta_{\varepsilon_d}(x_d)(-(H(x)a(\xi))_d)\int_0^\infty \int_{B^{d-1}} \lambda(\overline{y})j(\overline{y})f_+^\tau(s,\overline{y},\xi)f_+^b(s,\overline{y},\xi)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds,$$

we have

$$\begin{aligned} &|\Phi_0(t,x,\xi)-\Phi(t,x,\xi)|\\ &\leqslant \int_0^\infty \int_{B^{d-1}} j(\overline{y})|(H(\overline{y})a(\xi))_d\lambda(x) - (H(x)a(\xi))_d\lambda(\overline{y})|\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds\,\theta_{\varepsilon_d}(x_d)\\ &\leqslant C(\overline{\varepsilon}+\varepsilon_d)\theta_{\varepsilon_d}(x_d) \end{aligned}$$

(we used the fact that H and λ are Lipschitz continuous) and

$$\overline{T_5}^{\alpha,\varepsilon} \leqslant \int_0^T \int_{B_+^d} \int_{\mathbb{R}_\xi} j(x)g_-(t,x,\xi)\Phi(t,x,\xi)\,d\xi\,dx\,dt + C(\overline{\varepsilon}+\varepsilon_d). \quad (7.6.2)$$

7.6.3 Regularization of f_+^b

We now want to replace $f_+^b(s,\overline{y},\xi)$ in Φ by a regular approximation. Let $\mathrm{sgn}_{+,\delta}$ be a regular non-decreasing function, equal to 0 on $\mathbb{R}^-$, to 1 on $[\delta,\infty[$, such that $|(\mathrm{sgn}_{+,\delta})'| \leqslant C/\delta$ and $\mathrm{sgn}_{+,\delta} \to \mathrm{sgn}_+$ everywhere as $\delta \to 0$. We have, for all $(a,b,\xi)\in\mathbb{R}^3$,

$$\int_{\mathbb{R}_\xi} |\mathrm{sgn}_{+,\delta}(a-\xi) - \mathrm{sgn}_+(b-\xi)|\,d\xi \leqslant |a-b| + \delta. \quad (7.6.3)$$

Thus, defining

$$\begin{aligned}\Phi_\delta(t,x,\xi) &= \theta_{\varepsilon_d}(x_d)(-(H(x)a(\xi))_d)\\ &\quad\times\int_0^\infty\int_{B^{d-1}}\lambda(\overline{y})j(\overline{y})f_+^\tau(s,\overline{y},\xi)\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds,\end{aligned}$$

we have

$$\begin{aligned}&|\Phi(t,x,\xi)-\Phi_\delta(t,x,\xi)|\\ &\leqslant C\theta_{\varepsilon_d}(x_d)\int_0^\infty\int_{B^{d-1}}|\mathrm{sgn}_+(u_b(s,\overline{y})-\xi)-\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)|\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds\end{aligned}$$

and therefore, by (7.6.3),

$$\begin{aligned}&\left|\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}j(x)g_-(t,x,\xi)\Phi(t,x,\xi)\,d\xi\,dx\,dt-\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}j(x)g_-(t,x,\xi)\Phi_\delta(t,x,\xi)\,d\xi\,dx\,dt\right|\\ &\leqslant C\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}\int_0^\infty\int_{B^{d-1}}|\mathrm{sgn}_+(u_b(s,\overline{y})-\xi)-\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)|\\ &\qquad\qquad\times\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\theta_{\varepsilon_d}(x_d)\,d\overline{y}\,ds\,d\xi\,dx\,dt\\ &\leqslant C\int_0^T\int_{B_+^d}\int_0^\infty\int_{B^{d-1}}(|u_b(s,\overline{y})-u_b(t,\overline{x})|+\delta)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds\,\theta_{\varepsilon_d}(x_d)\,dx\,dt\\ &\leqslant C(\overline{\varepsilon}+\alpha+\delta)\int_0^T\int_{B_+^d}\int_0^\infty\int_{B^{d-1}}\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\theta_{\varepsilon_d}(x_d)\,d\overline{y}\,ds\,dx\,dt\\ &\leqslant C(\overline{\varepsilon}+\alpha+\delta)\end{aligned}$$

(we used the fact that u_b is Lipschitz continuous). We deduce from this last inequality and (7.6.2) that

$$\begin{aligned}&\overline{T_5}^{\alpha,\varepsilon}\\ &\leqslant\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}j(x)g_-(t,x,\xi)\Phi_\delta(t,x,\xi)\,d\xi\,dx\,dt+C(\overline{\varepsilon}+\varepsilon_d+\alpha+\delta)\\ &\leqslant\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}j(x)g_-(t,x,\xi)(H(x)a(\xi))_d\\ &\quad\times\left[\int_0^\infty\int_{B^{d-1}}\lambda(\overline{y})j(\overline{y})f_+^\tau(s,\overline{y},\xi)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds\right]\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)(-\theta_{\varepsilon_d}(x_d))\,d\xi\,dx\,dt\\ &\quad+C(\overline{\varepsilon}+\varepsilon_d+\alpha+\delta)\end{aligned}$$

Let $\Theta_{\varepsilon_d}(x_d)=\int_0^{x_d}\theta_{\varepsilon_d}(r)\,dr$ (we have $0\leqslant\Theta_{\varepsilon_d}\leqslant 1$ and $\Theta_{\varepsilon_d}=1$ on $[\varepsilon_d,+\infty[$),

$$\Gamma(t,x,\xi)=\left[\int_0^\infty\int_{B^{d-1}}\lambda(\overline{y})j(\overline{y})f_+^\tau(s,\overline{y},\xi)\theta_\alpha(t-s)\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})\,d\overline{y}\,ds\right]\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)(1-\Theta_{\varepsilon_d}(x_d))$$

and

$$S^{\alpha,\varepsilon}=\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi}j(x)g_-(t,x,\xi)(H(x)a(\xi))_d\partial_{x_d}\Gamma(t,x,\xi)\,d\xi\,dx\,dt.$$

The last estimate on $\overline{T_5}^{\alpha,\varepsilon}$ can be re-written

$$\overline{T_5}^{\alpha,\varepsilon}\leqslant S^{\alpha,\varepsilon}+C(\overline{\varepsilon}+\varepsilon_d+\alpha+\delta).\tag{7.6.4}$$

7.6.4 Estimate of $S^{\alpha,\varepsilon}$ and conclusion concerning the boundary term

The functions $f_+^\tau(s,\overline{y},\xi)$ and $\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)$ are non-negative and non-increasing with respect to ξ. Since $1-\Theta_{\varepsilon_d}\geqslant 0$, Γ is non-increasing with respect to ξ; it is also regular in (t,x). Moreover, we can see that $t\to\int_{B_+^d}\int_{\mathbb{R}_\xi}(jg_-)(t,x,\xi)\Gamma(t,x,\xi)\,d\xi\,dx$ is continuous (this is slightly more difficult to write than the continuity of (7.3.4), but similar). Hence, using $\Gamma(t,x,\xi)w_\beta(t)$ as a test function in (7.4.2), where $w_\beta(t)=\int_{t-T}^\infty\theta_\beta(s)\,ds$, and letting $\beta\to 0$ (then w_β converges to the characteristic function of $]-\infty,T]$ and w_β' converges to $-\delta_T$), we find

$$\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} jg_-(\partial_t\Gamma+(Ha)_{1\dots d-1}\cdot\nabla_{\overline{x}}\Gamma+(Ha)_d\partial_{x_d}\Gamma)+\int_{B_+^d}\int_{\mathbb{R}_\xi} jf_-^0\Gamma^{(t=0)}$$
$$-\int_{B_+^d}\int_{\mathbb{R}_\xi}(jg_-)^{(t=T)}\Gamma^{(t=T)}+\int_0^T\int_{B^{d-1}}\int_{\mathbb{R}_\xi} l(-a\cdot n)f_-^b\overline{\Gamma}+\int_0^T\int_{B^{d-1}}\int_{\mathbb{R}_\xi} lD\delta_{u_b}\overline{\Gamma}$$
$$+\eta\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} Z\delta_v\cdot\nabla\Gamma\leqslant 0, \quad (7.6.5)$$

where we have denoted $(Ha)_{1\dots d-1}$ the vector of $\mathbb{R}^{d-1}$ made of the $d-1$ first coordinates of Ha. But, since $\theta_\alpha(-s)=0$ for $s\geqslant 0$, we have $\Gamma^{(t=0)}=0$. Moreover, $f_-^b(t,\overline{x},\xi)\mathrm{sgn}_{+,\delta}(u_b(t,\overline{x})-\xi)=0$, so that $f_-^b\overline{\Gamma}\equiv 0$. We also have

$$\int_{\mathbb{R}_\xi}\delta_{u_b(t,\overline{x})}\overline{\Gamma}(t,\overline{x},\xi)=\Gamma(t,\overline{x},0,u_b(t,\overline{x}))=0.$$

Hence, in (7.6.5), the second, fourth and fifth terms are null and we deduce

$$\begin{aligned} S^{\alpha,\varepsilon} &\leqslant -\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} jg_-(\partial_t\Gamma+(Ha)_{1\dots d-1}\cdot\nabla_{\overline{x}}\Gamma)+\int_{B_+^d}\int_{\mathbb{R}_\xi}(jg_-)^{(t=T)}\Gamma^{(t=T)}\\ &\quad -\eta\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} Z\delta_v\cdot\nabla\Gamma. \end{aligned} \quad (7.6.6)$$

We have $\Gamma\geqslant 0$ and $g_-\leqslant 0$, so that

$$\int_{B_+^d}\int_{\mathbb{R}_\xi}(jg_-)^{(t=T)}\Gamma^{(t=T)}\leqslant 0. \quad (7.6.7)$$

We have

$$\begin{aligned} |\partial_t\Gamma(t,x,\xi)| &\leqslant C\left(\int_0^\infty|\theta_\alpha'(t-s)|\,ds+(\mathrm{sgn}_{+,\delta})'(u_b(t,\overline{x})-\xi)|\partial_t u_b(t,\overline{x})|\right)(1-\Theta_{\varepsilon_d}(x_d))\\ &\leqslant \left(\frac{C}{\alpha}+C(\mathrm{sgn}_{+,\delta})'(u_b(t,\overline{x})-\xi)\right)(1-\Theta_{\varepsilon_d}(x_d)). \end{aligned}$$

Since $\int_{\mathbb{R}_\xi}(\mathrm{sgn}_{+,\delta})'(a-\xi)\,d\xi=1$ for all $a\in\mathbb{R}$, this implies

$$\begin{aligned} -\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} jg_-\partial_t\Gamma &\leqslant \int_0^T\int_{B_+^d}\left(\frac{C}{\alpha}+C\int_{\mathbb{R}_\xi}(\mathrm{sgn}_{+,\delta})'(u_b(t,\overline{x})-\xi)\,d\xi\right)(1-\Theta_{\varepsilon_d}(x_d))\,dxdt\\ &\leqslant \left(\frac{C\varepsilon_d}{\alpha}+C\varepsilon_d\right) \end{aligned} \quad (7.6.8)$$

(indeed, $\int_0^\infty(1-\Theta_{\varepsilon_d}(x_d))\,dx_d\leqslant\varepsilon_d$ since $(1-\Theta_{\varepsilon_d}(x_d))=0$ for $x_d\geqslant\varepsilon_d$ and $0\leqslant 1-\Theta_{\varepsilon_d}\leqslant 1$).

In the same way,

$$\begin{aligned}|\nabla_{\overline{x}}\Gamma(t,x,\xi)| &\leqslant C\left(\int_{B^{d-1}}|\nabla_{\overline{x}}\overline{\gamma}_{\overline{\varepsilon}}(\overline{x}-\overline{y})|\,d\overline{y}+(\mathrm{sgn}_{+,\delta})'(u_b(t,\overline{x})-\xi)|\nabla_{\overline{x}}u_b(t,\overline{x})|\right)(1-\Theta_{\varepsilon_d}(x_d))\\ &\leqslant \left(\frac{C}{\overline{\varepsilon}}+C(\mathrm{sgn}_{+,\delta})'(u_b(t,\overline{x})-\xi)\right)(1-\Theta_{\varepsilon_d}(x_d)),\end{aligned} \tag{7.6.9}$$

and

$$-\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} jg_-(Ha)_{1\dots d-1}\cdot\nabla_{\overline{x}}\Gamma\leqslant\left(\frac{C\varepsilon_d}{\overline{\varepsilon}}+C\varepsilon_d\right). \tag{7.6.10}$$

Inequality (7.6.9) and the definition of $\mathrm{sgn}_{+,\delta}$ show that, for all (t,x,ξ),

$$|\nabla_{\overline{x}}\Gamma(t,x,\xi)|\leqslant\left(\frac{C}{\overline{\varepsilon}}+\frac{C}{\delta}\right).$$

Moreover,

$$|\partial_{x_d}\Gamma(t,x,\xi)|\leqslant C\theta_{\varepsilon_d}(x_d)\leqslant\frac{C}{\varepsilon_d}.$$

Hence, for all (t,x,ξ), $|\nabla\Gamma(t,x,\xi)|\leqslant\frac{C}{\overline{\varepsilon}}+\frac{C}{\delta}+\frac{C}{\varepsilon_d}$ and, Z being bounded in $\mathsf{L}^1((0,T)\times B_+^d)$,

$$-\eta\int_0^T\int_{B_+^d}\int_{\mathbb{R}_\xi} Z\delta_v\cdot\nabla\Gamma\leq\frac{C\eta}{\overline{\varepsilon}}+\frac{C\eta}{\delta}+\frac{C\eta}{\varepsilon_d}. \tag{7.6.11}$$

Gathering (7.6.7), (7.6.8), (7.6.10) and (7.6.11) in (7.6.6), we obtain

$$S^{\alpha,\varepsilon}\leq C\left(\varepsilon_d+\frac{\varepsilon_d}{\alpha}+\frac{\varepsilon_d}{\overline{\varepsilon}}+\frac{\eta}{\overline{\varepsilon}}+\frac{\eta}{\delta}+\frac{\eta}{\varepsilon_d}\right).$$

which gives, thanks to (7.6.1) and (7.6.4),

$$T_5^{\alpha,\varepsilon}\leq C\left(\overline{\varepsilon}+\varepsilon_d+\alpha+\delta+\frac{\varepsilon_d}{\alpha}+\frac{\varepsilon_d}{\overline{\varepsilon}}+\frac{\eta}{\overline{\varepsilon}}+\frac{\eta}{\delta}+\frac{\eta}{\varepsilon_d}\right). \tag{7.6.12}$$

7.7 Conclusion

We now sum up and conclude.

Combining (7.5.18) and (7.6.12) (recall that the estimates in Sections 7.5 and 7.6 concern, in fact, $\tilde{u}$ and $\tilde{v}$ — *i.e.* u and v transported), we find

$$\begin{aligned}\int_{B_+^d}\tilde{\lambda}(x)(\tilde{u}(T,x)-\tilde{v}(T,x))^+\,dx &\leqslant C\left(\overline{\varepsilon}+\varepsilon_d+\alpha+\delta+\frac{\varepsilon_d}{\alpha}+\frac{\varepsilon_d}{\overline{\varepsilon}}+\frac{\eta}{\overline{\varepsilon}}+\frac{\eta}{\varepsilon_d}+\frac{\eta}{\delta}\right)\\ &\quad+C\int_0^T\int_{B_+^d}(\tilde{u}(t,x)-\tilde{v}(t,x))^+\,dx\,dt.\end{aligned}$$

Minimizing on δ, α, $\overline{\varepsilon}$ and ε_d, we notice that an optimal choice of these parameters is $\delta=\eta^{1/2}$, $\varepsilon_d=\eta^{2/3}$, $\alpha=\overline{\varepsilon}=\eta^{1/3}$; we then re-transport this estimate on $\Omega\cap O$:

$$\int_{\Omega\cap O}\lambda(x)(u(T,x)-v(T,x))^+\,dx\leqslant C\eta^{1/3}+C\int_0^T\int_{\Omega\cap O}(u(t,x)-v(t,x))^+\,dx\,dt.$$

Summing on the local charts (recall that in the preceding inequality $\lambda = \lambda_i$ and $O = O_i$ for any $i \in \{1, \dots, n\}$) and using (7.3.6), we deduce

$$\int_\Omega (u(T,x) - v(T,x))^+ \, dx \leqslant C\eta^{1/3} + C \int_0^T \int_\Omega (u(t,x) - v(t,x))^+ \, dx \, dt.$$

This inequality being true for all $T \in [0, T_0]$, Gronwall's lemma applied to the continuous function $T \to \int_\Omega (u(T,x) - v(T,x))^+ \, dx$ ensures that there exists $C > 0$ such that, for all $T \in [0, T_0]$,

$$\int_\Omega (u(T,x) - v(T,x))^+ \, dx \leqslant C\eta^{1/3}. \tag{7.7.1}$$

Now, since u satisfies (7.2.1)—(7.2.2) for f_-, we see that $-u$ satisfies these equations for f_+ with $-u_0$, $-u_b$, s_*m and $s_*m_-^b$ instead of u_0, u_b, m and m_+^b (where s is the symmetry with respect to ξ). Similarly, $-v$ satisfies (7.2.3) for g_- with $-u_0$, $-u_b$ and s_*q instead of u_0, u_b and q. Hence, (7.7.1) applied to $-u$ and $-v$ gives

$$\int_\Omega (-u(T,x) + v(T,x))^+ \, dx = \int_\Omega (u(T,x) - v(T,x))^- \, dx \leqslant C\eta^{1/3}.$$

which concludes the proof of Theorem 7.1.1.

7.8 Appendix

7.8.1 Estimate of $T_5^{\alpha,\varepsilon}$ using boundary layers

If the solution u is regular, then $T_5^{\alpha,\varepsilon}$ can be estimated using boundary layer techniques. This is what we briefly explain here.

To simplify the exposition, we take $\Omega =]0, \infty[$ and recall some basic facts concerning boundary layers: if u is regular, then the parabolic approximation admits the decomposition $v(t,x) = u(t,x) + c(t, x/\eta^\gamma) + r^\eta(t,x)$, where $\gamma = 1/2$ or 1 depending if the boundary is characteristic or not, and r^η is a remainder (small, with respect to η, in L^∞ norm). Fix $t \in (0,T)$, set $w(y) = u(t,0) + c(t,y)$, $w_0 = u_b(t)$ and $w_\infty = u(t,0)$. Then, by properties of the layer c, w satisfies

$$\dot{w}(y) = A(w(y)) - A(w_\infty), \tag{7.8.1}$$

$$w(0) = w_0, \tag{7.8.2}$$

$$w(+\infty) = w_\infty. \tag{7.8.3}$$

Notice that, since (7.8.1) is an autonomous o.d.e., $\dot{w}$ vanishes on $[0, +\infty)$ if, and only if, w is constant (and then $w_0 = w_\infty$). Now, suppose $w_0 \neq w_\infty$. Then, since $\dot{w}$ does not vanish, it has a constant sign, which is actually the sign of $w_\infty - w_0$ since w is an orbit from w_0 to w_∞. To sum up, we have $\mathrm{sgn}(w_\infty - w_0)\dot{w}(y) \geq 0$ for all $y \geq 0$. In view of (7.8.1), this is equivalent to $\mathrm{sgn}(w_\infty - w_0)(A(w(y)) - A(w_\infty)) \geq 0$ for all $y \geq 0$ or still, since w is a bijection $[0, +\infty) \to [w_0, w_\infty)$,

$$\forall \kappa \in [w_0, w_\infty] \, , \; \mathrm{sgn}(w_\infty - w_0)(A(\kappa) - A(w_\infty)) \geq 0 \, . \tag{7.8.4}$$

Conversely, one can check that (7.8.4) is a sufficient condition to the existence of a solution to (7.8.1)-(7.8.2)-(7.8.3). Now, replacing w_0 and w_∞ by their respective values $u_b(t)$ and $u(t,0)$, (7.8.4) appears to be nothing but the BLN condition

$$\forall k \in [u_b(t), u(t,0)] \, , \; -\mathrm{sgn}(u(t,0) - u_b(t))(A(u(t,0)) - A(k)) \geq 0 \, .$$

In other words, the BLN condition is a necessary and sufficient condition to the existence of the boundary layer function c.

Let us now come back to the estimate of $T_5^{\alpha,\varepsilon}$: assuming that $\alpha = 0$ and $\lambda \equiv 1$, $T_5^{\alpha,\varepsilon}$ reduces in this setting to

$$\begin{aligned}\tilde{T}_5^{\varepsilon} &= \int_0^T \int_0^\infty \int_{\mathbb{R}_\xi} a(\xi)\mathrm{sgn}_+(u(t,0)-\xi)(-\mathrm{sgn}_-(v(t,x)-\xi))\theta_\varepsilon(x)\,d\xi\,dx\,dt \\ &= \int_0^T \int_0^\infty \mathrm{sgn}_+(u(t,0)-v(t,x))(A(u(t,0))-A(v(t,x)))\theta_\varepsilon(x)\,dx\,dt.\end{aligned}$$

Since $\zeta \to \mathrm{sgn}_+(u(t,0)-\zeta)(A(u(t,0))-A(\zeta))$ is Lipschitz continuous, $\tilde{T}_5^{\varepsilon}$ can be assimilated, up to an error of order $\eta+\varepsilon$, to

$$\int_0^T \int_0^\infty \mathrm{sgn}_+(-c(t,x/\eta^\gamma))(A(u(t,0))-A(u(t,0)+c(t,x/\eta^\gamma))\theta_\varepsilon(x)\,dx\,dt\,.$$

Since w is monotonous between w_0 and w_∞, $w_\infty - w(y)$ has the same sign than $w_\infty - w_0$. Reporting this result in (7.8.4) and replacing w, w_0 and w_∞ by $u(t,0)+c(t,y)$, $u_b(t)$ and $u(t,0)$ respectively we get

$$\mathrm{sgn}(-c(t,y))(A(u(t,0))-A(u(t,0)+c(t,y)) \le 0$$

for all $y \ge 0$, which shows that, up to an error of order $\eta+\varepsilon$, $T_5^{\alpha,\varepsilon}$ is nonpositive.

The basic idea in Section 7.6 is thus to compare $T_5^{\alpha,\varepsilon}$ to some nonpositive quantity, which is done as early as Subsection 7.6.1.

7.8.2 Technical results

The first lemma is classical, we do not prove it.

Lemma 7.8.1 *Let U be a bounded open set of $\mathbb{R}^d$ with Lipschitz continuous boundary and γ_ε be a regularizing kernel with support contained in the ball of radius $|\varepsilon|$. Then there exists C only depending on U such that, for all $w \in \mathsf{L}^1(U) \cap \mathsf{BV}(U)$,*

$$\int_U \int_U |w(x)-w(y)|\gamma_\varepsilon(x-y)\,dx\,dy \leqslant C|\varepsilon|(||w||_{\mathsf{L}^1(U)} + |w|_{\mathsf{BV}(U)}).$$

The second lemma is a technical result used in Section 7.5.

Lemma 7.8.2 *Let $D > 0$ and K be a compact subset of B^d. We take $\nu \in \mathbb{R}^d$ such that $|\nu| < \mathrm{dist}(K, \mathbb{R}^d \backslash B^d)$ and j a regular function on B^d. If $w \in \mathsf{BV}(B_+^d)$ then there exists C not depending on ν or w such that*

$$\int_K \int_{-D}^D |\nabla(j(x)\mathrm{sgn}_-(w(x)-\xi))^\nu|\,d\xi\,dx \leqslant C(1+|w|_{\mathsf{BV}(B_+^d)}).$$

Proof of Lemma 7.8.2
The proof is made in several steps. Let U be an open set relatively compact in B^d, containing K and such that $|\nu| \leqslant \mathrm{dist}(U, \mathbb{R}^d \backslash B^d)$. We prove the result of the lemma with U instead of K (the introduction of this open set is useful because we use classical results concerning BV functions on *open* sets).

Step 0: (a preliminary result) Let $r \in \mathsf{W}^{1,1}(B^d_+)$ and denote by R the extension of r to B^d by 0 outside B^d_+. Then $R \in \mathsf{BV}(B^d)$ and $|R|_{\mathsf{BV}(B^d)} \leqslant C||r||_{\mathsf{W}^{1,1}(B^d_+)}$. To see this, take $\phi \in (\mathsf{C}^\infty_\mathsf{c}(B^d))^d$; thanks to an integrate by parts, we have

$$\int_{B^d} R\mathrm{div}(\phi) = \int_{B^d_+} r\mathrm{div}(\phi) = -\int_{B^{d-1}} r\phi_d - \int_{B^d_+} \phi \cdot \nabla r.$$

The right-hand side of this equation is bounded by $C||r||_{\mathsf{W}^{1,1}(B^d_+)}||\phi||_\infty$, which proves the result (in fact, the preceding equation computes the gradient of R).

Step 1: let $\mathrm{sgn}_{-,\delta} : \mathbb{R} \to \mathbb{R}$ be a regular nondecreasing function, equal to 0 on $\mathbb{R}^+$, to -1 on $]-\infty, -\delta]$ and such that $\mathrm{sgn}_{-,\delta} \to \mathrm{sgn}_-$ as $\delta \to 0$. We prove the result when $w \in \mathsf{W}^{1,1}(B^d_+)$ and sgn_- is replaced by $\mathrm{sgn}_{-,\delta}$, with C not depending on δ.
We clearly have (since $\mathrm{sgn}_{-,\delta}$ is regular) $j\mathrm{sgn}_{-,\delta}(w-\xi) \in \mathsf{W}^{1,1}(B^d_+)$ and

$$\nabla(j\mathrm{sgn}_{-,\delta}(w-\xi)) = \nabla j\ \mathrm{sgn}_{-,\delta}(w-\xi) + j(\mathrm{sgn}_{-,\delta})'(w-\xi)\nabla w.$$

By Step 0, the extension W_ξ of $j\mathrm{sgn}_{-,\delta}(w-\xi)$ to B^d by 0 outside B^d_+ is in $\mathsf{BV}(B^d)$ and

$$|W_\xi|_{\mathsf{BV}(B^d)} \leqslant C + C||(\mathrm{sgn}_{-,\delta})'(w-\xi)\nabla w||_{L^1(B^d_+)},$$

where C does not depend on δ nor w.
Moreover, by choice of ν, $(j(\cdot)\mathrm{sgn}_{-,\delta}(w(\cdot)-\xi))^\nu = W_\xi \star \gamma_\nu$ and $\nabla(j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu = \nabla W_\xi \star \gamma_\nu$ on U. Thus,

$$||\nabla(j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu||_{L^1(U)} \leqslant |W_\xi|_{\mathsf{BV}(B^d)} \leqslant C + C\int_{B^d_+} (\mathrm{sgn}_{-,\delta})'(w-\xi)|\nabla w|$$

We now integrate with respect to ξ and use $\int_{-D}^{D}(\mathrm{sgn}_{-,\delta})'(s-\xi)\,d\xi \leqslant \int_{\mathbb{R}_\xi}(\mathrm{sgn}_{-,\delta})'(s-\xi)\,d\xi = 1$ for all $s \in \mathbb{R}$ to find

$$\int_U \int_{-D}^{D} |\nabla(j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu| \leqslant C + C\int_{B^d_+} |\nabla w|$$

which concludes this step.

Step 2: conclusion.
There exists $w_n \in W^{1,1}(B^d_+)$ which converge to w in $L^1(B^d_+)$ and such that $|w_n|_{\mathsf{BV}(B^d_+)} \to |w|_{\mathsf{BV}(B^d_+)}$.
Since $\mathrm{sgn}_{-,\delta}$ is regular, $j\mathrm{sgn}_{-,\delta}(w_n-\xi) \to j\mathrm{sgn}_{-,\delta}(w-\xi)$ in $\mathsf{L}^1(B^d_+)$ as $n \to \infty$ so that $(j\mathrm{sgn}_{-,\delta}(w_n-\xi))^\nu \to (j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu$ in $L^1(\mathbb{R}^d)$ as $n \to \infty$. Moreover, $\mathrm{sgn}_{-,\delta}(w-\xi) \to \mathrm{sgn}_-(w-\xi)$ in $L^1(B^d_+)$ as $\delta \to 0$ so that $(j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu \to (j\mathrm{sgn}_-(w-\xi))^\nu$ in $\mathsf{L}^1(\mathbb{R}^d)$ as $\delta \to 0$.
We deduce that

$$\begin{aligned}
\int_U |\nabla(j\mathrm{sgn}_-(w-\xi))^\nu| &= |(j\mathrm{sgn}_-(w-\xi))^\nu|_{\mathsf{BV}(U)} \\
&\leqslant \liminf_{\delta\to 0} |(j\mathrm{sgn}_{-,\delta}(w-\xi))^\nu|_{\mathsf{BV}(U)} \\
&\leqslant \liminf_{\delta\to 0}\liminf_{n\to\infty} |(j\mathrm{sgn}_{-,\delta}(w_n-\xi))^\nu|_{\mathsf{BV}(U)} \\
&= \liminf_{\delta\to 0}\liminf_{n\to\infty} \int_U |\nabla(j\mathrm{sgn}_{-,\delta}(w_n-\xi))^\nu|.
\end{aligned}$$

Integrating on $\xi \in [-D, D]$ and using Fatou's Lemma, the result of Step 1 gives

$$\begin{aligned}
\int_U \int_{-D}^{D} |\nabla (j\mathrm{sgn}_-(w-\xi))^\nu| &\leqslant \liminf_{\delta \to 0} \liminf_{n \to \infty} \int_{-D}^{D} \int_U |\nabla (j\mathrm{sgn}_{-,\delta}(w_n - \xi))^\nu| \\
&\leqslant \liminf_{\delta \to 0} \liminf_{n \to \infty} \left(C + C \int_{B_+^d} |\nabla w_n(x)| \, dx \right) \\
&\leqslant \liminf_{\delta \to 0} (C + C|w|_{\mathsf{BV}(B_+^d)}) = C + C|w|_{\mathsf{BV}(B_+^d)}
\end{aligned}$$

by choice of $(w_n)_{n \geqslant 1}$. ■

Partie V

Diffusion non-locale

Chapitre 8

Global solution and smoothing effect for a non-local regularization of an hyperbolic equation

Reference: J. Droniou, T. Gallouët and J. Vovelle. *J. Evol. Equ.* **3** (2003), no. 3, 499-521.

8.1 Introduction

We study the problem

$$\begin{cases} \partial_t u(t,x) + \partial_x(f(u))(t,x) + g[u(t,\cdot)](x) = 0 & t \in]0,\infty[\,,\ x \in \mathbb{R} \\ u(0,x) = u_0(x) & x \in \mathbb{R}, \end{cases} \tag{8.1.1}$$

where $f \in C^\infty(\mathbb{R})$ is such that $f(0) = 0$ (there is not loss of generality in assuming this), $u_0 \in L^\infty(\mathbb{R})$ and g is the non-local (in general) operator defined through the Fourier transform by

$$\mathcal{F}(g[u(t,\cdot)])(\xi) = |\xi|^\lambda \mathcal{F}(u(t,\cdot))(\xi)\,, \quad \text{with } \lambda \in]1,2].$$

Remark 8.1.1 *We could also very well study a multi-dimensional scalar equation, that is to say on $\mathbb{R}^N$ instead of $\mathbb{R}$. All the methods and results presented below would apply; but this would lead to more technical manipulations so, for the sake of clarity, we have chosen to fully describe only the mono-dimensional case.*

The interest of such an equation (namely Equation (8.1.1)) was pointed out to us by Paul Clavin in the context of pattern formation in detonation waves. The study of detonations leads, in a first approximation, to nonlinear hyperbolic equations. As it is well known, the solutions of such equations may develop discontinuities in finite time. A theory of existence and uniqueness of (entropy weak) solutions to Equation (8.1.1) with $g = 0$, in the L^∞ framework, is known since the work of Krushkov ([60], see also [80]). The case of a parabolic regularization (of a nonlinear hyperbolic equation) is often considered and used to prove the Krushkov result; it corresponds to (8.1.1) with $\lambda = 2$. In this case, existence and uniqueness of a solution is also well known along with a regularizing effect. However, it appears that the choice of $\lambda = 2$ is not quite natural, at least for the problem of detonation (see [24], [25], [23]) where it seems more natural to consider a nonlocal term as $g[u]$ with λ close to 1 but greater than 1 (although the case $\lambda = 1$ is also of interest but more complicated). This term corresponds to some spatial fractional derivative

of u of order λ. The main motivation of this paper is therefore to prove existence and uniqueness of the solution to (8.1.1) in the L^∞ framework as well as a regularizing effect (a regularizing effect which is well-known in the case $\lambda = 2$, as it is said above). In particular, the solution will be C^∞ in space and time for $t > 0$. We also prove the so called "maximum principle", namely the fact that the solution takes values between the maximum and the minimum values of the initial data, and a property of "L^1 contraction" on the solutions, which is the fact that, for any time, the L^1 norm of the difference of two solutions with different initial data is bounded by the L^1 norm (if it exists) of the difference of the initial data.
A major difficulty is due to the nonlocal character of $g[u]$ if $\lambda \in]1, 2[$; this prevents the classical way to prove the maximum principle, which leads to an L^∞ *a priori* bound on the solution (a crucial estimate to obtain global solutions). It is interesting to notice that the hypothesis $\lambda \leq 2$ is necessary for the maximum principle. Indeed, the maximum principle is no longer true in general for $\lambda > 2$. However, the regularizing effect is still true for $\lambda > 2$, a property which is probably not verified if $\lambda < 1$. The case $\lambda = 1$ is not so clear and needs an additional work. Indeed, for the study of detonation waves, our result has to be viewed as a preliminary result or, at least, as a study of a very simplified case. Realistic models are much more complicated. In particular, it seems that λ is actually depending on the unknown and, even if $\lambda > 1$, λ is probably not bounded from below by some $\lambda_0 > 1$. The possibility to generalize our result to such a case is not manifest.
We first prove (Section 4) the uniqueness of a "weak" solution (solution in the sense of Definition 8.3.1 below). Then, assuming the existence of a "weak" solution, we prove (Section 5) the regularizing effect (the equation is then satisfied in a classical sense). The results of these two sections are in fact true for any $\lambda > 1$. In Section 6, the existence result is given, using a splitting method. The use of splitting methods is classical, in particular to define numerical schemes, but is not usual to prove an existence result as it is done here. In this section, the central argument is the proof of the maximum principle (which is limited to $\lambda \leq 2$).

Here is our main result.

Theorem 8.1.1 *If $u_0 \in L^\infty(\mathbb{R})$, then there exists a unique solution u to (8.1.1) on $]0, \infty[$ (in the sense of Definition 8.3.1, see below). Moreover, this solution satisfies:*

i) $u \in C^\infty(]0, \infty[\times\mathbb{R})$ and all its derivatives are bounded on $]t_0, \infty[\times\mathbb{R}$ for all $t_0 > 0$,

ii) for all $t > 0$, $||u(t)||_{L^\infty(\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})}$ and, in fact, u takes its values between the essential lower and upper bounds of u_0,

iii) u satisfies $\partial_t u + \partial_x(f(u)) + g[u] = 0$ in the classical sense ($g[u]$ being properly defined by Proposition 8.5.2).

iv) $u(t) \to u_0$, as $t \to 0$, in $L^\infty(\mathbb{R})$ weak-$*$ and in $L^p_{\rm loc}(\mathbb{R})$ for all $p \in [1, \infty[$.

Remark 8.1.2 *In the course of our study of (8.1.1), we will also see that, if $u_0 \in L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$, then the solution u to (8.1.1) satisfies, for all $t > 0$: $||u(t)||_{L^1(\mathbb{R})} \leq ||u_0||_{L^1(\mathbb{R})}$.*
We will also see that (8.1.1) has a L^1 contraction property: if $(u_0, v_0) \in L^\infty(\mathbb{R})$ are such that $u_0 - v_0 \in L^1(\mathbb{R})$, then, denoting by u and v the solutions to (8.1.1) corresponding to initial conditions u_0 and v_0, we have, for all $t > 0$: $||u(t) - v(t)||_{L^1(\mathbb{R})} \leq ||u_0 - v_0||_{L^1(\mathbb{R})}$.

8.2 Properties of the kernel of g

Using the Fourier transform, we see that the semi-group generated by g is formally given by the convolution with the kernel (defined for $t > 0$ and $x \in \mathbb{R}$)

$$K(t,x) = \mathcal{F}^{-1}\left(e^{-t|\cdot|^\lambda}\right)(x) = \int_{\mathbb{R}} e^{2i\pi x\xi} e^{-t|\xi|^\lambda}\, d\xi = \mathcal{F}\left(e^{-t|\cdot|^\lambda}\right)(x).$$

The function $\xi \in \mathbb{R} \to e^{-t|\xi|^\lambda}$ being real-valued and even, K is real-valued (in the sequel, we consider only real-valued solutions to (8.1.1)).
The most important property of K is its nonnegativity. For the sake of completeness, we give here a sketch of the proof of this result, but notice that it is a well-known result since a rather long time now. We refer to the work of Lévy for example [65]. Also notice that we study the question of the non-negativity of the kernel K because it is the issue at stake in the analysis of a maximum principle for the equation $u_t + g[u] = 0$. From this point of view, we shall make reference to the work of Courrège and coworkers (see [16] and references therein) who give a characterization of a large class of pseudo-differential operators satisfying the positive maximum principle and also, more recently, to the work of Farkas, Jacob, Schilling [43] (see also Hoh [57]).

Lemma 8.2.1 *If $\lambda \in]0,2]$ then, for all $(t,x) \in]0,\infty[\times\mathbb{R}$, we have $K(t,x) \geq 0$.*

Proof of Lemma 8.2.1
If $\lambda = 2$, it is well-known that $K(t,x) = (\pi/t)^{1/2} e^{-\frac{\pi^2}{t}\xi^2}$, which implies the result. Assume now that $\lambda \in]0,2[$ and let $f(x) = A|x|^{-1-\lambda}\mathbf{1}_{\mathbb{R}\setminus]-1,1[}(x)$, with $A > 0$ such that $\int_{\mathbb{R}} f(x)\,dx = 1$. Since f is even with integral equal to 1, we have

$$\mathcal{F}(f)(\xi) = 1 + \int_{\mathbb{R}} (\cos(2\pi x\xi) - 1) f(x)\,dx = 1 + A|\xi|^\lambda \int_{|y|\geq|\xi|} \frac{\cos(2\pi y) - 1}{|y|^{1+\lambda}}\,dy.$$

Since $\cos(2\pi y) - 1 = \mathcal{O}(|y|^2)$ on the neighborhood of 0 and $\lambda < 2$, the dominated convergence theorem gives

$$\int_{|y|\geq|\xi|} \frac{\cos(2\pi y) - 1}{|y|^{1+\lambda}}\,dy \to I := \int_{\mathbb{R}} \frac{\cos(2\pi y) - 1}{|y|^{1+\lambda}}\,dy < 0 \quad \text{as } \xi \to 0.$$

Hence, $\mathcal{F}(f)(\xi) = 1 - c|\xi|^\lambda(1 + \omega(\xi))$ with $c = -AI > 0$ and $\lim_{\xi\to 0}\omega(\xi) = 0$.
Define $f_n(x) = n^{1/\lambda} f * f * \cdots * f(n^{1/\lambda}x)$, the convolution product being taken n times. By the properties of the Fourier transform with respect to the convolution product, we have, for all $\xi \in \mathbb{R}$,

$$\mathcal{F}(f_n)(\xi) = \left(\mathcal{F}(f)(n^{-1/\lambda}\xi)\right)^n = \left(1 - cn^{-1}|\xi|^\lambda(1 + \omega(n^{-1/\lambda}\xi))\right)^n \to e^{-c|\xi|^\lambda}$$

as $n \to \infty$. Since $(\mathcal{F}(f_n))_{n\geq 1}$ is bounded by 1 (the L^1-norm of f_n for all $n \geq 1$), this convergence is also true in $\mathcal{S}'(\mathbb{R})$ and, taking the inverse Fourier transform, we see that $f_n \to \mathcal{F}^{-1}(e^{-c|\cdot|^\lambda}) = K(c,\cdot)$ in $\mathcal{S}'(\mathbb{R})$ as $n \to \infty$. f_n being nonnegative for all n, we deduce that K is nonnegative on $\{c\}\times\mathbb{R}$; the homogeneity property (8.2.1) below concludes then the proof of the lemma. ■

Here are some other important properties of K:

$$\forall (t,x) \in]0,\infty[\times\mathbb{R}, \quad K(t,x) = \frac{1}{t^{1/\lambda}} K\left(1, \frac{x}{t^{1/\lambda}}\right). \tag{8.2.1}$$

$$\begin{gathered} K \text{ is } C^\infty \text{ on }]0,\infty[\times\mathbb{R} \text{ and, for all } m \geq 0, \text{ there exists } B_m \text{ such that} \\ \forall (t,x) \in]0,\infty[\times\mathbb{R}, \quad |\partial_x^m K(t,x)| \leq \frac{1}{t^{(1+m)/\lambda}} \frac{B_m}{(1 + t^{-2/\lambda}|x|^2)}. \end{gathered} \tag{8.2.2}$$

$$\begin{gathered} (K(t,\cdot))_{t>0} \text{ is, as } t \to 0, \text{ an approximate unit} \\ \text{(in particular, } ||K(t,\cdot)||_{L^1(\mathbb{R})} = 1 \text{ for all } t > 0). \end{gathered} \tag{8.2.3}$$

$$\exists \mathcal{K}_1 \text{ such that, for all } t > 0, \quad ||\partial_x K(t,\cdot)||_{L^1(\mathbb{R})} = \mathcal{K}_1 t^{-1/\lambda}. \tag{8.2.4}$$

$$\begin{gathered} \forall (a,b) \in]0,\infty[, \ K(a,\cdot) * K(b,\cdot) = K(a+b,\cdot) \\ \text{and } K(a,\cdot) * \partial_x K(b,\cdot) = \partial_x K(a+b,\cdot). \end{gathered} \tag{8.2.5}$$

Proof of these properties
Equation (8.2.1) is obtained thanks to the change of variable $\xi = t^{-1/\lambda}\eta$ in the integral defining K.

The regularity of K is an immediate application of the theorem of derivation under the integral sign. To prove the second part of (8.2.2), we write $\partial_x^m K(1,x) = \int_{\mathbb{R}} (2i\pi\xi)^m e^{-|\xi|^\lambda} e^{2i\pi x\xi}\, d\xi$; since $\lambda > 1$, the first two derivatives of $\xi \to \xi^m e^{-|\xi|^\lambda}$ are integrable on $\mathbb{R}$ and we can make two integrations by parts to obtain $\partial_x^m K(1,x) = \mathcal{O}(1/x^2)$ on $\mathbb{R}$; $\partial_x^m K(1,\cdot)$ being bounded on $\mathbb{R}$, we deduce the estimate of (8.2.2) for $t = 1$; the general case $t > 0$ comes from the case $t = 1$ and (8.2.1).
Since $K(1,\cdot) \geq 0$, we have $||K(1,\cdot)||_{L^1(\mathbb{R})} = \int_{\mathbb{R}} K(1,x)\, dx = \mathcal{F}(K(1,\cdot))(0) = e^{-|0|^\lambda} = 1$ and (8.2.3) is thus a consequence of (8.2.1).
The estimate (8.2.4) comes from the derivation of (8.2.1) and from the change of variable $y = t^{-1/\lambda}x$ in the computation of $||\partial_x K(1,\cdot/t^{1/\lambda})||_{L^1(\mathbb{R})}$.
The identity (8.2.5), which translates the fact that the convolution with $K(t)$ is the semi-group generated by g, can be directly checked via Fourier transform. ■

Let us also give some continuity results related to K.

Lemma 8.2.2 *i) If $u_0 \in L^1(\mathbb{R})$, then $t \in [0,\infty[\to K(t,\cdot) * u_0$ is continuous $[0,\infty[\to L^1(\mathbb{R})$ (with value u_0 at $t = 0$).*

ii) Let $T > 0$ and $(t_0, x_0) \in]0,T[\times\mathbb{R}$. If $v \in C_b(]0,T[\times\mathbb{R})$, then

a) for all $s_0 > 0$, $K(s,\cdot) * v(t,\cdot)(x) \to K(s_0,\cdot) * v(t_0,\cdot)(x_0)$ as $s \to s_0$, $t \to t_0$ and $x \to x_0$,

b) $K(s,\cdot) * v(t,\cdot)(x) \to v(t_0,x_0)$ as $s \to 0$, $t \to t_0$ and $x \to x_0$.

All these properties are either classical results of approximate units or consequences of the estimate in (8.2.2) (with $m = 0$) and of the dominated convergence theorem. We do not give a precise proof of these results.

8.3 Definition and first properties of the solutions

The idea, to study (8.1.1), is to search for a solution to $\partial_t u + g[u] = -\partial_x(f(u))$ using Duhamel's formula: a solution to this equation is formally given by $u(t,x) = K(t) * u_0(x) - \int_0^t K(t-s) * \partial_x(f(u(s,\cdot)))(x)\, ds$. By putting the derivative of $f(u)$ on K, we are led to the following definition.

Definition 8.3.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$ or $T = \infty$. A solution to (8.1.1) on $]0,T[$ is a function $u \in L^\infty(]0,T[\times\mathbb{R})$ which satisfies, for a.e. $(t,x) \in]0,T[\times\mathbb{R}$,*

$$u(t,x) = K(t,\cdot) * u_0(x) - \int_0^t \partial_x K(t-s,\cdot) * f(u(s,\cdot))(x)\, ds. \tag{8.3.1}$$

The following proposition shows that all the terms in (8.3.1) are well-defined.

Proposition 8.3.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$. If $v \in L^\infty(]0,T[\times\mathbb{R})$, then*

$$u : (t,x) \in]0,T[\times\mathbb{R} \to K(t,\cdot) * u_0(x) + \int_0^t \partial_x K(t-s,\cdot) * v(s,\cdot)\, ds$$

defines a function in $C_b(]0,T[\times\mathbb{R})$ and we have, for all $t_0 \in]0,T[$, all $x \in \mathbb{R}$ and all $t \in]0,T-t_0[$,

$$u(t_0+t,x) = K(t,\cdot) * u(t_0,\cdot)(x) + \int_0^t \partial_x K(t-s,\cdot) * v(t_0+s,\cdot)(x)\, ds. \tag{8.3.2}$$

Proof of Proposition 8.3.1
Step 1: first term of u.
Since $u_0 \in L^\infty(\mathbb{R})$ and, for $t > 0$, $K(t,\cdot) \in L^1(\mathbb{R})$, $K(t,\cdot) * u_0$ is well-defined and, by Young's inequalities for the convolution and (8.2.3), we have

$$\forall (t,x) \in]0,\infty[\times\mathbb{R},\ |K(t,\cdot) * u_0(x)| \leq ||u_0||_{L^\infty(\mathbb{R})}. \tag{8.3.3}$$

Let $t_0 \in]0,T[$ and $x_0 \in \mathbb{R}$; for all $0 < t_0 < T < \infty$, by (8.2.2), we can write

$$\forall (t,x,y) \in]t_0,T[\times\mathbb{R}\times\mathbb{R},\quad |K(t,x-y)| \leq \frac{C_1}{C_2 + |x-y|^2} \tag{8.3.4}$$

where $C_1 > 0$ and $C_2 > 0$ only depend on (t_0,T). We have $|y-x_0|^2 \leq 2|y-x|^2 + 2|x_0-x|^2$, so that $|y-x|^2 \geq \frac{1}{2}|y-x_0|^2 - |x_0-x|^2$. For all $x \in \mathbb{R}$ such that $|x-x_0|^2 \leq C_2/2$, for all $t \in]t_0,T[$ and all $y \in \mathbb{R}$, (8.3.4) gives

$$|K(t,x-y)| \leq \frac{C_1}{C_2 + \frac{1}{2}|x_0-y|^2 - |x_0-x|^2} \leq \frac{C_1}{(C_2/2) + \frac{1}{2}|x_0-y|^2} = F(y)$$

with $F \in L^1(\mathbb{R})$. Since u_0 is bounded and K is continuous, the theorem of continuity under the integral sign gives the continuity of $(t,x) \to K(t,\cdot) * u_0(x)$.

Step 2: the second term of u.
Define $G : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and $H : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ by: for all $x \in \mathbb{R}$,

$$G(t,x) = \partial_x K(t,x)\mathbf{1}_{]0,T[}(t) \ \text{ if } t > 0\,,\quad G(t,x) = 0 \ \text{ if } t \leq 0\,,$$
$$H(t,x) = v(t,x) \ \text{ if } t \in]0,T[\,,\quad H(t,x) = 0 \ \text{ if } t \in \mathbb{R}\backslash]0,T[.$$

We notice that $G \in L^1(\mathbb{R}\times\mathbb{R})$; indeed, by Fubini-Tonelli's theorem and (8.2.4),

$$\int_{\mathbb{R}\times\mathbb{R}} |G(t,x)|\,dxdt \leq \mathcal{K}_1 \int_0^T t^{-1/\lambda}\,dt = \frac{\lambda\mathcal{K}_1}{\lambda-1} T^{1-\frac{1}{\lambda}} < \infty. \tag{8.3.5}$$

The function H is clearly in $L^\infty(\mathbb{R}\times\mathbb{R})$, being bounded by $||v||_{L^\infty(]0,T[\times\mathbb{R})}$. Thus, denoting by $\star$ the convolution in $\mathbb{R}\times\mathbb{R}$, $G \star H$ is well-defined, bounded and uniformly continuous on $\mathbb{R}\times\mathbb{R}$; moreover, by (8.3.5),

$$||G\star H||_{C_b(\mathbb{R}\times\mathbb{R})} \leq ||G||_{L^1(\mathbb{R}\times\mathbb{R})}||H||_{L^\infty(\mathbb{R}\times\mathbb{R})} \leq \frac{\lambda\mathcal{K}_1}{\lambda-1} T^{1-\frac{1}{\lambda}} ||v||_{L^\infty(]0,T[\times\mathbb{R})}. \tag{8.3.6}$$

By Fubini's theorem, one checks that, for all $t \in]0,T[$ and all $x \in \mathbb{R}$, $G\star H(t,x) = \int_0^t \partial_x K(t-s,\cdot) * v(s,\cdot)(x)\,ds$, and the second term of u is thus continuous and bounded on $]0,T[\times\mathbb{R}$.
We notice that, thanks to (8.3.3) and (8.3.6),

$$||u||_{C_b(]0,T[\times\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})} + \frac{\lambda\mathcal{K}_1}{\lambda-1} T^{1-\frac{1}{\lambda}} ||v||_{L^\infty(]0,T[\times\mathbb{R})}. \tag{8.3.7}$$

Step 3: To prove (8.3.2), we make the change of variable $\tau = t_0 + s$ in the last term of this equation, we use Fubini's theorem (thanks to (8.2.4)) to permute the convolution by $K(t,\cdot)$ and the integral sign in $u(t_0,\cdot)$ and we apply (8.2.5). ■

As an immediate consequence of this proposition, we have:

Corollary 8.3.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$ or $T = \infty$. If u is a solution to (8.1.1) on $]0,T[$, then $u \in C_b(]0,T[\times\mathbb{R})$ and u satisfies (8.3.1) for all $(t,x) \in]0,T[\times\mathbb{R}$. Moreover, for all $t_0 \in]0,T[$ and all $(t,x) \in]0,T-t_0[\times\mathbb{R}$,*

$$u(t_0+t,x) = K(t,\cdot) * u(t_0,\cdot)(x) - \int_0^t \partial_x K(t-s,\cdot) * f(u(t_0+s,\cdot))(x)\,ds, \tag{8.3.8}$$

i.e. $u(t_0+\cdot,\cdot)$ is a solution to (8.1.1) on $]0,T-t_0[$ with $u(t_0,\cdot)$ instead of u_0.

To conclude this study of the first properties of the solutions, we prove item iv) of Theorem 8.1.1.
Proof of item iv) in Theorem 8.1.1
Suppose that u is a solution to (8.1.1) on $]0,T[$. Since $f(u)$ is bounded, we have, for all $(t,x) \in]0,T[\times\mathbb{R}$, by (8.2.4),

$$\left| \int_0^t \partial_x K(t-s,\cdot) * f(u(s,\cdot))(x)\, ds \right| \leq \mathcal{K}_1 ||f(u)||_\infty \int_0^t \frac{1}{(t-s)^{1/\lambda}}\, ds = C t^{1-\frac{1}{\lambda}}$$

where C does not depend on t; hence, the last term of (8.3.1) tends to 0 in $L^\infty(\mathbb{R})$ as $t \to 0$. By classical properties of the approximate units, the first term in the right-hand side of (8.3.1) converges as wanted to u_0 and the proof is complete. ■

8.4 Uniqueness of the solution

Theorem 8.4.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$ or $T = \infty$. There exists at most one solution to (8.1.1) on $]0,T[$ in the sense of Definition 8.3.1.*

Proof of Theorem 8.4.1
Step 1: we first prove a local uniqueness result. Denote by $\mathrm{Lip}_R(f)$ a lipschitz constant of f on $[-R,R]$. Let $T_1 > 0$. For all u and v solutions to (8.1.1) on $]0,T_1[$ bounded by R, by (8.3.1) and (8.2.4), we have

$$|u(t,x) - v(t,x)| \leq \frac{\lambda \mathcal{K}_1}{\lambda - 1} T_1^{1-\frac{1}{\lambda}} \mathrm{Lip}_R(f) ||u-v||_\infty = k(T_1,R)||u-v||_\infty.$$

There exists $T_0 > 0$ only depending on R such that, if $T_1 \leq T_0$, we have $k(T_1,R) < 1$; for $T_1 \leq T_0$, there exists therefore at most one solution to (8.1.1) on $]0,T_1[$ bounded by R.

Step 2: proof of the uniqueness result.
Let u and v be two solutions to (8.1.1) on $]0,T[$. Take $R = \max(||u||_\infty, ||v||_\infty)$; let T_0 be given by Step 1 for R. By Step 1, since u and v are bounded by R, $u = v$ on $]0, \inf(T,T_0)[\times\mathbb{R}$.
Let $T' = \sup\{t \in]0,T[\mid u = v \text{ on }]0,t[\times\mathbb{R}\} \geq \inf(T,T_0)$, and suppose that $T' < T$. By definition of T', and since u and v are continuous on $]0,T[\times\mathbb{R}$, we have $u(T',\cdot) = v(T',\cdot)$ on $\mathbb{R}$. By Corollary 8.3.1, $u(T'+\cdot,\cdot)$ and $v(T'+\cdot,\cdot)$ are two solutions to (8.1.1) on $]0,T-T'[$ with the same initial condition $u(T',\cdot) = v(T',\cdot)$. These solutions being bounded by R, Step 1 shows that $u(T'+\cdot,\cdot) = v(T'+\cdot,\cdot)$ on $]0, \inf(T_0, T-T')[\times\mathbb{R}$, which is a contradiction with the definition of T'. ■

8.5 Regularizing effect

8.5.1 Spatial regularity

If we formally differentiate (8.3.1) with respect to x, we see that the spatial derivatives of u satisfy integral equations; the following theorem gives some properties on these integral equations.

Proposition 8.5.1 *Let $M > 0$ and $F : (t,x,\zeta) \in]0,M[\times\mathbb{R}\times\mathbb{R} \to F(t,x,\zeta) \in \mathbb{R}$ be continuous; we suppose that $\partial_x F$, $\partial_\zeta F$, $\partial_\zeta \partial_x F$ and $\partial_\zeta \partial_\zeta F$ exist and are continuous on $]0,M[\times\mathbb{R}\times\mathbb{R}$; we also suppose that there exists $\omega :]0,\infty[\to \mathbb{R}^+$ such that, for all $L > 0$, F and these derivatives are bounded on $]0,M[\times\mathbb{R}\times[-L,L]$ by $\omega(L)$.*
Let $R_0 > 0$ and $R = (2+\mathcal{K}_1)R_0$. Then there exists $T_0 > 0$ only depending on (R_0,ω) such that, if $T = \inf(M,T_0)$ and $v_0 \in L^\infty(\mathbb{R})$ satisfies $||v_0||_{L^\infty(\mathbb{R})} \leq R_0$, there exists a unique $v \in C_b(]0,T[\times\mathbb{R})$ bounded by R and such that

$$v(t,x) = K(t,\cdot) * v_0(x) + \int_0^t \partial_x K(t-s,\cdot) * F(s,\cdot,v(s,\cdot))(x)\, ds. \tag{8.5.1}$$

Moreover, $\partial_x v \in C(]0,T[\times\mathbb{R})$ *and, for all* $a \in]0,T[$, $||\partial_x v||_{C_b(]a,T[\times\mathbb{R})} \leq Ra^{-1/\lambda}$.

Proof of Proposition 8.5.1
The idea is to use a fixed point theorem. Let, for $T \in]0,M[$, $E_T = \{v \in C_b(]0,T[\times\mathbb{R}) \mid \partial_x v \in C(]0,T[\times\mathbb{R})$ and $t^{1/\lambda}\partial_x v \in C_b(]0,T[\times\mathbb{R})\}$, endowed with its natural norm $||v||_{E_T} = ||v||_\infty + ||t^{1/\lambda}\partial_x v||_\infty$. We define, thanks to Proposition 8.3.1, $\Psi_T : C_b(]0,T[\times\mathbb{R}) \to C_b(]0,T[\times\mathbb{R})$ by

$$\Psi_T(v)(t,x) = K(t,\cdot) * v_0(x) + \int_0^t \partial_x K(t-s,\cdot) * F(s,\cdot,v(s,\cdot))(x)\, ds.$$

Step 1: the first term of $\Psi_T(v)$ belongs to E_T.
The estimate (8.2.2) allows to see, as in Step 1 of the proof of Proposition 8.3.1, that, by derivation and continuity under the integral sign, $K(t,\cdot) * v_0$ is derivable on $\mathbb{R}$ and that $(t,x) \in]0,T[\times\mathbb{R} \to \partial_x(K(t,\cdot) * v_0)(x) = \partial_x K(t,\cdot) * v_0(x)$ is continuous. By Young's inequalities and (8.2.4),

$$||\partial_x(K(t,\cdot) * v_0)||_{C_b(\mathbb{R})} \leq \mathcal{K}_1 t^{-1/\lambda} ||v_0||_{L^\infty(\mathbb{R})}, \tag{8.5.2}$$

which proves that $(t,x) \in]0,T[\times\mathbb{R} \to K(t,\cdot) * v_0(x)$ belongs to E_T.

Step 2: we prove that, if $v \in E_T$, the second term of $\Psi_T(v)$ belongs to E_T.
Define $H(t,x) = \int_0^t \partial_x K(t-s,\cdot) * F(s,\cdot,v(s,\cdot))(x)\, ds$. Let $t \in]0,T[$ and $s \in]0,t[$. The function $x \in \mathbb{R} \to F(s,x,v(s,x))$ is in $C_b^1(\mathbb{R})$. We can thus differentiate under the integral sign to see that $\partial_x K(t-s,\cdot) * F(s,\cdot,v(s,\cdot))$ is C^1 with derivative $\partial_x K(t-s,\cdot) * (\partial_x F(s,\cdot,v(s,\cdot)) + \partial_\zeta F(s,\cdot,v(s,\cdot))\partial_x v(s,\cdot))$. Moreover, for all $x \in \mathbb{R}$,

$$\begin{aligned} &|\partial_x K(t-s,\cdot) * (\partial_x F(s,\cdot,v(s,\cdot)) + \partial_\zeta F(s,\cdot,v(s,\cdot))\partial_x v(s,\cdot))(x)| \\ &\leq \frac{\mathcal{K}_1 ||\partial_x F(\cdot,\cdot,v(\cdot,\cdot))||_\infty}{(t-s)^{1/\lambda}} + \frac{\mathcal{K}_1 ||\partial_\zeta F(\cdot,\cdot,v(\cdot,\cdot))||_\infty ||v||_{E_T}}{s^{1/\lambda}(t-s)^{1/\lambda}}. \end{aligned} \tag{8.5.3}$$

This last function is integrable with respect to $s \in]0,t[$, and we can thus apply the theorem of derivation under the integral sign to see that

$$\begin{aligned} &\partial_x H(t,x) \\ &= \int_0^t \partial_x K(t-s,\cdot) * (\partial_x F(s,\cdot,v(s,\cdot)) + \partial_\zeta F(s,\cdot,v(s,\cdot))\partial_x v(s,\cdot))(x)\, ds. \end{aligned} \tag{8.5.4}$$

If $\partial_x v$ was bounded, the continuity of $\partial_x H$ would be a consequence of Proposition 8.3.1. We thus approximate $\partial_x v$ by bounded functions to conclude. Take $0 < \delta < T$ and define $w_\delta \in L^\infty(]0,T[\times\mathbb{R})$ by

$$w_\delta(t,x) = \partial_x F(t,x,v(t,x)) + \partial_\zeta F(t,x,v(t,x))\partial_x v(t,x)\mathbf{1}_{[\delta,T[}(t).$$

Denoting $A_\delta(t,x) = \int_0^t \partial_x K(t-s,\cdot) * w_\delta(s,\cdot)(x)\, ds$, (8.5.4) allows to see that, for all $t_0 \in]0,T[$, $A_\delta \to \partial_x H$ uniformly on $[t_0,T[\times\mathbb{R}$ as $\delta \to 0$; since, by Proposition 8.3.1, A_δ is continuous on $]0,T[\times\mathbb{R}$, we deduce that $\partial_x H$ is continuous on $]0,T[\times\mathbb{R}$. Moreover, by (8.5.4) and (8.5.3) and the change of variable $s = t\tau$ in the integrals on $]0,t[$, we have, for all $(t,x) \in]0,T[\times\mathbb{R}$,

$$\begin{aligned} &|\partial_x H(t,x)| \\ &\leq C_0\mathcal{K}_1 \left(||\partial_x F(\cdot,\cdot,v(\cdot,\cdot))||_\infty t^{1-\frac{1}{\lambda}} + ||\partial_\zeta F(\cdot,\cdot,v(\cdot,\cdot))||_\infty ||v||_{E_T} t^{1-\frac{2}{\lambda}} \right) \end{aligned} \tag{8.5.5}$$

where $C_0 = \max(\int_0^1 (1-\tau)^{-1/\lambda}\, d\tau, \int_0^1 \tau^{-1/\lambda}(1-\tau)^{-1/\lambda}\, d\tau)$, which proves that $H \in E_T$.
If v is bounded by R, the properties of F along with (8.5.2), (8.5.5) and (8.3.7), give

$$\begin{aligned} ||\Psi_T(v)||_{E_T} \leq\ & ||v_0||_{L^\infty(\mathbb{R})} + \frac{\lambda\mathcal{K}_1}{\lambda-1} T^{1-\frac{1}{\lambda}} \omega(R) \\ & + \mathcal{K}_1 ||v_0||_{L^\infty(\mathbb{R})} + C_0\mathcal{K}_1\omega(R)\left(T + T^{1-\frac{1}{\lambda}}||v||_{E_T}\right). \end{aligned} \tag{8.5.6}$$

Step 3: fixed point.
We take, as in the proposition, $R = (2 + \mathcal{K}_1)R_0$ and we denote, for $T > 0$, $B_T(R)$ the closed ball in E_T of center 0 and radius R. Let $T_0 > 0$ be such that

$$R_0 + \frac{\lambda \mathcal{K}_1}{\lambda - 1} T_0^{1-\frac{1}{\lambda}} \omega(R) + \mathcal{K}_1 R_0 + C_0 \mathcal{K}_1 \omega(R) \left(T_0 + T_0^{1-\frac{1}{\lambda}} R \right) \leq R \tag{8.5.7}$$

$$\mathcal{K}_1 \omega(R) \left(\frac{\lambda}{\lambda - 1} T_0^{1-\frac{1}{\lambda}} + \frac{\lambda}{\lambda - 1} T_0 + C_0 R T_0^{1-\frac{1}{\lambda}} + C_0 T_0^{1-\frac{1}{\lambda}} \right) < 1 \tag{8.5.8}$$

(by definition of R, such a T_0 exists and only depends on (R_0, ω)).

Let $T = \inf(M, T_0)$. Take $v_0 \in L^\infty(\mathbb{R})$ bounded by R_0. Thanks to (8.5.6) and (8.5.7), Ψ_T sends $B_T(R)$ into $B_T(R)$. Let $(u, v) \in B_T(R)$. u and v are bounded by R and we have thus, for all $(t, x) \in]0, T[\times\mathbb{R}$, by (8.2.4) and the properties of F,

$$|\Psi_T(u)(t,x) - \Psi_T(v)(t,x)| \leq \mathcal{K}_1 \frac{\lambda}{\lambda - 1} T^{1-\frac{1}{\lambda}} \omega(R) ||u - v||_\infty. \tag{8.5.9}$$

By (8.5.4) and the properties of F, we also have, for all $(t, x) \in]0, T[\times\mathbb{R}$,

$$\begin{aligned} & t^{1/\lambda} |\partial_x \Psi_T(u)(t,x) - \partial_x \Psi_T(v)(t,x)| \\ \leq \quad & \mathcal{K}_1 \omega(R) \left(\frac{\lambda}{\lambda - 1} T + C_0 T^{1-\frac{1}{\lambda}} ||u||_{E_T} + C_0 T^{1-\frac{1}{\lambda}} \right) ||u - v||_{E_T}. \end{aligned} \tag{8.5.10}$$

The properties (8.5.9), (8.5.10) and (8.5.8) ensure that Ψ_T is contracting on $B_T(R)$. Therefore, Ψ_T has a unique fixed point v in $B_T(R)$; v is a continuous and bounded solution to (8.5.1) such that $\partial_x v$ exists and is continuous on $]0, T[\times\mathbb{R}$. Moreover, since $v \in B_T(R)$, we have, for all $a \in]0, T[$ and all $(t, x) \in]a, T[\times\mathbb{R}$, $|\partial_x v(t,x)| \leq t^{-1/\lambda} ||v||_{E_T} \leq a^{-1/\lambda} R$, which is the estimate on $\partial_x v$ stated in the proposition.
The inequalities (8.5.9) and (8.5.8) ensure that Ψ_T is contracting on the ball in $C_b(]0, T[\times\mathbb{R})$ of center 0 and radius R. Thus, Ψ_T can have only one fixed point in this ball, which is the uniqueness result of the proposition. ■

Theorem 8.5.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$ or $T = \infty$. If u is a solution to (8.1.1) on $]0, T[$ in the sense of Definition 8.3.1, then u is indefinitely derivable with respect to x. Moreover, for all $n \geq 0$ and all $t_0 \in]0, T[$, we have*

i) $\partial_x^n u \in C_b(]t_0, T[\times\mathbb{R})$,

ii) for all $t \in]0, T - t_0[$,

$$\partial_x^n u(t_0 + t, \cdot) = K(t, \cdot) * \partial_x^n u(t_0, \cdot) - \int_0^t \partial_x K(t - s, \cdot) * \partial_x^n (f(u(t_0 + s, \cdot)))\, ds$$

iii) if $R \geq ||u||_{C_b(]0,T[\times\mathbb{R})}$, there exists C only depending on (R, t_0, n) such that $||\partial_x^n u||_{C_b(]t_0,T[\times\mathbb{R})} \leq C$.

Proof of Theorem 8.5.1
We prove, by induction on n, that : u has spatial derivatives of order up to n which are continuous and bounded by $C(R, t_0, n)$ on $]t_0, T[\times\mathbb{R}$ for all $t_0 \in]0, T[$, item ii) is satisfied on $]0, T - t_0[\times\mathbb{R}$ for all $t_0 \in]0, T[$ and

$$\partial_x^n (f(u)) = U_n + (1 - \delta_{n,0}) f'(u) \partial_x^n u + \delta_{n,0} f(u), \tag{8.5.11}$$

where $\delta_{n,0}$ is Krönecker's symbol, $U_0 = 0$ and, if $n \geq 1$, $U_n = G_n((\partial_x^k u)_{k \leq n-1})$ with G_n regular.

The validity of the property at the rank $n = 0$ is a consequence of Corollary 8.3.1. We suppose the induction hypothesis true up to a rank $n \geq 0$, and we prove it for the rank $n+1$.
Let $b_0 \in]0,T[$. Take $b \in]b_0,T[$ and define $F : (t,x,\zeta) :]0,T-b[\times\mathbb{R}\times\mathbb{R} \to -U_n(b+t,x)-(1-\delta_{n,0})f'(u)(b+t,x)\zeta - \delta_{n,0}f(\zeta)$. The function F satisfies the hypotheses of Proposition 8.5.1, with (by induction hypothesis) ω only depending on (R,b_0,n); we also have $||\partial_x^n u||_{C_b(]b_0,T[\times\mathbb{R})} \leq R_0$ where R_0 only depends on (R,b_0,n). Let T_0 only depending on (R_0,ω) (i.e. on (R,b_0,n)) be given by Proposition 8.5.1.
By induction hypothesis, $\partial_x^n u(b+\cdot,\cdot)$ is continuous and bounded by $R_0 \leq (2+\mathcal{K}_1)R_0$ and satisfies (8.5.1) on $]0,T-b[\times\mathbb{R}$ for the preceding F and with $v_0 = \partial_x^n u(b,\cdot)$ bounded by R_0. Proposition 8.5.1 shows thus that $\partial_x^{n+1}u$ exists and is continuous and bounded by $(2+\mathcal{K}_1)R_0 a^{-1/\lambda}$ on $]b+a,\inf(T,b+T_0)[\times\mathbb{R}$; this is true for all $b \in]b_0,T[$ and all $a \in]0,\inf(T-b,T_0)[$. Since T_0 does not depend on a or b, taking $t_0 \in]0,T[$, $b_0 = t_0/2$ and $a = \inf(t_0/2,T_0/2) < T-b_0$, we notice that the intervals $\{]b+a,\inf(T,b+T_0)[\,,\ b \in]b_0,T-a[\}$ cover $]b_0+a,T[\supset]t_0,T[$ and we deduce that $\partial_x^{n+1}u$ has the regularity and satisfies the estimates we wanted to obtain.
Let us prove the formula for $\partial_x^{n+1}u$. By induction hypothesis,

$$\partial_x^n u(t_0+t,\cdot) = K(t,\cdot) * \partial_x^n u(t_0,\cdot) - \int_0^t \partial_x K(t-s,\cdot) * \partial_x^n(f(u(t_0+\cdot)))\,ds. \tag{8.5.12}$$

But we have just proved that $\partial_x^n u(t_0,\cdot) \in C_b^1(\mathbb{R})$; thus, we can write

$$\partial_x(K(t,\cdot) * \partial_x^n u(t_0,\cdot)) = K(t,\cdot) * \partial_x^{n+1}u(t_0,\cdot). \tag{8.5.13}$$

The function $(t,x) \in]0,T-t_0[\times\mathbb{R} \to \partial_x^n(f(u))(t_0+t,x)$ and its first spatial derivative are continuous and bounded on $]0,T-t_0[\times\mathbb{R}$. The reasoning of Step 2 in the proof of Proposition 8.5.1 (with $\partial_x^n(f(u))(t_0+\cdot,\cdot)$ instead of $F(\cdot,\cdot,v(\cdot,\cdot))$) allows to compute the spatial derivative of the last term in (8.5.12) by derivation under the integral sign, and, thanks to (8.5.13), proves item ii) for $\partial_x^{n+1}u$.
Property (8.5.11) for the derivative of order $n+1$ simply comes from the derivation of this formula at rank n, and the induction is complete. ■

8.5.2 Temporal regularity

Preliminary: about the definition of g

The operator g has been formally defined by $\mathcal{F}(g[v])(\xi) = |\xi|^\lambda \mathcal{F}(v)(\xi)$; it can be shown that this definition makes sense for bounded functions, but we will not need it and we prefer to give here a simple formula for $g[v]$ which defines this operator on $C_b^\infty(\mathbb{R})$.

Proposition 8.5.2 *There exists $(g_1,g_2) \in (L^1(\mathbb{R}))^2$ such that, for all $v \in \mathcal{S}(\mathbb{R})$, $g[v] = g_1 * v + g_2 * v^{(4)}$. This formula allows thus to define $g[v]$ for $v \in C_b^\infty(\mathbb{R})$ (and this definition does not depend on the choice of g_1 and g_2 as above).*

Proof of Proposition 8.5.2
Let $\chi \in C_c^\infty(\mathbb{R})$ be even and equal to 1 on a neighborhood of 0. By linearity of $\mathcal{F}^{-1}$, if $v \in \mathcal{S}(\mathbb{R})$,

$$g[v] = \mathcal{F}^{-1}(|\cdot|^\lambda \chi \mathcal{F}(v)) + \mathcal{F}^{-1}(|\cdot|^\lambda(1-\chi)\mathcal{F}(v)) \tag{8.5.14}$$

(since $\mathcal{F}(v) \in \mathcal{S}(\mathbb{R})$, all these terms are well-defined as inverse Fourier transforms of integrable functions).
Let $h_1 : \xi \in \mathbb{R} \to |\xi|^\lambda \chi(\xi)$. The function h_1 is C^1 on $\mathbb{R}$, C^2 outside 0 and its first two derivatives are integrable on $\mathbb{R}$. We deduce, as in the proof of (8.2.2), that $\mathcal{F}^{-1}(h_1)(x) = \mathcal{O}(1/(1+|x|^2))$ on $\mathbb{R}$. Hence, $\mathcal{F}^{-1}(h_1) \in L^1(\mathbb{R})$ and we can write $\mathcal{F}(\mathcal{F}^{-1}(h_1) * v) = h_1 \mathcal{F}(v)$, that is to say $\mathcal{F}^{-1}(h_1\mathcal{F}(v)) = \mathcal{F}^{-1}(h_1) * v$.
Let $h_2 : \xi \in \mathbb{R} \to |\xi|^\lambda(1-\chi(\xi))$; the function $h_2^\star : \xi \in \mathbb{R} \to (2i\pi\xi)^{-4}h_2(\xi)$ is C^∞ and all its derivatives are integrable on $\mathbb{R}$ (the p-th derivative of $h_2^\star$ behaves, on a neighborhood of the infinity, as $|\xi|^{-4-p+\lambda}$ and $-4+\lambda < -1$ since $\lambda \leq 2$); thus, $\mathcal{F}^{-1}(h_2^\star) \in L^1(\mathbb{R})$ and

$$\mathcal{F}(\mathcal{F}^{-1}(h_2^\star) * v^{(4)})(\xi) = h_2^\star(\xi)\mathcal{F}(v^{(4)})(\xi) = (2i\pi\xi)^4 h_2^\star(\xi)\mathcal{F}(v)(\xi) = h_2(\xi)\mathcal{F}(v)(\xi),$$

that is to say $\mathcal{F}^{-1}(h_2\mathcal{F}(v)) = \mathcal{F}^{-1}(h_2^\star) * v^{(4)}$.
Identity (8.5.14) gives therefore $g[v] = g_1 * v + g_2 * v^{(4)}$, where $g_1 = \mathcal{F}^{-1}(h_1)$ and $g_2 = \mathcal{F}^{-1}(h_2^\star)$ are integrable on $\mathbb{R}$ (notice also that, since χ is even, h_1 and $h_2^\star$ are also even and real-valued, so that g_1 and g_2 are real-valued).
To prove that, if $v \in C_b^\infty(\mathbb{R})$, the definition of $g[v]$ by this formula does not depend on the choice of g_1 and g_2, we approximate v and its derivatives by functions in $C_c^\infty(\mathbb{R})$; this will be of no use to us in the sequel, so we do not detail this step. ■

The following proposition is quite natural, since K is the kernel associated to g. But the reasoning followed to obtain K was formal, so we must prove this result.

Proposition 8.5.3 *If $v \in C_b^\infty(\mathbb{R})$ then, for all $x \in \mathbb{R}$, $t \in]0,\infty[\to K(t,\cdot) * v(x)$ is C^1. Moreover, for all $t > 0$ and all $x \in \mathbb{R}$, we have $\frac{d}{dt}(K(\cdot,\cdot) * v(x))(t) = -g[K(t,\cdot) * v](x)$.*

Proof of Proposition 8.5.3
We notice that $g[K(t,\cdot) * v]$ makes sense since $K(t,\cdot) * v \in C_b^\infty(\mathbb{R})$.
Suppose first that $v \in \mathcal{S}(\mathbb{R})$. The functions $K(t,\cdot)$ and v are integrable on $\mathbb{R}$, so, by definition of K, $K(t,\cdot) * v = \mathcal{F}^{-1}(e^{-t|\cdot|^\lambda}\mathcal{F}(v))$. A derivation under the integral sign shows that $t \in]0,\infty[\to K(t,\cdot) * v(x)$ is C^1 and that, for all $t > 0$ and all $x \in \mathbb{R}$,

$$\frac{d}{dt}(K(\cdot,\cdot) * v(x))(t) = -\mathcal{F}^{-1}(|\cdot|^\lambda e^{-t|\cdot|^\lambda}\mathcal{F}(v))(x). \tag{8.5.15}$$

Taking g_1 and g_2 as in the proof of Proposition 8.5.2, we can check, since $K(t,\cdot)$ and all the derivatives of v are integrable, that $\mathcal{F}(g[K(t,\cdot) * v]) = |\cdot|^\lambda e^{-t|\cdot|^\lambda}\mathcal{F}(v)$, that is to say $g[K(t,\cdot) * v] = \mathcal{F}^{-1}(|\cdot|^\lambda e^{-t|\cdot|^\lambda}\mathcal{F}(v))$ and (8.5.15) concludes the proof if $v \in \mathcal{S}(\mathbb{R})$

Take now $v \in C_b^\infty(\mathbb{R})$. We can find a sequence $(v_n)_{n\geq 1} \in \mathcal{S}(\mathbb{R})$ whose derivatives are bounded in $L^\infty(\mathbb{R})$ and converge to the corresponding derivatives of v.
Let $x \in \mathbb{R}$; define $F_n : t \in]0,\infty[\to K(t,\cdot) * v_n(x)$ and $F : t \in]0,\infty[\to K(t,\cdot) * v(x)$. By the convergence of $(v_n)_{n\geq 1}$ and the dominated convergence theorem, we see that $(F_n)_{n\geq 1}$ converges to F on $]0,\infty[$ and is bounded in $L^\infty(]0,\infty[)$. Therefore, the convergence is also true in the sense of the distributions on $]0,\infty[$ and we have $F_n' \to F'$ in $\mathcal{D}'(]0,\infty[)$.
But, since $v_n \in \mathcal{S}(\mathbb{R})$, we have seen that F_n is C^1 and that $F_n'(t) = -g[K(t,\cdot) * v_n](x) = -g_1 * K(t,\cdot) * v_n(x) - g_2 * K(t,\cdot) * v_n^{(4)}(x)$. Hence, $(F_n')_{n\geq 1}$ converges to $-g_1 * K(t,\cdot) * v(x) - g_2 * K(t,\cdot) * v^{(4)}(x) = -g[K(t,\cdot) * v](x)$ and is bounded in $L^\infty(]0,\infty[)$, which proves that $F_n' \to -g[K(t,\cdot) * v](x)$ in $\mathcal{D}'(]0,\infty[)$. Identifying the limits of the derivatives of F_n, we find $F'(t) = -g[K(t,\cdot) * v](x)$ in $\mathcal{D}'(]0,\infty[)$; since $t \in]0,\infty[\to g[K(t,\cdot) * v](x) = g_1 * K(t,\cdot) * v(x) + g_2 * K(t,\cdot) * v^{(4)}(x)$ is continuous (Proposition 8.3.1 with $u_0 = g_1 * v$ or $u_0 = g_2 * v^{(4)}$), we deduce that $F : t \in]0,\infty[\to K(t,\cdot) * v(x)$ is in fact C^1 on $]0,\infty[$, which concludes the proof. ■

Proof of the temporal regularity

Lemma 8.5.1 *Let $u_0 \in L^\infty(\mathbb{R})$ and $T > 0$ or $T = \infty$. If u is a solution to (8.1.1) on $]0,T[$ in the sense of Definition 8.3.1, then u is derivable with respect to t and $\partial_t u + \partial_x(f(u)) + g[u] = 0$ on $]0,T[\times\mathbb{R}$.*

Proof of Lemma 8.5.1
We can suppose that T is finite. Let $t_0 > 0$, $t \in]t_0,T[$ and $s \in]0,t[$. Using (8.3.4) (and an equivalent estimate for $\partial_x K$, obtained thanks to (8.2.2)), since $f(u(t_0+s,\cdot))$ is in $C_b^1(\mathbb{R})$, we see that $\partial_x K(t-s,\cdot) * f(u(t_0+s,\cdot)) = K(t-s,\cdot) * \partial_x(f(u(t_0+s,\cdot)))$.
Defining $v : (t,x) \in]0,T-t_0[\times\mathbb{R} \to -\partial_x(f(u))(t_0+t,x) \in \mathbb{R}$ (which is continuous and bounded, and has all its spatial derivatives continuous and bounded — see Theorem 8.5.1), we write, by (8.3.8),

$$u(t_0+t,x) = K(t,\cdot) * u(t_0,\cdot) + \int_0^t K(t-s,\cdot) * v(s,\cdot)(x)\,ds. \tag{8.5.16}$$

Since $u(t_0,\cdot) \in C_b^\infty(\mathbb{R})$, Proposition 8.5.3 shows that $(t,x) \in]0,T-t_0[\times\mathbb{R} \to K(t,\cdot)*u(t_0,\cdot)(x)$ is derivable with respect to t, and has $-g[K(t,\cdot)*u(t_0,\cdot)](x)$ as derivative.

Proving the derivability of the second term of the right-hand side of (8.5.16) is more troublesome (because $K(t-s,\cdot)$ explodes as $s \to t$). Fix $x \in \mathbb{R}$ and $\delta_0 \in]0,T-t_0[$. Let $\delta \in]0,\delta_0[$ and, for $t \in]\delta_0, T-t_0[$, $H_\delta(t) = \int_0^{t-\delta} K(t-s,\cdot)*v(s,\cdot)(x)\,ds$. Also denote $H(t) = \int_0^t K(t-s,\cdot)*v(s,\cdot)(x)\,ds$. We have $|H_\delta(t) - H(t)| \leq \delta ||v||_{C_b(]0,T-t_0[\times\mathbb{R})}\,ds$, so that $H_\delta \to H$ uniformly on $]\delta_0, T-t_0[$ as $\delta \to 0$.
The function $\phi : (t,s) \in \{(t',s') \in]\delta, T-t_0[\times]0,T-t_0[\mid s' < t'-\delta/2\} \to K(t-s,\cdot)*v(s,\cdot)(x)$ is continuous (Lemma 8.2.2 ii)-a)) and bounded. By Proposition 8.5.3, ϕ is derivable with respect to t and $\partial_t\phi(t,s) = -g[K(t-s,\cdot)*v(s,\cdot)](x) = -g_1 * K(t-s,\cdot)*v(s,\cdot)(x) - g_2 * K(t-s,\cdot)*\partial_x^4 v(s,\cdot)(x)$; this formula and Lemma 8.2.2 ii)-a) show that $\partial_t\phi$ is continuous and bounded (because, by continuity under the integral sign, $(s,x) \to g_1 * v(s,\cdot)(x)$ and $(s,x) \to g_2 * \partial_x^4 v(s,\cdot)(x)$ are continuous and bounded on $]0,T-t_0[\times\mathbb{R}$). These properties allow to prove that H_δ is C^1 on $]\delta_0, T-t_0[$ and that

$$H_\delta'(t) = K(\delta,\cdot)*v(t-\delta,\cdot)(x) - \int_0^{t-\delta} g[K(t-s,\cdot)*v(s,\cdot)](x)\,ds.$$

The function $(s,x) \in]0,t[\times\mathbb{R} \to g[K(t-s,\cdot)*v(s,\cdot)](x) = g_1 * K(t-s,\cdot)*v(s,\cdot)(x) + g_2 * K(t-s,\cdot)*\partial_x^4 v(s,\cdot)(x)$ is continuous and bounded (Lemma 8.2.2 ii)-a)); thus, by Lemma 8.2.2 ii)-b), we see that H_δ' converges on $]\delta_0, T-t_0[$ to

$$F : t \in]0,T-t_0[\to v(t,x) - \int_0^t g[K(t-s,\cdot)*v(s,\cdot)](x)\,ds$$

while remaining bounded in $L^\infty(]\delta_0, T-t_0[)$. Since H_δ uniformly converges on $]\delta_0, T-t_0[$ to H, we deduce that $H' = F$ in $\mathcal{D}'(]\delta_0, T-t_0[)$. Since $g[K(t-s,\cdot)*v(s,\cdot)] = g_1 * K(t-s,\cdot)*v(s,\cdot) + g_2 * K(t-s,\cdot)*\partial_x^4 v(s,\cdot)$, the same reasoning as in Step 2 of the proof of Proposition 8.3.1 (with K instead of $\partial_x K$ and $(t,x) \to g_1 * v(t,\cdot)(x)$ or $(t,x) \to g_2 * \partial_x^4 v(t,\cdot)(x)$ instead of v) shows that F is in fact continuous. Hence, δ_0 being arbitrary, H is C^1 on $]0,T-t_0[$ and $H' = F$.

Coming back to (8.5.16), we see that $u(t_0+\cdot,\cdot)$ is derivable with respect to t on $]0,T-t_0[\times\mathbb{R}$ and that

$$\begin{aligned}\partial_t u(t_0+t,x) &= -g[K(t,\cdot)*u(t_0,\cdot)](x) - \partial_x(f(u))(t_0+t,x) \\ &\quad - \int_0^t g_1 * K(t-s,\cdot)*v(s,\cdot)(x) + g_2 * K(t-s,\cdot)*\partial_x^4 v(s,\cdot)(x)\,ds.\end{aligned} \tag{8.5.17}$$

The time t_0 being arbitrary, this gives the temporal derivability of u on $]0,T[\times\mathbb{R}$. We now prove that the right-hand side of (8.5.17) is $-\partial_x(f(u)) - g[u]$.
Let $t \in]0,T-t_0[$. By Fubini, we have

$$\begin{aligned}&\int_0^t g_1 * K(t-s,\cdot)*v(s,\cdot) + g_2 * K(t-s,\cdot)*\partial_x^4 v(s,\cdot)\,ds \\ &= g_1 * \left[\int_0^t K(t-s,\cdot)*v(s,\cdot)\,ds\right] + g_2 * \left[\int_0^t K(t-s,\cdot)*\partial_x^4 v(s,\cdot)\,ds\right].\end{aligned} \tag{8.5.18}$$

With the same reasoning as in Step 2 of the proof of Proposition 8.5.1 (with K instead of $\partial_x K$ and v instead of $F(\cdot,\cdot,v(\cdot,\cdot))$), we prove by induction that $(t,x) \in]0,T-t_0[\times\mathbb{R} \to \int_0^t K(t-s,\cdot)*v(s,\cdot)(x)\,ds$ is indefinitely derivable with respect to x, has all its spatial derivatives continuous and bounded on $]0,T-t_0[\times\mathbb{R}$ and satisfies, for all $m \geq 0$,

$$\partial_x^m \left(\int_0^t K(t-s,\cdot)*v(s,\cdot)\,ds\right) = \int_0^t K(t-s,\cdot)*\partial_x^m v(s,\cdot)\,ds.$$

Thus, by (8.5.18),

$$\begin{aligned}\int_0^t g_1 * K(t-s,\cdot) * v(s,\cdot) + g_2 * K(t-s,\cdot) * \partial_x^4 v(s,\cdot)\, ds \\ = g_1 * \left[\int_0^t K(t-s,\cdot) * v(s,\cdot)\, ds\right] + g_2 * \partial_x^4 \left[\int_0^t K(t-s,\cdot) * v(s,\cdot)\, ds\right] \\ = g\left[\int_0^t K(t-s,\cdot) * v(s,\cdot)\, ds\right].\end{aligned}$$

This equation, combined with (8.5.17) and (8.5.16), shows that u satisfies $\partial_t u + \partial_x(f(u)) + g[u] = 0$ on $]t_0, T[\times\mathbb{R}$ for all $t_0 > 0$, which concludes the proof. ■

Item i) of Theorem 8.1.1 is a direct consequence of Theorem 8.5.1, Lemma 8.5.1 and Proposition 8.5.2 (as well as the theorem of continuity under the integral sign and Young's inequalities which show that, if $v \in C_b(]t_0, T[\times\mathbb{R})$ and $w \in L^1(\mathbb{R})$, then $(t,x) \in]t_0, T[\times\mathbb{R} \to w * v(t,\cdot)(x)$ is continuous and bounded).

8.6 L^∞ estimate and global existence

We construct here a solution to (8.1.1) on $]0,\infty[$ which is bounded by $||u_0||_{L^\infty(\mathbb{R})}$ and satisfies the maximum principle, thus concluding the proof of Theorem 8.1.1.
We assume, in the three following subsections, that $u_0 \in C_c^\infty(\mathbb{R})$ (in fact, we just need $u_0 \in L^1(\mathbb{R}) \cap BV(\mathbb{R})$).

8.6.1 Construction of an approximate solution by a splitting method

Let $\delta > 0$. We construct, by induction, a function $u^\delta : [0,\infty[\times\mathbb{R} \to \mathbb{R}$ the following way: we let $u^\delta(0,\cdot) = u_0$ and, for all $n \geq 0$, we define

- u^δ on $]2n\delta, (2n+1)\delta] \times \mathbb{R}$ as the solution to $\partial_t u^\delta + 2g[u^\delta] = 0$ ([1]) with initial condition $u^\delta(2n\delta,\cdot)$, that is to say $u^\delta(t,x) = K(2(t-2n\delta),\cdot) * u^\delta(2n\delta,\cdot)(x)$ for $(t,x) \in]2n\delta, (2n+1)\delta] \times \mathbb{R}$.
- u^δ on $](2n+1)\delta, 2(n+1)\delta] \times \mathbb{R}$ as the (entropy) solution to $\partial_t u^\delta + 2\partial_x(f(u^\delta)) = 0$ with initial condition $u^\delta((2n+1)\delta,\cdot)$.

Since $||K(t,\cdot)||_{L^1(\mathbb{R})} = 1$ for all $t > 0$, the regularizing operator does not increase the L^∞ norm (in fact, K being nonnegative, the maximum principle is satisfied), the L^1 norm and the BV semi-norm; it is a well-known result that the hyperbolic operator has the same properties. Moreover, the solutions to both equations are continuous with values in $L^1(\mathbb{R})$ (this is what states Lemma 8.2.2-i) for the regularizing equation). We have therefore defined $u^\delta \in C([0,\infty[; L^1(\mathbb{R}))$ such that $u^\delta(0,\cdot) = u_0$,

$$\begin{aligned}\forall t \geq 0\,,\ ||u^\delta(t,\cdot)||_{L^\infty(\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})}\,,\ ||u^\delta(t,\cdot)||_{L^1(\mathbb{R})} \leq ||u_0||_{L^1(\mathbb{R})} \\ \text{and } |u^\delta(t,\cdot)|_{BV(\mathbb{R})} \leq ||u_0'||_{L^1(\mathbb{R})},\end{aligned} \tag{8.6.1}$$

(in fact, u^δ takes its values between the minimum and maximum values of u_0) and, for all $n \geq 0$,

$$\begin{aligned}u^\delta(t,\cdot) = K(2(t-2n\delta),\cdot) * u^\delta(2n\delta,\cdot) \text{ for all } t \in]2n\delta, (2n+1)\delta], \\ u^\delta \text{ satisfies } \partial_t u^\delta + 2\partial_x(f(u^\delta)) = 0 \text{ on }](2n+1)\delta, 2(n+1)\delta] \times \mathbb{R}.\end{aligned} \tag{8.6.2}$$

[1] The factor 2 comes from the fact that we solve the regularizing equation (and the hyperbolic equation) on half of the total time, so we must give it twice more weight.

By (8.2.2) (which also gives, through (8.2.1), estimates on the time derivatives of K) and the fact that $u^\delta(2n\delta,\cdot) \in L^\infty(\mathbb{R})$, we see that, for all $n \geq 0$, u^δ is C^∞ on $]2n\delta,(2n+1)\delta] \times \mathbb{R}$. Moreover,

$$||\partial_x u^\delta((2n+1)\delta,\cdot)||_\infty = ||\partial_x K(2\delta,\cdot) * u^\delta(2n\delta,\cdot)||_\infty \leq \mathcal{K}_1 ||u_0||_\infty (2\delta)^{-1/\lambda}.$$

Hence, the time of regularity of u^δ on $](2n+1)\delta, 2(n+1)\delta] \times \mathbb{R}$ is at least

$$T^* \geq \frac{1}{2||f''(u^\delta((2n+1)\delta))\partial_x u^\delta((2n+1)\delta)||_{L^\infty(\mathbb{R})}} \geq C_0 \delta^{1/\lambda}$$

where C_0 does not depend on δ or n (we have used (8.6.1) to bound $f''(u^\delta((2n+1)\delta))$). For δ small enough, this time of regularity is thus greater than δ.
The parameter δ being destined to tend to 0, we can always suppose that it is small enough (let us say $\delta \leq \delta_0$) in order that u^δ is regular on $]2n\delta,(2n+1)\delta] \times \mathbb{R}$ and on $[(2n+1)\delta, 2(n+1)\delta] \times \mathbb{R}$ for all $n \geq 0$ (the BV estimate of (8.6.1) turns then into a L^1 estimate on the first spatial derivative).

Remark 8.6.1 *It is also possible to construct u^δ via a classical splitting method, i.e. to solve the regularizing equation (without the factor 2) on $[k\delta,(k+1)\delta]$ and then use the value thus obtained at $t=(k+1)\delta$ to solve the hyperbolic equation (still without the factor 2) on $[k\delta,(k+1)\delta]$ once again (and not on $[(k+1)\delta,(k+2)\delta]$). All the following reasoning can be done with such a construction; however, since the function thus defined is not continuous on $[0,\infty[$, more work is to be done.*

8.6.2 Compactness result on the sequence $(u^\delta)_{\delta>0}$

Proposition 8.6.1 *For all compact subset Q of $\mathbb{R}$ and all $T>0$, $\{u^\delta\ ,\ \delta \in]0,\delta_0]\}$ is relatively compact in $C([0,T];L^1(Q))$.*

Proof of Proposition 8.6.1
Let Q be a compact subset of $\mathbb{R}$ and $T>0$. For all $t \in [0,T]$, we have, by (8.6.1), $||u^\delta(t)||_{L^1(\mathbb{R})\cap BV(\mathbb{R})} \leq ||u_0||_{W^{1,1}(\mathbb{R})}$ (we omit the space variable in u^δ); thus, by Helly's Theorem, $\{u^\delta(t,\cdot)\ ,\ \delta \in]0,\delta_0]\}$ is relatively compact in $L^1(Q)$
We will prove the equicontinuity of $\{u^\delta\ ,\ \delta \in]0,\delta_0]\}$ in $C([0,\infty[;L^1(\mathbb{R}))$; this implies the equicontinuity in $C([0,T];L^1(Q))$ and, thanks to Ascoli-Arzela's theorem, concludes the proof of the proposition.

It is classical that the solution to an hyperbolic equation is lipschitz-continuous $[0,\infty[\to L^1(\mathbb{R})$. Thanks to (8.6.1), we see that the lipschitz constant of u^δ on $[(2n+1)\delta, 2(n+1)\delta]$ does not depend on δ or $n \geq 0$: there exists C_0 such that, for all $\delta \in]0,\delta_0]$, for all $n \geq 0$ and all $(t,s) \in [(2n+1)\delta, 2(n+1)\delta]$,

$$||u^\delta(t) - u^\delta(s)||_{L^1(\mathbb{R})} \leq C_0 |t-s|. \tag{8.6.3}$$

Taking into account that $u^\delta(s,\cdot) \in W^{1,1}(\mathbb{R})$ and the estimates of (8.6.1), some classical cuttings of integrals involving approximate units give, for all $\delta \in]0,\delta_0]$, all $t>0$, all $s \geq 0$ and all $\eta > 0$,

$$||K(t) * u^\delta(s) - u^\delta(s)||_{L^1(\mathbb{R})} \leq 2||u_0||_{L^1(\mathbb{R})} \int_{|y|\geq\eta} K(t,y)\,dy + \eta ||u_0'||_{L^1(\mathbb{R})}. \tag{8.6.4}$$

Let us now prove the equicontinuity of $\{u^\delta\ ,\ \delta \in]0,\delta_0]\}$ in $C([0,\infty[;L^1(\mathbb{R}))$. Let $\delta \in]0,\delta_0]$ and $0 \leq t < s$. Let $p \leq q$ be integers such that $p\delta \leq t < (p+1)\delta$ and $q\delta \leq s < (q+1)\delta$; because of the different behaviours of u^δ (see (8.6.2)), we must separate the cases depending on the parity of p and q; since all these cases are similar, we study only one, for example p even and q odd.
The idea, to estimate $||u^\delta(s,\cdot) - u^\delta(t,\cdot)||_{L^1(\mathbb{R})}$ is to go from $u^\delta(q\delta)$ to $u^\delta((p+1)\delta)$ by the following technique: on the intervals where u^δ satisfies the regularizing equation, we use the formula $u^\delta(k\delta) = K(2\delta) * u^\delta((k-1)\delta)$ (hence for k odd) and, on the intervals where u^δ satisfies the hyperbolic problem,

we write $u^\delta(k\delta) = u^\delta((k-1)\delta) + (u^\delta(k\delta) - u^\delta((k-1)\delta))$ (k even), the second term being estimated by (8.6.3).
Applying this idea, using the semi-group property of the convolution by $K(t)$ and recalling that q is odd in our example, an induction allows to see that, for all $l \in [0, (q-1)/2]$,

$$u^\delta(q\delta) = K(2l\delta) * u^\delta((q-2l)\delta) + \sum_{j=1}^{l} K(2j\delta) * (u^\delta((q-2j+1)\delta) - u^\delta((q-2j)\delta))$$

(if $l = 0$, $\sum_{j=1}^{l}(...)$ is null and $K(2l\delta) * u^\delta((q-2l)\delta)$ is replaced by $u^\delta(q\delta)$). Taking $l = (q-p-1)/2 \in [0, (q-1)/2]$ (recall that q is odd and p is even and inferior to q, thus $p+1 \leq q$) in this formula, we obtain

$$\begin{aligned} u^\delta(s) = u^\delta(s) - u^\delta(q\delta) + K((q-p-1)\delta) * u^\delta((p+1)\delta) \\ + \sum_{j=1}^{(q-p-1)/2} K(2j\delta) * (u^\delta((q-2j+1)\delta) - u^\delta((q-2j)\delta)). \end{aligned}$$

Since p is even, by definition of u^δ on $]p\delta, (p+1)\delta]$ and (8.2.5), we have $u^\delta((p+1)\delta) = K(2((p+1)\delta - t)) * (K(2(t-p\delta)) * u^\delta(p\delta)) = K(2((p+1)\delta - t)) * u^\delta(t)$. We can therefore write

$$\begin{aligned} u^\delta(s) - u^\delta(t) &= u^\delta(s) - u^\delta(q\delta) \\ &\quad + K((q-p-1)\delta + 2((p+1)\delta - t)) * u^\delta(t) - u^\delta(t) \\ &\quad + \sum_{j=1}^{(q-p-1)/2} K(2j\delta) * (u^\delta((q-2j+1)\delta) - u^\delta((q-2j)\delta)). \end{aligned}$$

On $[q\delta, s] \subset [q\delta, (q+1)\delta]$ and each $[(q-2j)\delta, (q-2j+1)\delta]$ for $j \in [1, (q-p-1)/2]$, u^δ satisfies the hyperbolic problem; thus, by (8.6.3) and (8.6.4), we have, for all $\eta > 0$,

$$\begin{aligned} ||u^\delta(s) - u^\delta(t)||_{L^1(\mathbb{R})} &\leq C_0|s - q\delta| + \frac{q-p-1}{2} C_0\delta \\ &\quad + 2||u_0||_{L^1(\mathbb{R})} \int_{|y|\geq\eta} K((q-p-1)\delta + 2((p+1)\delta - t), y)\, dy + \eta||u_0'||_{L^1(\mathbb{R})}. \end{aligned}$$

But, since $t < (p+1)\delta \leq q\delta \leq s$, we have $(q-p-1)\delta = q\delta - (p+1)\delta \leq s - t$, $2((p+1)\delta - t) \leq 2(s-t)$ and $s - q\delta \leq s - t$. Using these bounds in the preceding inequality, we obtain, for all $\delta \in]0, \delta_0]$, for all $0 \leq t < s$ and for all $\eta > 0$,

$$\begin{aligned} ||u^\delta(s) - u^\delta(t)||_{L^1(\mathbb{R})} &\leq \frac{3C_0}{2}|s-t| \\ &\quad + 2||u_0||_{L^1(\mathbb{R})} \sup_{\tau\in]0,3|s-t|]} \int_{|y|\geq\eta} K(\tau, y)\, dy + \eta||u_0'||_{L^1(\mathbb{R})} \end{aligned} \tag{8.6.5}$$

(the same kind of formula can be obtained in the cases where p and q have other parities than the ones considered here).
Since, for all $\eta > 0$, $\sup_{\tau\in]0,3|s-t|]} \int_{|y|\geq\eta} K(\tau, y)\, dy \to 0$ as $|s-t| \to 0$ (property of an approximate unit), (8.6.5) gives the desired equicontinuity and concludes the proof of Proposition 8.6.1. ■

8.6.3 Passing to the limit $\delta \to 0$

By Proposition 8.6.1, we can suppose, up to a subsequence, that, for all $T > 0$ and all Q compact subset of $\mathbb{R}$, $u^\delta \to u$ in $C([0,T]; L^1(Q))$ as $\delta \to 0$. For all $t \geq 0$, $u^\delta(t) \to u(t)$ in $L^1_{\text{loc}}(\mathbb{R})$, hence almost everywhere on $\mathbb{R}$ up to a subsequence. We deduce thus from (8.6.1) and Fatou's lemma that, for all $t > 0$, $||u(t)||_{L^1(\mathbb{R})} \leq ||u_0||_{L^1(\mathbb{R})}$, that

$$\forall t \geq 0\,, \quad ||u(t)||_{L^\infty(\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})} \tag{8.6.6}$$

and that, as u^δ, the function u takes its values between the minimum and maximum values of u_0. Still using Fatou's lemma on subsequences (depending on s and t), we see that (8.6.5) is satisfied for all $(s,t) \in [0,\infty[$ with u instead of u^δ; hence, $u \in C([0,\infty[;L^1(\mathbb{R}))$ and, since $u^\delta(0) = u_0$ for all $\delta > 0$, we have $u(0) = u_0$.

We now show that u satisfies (8.1.1) if we use a formulation involving test functions.

Proposition 8.6.2 *For all $\gamma \in C_c^\infty(]0,\infty[)$ and all $\varphi \in \mathcal{S}(\mathbb{R})$, we have*

$$\int_{\mathbb{R}^+\times\mathbb{R}} u(t,x)\gamma'(t)\varphi(x) + f(u(t,x))\gamma(t)\varphi'(x) - u(t,x)\gamma(t)g[\varphi](x)\,dxdt = 0. \tag{8.6.7}$$

Proof of Proposition 8.6.2
Let $\delta \in]0,\delta_0]$. If p is an odd integer, u^δ is a regular solution to $\partial_t u^\delta + 2\partial_x(f(u^\delta)) = 0$ on $[p\delta,(p+1)\delta]\times\mathbb{R}$. Multiplying this equation by $\gamma(t)\varphi(x)$ and integrating by parts (recall that u^δ is bounded and that $\varphi \in \mathcal{S}(\mathbb{R})$), we find

$$\begin{aligned} 0 = &-\int_{p\delta}^{(p+1)\delta}\int_{\mathbb{R}} u^\delta(t,x)\gamma'(t)\varphi(x) + 2f(u^\delta(t,x))\gamma(t)\varphi'(x)\,dtdx \\ &+\int_{\mathbb{R}} u^\delta((p+1)\delta,x)\gamma((p+1)\delta)\varphi(x)\,dx - \int_{\mathbb{R}} u^\delta(p\delta,x)\gamma(p\delta)\varphi(x)\,dx. \end{aligned} \tag{8.6.8}$$

If p is an even integer, then $u^\delta(t) = K(2(t-p\delta)) * u^\delta(p\delta)$ on $]p\delta,(p+1)\delta]\times\mathbb{R}$. Since $u^\delta(t) \in L^1(\mathbb{R})$ for all $t \geq 0$, Fubini's theorem allows to write

$$\mathcal{F}^{-1}(u^\delta(t)) = \mathcal{F}^{-1}(K(2(t-p\delta)))\mathcal{F}^{-1}(u^\delta(p\delta)) = e^{-2(t-p\delta)|\cdot|^\lambda}\mathcal{F}^{-1}(u^\delta(p\delta)).$$

Writing $\varphi = \mathcal{F}^{-1}(\mathcal{F}(\varphi))$ and $g[\varphi] = \mathcal{F}^{-1}(|\cdot|^\lambda\mathcal{F}(\varphi))$, since $u^\delta(t) \in L^1(\mathbb{R})$ for all $t \geq 0$, by Fubini's theorem we can put the inverse Fourier transform on u^δ and we thus check that

$$\begin{aligned} &\int_{p\delta}^{(p+1)\delta}\int_{\mathbb{R}} u^\delta(t,x)\gamma'(t)\varphi(x) - 2u^\delta(t,x)\gamma(t)g[\varphi](x)\,dxdt \\ &= \int_{\mathbb{R}} u^\delta((p+1)\delta,x)\gamma((p+1)\delta)\varphi(x) - u^\delta(p\delta,x)\gamma(p\delta)\varphi(x)\,dx. \end{aligned} \tag{8.6.9}$$

Summing (8.6.8) on all odd integers p and (8.6.9) on all even integers p (notice that, since the support of γ is compact, these sums are finite), the boundary terms disappear (even for $p = 0$ since $\gamma(0) = 0$) and we find

$$\begin{aligned} &\int_{\mathbb{R}^+}\int_{\mathbb{R}} u^\delta(t,x)\varphi(x)\gamma'(t)\,dxdt \\ &+\int_{\mathbb{R}^+}\int_{\mathbb{R}} 2f(u^\delta(t,x))\varphi'(x)\gamma(t)(1-\chi_\delta(t))\,dx\,dt \\ &-\int_{\mathbb{R}^+}\int_{\mathbb{R}} 2u^\delta(t,x)g[\varphi](x)\gamma(t)\chi_\delta(t)\,dx\,dt = 0 \end{aligned} \tag{8.6.10}$$

where χ_δ is the characteristic function of $\cup_{\text{even } p}\,]p\delta,(p+1)\delta]$.
Taking $T \geq \max(\text{supp}(\gamma))$, we have, for all $A \geq 0$, thanks to (8.6.1) and (8.6.6)

$$\begin{aligned} &\left|\int_{\mathbb{R}^+}\int_{\mathbb{R}} u^\delta(t,x)g[\varphi](x)\gamma(t)\,2\chi_\delta(t)\,dx\,dt - \int_{\mathbb{R}^+}\int_{\mathbb{R}} u(t,x)g[\varphi](x)\gamma(t)\,dx\,dt\right| \\ &\leq \left|\int_0^T\int_{-A}^A \left(2\chi_\delta(t)u^\delta(t,x) - u(t,x)\right) g[\varphi](x)\gamma(t)\,dx\,dt\right| \qquad (8.6.11) \\ &\quad +3||u_0||_{L^\infty(\mathbb{R})}T||\gamma||_{L^\infty(\mathbb{R}^+)}\int_{\mathbb{R}\setminus[-A,A]}|g[\varphi](x)|\,dx \qquad (8.6.12) \end{aligned}$$

Since $g[\varphi]$ is bounded on $\mathbb{R}$, $u^\delta \to u$ in $C([0,T];L^1([-A,A]))$ and $\chi_\delta \to 1/2$ in $L^\infty(]0,\infty[)$ weak-$*$ as $\delta \to 0$, we see that (8.6.11) tends to 0 as $\delta \to 0$. By Proposition 8.5.2, we have $g[\varphi] = g_1 * \varphi + g_2 * \varphi^{(4)} \in L^1(\mathbb{R})$; hence, (8.6.12) tends to 0 as $A \to \infty$. We deduce thus that, as $\delta \to 0$,

$$\int_{\mathbb{R}+}\int_{\mathbb{R}} u^\delta(t,x) g[\varphi](x)\gamma(t)\, 2\chi_\delta(t)\, dx\, dt \to \int_{\mathbb{R}+}\int_{\mathbb{R}} u(t,x) g[\varphi](x)\gamma(t)\, dx\, dt.$$

The flux function f is lipschitz-continuous on $[-||u_0||_{L^\infty(\mathbb{R})}, ||u_0||_{L^\infty(\mathbb{R})}]$ so that, by (8.6.1) and (8.6.6), $f(u^\delta) \to f(u)$ in $C([0,T];L^1(Q))$ for all $T > 0$ and all compact subset Q of $\mathbb{R}$. Therefore, with the same kind of reasoning as before, we can pass to the limit $\delta \to 0$ in (8.6.10) to conclude that u satisfies (8.6.7). ■

We now prove that u is in fact a solution to (8.1.1) in the sense of Definition 8.3.1.
Recall that $u \in C([0,\infty[;L^1(\mathbb{R}))$. By (8.6.6) and the local lipschitz-continuity of f, we have $f(u) \in C([0,\infty[;L^1(\mathbb{R}))$ (recall that $f(0) = 0$). We deduce, since $\mathcal{F}^{-1} : L^1(\mathbb{R}) \to C_b(\mathbb{R})$ is continuous, that $t \to \mathcal{F}^{-1}(u(t))$ and $t \to \mathcal{F}^{-1}(f(u(t)))$ are in $C([0,\infty[;C_b(\mathbb{R})) \subset C([0,\infty[\times\mathbb{R})$. Hence, for all $\gamma \in C_c^\infty(]0,\infty[)$, the function

$$w(\xi) = \int_{\mathbb{R}+} \mathcal{F}^{-1}(u(t))(\xi)\gamma'(t) + 2i\pi\xi\mathcal{F}^{-1}(f(u(t)))(\xi)\gamma(t) - \mathcal{F}^{-1}(u(t))(\xi)|\xi|^\lambda\gamma(t)\, dt\,,$$

is continuous on $\mathbb{R}$. Let $\psi \in C_c^\infty(\mathbb{R})$; applying (8.6.7) with $\varphi = \mathcal{F}^{-1}(\psi) \in \mathcal{S}(\mathbb{R})$ and using Fubini's theorem, we have $\int_{\mathbb{R}} w\psi = 0$; the function ψ being arbitrary, this implies $w \equiv 0$. Since this is true for all $\gamma \in C_c^\infty(]0,\infty[)$, we deduce that, for all $\xi \in \mathbb{R}$, $\frac{d}{dt}(\mathcal{F}^{-1}(u(\cdot))(\xi)) = -|\xi|^\lambda\mathcal{F}^{-1}(u(\cdot))(\xi) + 2i\pi\xi\mathcal{F}^{-1}(f(u(\cdot)))(\xi)$ in $\mathcal{D}'(]0,\infty[)$. The right-hand side of this equation is a continuous function, and the equation is therefore a classical ODE; thus, for all $\xi \in \mathbb{R}$ and all $t \geq 0$,

$$\begin{aligned}
\mathcal{F}^{-1}(u(t))(\xi) &= e^{-t|\xi|^\lambda}\mathcal{F}^{-1}(u_0)(\xi) + \int_0^t 2i\pi\xi e^{-(t-s)|\xi|^\lambda}\mathcal{F}^{-1}(f(u(s)))(\xi)\, ds \\
&= \mathcal{F}^{-1}(K(t) * u_0)(\xi) - \int_0^t \mathcal{F}^{-1}(\partial_x K(t-s))\mathcal{F}^{-1}(f(u(s)))(\xi)\, ds \\
&= \mathcal{F}^{-1}(K(t) * u_0)(\xi) - \int_0^t \mathcal{F}^{-1}(\partial_x K(t-s) * f(u(s)))(\xi)\, ds
\end{aligned}$$

By (8.2.4) and since $f(u) \in C([0,\infty[;L^1(\mathbb{R}))$, Fubini's theorem gives then

$$\mathcal{F}^{-1}(u(t))(\xi) = \mathcal{F}^{-1}(K(t) * u_0)(\xi) - \mathcal{F}^{-1}\left(\int_0^t \partial_x K(t-s) * f(u(s))\, ds\right)(\xi).$$

$\mathcal{F}^{-1}$ being injective on $L^1(\mathbb{R})$, we deduce that u satisfies (8.3.1) on $]0,\infty[\times\mathbb{R}$.

Here is a summary of what we have proved so far in this section.

Proposition 8.6.3 *If $u_0 \in C_c^\infty(\mathbb{R})$, then there exists a solution to (8.1.1) on $]0,\infty[$ which is bounded by $||u_0||_{L^\infty(\mathbb{R})}$ and takes its values between the minimum and maximum values of u_0.*

8.6.4 Conclusion

We now prove that if $u_0 \in L^\infty(\mathbb{R})$, then there exists a solution to (8.1.1) on $]0,\infty[$ which satisfies item ii) in Theorem 8.1.1, which concludes the proof of this theorem.

Let $u_0 \in L^\infty(\mathbb{R})$ and take $(u_0^n)_{n\geq 0} \in C_c^\infty(\mathbb{R})$ which converges a.e. on $\mathbb{R}$ to u_0 and such that, for all $n \geq 1$, u_0^n takes its values between the essential lower and upper bounds of u_0; in particular, $||u_0^n||_{L^\infty(\mathbb{R})} \leq$

$||u_0||_{L^\infty(\mathbb{R})}$ for all $n \geq 0$. Denote by u^n a solution, given by Proposition 8.6.3, to (8.1.1) on $]0,\infty[$ with initial condition u_0^n instead of u_0; u^n is bounded by $||u_0^n||_{L^\infty(\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})}$. This bound and Theorem 8.5.1 show that, for all $t_0 > 0$ and all $m \geq 0$, $(\partial_x^m u^n)_{n\geq 1}$ is bounded on $]t_0,\infty[\times\mathbb{R}$; by Lemma 8.5.1 and Proposition 8.5.2, these bounds on the spatial derivatives imply that $(\partial_t u^n)_{n\geq 1}$ is also bounded on $]t_0,\infty[\times\mathbb{R}$.
Hence, by Ascoli-Arzela's theorem, up to a subsequence, we can suppose that there exists u such that $u^n \to u$ on $]0,\infty[\times\mathbb{R}$. Since $||u^n||_{L^\infty(]0,\infty[\times\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})}$, we also have $||u||_{L^\infty(]0,\infty[\times\mathbb{R})} \leq ||u_0||_{L^\infty(\mathbb{R})}$ and, in fact, u takes (as each u^n) its values between the essential lower and upper bounds of u_0. The function u^n satisfies (8.3.1) with u_0^n instead of u_0; passing to the limit $n \to \infty$ in this equation, thanks to the dominated convergence theorem, we see that u is a solution to (8.1.1) on $]0,\infty[$.

Remark 8.6.2 *Since both the hyperbolic and regularizing equations satisfy the properties given in Remark 8.1.2, it is quite obvious, on our construction of a solution to (8.1.1), that (8.1.1) also satisfies the properties stated in Remark 8.1.2 (because we can always choose approximations of the initial conditions by regular data which satisfy the hypotheses of this remark).*

Chapitre 9

Généralisations du chapitre 8

9.1 En dimension supérieure à 1

Nous indiquons ici quelles sont les modifications à apporter au chapitre 8 pour traiter le cas d'une équation posée dans $\mathbb{R}^N$ avec $N \geq 1$, c'est à dire

$$\begin{cases} \partial_t u(t,x) + \operatorname{div}(f(u))(t,x) + g[u](t,x) = 0 & t > 0\,,\ x \in \mathbb{R}^N\,, \\ u(0,x) = u_0(x) & x \in \mathbb{R}^N \end{cases} \tag{9.1.1}$$

avec $f \in C^\infty(\mathbb{R};\mathbb{R}^N)$, $u_0 \in L^\infty(\mathbb{R}^N)$ et g toujours défini en Fourier spatial par $\widehat{g[u]}(\xi) = |\xi|^\lambda \widehat{u}(\xi)$ avec $1 < \lambda \leq 2$.

9.1.1 Positivité du noyau

Nous prouvons ici que, en toute dimension $N \geq 1$, le noyau associé à g est positif. Le cas $\lambda = 2$ étant connu, nous nous limiterons à $1 < \lambda < 2$. Rappelons que l'on a

$$K(t,x) = \mathcal{F}^{-1}(e^{-t|\cdot|^\lambda})(x) = \int_{\mathbb{R}^N} e^{2i\pi\langle x,\xi\rangle} e^{-t|\xi|^\lambda}\, d\xi.$$

Le lemme suivant est une version allégée du théorème de Lévy.

Lemme 9.1.1 *Soient $(f_n)_{n\geq 1}$ des fonctions positives intégrables sur $\mathbb{R}^N$. Soit $F : \mathbb{R}^N \to \mathbb{R}$ une fonction. Si $\mathcal{F}(f_n) \to F$ partout sur $\mathbb{R}^N$ lorsque $n \to \infty$ ([1]), alors $F \in L^\infty(\mathbb{R}^N)$ et $\mathcal{F}^{-1}(F) \geq 0$ dans $\mathcal{S}'(\mathbb{R}^N)$. En particulier, si F est intégrable, alors $\mathcal{F}^{-1}(F) \geq 0$ sur $\mathbb{R}^N$.*

Preuve du lemme 9.1.1
Les f_n étant positives, $\mathcal{F}(f_n)(0) = \int_{\mathbb{R}^N} f_n(x)\,dx = ||f_n||_{L^1(\mathbb{R}^N)}$. On constate donc que $||f_n||_{L^1(\mathbb{R}^N)} \to F(0)$, et que $(f_n)_{n\geq 1}$ est bornée dans $L^1(\mathbb{R}^N)$. On en déduit que $(\mathcal{F}(f_n))_{n\geq 1}$ est bornée dans $L^\infty(\mathbb{R}^N)$; comme cette suite converge ponctuellement vers F, on a, d'une part, $F \in L^\infty(\mathbb{R}^N)$ et, d'autre part, la convergence de $(\mathcal{F}(f_n))_{n\geq 1}$ vers F dans $\mathcal{S}'(\mathbb{R}^N)$ (convergence dominée).
On voit donc que $f_n = \mathcal{F}^{-1}(\mathcal{F}(f_n)) \to \mathcal{F}^{-1}(F)$ dans $\mathcal{S}'(\mathbb{R}^N)$ et, puisque les f_n sont positives, cela nous donne bien $\mathcal{F}^{-1}(F) \geq 0$ dans $\mathcal{S}'(\mathbb{R}^N)$ (i.e. $\langle \mathcal{F}^{-1}(F), \varphi\rangle \geq 0$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^N)$ positive). Si F est intégrable, $\mathcal{F}^{-1}(F)$ est une fonction (continue bornée) et la positivité dans $\mathcal{S}'(\mathbb{R}^N)$ nous donne la positivité au sens classique. ■

[1] En fait, il suffit d'avoir la convergence en 0 et presque partout ailleurs.

Lemme 9.1.2 *S'il existe $f \in L^1(\mathbb{R}^N)$ positive telle que, pour tout $\xi \in \mathbb{R}^N$,*

$$\mathcal{F}(f)(\xi) = 1 - c|\xi|^\lambda(1+\omega(\xi))\,, \qquad \text{avec } c>0 \text{ et } \omega(\xi)\to 0 \text{ lorsque } \xi \to 0, \tag{9.1.2}$$

alors K est positif.

Preuve du lemme 9.1.2
Considérons, pour $n \geq 1$, la fonction positive $f_n(x) = n^{N/\lambda} f * f * \cdots * f(n^{1/\lambda}x)$, où la convolution a été prise n fois. Par les résultats de convolution, $f_n \in L^1(\mathbb{R}^N)$. On a, par changement de variable,

$$\begin{aligned}
\mathcal{F}(f_n)(\xi) &= \int_{\mathbb{R}^N} n^{N/\lambda} e^{-2i\pi\langle x,\xi\rangle}(f*\cdots*f)(n^{1/\lambda}x)\,dx \\
&= \int_{\mathbb{R}^N} e^{-2i\pi\langle y, n^{-1/\lambda}\xi\rangle}(f*\cdots*f)(y)\,dy \\
&= \mathcal{F}(f*\cdots*f)(n^{-1/\lambda}\xi).
\end{aligned}$$

Ainsi, par les propriétés de la transformée de Fourier,

$$\mathcal{F}(f_n)(\xi) = \Big(\mathcal{F}(f)(n^{-1/\lambda}\xi)\Big)^n.$$

Soit $\xi \in \mathbb{R}^N$; pour n assez grand, par (9.1.2), on a $\mathcal{F}(f)(n^{-1/\lambda}\xi) > 0$ donc $\mathcal{F}(f_n)(\xi) > 0$ et

$$\begin{aligned}
\ln\left(\mathcal{F}(f_n)(\xi)\right) &= n\ln\Big(\mathcal{F}(f)(n^{-1/\lambda}\xi)\Big) \\
&= n\ln\left(1 - c\left|\frac{\xi}{n^{1/\lambda}}\right|^\lambda\left(1+\omega\left(\frac{\xi}{n^{1/\lambda}}\right)\right)\right) \\
&= n\ln(1-X_n)
\end{aligned}$$

avec $X_n \to 0$ lorsque $n\to\infty$. Par développement limité du logarithme au voisinage de 1, on peut écrire $\ln(\mathcal{F}(f_n)(\xi)) = -nX_n + n\mathcal{O}(X_n^2)$. Or $nX_n = c|\xi|^\lambda(1+\omega(n^{-1/\lambda}\xi)) \to c|\xi|^\lambda$ et $nX_n^2 = c|\xi|^\lambda(1+\omega(n^{-1/\lambda}\xi))X_n \to 0$ lorsque $n\to\infty$. On en déduit que $\ln(\mathcal{F}(f_n)(\xi)) \to -c|\xi|^\lambda$, i.e. que $\mathcal{F}(f_n)(\xi) \to e^{-c|\xi|^\lambda}$ lorsque $n\to\infty$.
Puisque les f_n sont positives, le lemme 9.1.1, nous donne $\mathcal{F}^{-1}(e^{-c|\cdot|^\lambda}) = K(c,\cdot) \geq 0$. Comme $K(c,\cdot) = c^{-N/\lambda}K(1,\cdot/c^{1/\lambda})$ (changement de variable immédiat dans la définition de K), cela implique $K(1,\cdot)\geq 0$ soit, pour tout $t>0$, $K(t,\cdot) = t^{-N/\lambda}K(1,\cdot/t^{1/\lambda}) \geq 0$. ■

Il reste donc à trouver f positive intégrable dont la transformée de Fourier a le comportement voulu en 0. On pose $f(x) = A|x|^{-N-\lambda}\mathbf{1}_{\{|x|\geq 1\}}(x)$ avec $A>0$ de sorte que $\int_{\mathbb{R}^N} f(x)\,dx = 1$. Comme f est paire, on a, pour $\xi\in\mathbb{R}^N$,

$$\begin{aligned}
\mathcal{F}(f)(\xi) &= \int_{\mathbb{R}^N}\cos(2\pi\langle x,\xi\rangle)f(x)\,dx \\
&= \int_{\mathbb{R}^N} f(x)\,dx + \int_{\mathbb{R}^N}(\cos(2\pi\langle x,\xi\rangle)-1)f(x)\,dx \\
&= 1 + A\int_{|x|\geq 1}\frac{\cos(2\pi\langle x,\xi\rangle)-1}{|x|^{\lambda+N}}\,dx.
\end{aligned}$$

Un changement de variable $y = |\xi|x$ donne donc, lorsque $\xi\neq 0$,

$$\mathcal{F}(f)(\xi) = 1 + A|\xi|^\lambda\int_{|y|\geq|\xi|}\frac{\cos\left(2\pi\langle y,\frac{\xi}{|\xi|}\rangle\right)-1}{|y|^{\lambda+N}}\,dy. \tag{9.1.3}$$

On peut modifier l'intégrande de cette dernière intégrale de sorte à obtenir une fonction qui ne dépende pas de ξ; en effet, en prenant R une rotation de $\mathbb{R}^N$ qui amène le vecteur ξ sur le vecteur $|\xi|e_1$ (où e_1 est le premier vecteur de la base canonique de $\mathbb{R}^N$), par changement de variable et puisque R^T préserve la norme, on a

$$\begin{aligned}\int_{|y|\geq|\xi|} \frac{\cos\left(2\pi\langle y, \frac{\xi}{|\xi|}\rangle\right)-1}{|y|^{\lambda+N}}\,dy &= \int_{|R^Tz|\geq|\xi|} \frac{\cos\left(2\pi\langle R^Tz, \frac{\xi}{|\xi|}\rangle\right)-1}{|R^Tz|^{\lambda+N}}\,dz \\ &= \int_{|z|\geq|\xi|} \frac{\cos\left(2\pi\langle z, R\frac{\xi}{|\xi|}\rangle\right)-1}{|z|^{\lambda+N}}\,dz \\ &= \int_{|z|\geq|\xi|} \frac{\cos\left(2\pi\langle z, e_1\rangle\right)-1}{|z|^{\lambda+N}}\,dz.\end{aligned}$$

(9.1.3) s'écrit alors

$$\mathcal{F}(f)(\xi) = 1 + A|\xi|^\lambda \int_{|z|\geq|\xi|} \frac{\cos(2\pi\langle z, e_1\rangle)-1}{|z|^{\lambda+N}}\,dz. \tag{9.1.4}$$

Mais $\cos(2\pi\langle z, e_1\rangle) - 1 = \mathcal{O}((2\pi\langle z, e_1\rangle)^2) = \mathcal{O}(|z|^2)$ au voisinage de 0, et la fonction $z \to \frac{\cos(2\pi\langle z,e_1\rangle)-1}{|z|^{\lambda+N}}$ est donc intégrable sur $\mathbb{R}^N$ (elle est majorée par $2/|z|^{\lambda+N}$ (avec $\lambda+N>N$) au voisinage de l'infini et par $C/|z|^{\lambda+N-2}$ (avec $\lambda+N-2<N$ puisque $\lambda<2$) au voisinage de 0). Ainsi,

$$\int_{|z|\geq|\xi|} \frac{\cos(2\pi\langle z, e_1\rangle)-1}{|z|^{\lambda+N}}\,dz \to I := \int_{\mathbb{R}^N} \frac{\cos(2\pi\langle z, e_1\rangle)-1}{|z|^{\lambda+N}}\,dz \quad \text{lorsque } \xi \to 0.$$

(9.1.4) nous montre donc que, pour $\xi \in \mathbb{R}^N\backslash\{0\}$,

$$\mathcal{F}(f)(\xi) = 1 + A|\xi|^\lambda(I + \omega(\xi))$$

où $\omega(\xi) \to 0$ lorsque $\xi \to 0$. I étant strictement négative (c'est l'intégrale d'une fonction négative non nulle presque partout), on en déduit que f vérifie (9.1.2) avec $c = -AI > 0$ ((9.1.2) est trivialement vérifiée en $\xi = 0$), et f est donc bien la fonction que nous cherchions.

9.1.2 Intégrabilité de transformées de Fourier

Les résultats que nous prouvons ici serviront directement à montrer que le noyau associé à g est aussi intégrable quand $N \geq 2$.

Lemme 9.1.3 *Soit $k \in \mathbb{N}$. On suppose que $G : \mathbb{R}^N \to \mathbb{R}$ vérifie: pour tout $\alpha \in \mathbb{N}^N$ de longueur inférieure ou égale à k, $\partial^\alpha G \in L^1(\mathbb{R}^N)$ (les dérivées étant prises au sens des distributions). Alors il existe $C > 0$ tel que, pour tout $x \in \mathbb{R}^N$,*

$$|\mathcal{F}^{-1}(G)(x)| \leq \frac{C}{1+|x|^k}.$$

En particulier, si $k = N+1$, alors $\mathcal{F}^{-1}(G) \in L^1(\mathbb{R}^N)$.

Preuve du lemme 9.1.3
On a $\mathcal{F}^{-1}(\partial^\alpha G)(x) = (-2i\pi x)^\alpha \mathcal{F}^{-1}(G)(x)$. Lorsque $x \in \mathbb{R}^N\backslash\{0\}$, en notant $P(x) = x/|x|$ la projection sur la sphère unité, on a $x^\alpha = |x|^{|\alpha|}P(x)^\alpha$, et on déduit de ce qui précède que

$$\sup_{\alpha\in\mathbb{N}^N,\ |\alpha|=k} |\mathcal{F}^{-1}(\partial^\alpha G)(x)| \geq (2\pi|x|)^k |\mathcal{F}^{-1}(G)(x)| \sup_{\alpha\in\mathbb{N}^N,\ |\alpha|=k} |P(x)^\alpha|. \tag{9.1.5}$$

Or

$$\inf_{y\in S^{N-1}} \sup_{\alpha\in\mathbb{N}^N,\ |\alpha|=k} |y^\alpha| \geq N^{-k/2}.$$

En effet, pour tout $y \in S^{N-1}$, il existe $j \in [1,N]$ tel que $|y_j| \geq 1/\sqrt{N}$ (on doit avoir $|y|^2 = \sum_{j=1}^N y_j^2 = 1$); on prend alors $\alpha_y = (0, \cdots, 0, k, 0, \cdots, 0)$, où le k est situé en j-ième position, et on trouve

$$\sup_{\alpha \in \mathbb{N}^N,\ |\alpha|=k} |y^\alpha| \geq |y^{\alpha_y}| = |y_j^k| \geq \frac{1}{\sqrt{N}^k}.$$

Ainsi, (9.1.5) donne, pour tout $x \neq 0$,

$$\sup_{\alpha \in \mathbb{N}^N,\ |\alpha|=k} ||\partial^\alpha G||_{L^1(\mathbb{R}^N)} \geq N^{-k/2}(2\pi|x|)^k |\mathcal{F}^{-1}(G)(x)|$$

(on a utilisé le fait que la transformée de Fourier inverse d'une fonction intégrable est majorée par la norme L^1 de cette fonction), soit, en notant $C_0 = N^{k/2}(2\pi)^{-k} \sup_{\alpha \in \mathbb{N}^N,\ |\alpha|=k} ||\partial^\alpha G||_{L^1(\mathbb{R}^N)}$,

$$|\mathcal{F}^{-1}(G)(x)| \leq \frac{C_0}{|x|^k}$$

pour tout $x \neq 0$.
Mais G est intégrable sur $\mathbb{R}^N$, donc $\mathcal{F}^{-1}(G)$ est bornée (par $||G||_{L^1(\mathbb{R}^N)}$) sur $\mathbb{R}^N$. On a alors, pour tout $x \in \mathbb{R}^N$,

$$|\mathcal{F}^{-1}(G)(x)| \leq ||G||_{L^1(\mathbb{R}^N)} \mathbf{1}_{B(0,1)}(x) + \frac{C_0}{|x|^k} \mathbf{1}_{\mathbb{R}^N \setminus B(0,1)}(x).$$

Comme $1 \leq \frac{2}{1+|x|^k}$ si $x \in B(0,1)$ et $\frac{1}{|x|^k} \leq \frac{2}{1+|x|^k}$ si $x \notin B(0,1)$, on en déduit que, pour tout $x \in \mathbb{R}^N$,

$$|\mathcal{F}^{-1}(G)(x)| \leq \frac{\sup(2||G||_{L^1(\mathbb{R}^N)}, 2C_0)}{1+|x|^k}$$

ce qui conclut la preuve. ■

Lemme 9.1.4 *Soit $\lambda > 1$, $\nu \in \mathbb{N}^N$ et $G : \xi \in \mathbb{R}^N \to \xi^\nu e^{-|\xi|^\lambda}$. Alors, pour tout $\alpha \in \mathbb{N}^N$ de longueur inférieure ou égale à $N+1$, on a $\partial^\alpha G \in L^1(\mathbb{R}^N)$ (dérivées au sens des distributions).*

Preuve du lemme 9.1.4
Commençons par traiter le cas $|\nu| = 0$, le cas général s'en déduisant facilement. On note donc $H(\xi) = e^{-|\xi|^\lambda}$. H est clairement de classe C^∞ hors de 0.
Prouvons par récurrence sur k que, pour tout $\alpha \in \mathbb{N}^N$ de longueur $k \geq 1$, il existe $n \geq 1$ et des fonctions $(P_l)_{l \in [1,n]}$ de classe C^∞ sur $\mathbb{R}^N \setminus \{0\}$ telles que, pour tout $l \in [1,n]$, P_l est homogène de rapport $\mu_l \geq \lambda - k$ et, pour tout $\xi \in \mathbb{R}^N \setminus \{0\}$, on a

$$\partial^\alpha H(\xi) = e^{-|\xi|^\lambda} \sum_{l=1}^n P_l(\xi).$$

Pour $k = 1$, on écrit

$$\partial_j H(\xi) = e^{-|\xi|^\lambda}(-\lambda)|\xi|^{\lambda-1} \frac{\xi_j}{|\xi|}$$

et on constate que $\xi \to (-\lambda)|\xi|^{\lambda-1}\frac{\xi_j}{|\xi|}$ est homogène de rapport $\lambda - 1$.
On suppose l'hypothèse de récurrence vérifiée en $k \geq 1$ et on va la montrer pour $k+1$. Si $\alpha \in \mathbb{N}^N$ est de longueur $k+1$, on écrit $\alpha = \beta + e_j$ où β est de longueur k et $e_j = (0, \cdots, 0, 1, 0, \cdots, 0)$ (le 1 est en j-ième position). On a alors, en dérivant par rapport à ξ_j la formule donnée par hypothèse de récurrence pour $\partial^\beta H$, lorsque $\xi \neq 0$,

$$\partial^\alpha H(\xi) = e^{-|\xi|^\lambda} \sum_{l=1}^n \partial_j P_l(\xi) - e^{-|\xi|^\lambda} \lambda |\xi|^{\lambda-2} \xi_j \sum_{l=1}^n P_l(\xi), \tag{9.1.6}$$

où chaque P_l est régulier hors de 0 et homogène de rapport $\mu_l \geq \lambda - k$. Pour tout $l \in [1,n]$, les dérivées premières de P_l sont homogènes de rapport $\mu_l - 1$; $\xi \to |\xi|^{\lambda-2}\xi_j$ est régulière hors de 0 et homogène de rapport $\lambda - 1$. Les fonctions intervenant en facteur de $e^{-|\cdot|^\lambda}$ dans (9.1.6) sont donc des fonctions régulières hors de 0 et homogènes avec, pour chacune, un rapport inclus dans $\{\mu_l - 1, \mu_l + \lambda - 1\,;\ l \in [1,n]\}$. Comme $\lambda \geq 0$ et $\mu_l - 1 \geq \lambda - (k+1)$ pour tout $l \in [1,n]$, cela conclut la récurrence.

Une fonction homogène continue sur $\mathbb{R}^N \backslash \{0\}$ ayant une croissance au plus polynômiale à l'infini, on constate donc que toutes les dérivées de H sont intégrables sur $\mathbb{R}^N \backslash B_N(1)$ (elles décroissent en fait plus vite que tout polynôme à l'infini).
Une fonction continue sur $\mathbb{R}^N \backslash \{0\}$ et homogène de rapport μ est un $\mathcal{O}(|\xi|^\mu)$ au voisinage de 0. Ainsi, pour tout $\alpha \in \mathbb{N}^N$, au voisinage de 0, $\partial^\alpha H(\xi) = \mathcal{O}(|\xi|^{\mu_1} + \cdots + |\xi|^{\mu_n})$ avec, pour tout $l \in [1,n]$, $\mu_l \geq \lambda - |\alpha|$; on en déduit que, pour tout $\alpha \in \mathbb{N}^N$, au voisinage de 0,

$$\partial^\alpha H(\xi) = \mathcal{O}(|\xi|^{\lambda - |\alpha|}). \tag{9.1.7}$$

Comme $\lambda > 1$, cette propriété nous permet de voir classiquement, que, pour tout $\alpha \in \mathbb{N}^N$ de longueur inférieure ou égal à $N+1$, $\partial^\alpha H$ est aussi la dérivée au sens des distributions de H, et qu'elle est intégrable sur $B_N(1)$ (car $\lambda - |\alpha| \geq \lambda - N - 1 > -N$).

Prenons maintenant $\nu \in \mathbb{N}^N$ de longueur strictement positive et notons $G(\xi) = \xi^\nu H(\xi)$. G est C^∞ hors de 0 et, pour tout $\alpha \in \mathbb{N}^N$,

$$\partial^\alpha G(\xi) = \sum_{\beta \leq \alpha} \partial^{\alpha-\beta}(\xi^\nu)\partial^\beta H(\xi).$$

Comme les dérivées de H décroissent plus vite que tout polynôme à l'infini, et comme $\partial^{\alpha-\beta}(\xi^\nu)$ a une croissance polynomiale à l'infini, les dérivées de G sont intégrables sur $\mathbb{R}^N \backslash B_N(1)$.
Par (9.1.7) et comme $\partial^{\alpha-\beta}(\xi^\nu)$ est bornée au voisinage de 0, on a aussi $\partial^\alpha G(\xi) = \sum_{\beta \leq \alpha} \mathcal{O}(|\xi|^{\lambda-|\beta|}) = \mathcal{O}(|\xi|^{\lambda-|\alpha|})$ ([2]) au voisinage de 0. Ceci nous permet de voir, comme pour H, que si α est de longueur inférieure à $N+1$, alors $\partial^\alpha G$ est la dérivée au sens des distributions de G, et qu'elle est intégrable sur $\mathbb{R}^N$. ■

9.1.3 Modifications à apporter aux preuves

Le principal changement concerne les propriétés du noyau K de g. Sa positivité a déjà été établie dans la partie 9.1.1. Les autres propriétés essentielles sont:

$$\forall (t,x) \in]0,\infty[\times\mathbb{R}, \quad K(t,x) = \frac{1}{t^{N/\lambda}} K\left(1, \frac{x}{t^{1/\lambda}}\right). \tag{9.1.8}$$

$$\begin{gathered} K \text{ est } C^\infty \text{ sur }]0,\infty[\times\mathbb{R}^N \text{ et, pour tout } \nu \in \mathbb{N}^N, \text{ il existe } B_\nu \text{ tel que} \\ \forall (t,x) \in]0,\infty[\times\mathbb{R}^N, \quad |\partial_x^\nu K(t,x)| \leq \frac{1}{t^{(N+|\nu|)/\lambda}} \frac{B_\nu}{(1 + t^{-(N+1)/\lambda}|x|^{N+1})}. \end{gathered} \tag{9.1.9}$$

$$\begin{gathered} (K(t,\cdot))_{t>0} \text{ est, pour } t \to 0, \text{ une approximation de l'unité} \\ \text{(en particulier, } ||K(t,\cdot)||_{L^1(\mathbb{R})} = 1 \text{ pour tout } t > 0). \end{gathered} \tag{9.1.10}$$

$$\exists \mathcal{K}_1 \text{ tel que, pour tout } t > 0, \quad ||\nabla_x K(t,\cdot)||_{L^1(\mathbb{R}^N)} = \mathcal{K}_1 t^{-1/\lambda}. \tag{9.1.11}$$

$$\begin{gathered} \forall (a,b) \in]0,\infty[, \ K(a,\cdot) * K(b,\cdot) = K(a+b,\cdot) \\ \text{et } K(a,\cdot) * \nabla_x K(b,\cdot) = \nabla_x K(a+b,\cdot). \end{gathered} \tag{9.1.12}$$

Preuve de ces propriétés
(9.1.8) s'obtient par changement de variable $\xi = t^{-1/\lambda}\eta$ dans l'intégrale qui définit K.

[2] On a en fait même mieux: on peut montrer que $\partial^\alpha G(\xi)$ est un $\mathcal{O}(|\xi|^{\lambda+|\nu|-|\alpha|})$ au voisinage de 0, et donc prendre des α plus longs que pour H.

La régularité de K est une conséquence directe de la dérivation sous l'intégrale. Pour prouver l'estimation, on écrit

$$\partial_x^\nu K(1,x) = \int_{\mathbb{R}^N} (2i\pi\xi)^\nu e^{-|\xi|^\lambda} e^{2i\pi\langle x,\xi\rangle}\, d\xi = \mathcal{F}^{-1}\left((2i\pi\cdot)^\nu e^{-|\cdot|^\lambda}\right)(x).$$

L'estimation pour $t=1$ découle alors des lemmes 9.1.4 et 9.1.3, et le cas général est une conséquence du cas $t=1$ et de (9.1.8).
$K(1,\cdot)$ est positif et intégrable sur $\mathbb{R}^N$ (voir (9.1.9)), et on peut ainsi écrire

$$||K(1,\cdot)||_{L^1(\mathbb{R}^N)} = \int_{\mathbb{R}^N} K(1,x)\, dx = \mathcal{F}(K(1,\cdot))(0) = e^{-|0|^\lambda} = 1.$$

(9.1.10) est donc une consequence de (9.1.8).
Par (9.1.9), $\nabla_x K(1,\cdot)$ est intégrable sur $\mathbb{R}^N$. (9.1.11) vient alors de la dérivation de (9.1.8) et du changement de variable $y = t^{-1/\lambda}x$ dans le calcul de $||\nabla_x K(1,\cdot/t^{1/\lambda})||_{L^1(\mathbb{R}^N)}$.
Les identités (9.1.12) peuvent être directement vérifiées en prenant la transformée de Fourier (toutes les fonctions sont intégrables). ■

On peut alors reprendre les techniques du chapitre 8 en invoquant ces propriétés (pour majorer K par des fonctions intégrables, estimer $||\nabla_x K(t,\cdot)||_{L^1(\mathbb{R}^N)}$, etc...) et en remplacant les ∂_x par des ∇_x. La seule différence notable est la représentation de g (Proposition 8.5.2), qui doit être modifiée ainsi:

Proposition 9.1.1 *Soit $m \in \mathbb{N}$ tel que $2m > N+\lambda$. Il existe $(g_1,g_2) \in L^1(\mathbb{R}^N)$ tels que, pour tout $v \in \mathcal{S}(\mathbb{R}^N)$, $g[v] = g_1 * v + g_2 * \Delta^m v$. Cette formule permet donc de définir $g[v]$ lorsque $v \in C_b^\infty(\mathbb{R}^N)$.*

Preuve de la proposition 9.1.1
Soit $\chi \in C_c^\infty(\mathbb{R}^N)$ paire et égale à 1 au voisinage de 0. On a, pour $v \in \mathcal{S}(\mathbb{R}^N)$,

$$g[v] = \mathcal{F}^{-1}(|\cdot|^\lambda \chi \mathcal{F}(v)) + \mathcal{F}^{-1}(|\cdot|^\lambda (1-\chi)\mathcal{F}(v)) \tag{9.1.13}$$

(puisque $\mathcal{F}(v) \in \mathcal{S}(\mathbb{R}^N)$, tous ces termes sont bien définis en tant que transformées de Fourier inverses de fonctions intégrables).
Soit $h_1 : \xi \in \mathbb{R}^N \to |\xi|^\lambda \chi(\xi)$. Les dérivées d'ordre inférieur à $N+1$ de h_1 sont integrables sur $\mathbb{R}^N$ (les dérivées d'ordre p se comportent comme $|\xi|^{\lambda-p}$ près de 0, et $\lambda - p > -N$ lorsque $p \leq N+1$), ce qui implique $\mathcal{F}^{-1}(h_1) \in L^1(\mathbb{R}^N)$. On peut donc écrire $\mathcal{F}(\mathcal{F}^{-1}(h_1) * v) = h_1 \mathcal{F}(v)$, c'est à dire $\mathcal{F}^{-1}(h_1 \mathcal{F}(v)) = \mathcal{F}^{-1}(h_1) * v$.
Soit $h_2 : \xi \in \mathbb{R}^N \to |\xi|^\lambda (1-\chi(\xi))$; la fonction $h_2^\star : \xi \in \mathbb{R}^N \to |2\pi\xi|^{-2m} h_2(\xi)$ est C^∞ et toutes ses dérivées sont intégrables sur $\mathbb{R}^N$ (les dérivées d'ordre p sont, au voisinage de l'infini, des $\mathcal{O}(|\xi|^{-2m-p+\lambda})$ et $-2m+\lambda < -N$); ainsi, $\mathcal{F}^{-1}(h_2^\star) \in L^1(\mathbb{R}^N)$ et

$$\mathcal{F}(\mathcal{F}^{-1}(h_2^\star) * \Delta^m v)(\xi) = h_2^\star(\xi)\mathcal{F}(\Delta^m v)(\xi) = |2\pi\xi|^{2m} h_2^\star(\xi)\mathcal{F}(v)(\xi) = h_2(\xi)\mathcal{F}(v)(\xi),$$

c'est à dire $\mathcal{F}^{-1}(h_2\mathcal{F}(v)) = \mathcal{F}^{-1}(h_2^\star) * \Delta^m v$.
L'identité (9.1.13) donne donc $g[v] = g_1 * v + g_2 * \Delta^m v$, où $g_1 = \mathcal{F}^{-1}(h_1)$ et $g_2 = \mathcal{F}^{-1}(h_2^\star)$ sont integrables sur $\mathbb{R}^N$ (remarquons que, puisque χ est paire, h_1 et $h_2^\star$ sont paires à valeurs réelles, de sorte que g_1 et g_2 sont aussi à valeurs réelles). ■

9.2 Equations hyperboliques plus générales

9.2.1 Introduction, hypothèses

Nous montrons ici que les techniques du chapitre 8 permettent aussi de trouver une unique solution régulière et globale à

$$\begin{cases} \partial_t u(t,x) + \operatorname{div}(f(t,x,u(t,x))) + g[u(t,\cdot)](x) = h(t,x,u(t,x)) & t>0\,,\ x \in \mathbb{R}^N \\ u(0,x) = u_0(x) & x \in \mathbb{R}^N\,, \end{cases} \tag{9.2.1}$$

où $u_0 \in L^\infty(\mathbb{R}^N)$, $f \in C^\infty([0,\infty[\times\mathbb{R}^N\times\mathbb{R})^N$ et $h \in C^\infty([0,\infty[\times\mathbb{R}^N\times\mathbb{R})$ satisfont

$$\begin{array}{c}\forall T>0\,,\ \forall R>0\,,\ \forall k\in\mathbb{N},\ \exists C_{T,R,k}>0 \text{ tel que, } \forall (t,x,\zeta)\in[0,T]\times\mathbb{R}^N\times[-R,R]\,,\\ \forall \alpha\in\mathbb{N}^{N+2} \text{ avec } |\alpha|\le k\,,\\ |\partial^\alpha f(t,x,\zeta)|+|\partial^\alpha h(t,x,\zeta)|\le C_{T,R,k}\,,\end{array} \tag{9.2.2}$$

$$\forall T>0\,,\ \left(h-\textstyle\sum_{i=1}^N \partial_{x_i} f_i\right)(\cdot,\cdot,0)_{|[0,T]\times\mathbb{R}^N} \text{ est à support compact dans } [0,T]\times\mathbb{R}^N\,, \tag{9.2.3}$$

$$\begin{array}{c}\forall T>0\,,\ \exists C'_T>0 \text{ et } \theta_T\in L^1(\mathbb{R}^N;[0,1]),\ \text{croissante par rapport à } T, \text{ tels que,}\\ \forall (t,x,\zeta)\in[0,T]\times\mathbb{R}^N\times\mathbb{R},\\ \left|h(t,x,\zeta)-\textstyle\sum_{i=1}^N \partial_{x_i} f_i(t,x,\zeta)\right|\le C'_T(\theta_T(x)+|\zeta|)\\ \text{et } \left|\nabla_x\left(h(t,x,\zeta)-\textstyle\sum_{i=1}^N \partial_{x_i} f_i(t,x,\zeta)\right)\right|\le C'_T(\theta_T(x)+|\zeta|).\end{array} \tag{9.2.4}$$

On prendra aussi (cela ne coûte rien) les $C_{T,R,k}$ et C'_T croissants par rapport à T et R.

La solution faible de (9.2.1) est à entendre au sens: pour tout $T>0$, $u\in L^\infty(]0,T[\times\mathbb{R}^N)$ et vérifie, pour presque tout $(t,x)\in]0,T[\times\mathbb{R}^N$,

$$u(t,x)=K(t,\cdot)*u_0(x)-\int_0^t \nabla K(t-s,\cdot)*f(s,\cdot,u(s,\cdot))(x)\,ds+\int_0^t K(t-s,\cdot)*h(s,\cdot,u(s,\cdot))(x)\,ds \tag{9.2.5}$$

(le deuxième produit de convolution cache un produit scalaire).

L'hypothèse (9.2.2) n'est "chère" que dans le sens où elle demande des estimations uniformes en espace; elle est cependant compréhensible puisque l'on cherche des solutions bornées sur $\mathbb{R}^N$ en entier.
L'hypothèse (9.2.3) est forte, mais elle peut être allégée et est surtout simplificatrice; elle permet de justifier certaines intégrations par parties que nous allons faire.
L'hypothèse (9.2.4) est cruciale et on ne peut totalement s'en débarasser: en effet, $h-\sum_i \partial_i f_i$ représente une source pour (9.2.1) (ou le problème hyperbolique associé), et une hypothèse de croissance au plus linéaire par rapport à l'inconnue est donc nécessaire pour avoir des solutions globales en temps. On peut s'en convaincre en regardant le cas $u_0=1$, $f=0$ et $h(u)=u^2$; on veut alors

$$u(t,x)=1+\int_0^t K(t-s,\cdot)*u^2(s,\cdot)(x)\,ds$$

et on se rend compte que l'on peut trouver une solution indépendante de x à ce problème, à savoir $u(t)=\frac{1}{1-t}$ (solution de $u(t)=1+\int_0^t u^2(s)\,ds$, c'est à dire de $u'=u^2$). Par unicité (voir plus bas, cette solution est bien dans $L^\infty(]0,T[\times\mathbb{R}^N)$ pour tout $T<1$), il s'agit de la solution cherchée, et on voit qu'elle a un temps d'existence fini.
Des hypothèses de la forme (9.2.2) et (9.2.4) apparaissent aussi dans [60]. Les remarques 9.2.2 et 9.2.3 donnent d'autres jeux d'hypothèses qui permettent de résoudre le problème.

Remarque 9.2.1 *A noter qu'on ne suppose pas de croissance linéaire sur f ou h, voir (9.2.2) (la croissance linéaire est juste demandée pour une expression entre h et certaines dérivées de f par rapport à x); si l'on avait ajouté une hypothèse de ce genre, la méthode de splitting serait totalement inutile: on pourrait directement voir sur (9.2.5) que la solution n'explose pas en temps fini, et qu'elle est donc globale en temps.*

9.2.2 Premières constatations

On voit d'abord aisément que tous les résultats des sections 8.3, 8.4 et 8.5 se transposent immédiatement (le terme ajouté dans (9.2.5) par rapport à (8.3.1), correspondant à la présence de h, est d'ordre inférieur — il fait intervenir K au lieu de ∇K — et ne pose donc aucun problème supplémentaire). On utilise

simplement l'hypothèse (9.2.2) pour s'assurer que, lorsque u est bornée, il en est de même pour $(t,x) \to f(t,x,u(t,x))$ et $(t,x) \to h(t,x,u(t,x))$ (voir section 8.3), et que $f(t,x,\cdot)$ et $h(t,x,\cdot)$ sont lipschitziennes, uniformément par rapport à $t \in [0,T]$ et $x \in \mathbb{R}^N$, sur tout borné (voir section 8.4). Cette hypothèse sert aussi dans la section 8.5, pour appliquer récursivement la proposition 8.5.1 et obtenir une borne sur toutes les dérivées de u.
Ainsi, on obtient l'unicité et la régularité de la solution faible à (9.2.1) (avec estimations sur les dérivées ne dépendant que de la norme L^∞ de la solution — ces estimations n'étant cependant valables, en temps, que sur tout compact de $]0,\infty[$, et non pas comme dans le chapitre 8 sur $]t_0,\infty[$ dès que $t_0 > 0$, puisqu'on cherche des solutions u qui ne sont bornées que localement en temps). Il reste à voir que la méthode de splitting permet de construire une solution globale; attention cependant: comme on l'a déjà indiqué, la solution ne sera pas (en général) dans $L^\infty(]0,\infty[\times\mathbb{R}^N)$, mais uniquement dans $L^\infty(]0,T[\times\mathbb{R}^N)$ pour tout $T > 0$ (ce qui est compréhensible, puisque l'on a mis une source: lorsque $u_0 = 1$, $f = 0$ et $h(t,x,u) = u$, la solution est $u(t,x) = e^t$).

9.2.3 Partie hyperbolique

Nous étudions ici la méthode des caractéristiques, pour obtenir des solutions locales régulières à

$$\partial_t u(t,x) + \operatorname{div}(f(t,x,u(t,x))) = h(t,x,u(t,x))\,, \quad u(0,\cdot) \in C^1_b(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N). \tag{9.2.6}$$

La principale difficulté de cette partie consiste à obtenir des estimations assez fines sur les normes de la solution régulière, afin de voir que ces estimations ne provoquent pas une explosion lorsqu'elles seront cumulées lors du splitting.

Méthode des caractéristiques

En développant (9.2.6) et en notant u_0 la condition initiale, cette équation s'écrit

$$\begin{cases} \partial_t u(t,x) + F(t,x,u(t,x)) \cdot \nabla u(t,x) = H(t,x,u(t,x)) & t > 0\,,\ x \in \mathbb{R}^N\,, \\ u(0,x) = u_0(x) & x \in \mathbb{R}^N \end{cases} \tag{9.2.7}$$

avec $F = (\partial_\zeta f)^T$ ($\partial_\zeta f$ est un vecteur ligne, F est un vecteur colonne) et $H = h - \sum_{i=1}^N \partial_{x_i} f_i$ régulières et, pour tout $T > 0$,

$$|H(t,x,\zeta)| \le C'_T(\theta_T(x) + |\zeta|) \tag{9.2.8}$$

dès que $(t,x,\zeta) \in [0,T] \times \mathbb{R}^N \times \mathbb{R}$ (voir (9.2.4)).

En suivant la méthode des caractéristiques, on est amené à poser

$$\begin{cases} \dfrac{d\psi}{dt}(t,x) = F(t,\psi(t,x),w(t,x)) \\ \dfrac{dw}{dt}(t,x) = H(t,\psi(t,x),w(t,x)) \\ \psi(0,x) = x \\ w(0,x) = u_0(x). \end{cases} \tag{9.2.9}$$

Si on sait prouver que, pour tout $t \in [0,T]$, $\psi_t = \psi(t,\cdot)$ est un difféomorphisme de $\mathbb{R}^N$, alors en posant $u(t,x) = w(t,\psi_t^{-1}(x))$, i.e. $u(t,\psi(t,x)) = w(t,x)$, on définit sur $[0,T] \times \mathbb{R}^N$ une fonction régulière (car $(t,x) \to (t,\psi(t,x))$ est régulière, bijective de $[0,T] \times \mathbb{R}^N$ dans lui-même et son déterminant jacobien est

$J\psi_t(x) \neq 0$, donc l'inverse de cette fonction est aussi régulière ([3])) qui vérifie

$$\begin{aligned}
&\partial_t u(t,\psi(t,x)) + F(t,\psi(t,x),u(t,\psi(t,x))) \cdot \nabla u(t,\psi(t,x)) \\
&\qquad = \partial_t u(t,\psi(t,x)) + F(t,\psi(t,x),w(t,x)) \cdot \nabla u(t,\psi(t,x)) \\
&\qquad = \partial_t u(t,\psi(t,x)) + \left(\frac{d\psi}{dt}(t,x)\right) \cdot \nabla u(t,\psi(t,x)) \\
&\qquad = \frac{d}{dt}\left(u(t,\psi(t,x))\right) \\
&\qquad = \frac{d}{dt}(w(t,x)) \\
&\qquad = H(t,\psi(t,x),w(t,x)) \\
&\qquad = H(t,\psi(t,x),u(t,\psi(t,x))).
\end{aligned}$$

Ceci étant vrai pour tout $x \in \mathbb{R}^N$ et tout $t \in [0,T]$, avec $\psi(t,\cdot)$ bijection de $\mathbb{R}^N$, on en déduit que u est solution de (9.2.7) sur $[0,T] \times \mathbb{R}^N$. Nous allons donc maintenant prouver qu'il existe $T > 0$ (et estimer sa taille) tel que, pour tout $x \in \mathbb{R}^N$, la solution de (9.2.9) existe sur $[0,T]$ et tel que ψ_t reste un difféomorphisme de $\mathbb{R}^N$.

On montre tout d'abord que les solutions de (9.2.9) sont globales (existent sur $[0,\infty[$). Par (9.2.8), on a, pour tout $t \in [0,T[$,

$$\left|\frac{dw}{dt}(t,x)\right| \leq C'_T(\theta_T(x) + |w(t,x)|) \leq C'_T(1 + |w(t,x)|),$$

d'où $|w(t,x)| \leq |w(0,x)| + C'_T T + C'_T \int_0^t |w(s,x)|\, ds$ et, par Gronwall,

$$|w(t,x)| \leq (|u_0(x)| + C'_T T) e^{C'_T T}. \tag{9.2.10}$$

En utilisant ceci et (9.2.2), on déduit $|\frac{d\psi}{dt}(t,x)| \leq L$ avec L indépendant de $t \in [0,T[$, donc $|\psi(t,x)| \leq |x| + LT$. Ainsi, (ψ, w) ne peut exploser en temps fini et la solution de (9.2.9) est globale.
Au passage, (9.2.10) donne, lorsque $T \leq 1$ (ce que l'on suppose à partir de maintenant),

$$\forall (t,x) \in [0,T] \times \mathbb{R}^N\ ,\ |w(t,x)| \leq (||u_0||_\infty + C'_1 T) e^{C'_1 T} \leq (||u_0||_\infty + C'_1) e^{C'_1} =: R(u_0). \tag{9.2.11}$$

Soit $T \leq 1$. En dérivant (9.2.9) par rapport à x, on a

$$\begin{cases}
\dfrac{d(\partial_x \psi)}{dt}(t,x) = \partial_x F(t,\psi(t,x),w(t,x))\partial_x\psi(t,x) + \partial_\zeta F(t,\psi(t,x),w(t,x))\partial_x w(t,x) \\
\dfrac{d(\partial_x w)}{dt}(t,x) = \partial_x H(t,\psi(t,x),w(t,x))\partial_x\psi(t,x) + \partial_\zeta H(t,\psi(t,x),w(t,x))\partial_x w(t,x) \\
\partial_x\psi(0,x) = Id \\
\partial_x w(0,x) = \partial_x u_0(x).
\end{cases} \tag{9.2.12}$$

Par (9.2.11) et (9.2.2), on en déduit, pour tout $t \in [0,T]$ et $x \in \mathbb{R}^N$,

$$\left|\frac{d(\partial_x\psi)}{dt}(t,x)\right| \leq C_{1,R(u_0),2}|\partial_x\psi(t,x)| + C_{1,R(u_0),2}|\partial_x w(t,x)|$$
$$\left|\frac{d(\partial_x w)}{dt}(t,x)\right| \leq NC_{1,R(u_0),2}|\partial_x\psi(t,x)| + NC_{1,R(u_0),2}|\partial_x w(t,x)|$$

[3]Pour être précis, ce raisonnement ne donne pas vraiment la régularité de $(t,x) \to \psi_t^{-1}(x)$ jusqu'en $t = 0$; cependant, nous résoudrons toujours l'équation hyperbolique à partir d'un temps initial $t_0 > 0$ (voir partie 9.2.5) de sorte que l'on pourra faire le même raisonnement pour des temps dans un *voisinage* de $[t_0, t_0 + T]$ et obtenir donc la régularité jusqu'à cet instant initial t_0 inclus.

soit

$$\left|\frac{d}{dt}\begin{pmatrix}\partial_x\psi\\ \partial_x w\end{pmatrix}(t,x)\right| \leq (N+1)C_{1,R(u_0),2}\left|\begin{pmatrix}\partial_x\psi\\ \partial_x w\end{pmatrix}(t,x)\right|.$$

Gronwall donne donc

$$\left|\begin{pmatrix}\partial_x\psi\\ \partial_x w\end{pmatrix}(t,x)\right| \leq \left|\left|\begin{pmatrix}Id\\ \partial_x u_0\end{pmatrix}\right|\right|_\infty e^{(N+1)C_{1,R(u_0),2}T} \leq (1+||\nabla u_0||_\infty)e^{(N+1)C_{1,R(u_0),2}} =: M_1.$$

Ceci montre que ψ et w sont M_1-lipschitziennes par rapport à $x \in \mathbb{R}^N$, uniformément pour $t \in [0,T]$. On a

$$\psi(t,x) = x + \int_0^t F(s,\psi(s,x),w(s,x))\,ds \tag{9.2.13}$$

et par (9.2.2), (9.2.11) et le caractère lipschitzien uniforme de ψ et w mentionné à l'instant, on voit que $(s,x) \to F(s,\psi(s,x),w(s,x))$ est $M(u_0)$-lipschitzien par rapport à $x \in \mathbb{R}^N$, uniformément pour $s \in [0,T]$, avec

$$M(u_0) = 2C_{1,R(u_0),2}M_1 = 2C_{1,R(u_0),2}(1+||\nabla u_0||_\infty)e^{(N+1)C_{1,R(u_0),2}}. \tag{9.2.14}$$

Ainsi, $(t,x) \in [0,T]\times\mathbb{R}^N \to \int_0^t F(s,\psi(s,x),w(s,x))\,ds$ est $M(u_0)T$-lipschitzien par rapport à x, uniformément pour $t \in [0,T]$, ce qui donne en particulier

$$\left|\partial_x \int_0^t F(s,\psi(s,x),w(s,x))\,ds\right| \leq M(u_0)T. \tag{9.2.15}$$

Le théorème d'inversion globale lipschitzienne (un simple point fixe contractant sur $\psi(t,x) = y$ à partir de (9.2.13)) permet donc de voir que, si $T \leq 1$ et $M(u_0)T < 1$, ψ_t est, pour tout $t \in [0,T]$, un homéomorphisme de $\mathbb{R}^N$; il est bon de se rappeler au passage que ce théorème dit aussi que ψ_t^{-1} est $\frac{1}{1-M(u_0)T}$-lipschitzienne. De plus, (9.2.15) et

$$\partial_x\psi_t(x) = Id + \partial_x \int_0^t F(s,\psi(s,x),w(s,x))\,ds$$

montrent, sous les mêmes conditions, que $\partial_x\psi_t(x)$ est inversible dans $\mathcal{L}(\mathbb{R}^N)$ pour tout $x \in \mathbb{R}^N$, et que ψ_t est donc un difféomorphisme de $\mathbb{R}^N$.

Ces considérations établissent que, sous les hypothèses $T \leq 1$ et $M(u_0)T < 1$ avec $M(u_0)$ défini par (9.2.14), on a une solution régulière à (9.2.6) sur $[0,T]\times\mathbb{R}^N$, obtenue en posant $u(t,x) = w(t,\psi_t^{-1}(x))$ avec (ψ,w) solution de (9.2.9).

Estimations

Nous allons maintenant prouver des estimations L^∞, L^1 et BV sur la solution construite par les caractéristiques. On prend $u_0 \in C_b^1(\mathbb{R}^N)\cap W^{1,1}(\mathbb{R}^N)\cap C^\infty(\mathbb{R}^N)$ et $T \leq 1$ tel que $M(u_0)T < 1$.

CONDITION INITIALE À SUPPORT COMPACT

On suppose dans un premier temps que la condition initiale u_0 est à support compact. Par (9.2.13), on a

$$|\psi(t,x) - x| \leq \int_0^t |F(s,\psi(s,x),w(s,x))|\,ds$$

et, en utilisant (9.2.11) et (9.2.2), on trouve donc $|\psi(t,x) - x| \leq C_{1,R(u_0),1}T$, d'où $|\psi(t,x)| \geq |x| - C_{1,R(u_0),1}T$. Prenons A tel que la boule de rayon A contienne le support de $H(t,\cdot,0)$ pour tout $t \in [0,T]$ (voir (9.2.3)), ainsi que le support de u_0; alors, pour tout $|x| \geq A + C_{1,R(u_0),1}T$ et tout $t \in [0,T]$, on a $|\psi(t,x)| \geq A$, donc $H(t,\psi(t,x),0) = 0$; cela montre que, lorsque $|x| \geq A + C_{1,R(u_0),1}T$, la seconde

composante de la solution de (9.2.9) sur $[0,T]$ est $w(t,x)=0$. Ainsi, la boule de rayon $A+C_{1,R(u_0),1}T$ contient le support de $w(t,\cdot)$ pour tout $t\in[0,T]$.
Comme ψ_t est $(1+M(u_0)T)$-lipschiztienne sur $\mathbb{R}^N$ et $(\psi_t^{-1}(0))_{t\in[0,T]}$ est bornée ($t\to\psi_t^{-1}(0)$ est continue), l'équation $u(t,x)=w(t,\psi_t^{-1}(x))$ nous montre que $u(t,\cdot)$ reste à support compact: pour avoir $u(t,x)\neq 0$, il faut $|\psi_t^{-1}(x)|\leq A+C_{1,R(u_0),1}T$, donc

$$\begin{aligned}|x| &= |\psi_t(\psi_t^{-1}(x))-\psi_t(\psi_t^{-1}(0))|\\ &\leq (1+M(u_0)T)(|\psi_t^{-1}(x)|+|\psi_t^{-1}(0)|)\\ &\leq (1+M(u_0)T)\left(A+C_{1,R(u_0),1}T+\sup_{t\in[0,T]}|\psi_t^{-1}(0)|\right).\end{aligned}$$

Par $u(t,x)=w(t,\psi_t^{-1}(x))$ et (9.2.11), on a, lorsque $t\in[0,T]$,

$$||u(t)||_{L^\infty(\mathbb{R}^N)}\leq(||u_0||_\infty+C_1'T)e^{C_1'T}. \tag{9.2.16}$$

On procède maintenant comme dans [3]. Soit sgn_η une approximation régulière, nulle en 0, de la fonction signe, et $|\zeta|_\eta=\int_0^\zeta \mathrm{sgn}_\eta(\tau)\,d\tau$ l'approximation correspondante de la valeur absolue. En multipliant (9.2.7) par $\mathrm{sgn}_\eta(u)$ et en intégrant sur $\mathbb{R}^N$ (cette intégration et la manipulation qui suit sur la dérivée temporelle sont valables puisque u est à support compact et toutes les fonctions sont régulières), on trouve

$$\partial_t\int_{\mathbb{R}^N}|u(t,x)|_\eta\,dx+\int_{\mathbb{R}^N}\mathrm{sgn}_\eta(u(t,x))F(t,x,u(t,x))\cdot\nabla u(t,x)\,dx=\int_{\mathbb{R}^N}H(t,x,u(t,x))\mathrm{sgn}_\eta(u(t,x))\,dx$$

soit, après intégration en temps,

$$\begin{aligned}\int_{\mathbb{R}^N}|u(t,x)|_\eta\,dx &= \int_{\mathbb{R}^N}|u_0(x)|_\eta\,dx-\int_0^t\int_{\mathbb{R}^N}\mathrm{sgn}_\eta(u(s,x))F(s,x,u(s,x))\cdot\nabla u(s,x)\,dx\,ds\\ &\quad+\int_0^t\int_{\mathbb{R}^N}H(s,x,u(s,x))\mathrm{sgn}_\eta(u(s,x))\,dx\,ds.\end{aligned}$$

Mais en posant

$$G_\eta(s,x,\zeta)=\int_0^\zeta\mathrm{sgn}_\eta(\tau)F(s,x,\tau)\,d\tau$$

on a $\mathrm{sgn}_\eta(u(s,x))F(s,x,u(s,x))\cdot\nabla u(s,x)=\mathrm{div}(G_\eta(s,x,u(s,x)))-\mathrm{div}_x(G_\eta)(s,x,u(s,x))$ et

$$\int_{\mathbb{R}^N}\mathrm{div}(G_\eta(s,x,u(s,x)))\,dx=0 \tag{9.2.17}$$

puisque $G_\eta(s,x,u(s,x))=0$ pour x assez grand (comme $G_\eta(s,x,0)=0$, il suffit de prendre x hors du support de $u(s,\cdot)$); donc

$$\int_{\mathbb{R}^N}|u(t,x)|_\eta\,dx=\int_{\mathbb{R}^N}|u_0(x)|_\eta\,dx+\int_0^t\int_{\mathbb{R}^N}\mathrm{div}_x(G_\eta)(s,x,u(s,x))+H(s,x,u(s,x))\mathrm{sgn}_\eta(u(s,x))\,dx\,ds.$$

Par définition de G_η, (9.2.2) et (9.2.8), on a

$$\begin{aligned}|\mathrm{div}_x(G_\eta)(s,x,u(s,x))+H(s,x,u(s,x))\mathrm{sgn}_\eta(u(s,x))| &\leq C_{1,||u||_\infty,2}|u(s,x)|+C_1'(\theta_1(x)+|u(s,x)|)\\ &= (C_{1,||u||_\infty,2}+C_1')|u(s,x)|+C_1'\theta_1(x)\end{aligned}$$

(la norme infinie de u étant prise sur $[0,T]\times\mathbb{R}^N$) et on obtient finalement, en faisant $\eta\to 0$,

$$\int_{\mathbb{R}^N}|u(t,x)|\,dx\leq\int_{\mathbb{R}^N}|u_0(x)|\,dx+C_1'||\theta_1||_{L^1(\mathbb{R}^N)}T+(C_{1,||u||_\infty,2}+C_1')\int_0^t\int_{\mathbb{R}^N}|u(s,x)|\,dx\,ds$$

soit, par Gronwall, pour $t \in [0,T]$,

$$||u(t)||_{L^1(\mathbb{R}^N)} \leq (||u_0||_{L^1(\mathbb{R}^N)} + C_1'||\theta_1||_{L^1(\mathbb{R}^N)}T)e^{(C_{1,||u||_\infty,2}+C_1')T}. \tag{9.2.18}$$

On prend maintenant le gradient de (9.2.7) par rapport à x, en notant que

$$F(t,x,u(t,x)) \cdot \nabla u(t,x) = \sum_{i=1}^N F_i(t,x,u(t,x))\partial_i u(t,x),$$

et on trouve

$$\begin{aligned}
&\partial_t(\nabla u)(t,x) + \sum_{i=1}^N \partial_\zeta F_i(t,x,u(t,x))\partial_i u(t,x)\nabla u(t,x) + \sum_{i=1}^N F_i(t,x,u(t,x))\partial_i \nabla u(t,x) \\
&= \quad \nabla_x H(t,x,u(t,x)) + \partial_\zeta H(t,x,u(t,x))\nabla u(t,x) - \sum_{i=1}^N \nabla_x F_i(t,x,u(t,x))\partial_i u(t,x).
\end{aligned} \tag{9.2.19}$$

Soit $I_\eta(\xi) = \int_0^{|\xi|} \mathrm{sgn}_\eta(\tau)\,d\tau$ une approximation de la norme euclidienne. Puisque sgn_η est régulière nulle en 0 et $\nabla I_\eta(\xi) = \mathrm{sgn}_\eta(|\xi|)\frac{\xi}{|\xi|}$ hors de 0, on constate que I_η est de classe C^1. On peut donc multiplier (9.2.19) scalairement par $\nabla I_\eta(\nabla u) = \mathrm{sgn}_\eta(|\nabla u|)\frac{\nabla u}{|\nabla u|}$ pour trouver

$$\begin{aligned}
&\partial_t(I_\eta(\nabla u))(t,x) + \sum_{i=1}^N \partial_\zeta F_i(t,x,u(t,x))\partial_i u(t,x)\mathrm{sgn}_\eta(|\nabla u(t,x)|)|\nabla u(t,x)| \\
&+ \sum_{i=1}^N F_i(t,x,u(t,x))\nabla I_\eta(\nabla u(t,x)) \cdot \partial_i \nabla u(t,x) \\
&= \quad \nabla_x H(t,x,u(t,x)) \cdot \mathrm{sgn}_\eta(|\nabla u(t,x)|)\frac{\nabla u(t,x)}{|\nabla u(t,x)|} + \partial_\zeta H(t,x,u(t,x))\mathrm{sgn}_\eta(|\nabla u(t,x)|)|\nabla u(t,x)| \\
&\quad - \sum_{i=1}^N \nabla_x F_i(t,x,u(t,x))\partial_i u(t,x) \cdot \mathrm{sgn}_\eta(|\nabla u(t,x)|)\frac{\nabla u(t,x)}{|\nabla u(t,x)|}.
\end{aligned}$$

Mais $\partial_i(I_\eta(\nabla u)) = \nabla I_\eta(\nabla u) \cdot \partial_i \nabla u$ donc

$$\begin{aligned}
&F_i(t,x,u(t,x))\nabla I_\eta(\nabla u(t,x)) \cdot \partial_i \nabla u(t,x) \\
&\qquad = \quad F_i(t,x,u(t,x))\partial_i[I_\eta(\nabla u(t,x))] \\
&\qquad = \quad \partial_i[F_i(t,x,u(t,x))I_\eta(\nabla u(t,x))] - \partial_i[F_i(t,x,u(t,x))]I_\eta(\nabla u(t,x)) \\
&\qquad = \quad \partial_i[F_i(t,x,u(t,x))I_\eta(\nabla u(t,x))] - \partial_i F_i(t,x,u(t,x))I_\eta(\nabla u(t,x)) \\
&\qquad\quad -\partial_\zeta F_i(t,x,u(t,x))\partial_i u(t,x)I_\eta(\nabla u(t,x)).
\end{aligned}$$

Puisque $\nabla u(t,\cdot)$ et $I_\eta(\nabla u(t,\cdot))$ sont à support compact (car $I_\eta(0) = 0$), tous les termes considérés sont nuls hors d'un compact de $\mathbb{R}^N$; on peut donc intégrer sur $\mathbb{R}^N$ et utiliser une intégration par parties pour

éliminer $\partial_i[F_i(t,x,u(t,x))I_\eta(\nabla u(t,x))]$:

$$\begin{aligned}
&\partial_t \int_{\mathbb{R}^N} I_\eta(\nabla u)(t,x)\,dx \\
&+\sum_{i=1}^N \int_{\mathbb{R}^N} \partial_\zeta F_i(t,x,u(t,x))\partial_i u(t,x)\left[\mathrm{sgn}_\eta(|\nabla u(t,x)|)|\nabla u(t,x)| - I_\eta(\nabla u(t,x))\right]\,dx \\
&\quad = \int_{\mathbb{R}^N} \nabla_x H(t,x,u(t,x))\cdot \mathrm{sgn}_\eta(|\nabla u(t,x)|)\frac{\nabla u(t,x)}{|\nabla u(t,x)|}\,dx \\
&\qquad + \int_{\mathbb{R}^N} \partial_\zeta H(t,x,u(t,x))\mathrm{sgn}_\eta(|\nabla u(t,x)|)|\nabla u(t,x)|\,dx \\
&\qquad - \sum_{i=1}^N \int_{\mathbb{R}^N} \nabla_x F_i(t,x,u(t,x))\partial_i u(t,x)\cdot \mathrm{sgn}_\eta(|\nabla u(t,x)|)\frac{\nabla u(t,x)}{|\nabla u(t,x)|}\,dx \\
&\qquad + \sum_{i=1}^N \int_{\mathbb{R}^N} \partial_i F_i(t,x,u(t,x))I_\eta(\nabla u(t,x))\,dx. \qquad (9.2.20)
\end{aligned}$$

On intègre ensuite en temps. Lorsque $\eta \to 0$, on a $I_\eta \to |\cdot|$ et $\mathrm{sgn}_\eta(|\cdot|)|\cdot| - I_\eta(\cdot) \to 0$, tout en restant majorées par $|\cdot|$ et $2|\cdot|$. On peut donc passer à la limite par Lebesgue (rappelons que les intégrales précédentes sont réduites à des compacts fixes, puisque les fonctions intervenant s'annulent toutes pour x assez grand), et on arrive à

$$\begin{aligned}
\int_{\mathbb{R}^N} |\nabla u(t,x)|\,dx \quad \le \quad & \int_{\mathbb{R}^N} |\nabla u_0(x)|\,dx + \int_0^t \int_{\mathbb{R}^N} |\nabla_x H(s,x,u(s,x))|\,dx\,ds \qquad (9.2.21)\\
&+ \int_0^t \int_{\mathbb{R}^N} |\partial_\zeta H(s,x,u(s,x))|\,|\nabla u(s,x)|\,dx\,ds \\
&+ \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^N} |\nabla_x F_i(s,x,u(s,x))|\,|\partial_i u(s,x)|\,dx\,ds \\
&+ \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^N} |\partial_i F_i(s,x,u(s,x))|\,|\nabla u(s,x)|\,dx\,ds.
\end{aligned}$$

On utilise maintenant (9.2.4), (9.2.2) et (9.2.18) pour obtenir

$$\begin{aligned}
\int_{\mathbb{R}^N} |\nabla u(t,x)|\,dx \quad \le \quad & \int_{\mathbb{R}^N} |\nabla u_0(x)|\,dx + C_1' \int_0^t \int_{\mathbb{R}^N} (|\theta_1(x)| + |u(s,x)|)\,dx\,ds \\
&+ NC_{1,||u||_\infty,2} \int_0^t \int_{\mathbb{R}^N} |\nabla u(s,x)|\,dx\,ds + C_{1,||u||_\infty,2} \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^N} |\partial_i u(s,x)|\,dx\,ds \\
&+ NC_{1,||u||_\infty,2} \int_0^t \int_{\mathbb{R}^N} |\nabla u(s,x)|\,dx\,ds \\
\le \quad & ||\nabla u_0||_{L^1(\mathbb{R}^N)} + C_1'(||\theta_1||_{L^1(\mathbb{R}^N)} + \widetilde{R}_1(u_0))T \\
&+ 3NC_{1,||u||_\infty,2} \int_0^t \int_{\mathbb{R}^N} |\nabla u(s,x)|\,dx\,ds
\end{aligned}$$

où $\widetilde{R}_1(u_0) = (||u_0||_{L^1(\mathbb{R}^N)} + C_1'||\theta_1||_{L^1(\mathbb{R}^N)})e^{C_{1,||u||_\infty,2}+C_1'}$ (rappelons que $t \le T \le 1$); on conclut alors toujours grâce à Gronwall, pour $t \in [0,T]$ et avec $||u||_\infty = ||u||_{L^\infty([0,T]\times\mathbb{R}^N)}$,

$$||\nabla u(t)||_{L^1(\mathbb{R}^N)} \le \left(||\nabla u_0||_{L^1(\mathbb{R}^N)} + C_1'(||\theta_1||_{L^1(\mathbb{R}^N)} + \widetilde{R}_1(u_0))T\right) e^{3NC_{1,||u||_\infty,2}T}. \qquad (9.2.22)$$

CONDITION INITIALE DANS $C_b^1(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$

Nous souhaitons maintenant montrer que (9.2.16), (9.2.18) et (9.2.22) restent valables (en remplaçant, dans ces deux dernières inégalités, $||u||_\infty$ par l'estimation venant de (9.2.16)), toujours lorsque $T \leq 1$ vérifie $M(u_0)T < 1$, quand u_0 est seulement dans $C_b^1(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$.
Pour cela, on approche u_0 localement uniformément par $u_{0,n} \in C_c^\infty(\mathbb{R}^N)$ telle que

$$\begin{array}{c} ||u_{0,n}||_\infty \leq ||u_0||_\infty \,,\ ||u_{0,n}||_{L^1(\mathbb{R}^N)} \leq ||u_0||_{L^1(\mathbb{R}^N)} \,, \\ \limsup_{n\to\infty} ||\nabla u_{0,n}||_\infty \leq ||\nabla u_0||_\infty \ \text{ et } \ \limsup_{n\to\infty} ||\nabla u_{0,n}||_{L^1(\mathbb{R}^N)} \leq ||\nabla u_0||_{L^1(\mathbb{R}^N)}. \end{array} \tag{9.2.23}$$

Une telle suite existe bien: on prend $\gamma \in C_c^\infty(\mathbb{R}^N;[0,1])$ telle que $\gamma(0) = 1$ et on pose $\gamma_n(x) = \gamma(x/n)$; alors $\gamma_n u_0$ est dans $C_c^\infty(\mathbb{R}^N)$, $\gamma_n u_0 \to u_0$ localement uniformément, $|\gamma_n u_0| \leq |u_0|$ et $||\nabla(\gamma_n u_0)||_p \leq ||u_0||_p ||\nabla\gamma_n||_\infty + ||\nabla u_0||_p$ ($p = 1$ ou ∞) avec $||\nabla\gamma_n||_\infty \to 0$ quand $n \to \infty$, puisque $\nabla\gamma_n = \frac{1}{n}\nabla\gamma(\frac{\cdot}{n})$.

Vu (9.2.14), (9.2.11) et (9.2.23), on a $\limsup_{n\to\infty} M(u_{0,n}) \leq M(u_0)$. Ainsi, pour n assez grand, $M(u_{0,n})T < 1$ et (9.2.16), (9.2.18) et (9.2.22) sont valables sur $[0,T]$ avec u^n et $u_{0,n}$ à la place de u et u_0, où u^n est la solution de (9.2.7) avec $u_{0,n}$ comme condition initiale.

Soit (ψ^n, w^n) la solution de (9.2.9) avec $u_{0,n}$ au lieu de u_0. Cette solution existe globalement et, puisque $u_{0,n} \to u_0$ localement uniformément, le théorème de dépendance continue des solutions d'EDO par rapport aux conditions initiales nous dit que $(\psi^n, w^n) \to (\psi, w)$ localement uniformément sur $[0,T]\times\mathbb{R}^N$, où (ψ, w) est la solution de (9.2.9) correspondant à u_0.
Soit $t \in [0,T]$ et $a_n = (\psi_t^n)^{-1}(0)$, c'est à dire, par (9.2.13), $a_n = -\int_0^t F(s, \psi^n(s,a_n), w^n(s,a_n))\,ds$. L'estimation (9.2.11) donne $||w^n||_{L^\infty([0,T]\times\mathbb{R}^N)} \leq R(u_{0,n}) \leq R(u_0)$ et, par (9.2.2), montre que $(a_n)_{n\geq 1}$ est bornée. On sait de plus que $(\psi_t^n)^{-1}$ est lipschitzienne de constante inférieure à $\frac{1}{1-M(u_{0,n})T}$; cela signifie, puisque $\limsup_{n\to\infty} M(u_{0,n}) \leq M(u_0)$ et $M(u_0)T < 1$, que $((\psi_t^n)^{-1})_{n\geq 1}$ est uniformément lipschitzienne, et donc localement bornée puisque bornée en $x = 0$. Le théorème d'Ascoli-Arzela nous dit alors que, à une sous-suite près, $(\psi_t^n)^{-1}$ converge localement uniformément vers une fonction Ξ. Mais par convergence locale uniforme de ψ_t^n vers ψ_t, on peut passer à la limite dans $\psi_t^n \circ (\psi_t^n)^{-1} = Id$ pour voir que, nécessairement, on a $\Xi = \psi_t^{-1}$.
Ainsi, pour tout $t \in [0,T]$, $w^n(t,\cdot) \to w(t,\cdot)$ et $(\psi_t^n)^{-1} \to \psi_t^{-1}$ localement uniformément. On en déduit que $u^n(t,x) = w^n(t,(\psi_t^n)^{-1}(x)) \to w(t,\psi_t^{-1}(x)) = u(t,x)$, autrement dit que $u^n \to u$ ponctuellement (en fait localement uniformément, mais cela ne nous sera pas utile).

On peut alors directement passer à la limite dans (9.2.16) appliqué à u^n, en se souvenant que $||u_{0,n}||_\infty \leq ||u_0||_\infty$, pour voir que cette estimation est encore valable pour u. De même, Fatou permet de passer à la limite dans (9.2.18) appliqué à u^n (quitte à remplacer $||u||_\infty$ dans $C_{1,||u||_\infty,2}$ par la majoration venant de (9.2.16)). Pour passer à la limite dans (9.2.22), on utilise juste la relation bien connue des spécialistes des espaces BV:

$$||\nabla u(t)||_{L^1(\mathbb{R}^N)} = |u(t)|_{BV(\mathbb{R}^N)} \leq \liminf_{n\to\infty} |u^n(t)|_{BV(\mathbb{R}^N)} = \liminf_{n\to\infty} ||\nabla u^n(t)||_{L^1(\mathbb{R}^N)} \,,$$

valable puisque $u^n(t) \to u(t)$ dans $L^1_{\rm loc}(\mathbb{R}^N)$ par convergence dominée (convergence ponctuelle et borne L^∞); rappelons que cette relation consiste simplement à dire que

$$\begin{aligned} ||\nabla u(t)||_{L^1(\mathbb{R}^N)} &= \sup\left\{ \int_{\mathbb{R}^N} \nabla u(t,x)\cdot\Gamma(x)\,dx \,;\ \Gamma \in C_c^\infty(\mathbb{R}^N)^N \,,\ |\Gamma| \leq 1 \right\} \\ &= \sup\left\{ -\int_{\mathbb{R}^N} u(t,x)\mathrm{div}\Gamma(x)\,dx \,;\ \Gamma \in C_c^\infty(\mathbb{R}^N)^N \,,\ |\Gamma| \leq 1 \right\} \end{aligned}$$

et à passer à la limite sur chaque inégalité $-\int_{\mathbb{R}^N} u^n(t,x)\mathrm{div}\Gamma(x)\,dx \leq ||\nabla u^n(t)||_{L^1(\mathbb{R}^N)}$.

Remarque 9.2.2 *Les hypothèses (9.2.2), (9.2.3) et (9.2.4) sont des exemples qui marchent, mais en étudiant les arguments qui précèdent, on peut en trouver d'autres. En particulier, on peut se débarasser de (9.2.3) en supposant que* $\theta_T(x)|x|^{N-1} \to 0$ *lorsque* $|x| \to \infty$. *En effet, en conservant* θ_T *dans le raisonnement qui mène à (9.2.10), on voit que* $w(t,\cdot)$ *décroit plus vite que* $|x|^{-(N-1)}$ *à l'infini (lorsque* u_0 *est à support compact). Comme* ψ_t^{-1} *est globalement lipschitzienne,* $u(t,\cdot)$ *décroit aussi plus vite que* $|x|^{-(N-1)}$ *à l'infini, ce qui permet de justifier (9.2.17) en écrivant Stokes sur une boule de rayon* R *et en faisant* $R \to \infty$ *(car* $|G(t,x,u(t,x))| \leq C_{1,||u||_\infty,1}|u(t,x)|$*). De la même manière, (9.2.12) et Gronwall (plus les estimations uniformes que l'on a sur les dérivées spatiales de* ψ *et* w*) montrent que* $\partial_x w(t,\cdot)$ *décroit aussi plus vite que* $|x|^{-(N-1)}$*, ce qui se répercute encore sur* $\nabla u(t,\cdot)$ *(puisque* $\partial_x \psi_t^{-1}$ *est borné sur* $\mathbb{R}^N$*) et permet de justifier l'intégration par parties qui élimine* $\partial_i[F_i(t,x,u(t,x))I_\eta(\nabla u(t,x))]$ *dans (9.2.20).*

9.2.4 Partie non-locale

Il n'y a rien à faire, car tout ce qui concerne

$$\partial_t u(t,x) + g[u(t,\cdot)](x) = 0\,, \quad u(0,x) \in L^\infty(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N) \tag{9.2.24}$$

est déjà connu et découle de $u(t,x) = K(t,\cdot) * u_0(x)$. Nous rappelons simplement les estimations qui nous seront utiles:

$$||u(t)||_{L^\infty(\mathbb{R}^N)} \leq ||u_0||_{L^\infty(\mathbb{R}^N)} \tag{9.2.25}$$
$$||u(t)||_{L^1(\mathbb{R}^N)} \leq ||u_0||_{L^1(\mathbb{R}^N)} \tag{9.2.26}$$
$$||\nabla u(t)||_{L^1(\mathbb{R}^N)} \leq ||\nabla u_0||_{L^1(\mathbb{R}^N)} \tag{9.2.27}$$
$$||\nabla u(t)||_{L^\infty(\mathbb{R}^N)} \leq ||\nabla K(t)||_{L^1(\mathbb{R}^N)} ||u_0||_{L^\infty(\mathbb{R}^N)} = \mathcal{K}_1 ||u_0||_{L^\infty(\mathbb{R}^N)} t^{-1/\lambda}. \tag{9.2.28}$$

9.2.5 Conclusion

On commence par un lemme, puis on voit que toute la section 8.6 peut être immédiatement adaptée grâce aux résultats tout juste prouvés.

Lemme 9.2.1 *Soit* $(a_k)_{k\geq 0}$ *une suite (éventuellement finie) de réels positifs qui vérifie, pour un certain* B *positif,* $a_{k+1} \leq (a_k + B)e^B$. *Alors, pour tout* $k \geq 1$, *on a* $a_k \leq a_0 e^{kB} + e^{(k+1)B}$.

Preuve du lemme 9.2.1
On montre par récurrence sur k que $a_k \leq a_0 e^{kB} + B\sum_{l=1}^k e^{lB}$. Comme $\sum_{l=1}^k e^{lB} = e^B \frac{e^{kB}-1}{e^B-1} \leq \frac{e^{(k+1)B}}{e^B-1}$ et $\frac{B}{e^B-1} \leq 1$, cela prouvera le lemme. Le rang $k=1$ est évident. Si l'hypothèse est vérifiée au rang k, alors

$$a_{k+1} \leq (a_k + B)e^B \leq a_0 e^{kB} e^B + Be^B \sum_{l=1}^k e^{lB} + Be^B = a_0 e^{(k+1)B} + B\sum_{l=2}^{k+1} e^{lB} + Be^B$$

et le lemme est prouvé. ■

Construction via splitting

On reproduit maintenant sur le problème (9.2.1) la construction par splitting de la section 8.6. On se donne $\delta > 0$ et, partant de u_0 condition initiale régulière (dans $L^\infty(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$), on commence par résoudre (9.2.24) sur $[0,\delta]$ avec un coefficient 2 devant g; cela donne une fonction u^δ qui, à chaque instant, est dans $C_b^1(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$. On peut alors, partant de $t=\delta$ (et non de $t=0$) avec la condition initiale $u^\delta(\delta)$, résoudre (9.2.6), avec un coefficient 2 devant f et h (ce qui n'a pour effet que de multiplier les constantes C et C' par 2, pour ce qui nous intéresse), sur un certain intervalle de temps, et on obtient une fonction (prolongeant u^δ) qui est, à chaque instant, dans $L^\infty(\mathbb{R}^N) \cap W^{1,1}(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$ (voir sous-section 9.2.3); supposons que cet intervalle de temps soit au

moins $[\delta, 2\delta]$. On peut alors partir de $t = 2\delta$ et de la condition initiale $u^\delta(2\delta)$ pour résoudre à nouveau (9.2.24). On continue ainsi de proche en proche; la méthode sera bloquée si, à un instant $n\delta$ avec n impair, la solution de (9.2.6) partant de $t = n\delta$ avec $u^\delta(n\delta)$ comme condition initiale n'est pas définie sur $[n\delta, (n+1)\delta]$; nous allons montrer que, si on se fixe $T_0 > 0$, alors, quitte à prendre δ assez petit, la construction se poursuit au moins jusqu'au temps T_0; nous allons simultanément établir des estimations sur la fonction u^δ construite.

Soit $T_0 > 0$ et $\delta \leq 1$; on note $n\delta < T_0$ le point où l'on est éventuellement bloqué. En posant

$$a_0 = ||u_0||_{L^\infty(\mathbb{R}^N)} \quad \text{et} \quad a_k = \sup_{t\in[(k-1)\delta, k\delta]} ||u^\delta(t)||_{L^\infty(\mathbb{R}^N)},$$

la construction de u^δ, (9.2.16) et (9.2.25) donnent, pour tout $k \in [0, n-1]$,

$$a_{k+1} \leq \left(a_k + 2C'_{T_0+1}\delta\right) e^{2C'_{T_0+1}\delta}$$

(on remarque que, pour tout $k \in [0, n-1]$, les C'_1 correspondant à $f(k\delta + \cdot, \cdot, \cdot)$ et $h(k\delta + \cdot, \cdot, \cdot)$ — c'est à dire, modulo un facteur 2, aux fonctions intervenant effectivement dans nos résolutions successives de (9.2.6) — sont majorés par le C'_{T_0+1} correspondant à f et h). Le lemme 9.2.1 montre alors que, pour tout $k \in [1, n]$,

$$a_k \leq ||u_0||_{L^\infty(\mathbb{R}^N)} e^{2C'_{T_0+1}T_0} + e^{2C'_{T_0+1}(T_0+1)}$$

(on a $(k+1)\delta \leq (n+1)\delta \leq T_0 + 1$) et donc que

$$\forall t \in [0, n\delta],\ ||u^\delta(t)||_{L^\infty(\mathbb{R}^N)} \leq ||u_0||_{L^\infty(\mathbb{R}^N)} e^{2C'_{T_0+1}T_0} + e^{2C'_{T_0+1}(T_0+1)} =: \Lambda_1(u_0, T_0). \tag{9.2.29}$$

Injectée dans (9.2.28) traduite dans l'intervalle $[(n-1)\delta, n\delta]$ (où u_0 doit donc être remplacé par $u^\delta((n-1)\delta)$), cette estimation donne $\Lambda_2(u_0, T_0)$ indépendant de δ et n tel que

$$||\nabla u^\delta(n\delta)||_{L^\infty(\mathbb{R}^N)} \leq \Lambda_2(u_0, T_0)\delta^{-1/\lambda}.$$

Le $M(u^\delta(n\delta))$ correspondant à la résolution de (9.2.6) à partir de $t = n\delta$ est alors majoré (voir (9.2.14)) par $\Lambda_3(u_0, T_0)\delta^{-1/\lambda}$, avec $\Lambda_3(u_0, T_0)$ indépendant de δ et n.
Prenons donc $\delta \leq 1$ tel que $\Lambda_3(u_0, T_0)\delta^{-1/\lambda}\delta < 1$; ce choix ne dépend que de u_0 et T_0 fixés au début. On constate, par les résultats de la partie 9.2.3, que la solution de (9.2.6) partant de $t = n\delta$ et de la condition initiale $u^\delta(n\delta)$ est définie au moins sur $[n\delta, (n+1)\delta]$ et $n\delta$ n'était donc pas un point de blocage; avec ce choix de δ assez petit, u^δ est donc définie au moins sur $[0, T_0] \times \mathbb{R}^N$.

La relation (9.2.29) donne une estimation L^∞ sur u^δ (on peut prendre pour δ une fraction entière de T_0 de sorte que, vu ce qui précède, (9.2.29) est valable pour $n\delta = T_0$). Utilisée dans (9.2.18) à chaque étape du splitting, cette estimation montre, avec (9.2.26), que, pour tout $k \geq 0$ tel que $k\delta < T_0$,

$$\sup_{t\in[k\delta,(k+1)\delta]} ||u^\delta(t)||_{L^1(\mathbb{R}^N)} \leq \left(\sup_{t\in[(k-1)\delta,k\delta]} ||u^\delta(t)||_{L^1(\mathbb{R}^N)} + \Gamma\delta\right) e^{\Gamma\delta}$$

avec $\sup_{t\in[(0-1)\delta,0\delta]} ||u^\delta(t)||_{L^1(\mathbb{R}^N)} = ||u_0||_{L^1(\mathbb{R}^N)}$ et Γ ne dépendant que de (u_0, T_0) (on utilise le fait que le $||\theta_1||_{L^1(\mathbb{R}^N)}$ correspondant à $f(k\delta + \cdot, \cdot, \cdot)$ et $h(k\delta + \cdot, \cdot, \cdot)$ est majoré par le $||\theta_{T_0+1}||_{L^1(\mathbb{R}^N)}$ correspondant à f et h); le lemme 9.2.1 donne ainsi

$$\forall t \in [0, T_0],\ ||u^\delta(t)||_{L^1(\mathbb{R}^N)} \leq ||u_0||_{L^1(\mathbb{R}^N)} e^{\Gamma T_0} + e^{\Gamma(T_0+1)}. \tag{9.2.30}$$

De la même manière, (9.2.29), (9.2.30), (9.2.22) et (9.2.27) donnent Γ' ne dépendant que de (u_0, T_0) tel que

$$\forall t \in [0, T_0],\ ||\nabla u^\delta(t)||_{L^1(\mathbb{R}^N)} \leq ||\nabla u_0||_{L^1(\mathbb{R}^N)} e^{\Gamma' T_0} + e^{\Gamma'(T_0+1)}. \tag{9.2.31}$$

Compacité

(9.2.29), (9.2.30) et (9.2.31) peuvent se résumer en: il existe $\Lambda_4(u_0,T_0)$ tel que, pour tout δ assez petit et tout $t\in[0,T_0]$,

$$||u^\delta(t)||_{L^\infty(\mathbb{R}^N)}+||u^\delta(t)||_{L^1(\mathbb{R}^N)}+||\nabla u^\delta(t)||_{L^1(\mathbb{R}^N)}\leq\Lambda_4(u_0,T_0). \tag{9.2.32}$$

Il suffit donc d'établir l'équicontinuité de $u^\delta:[0,T_0]\to L^1(\mathbb{R}^N)$ pour avoir sa convergence, à une sous-suite près, dans $C([0,T_0];L^1_{\rm loc}(\mathbb{R}^N))$.

Sur les parties où u^δ vérifie l'équation hyperbolique (rappelons que u^δ est régulière sur ces parties), on a

$$\partial_t u^\delta(t,x)=2H(t,x,u^\delta(t,x))-2F(t,x,u^\delta(t,x))\cdot\nabla u^\delta(t,x)$$

et (9.2.32) associé à (9.2.8) et (9.2.2) donne donc $\Lambda_5(u_0,T_0)$ indépendant de δ tel que

$$||\partial_t u^\delta(t)||_{L^1(\mathbb{R}^N)}\leq\Lambda_5(u_0,T_0)$$

soit, par Taylor, pour tout $(t,s)\in[k\delta,(k+1)\delta]$ avec k impair,

$$||u^\delta(t)-u^\delta(s)||_{L^1(\mathbb{R}^N)}\leq\Lambda_5(u_0,T_0)|t-s|. \tag{9.2.33}$$

Sur $[k\delta,(k+1)\delta]$ avec k pair, on a $u^\delta(t)=K(2(t-k\delta))*u^\delta(k\delta)$. L'équicontinuité s'obtient alors à partir de (9.2.32) et (9.2.33) comme en page 157.

Conclusion

On obtient donc une fonction u limite de u^δ dans $C([0,T_0];L^1_{\rm loc}(\mathbb{R}^N))$, pour tout $T_0>0$. Vu (9.2.32), u est en fait à valeurs dans $L^1(\mathbb{R}^N)$ et, puisque u^δ est équicontinue $[0,T_0]\to L^1(\mathbb{R}^N)$, on déduit finalement que $u\in C([0,T_0];L^1(\mathbb{R}^N))$. (9.2.32) montre aussi que $u\in L^\infty(]0,T_0[\times\mathbb{R}^N)$.
La convergence de u^δ vers u permet de prouver le pendant de la proposition 8.6.2, à savoir: pour tout $\gamma\in C_c^\infty(]0,T_0[)$ et tout $\varphi\in\mathcal{S}(\mathbb{R}^N)$,

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^N}u(t,x)\gamma'(t)\varphi(x)+\gamma(t)f(t,x,u(t,x))\cdot\nabla\varphi(x)-u(t,x)\gamma(t)g[\varphi](x)\,dtdx\\ =\quad-\int_0^\infty\int_{\mathbb{R}^N}h(t,x,u(t,x))\gamma(t)\varphi(x)\,dtdx.\end{aligned} \tag{9.2.34}$$

Etant donné que l'on n'a pas supposé que $f(t,\cdot,0)$ est intégrable sur $\mathbb{R}^N$, il faut un peu adapter le raisonnement du reste de la section 8.6.3. On pose $\widetilde{f}(t,x,\zeta)=f(t,x,\zeta)-f(t,x,0)$; comme $(K(t),\nabla K(t))\in L^1(\mathbb{R}^N)$ et $(f(t,\cdot,0),\sum_i\partial_i f_i(t,\cdot,0))\in L^\infty(\mathbb{R}^N)$ pour tout $t>0$, deux dérivations sous l'intégrale montrent que, pour $0<s<t$ et $x\in\mathbb{R}^N$,

$$\nabla K(t-s,\cdot)*f(s,\cdot,0)(x)=\mathrm{div}(K(t-s,\cdot)*f(s,\cdot,0))(x)=K(t-s,\cdot)*\left(\sum_{i=1}^N\partial_i f_i(s,\cdot,0)\right)(x).$$

Ainsi, u est solution de (9.2.5) si et seulement si u vérifie

$$u(t,x)=K(t)*u_0(x)-\int_0^t\nabla K(t-s,\cdot)*\widetilde{f}(s,\cdot,u(s,\cdot))(x)\,ds+\int_0^t K(t-s,\cdot)*\widetilde{h}(s,\cdot,u(s,\cdot))(x)\,ds \tag{9.2.35}$$

où $\widetilde{h}(t,x,\zeta)=h(t,x,\zeta)-\sum_{i=1}^N\partial_i f_i(t,x,0)$ (cela consiste à changer f en $\widetilde{f}$ dans (9.2.1) et à ajuster le second membre en conséquence).

Une intégration par parties (utilisant le fait que $f(t,\cdot,0)$ est bornée) donne, lorsque $\varphi \in \mathcal{S}(\mathbb{R}^N)$,

$$\int_{\mathbb{R}} f(t,x,0)\cdot\nabla\varphi(x)\,dx = -\int_{\mathbb{R}^N}\sum_{i=1}^N \partial_i f_i(t,x,0)\varphi(x)\,dx$$

et (9.2.34) se lit donc

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^N} u(t,x)\gamma'(t)\varphi(x)+\gamma(t)\widetilde{f}(t,x,u(t,x))\cdot\nabla\varphi(x)-u(t,x)\gamma(t)g[\varphi](x)\,dtdx&\\ = \quad -\int_0^\infty\int_{\mathbb{R}^N}\widetilde{h}(t,x,u(t,x))\gamma(t)\varphi(x)\,dtdx.&\end{aligned}\tag{9.2.36}$$

Par (9.2.2) et puisque u est bornée sur $[0,T_0]\times\mathbb{R}^N$, on a

$$|\widetilde{f}(t,x,u(t,x))| = |f(t,x,u(t,x))-f(t,x,0)| \le C_{T_0,||u||_\infty,1}|u(t,x)|$$

et donc $\widetilde{f}(t,\cdot,u(t,\cdot))\in L^1(\mathbb{R}^N)^N$ pour tout $t\in[0,T_0]$; de plus, pour tout $(t,s)\in[0,T_0]$, toujours en utilisant (9.2.2), on a

$$\begin{aligned}|\widetilde{f}(s,x,u(s,x))-\widetilde{f}(t,x,u(t,x))| &\le |\widetilde{f}(s,x,u(s,x))-\widetilde{f}(s,x,u(t,x))|+|\widetilde{f}(s,x,u(t,x))-\widetilde{f}(t,x,u(t,x))|\\ &\le C_{T_0,||u||_\infty,1}|u(s,x)-u(t,x)|+\left|\int_s^t|\partial_t\widetilde{f}(\tau,x,u(t,x))|\,d\tau\right|.\end{aligned}$$

Mais

$$\partial_t\widetilde{f}(\tau,x,u(t,x)) = \partial_t f(\tau,x,u(t,x))-\partial_t f(\tau,x,0)$$

donc, par (9.2.2), $|\partial_t\widetilde{f}(\tau,x,u(t,x))|\le C_{T_0,||u||_\infty,2}|u(t,x)|$ et ainsi

$$|\widetilde{f}(s,x,u(s,x))-\widetilde{f}(t,x,u(t,x))|\le C_{T_0,||u||_\infty,1}|u(s,x)-u(t,x)|+C_{T_0,||u||_\infty,2}|t-s|\,|u(t,x)|.$$

Puisque $u\in C([0,T_0];L^1(\mathbb{R}^N))$, on déduit de ceci que $t\to\widetilde{f}(t,\cdot,u(t,\cdot))$ est dans $C([0,T_0];L^1(\mathbb{R}^N)^N)$. On a $\widetilde{h}(t,x,u(t,x)) = h(t,x,u(t,x))-h(t,x,0)+h(t,x,0)-\sum_{i=1}^N\partial_i f_i(t,x,0)$. De la même manière que précédemment, on voit que $t\to h(t,\cdot,u(t,\cdot))-h(t,\cdot,0)$ est dans $C([0,T_0];L^1(\mathbb{R}^N))$. Par (9.2.4), la fonction $h(t,\cdot,0)-\sum_{i=1}^N\partial_i f_i(t,\cdot,0)$ est majorée en valeur absolue par $C'_{T_0}\theta_{T_0}\in L^1(\mathbb{R}^N)$ et la convergence dominée montre que $t\to h(t,\cdot,0)-\sum_{i=1}^N\partial_i f_i(t,\cdot,0)$ est aussi dans $C([0,T_0];L^1(\mathbb{R}^N))$.

Pour résumer, les fonctions u, $\widetilde{f}(\cdot,\cdot,u)$ et $\widetilde{h}(\cdot,\cdot,u)$ sont toutes continues sur $[0,T_0]$ à valeurs dans $L^1(\mathbb{R}^N)$. On peut alors conclure comme dans le chapitre 8. Soit $\gamma\in C_c^\infty(]0,T_0[)$; puisque $\mathcal{F}^{-1}:L^1(\mathbb{R}^N)\to C_b(\mathbb{R}^N)$ est continue, la fonction

$$\begin{aligned}w:\xi\in\mathbb{R}^N \quad\to\quad &\int_0^{T_0}\gamma'(t)\mathcal{F}^{-1}(u(t))(\xi)+\gamma(t)(2i\pi)\mathcal{F}^{-1}(\widetilde{f}(t,\cdot,u(t)))(\xi)\cdot\xi-\gamma(t)|\xi|^\lambda\mathcal{F}^{-1}(u(t))(\xi)\\ &+\gamma(t)\mathcal{F}^{-1}(\widetilde{h}(t,\cdot,u(t)))(\xi)\,dt\end{aligned}$$

est continue. Par Fubini, en appliquant (9.2.36) à $\varphi=\mathcal{F}^{-1}(\psi)$ pour un $\psi\in\mathcal{D}(\mathbb{R}^N)$, on a

$$\begin{aligned}\int_{\mathbb{R}^N}\int_0^\infty \gamma'(t)\mathcal{F}^{-1}(u(t))(\xi)\psi(\xi)+\gamma(t)(2i\pi)\mathcal{F}^{-1}(\widetilde{f}(t,\cdot,u(t)))(\xi)\cdot\xi\,\psi(\xi)&\\ -\gamma(t)|\xi|^\lambda\mathcal{F}^{-1}(u(t))(\xi)\psi(\xi)+\gamma(t)\mathcal{F}^{-1}(\widetilde{h}(t,\cdot,u(t)))(\xi)\psi(\xi)\,dtd\xi = 0.&\end{aligned}$$

Ceci étant vrai pour tout $\psi\in\mathcal{D}(\mathbb{R}^N)$, cela montre que $w(\xi)=0$ pour tout $\xi\in\mathbb{R}^N$. L'annulation de $w(\xi)$, valable pour tout $\gamma\in C_c^\infty(]0,T_0[)$, nous dit que, au sens des distributions sur $]0,T_0[$,

$$\frac{d}{dt}\mathcal{F}^{-1}(u(t))(\xi) = -|\xi|^\lambda\mathcal{F}^{-1}(u(t))(\xi)+(2i\pi)\mathcal{F}^{-1}(\widetilde{f}(t,\cdot,u(t)))(\xi)\cdot\xi+\mathcal{F}^{-1}(\widetilde{h}(t,\cdot,u(t)))(\xi).$$

Toutes les fonctions intervenant ici sont continues sur $[0,T_0]\times\mathbb{R}^N$ (car u, $\widetilde{f}(\cdot,\cdot,u)$ et $\widetilde{h}(\cdot,\cdot,u)$ sont continues $[0,T_0]\to L^1(\mathbb{R}^N)$ et $\mathcal{F}^{-1}$ est continue $L^1(\mathbb{R}^N)\to C_b(\mathbb{R}^N)$), donc cette égalité est une équation différentielle classique qu'on intègre en

$$\begin{aligned}
\mathcal{F}^{-1}(u(t))(\xi) &= e^{-t|\xi|^\lambda}\mathcal{F}^{-1}(u_0)(\xi)+\int_0^t (2i\pi)e^{-(t-s)|\xi|^\lambda}\mathcal{F}^{-1}(\widetilde{f}(s,\cdot,u(s)))(\xi)\cdot\xi\,ds\\
&\quad+\int_0^t e^{-(t-s)|\xi|^\lambda}\mathcal{F}^{-1}(\widetilde{h}(s,\cdot,u(s)))(\xi)\,ds\\
&= \mathcal{F}^{-1}(K(t))(\xi)\mathcal{F}^{-1}(u_0)(\xi)+\int_0^t (2i\pi)\mathcal{F}^{-1}(K(t-s))(\xi)\xi\cdot\mathcal{F}^{-1}(\widetilde{f}(s,\cdot,u(s)))(\xi)\,ds\\
&\quad+\int_0^t \mathcal{F}^{-1}(K(t-s))(\xi)\mathcal{F}^{-1}(\widetilde{h}(s,\cdot,u(s)))(\xi)\,ds.
\end{aligned}$$

On a assez d'intégrabilité sur les diverses quantités en jeu pour conclure que

$$\begin{aligned}
\mathcal{F}^{-1}(u(t))(\xi) &= \mathcal{F}^{-1}(K(t)*u_0)(\xi)-\int_0^t \mathcal{F}^{-1}(\nabla K(t-s))(\xi)\cdot\mathcal{F}^{-1}(\widetilde{f}(s,\cdot,u(s)))(\xi)\,ds\\
&\quad+\int_0^t \mathcal{F}^{-1}(K(t-s))(\xi)\mathcal{F}^{-1}(\widetilde{h}(s,\cdot,u(s)))(\xi)\,ds\\
&= \mathcal{F}^{-1}(K(t)*u_0)(\xi)-\int_0^t \mathcal{F}^{-1}(\nabla K(t-s)*\widetilde{f}(s,\cdot,u(s)))(\xi)\,ds\\
&\quad+\int_0^t \mathcal{F}^{-1}(K(t-s)*\widetilde{h}(s,\cdot,u(s)))(\xi)\,ds\\
&= \mathcal{F}^{-1}(K(t)*u_0)(\xi)-\mathcal{F}^{-1}\left(\int_0^t \nabla K(t-s)*\widetilde{f}(s,\cdot,u(s))\,ds\right)(\xi)\\
&\quad+\mathcal{F}^{-1}\left(\int_0^t K(t-s)*\widetilde{h}(s,\cdot,u(s))\,ds\right)(\xi)
\end{aligned}$$

et l'injectivité de $\mathcal{F}^{-1}$ montre bien que u vérifie (9.2.35), et donc (9.2.5)

La conclusion lorsque u_0 n'est pas régulière se fait exactement comme dans le chapitre 8: (9.2.29) implique, pour une donnée initiale régulière et lorsqu'on fait $\delta\to 0$, une estimation $L^\infty(]0,T_0[\times\mathbb{R}^N)$ sur la solution précédemment construite, estimation qui ne dépend que de la norme infinie de la donnée initiale; donc, si on prend $u_0\in L^\infty(\mathbb{R}^N)$ et qu'on l'approche presque partout par des $u_{0,n}$ réguliers bornés dans $L^\infty(\mathbb{R}^N)$, on voit que les solutions u^n correspondantes restent bornées dans $L^\infty(]0,T_0[\times\mathbb{R}^N)$ pour tout $T_0>0$; cette borne nous donne (via les estimations mentionnées en sous-section 9.2.2) une borne sur les dérivées de u^n et donc, à une sous-suite près, une convergence ponctuelle vers un u qui est dans $L^\infty(]0,T_0[\times\mathbb{R}^N)$ pour tout $T_0>0$; on peut alors aisément passer à la limite, par convergence dominée, dans (9.2.5) appliqué à u^n pour prouver que u est solution de (9.2.1) avec condition initiale u_0.

Remarque 9.2.3 *Comme signalé précédemment, on peut aussi prouver l'existence, l'unicité et la régularité d'une solution à (9.2.1) avec des hypothèses différentes sur f et h. Par exemple, si f ne dépend pas de x, il suffit que h vérifie (9.2.2) (cette hypothèse n'impose alors rien sur f) et*

$$|h(t,x,\zeta)|\le K_T(1+|\zeta|)\ \textit{sur}\ [0,T]\times\mathbb{R}^N\times\mathbb{R} \tag{9.2.37}$$

pour que le problème soit soluble. En effet, en prenant $\gamma_\varepsilon\in C_c^\infty(\mathbb{R}^N)$ majorée par 1 et qui tend vers 1 quand $\varepsilon\to 0$, on peut considérer $h_\varepsilon(t,x,\zeta)=\gamma_\varepsilon(x)h(t,x,\zeta)$; cette fonction ne vérifie pas (9.2.4) mais on peut néanmoins reproduire la méthode précédente (le problème étant, à ε fixé, de borner le terme $\nabla_x H=\nabla_x(\gamma_\varepsilon h)$ dans (9.2.21), ce qui se fait en utilisant la borne L^∞ que l'on a sur la fonction u

considérée); on peut donc construire une solution u_ε à (9.2.1) avec h_ε au lieu de h, et (9.2.37) montre aisément (en relisant (9.2.11)) que les estimations L^∞ sur cette solution ne dépendent pas de ε; ces estimations donnant des bornes sur les dérivées de u_ε, on obtient alors suffisamment de compacité pour passer à la limite dans l'équation (9.2.5) satisfaite par u_ε et trouver une solution à (9.2.1).

Chapitre 10

Vanishing non-local regularization of a scalar conservation law

Reference: J. Droniou. *Electron. J. Differential Equations* **2003** (2003), no. 117, 1-20.

Abstract We prove that the solution to the regularization of a scalar conservation law by a fractional power of the Laplacian converges, as the regularization vanishes, to the entropy solution of the hyperbolic problem. We also give an error estimate when the initial condition has bounded variation.

Keywords: scalar conservation law, vanishing regularization, fractal operator, error estimate.

Mathematics Subject Classification: 35L65, 35S30, 35A35, 35B20

10.1 Introduction

We consider the problem

$$\begin{cases} \partial_t u^\varepsilon(t,x) + \operatorname{div}(f(u^\varepsilon))(t,x) + \varepsilon g[u^\varepsilon(t,\cdot)](x) = 0\,, & t>0\,,\ x\in\mathbb{R}^N\,,\\ u^\varepsilon(0,x) = u_0(x)\,, & x\in\mathbb{R}^N\,, \end{cases} \tag{10.1.1}$$

where $f=(f_1,\dots,f_N)\in (C^\infty(\mathbb{R}))^N$, $u_0\in L^\infty(\mathbb{R}^N)$ and g is the non-local operator defined through Fourier transform by

$$\mathcal{F}(g[u^\varepsilon(t,\cdot)])(\xi) = |\xi|^\lambda \mathcal{F}(u^\varepsilon(t,\cdot))(\xi)\,, \quad \text{with } \lambda\in]1,2]. \tag{10.1.2}$$

In the case $\varepsilon=0$, this equation reduces to the classical scalar conservation law

$$\begin{cases} \partial_t u(t,x) + \operatorname{div}(f(u))(t,x) = 0\,, & t>0\,,\ x\in\mathbb{R}^N\,,\\ u(0,x) = u_0(x)\,, & x\in\mathbb{R}^N. \end{cases} \tag{10.1.3}$$

Existence and uniqueness of a solution to this equation, in the L^∞ framework, has been established by Krushkov [60]; it relies on so-called "entropy solutions", which must satisfy particular inequalities. The case $\lambda=2$ and $\varepsilon>0$ in (10.1.1) corresponds to $g[u^\varepsilon(t,\cdot)](x) = -(2\pi)^2\Delta u^\varepsilon(t,x)$ and is called the parabolic regularization of (10.1.3). In this situation, existence, uniqueness and regularity of solutions to this equation are well-known (see e.g. [62]), and an entropy solution of (10.1.3) can be obtained by proving that, as $\varepsilon\to 0$, the solution to this parabolic regularization converges to a function which satifies the entropy inequalities of (10.1.3).
For general $\lambda\in]1,2]$ and $\varepsilon>0$, in which case g is a fractional power of the Laplacian, the study of (10.1.1) less classical, though motivated by physical problems of detonation (see [25], [23] for example),

hydrodynamics, molecular biology, etc... (see the introduction of [7] and references therein). A number of papers ([7], [8]...) have studied this equation (also called "fractal conservation law"). The existence results in [7] for (10.1.1) give global solutions in the case $N = 1$, but which are not very regular, or local (in time) solutions for general $N \geq 1$ and small initial data, but still not regular (in Morrey spaces). In [8] or [9], the authors consider a parabolic regularization of (10.1.1), that is to say they add a Laplacian operator to the equation; thanks to this second order operator, a global solution is obtained and regularity results can be proved. These papers are mainly interested in asymptotic behaviours for this equation.
However, one could consider (10.1.1) as a (possible) regularization of (10.1.3), without having to add another term. In this case, a natural space for the initial data is $L^\infty(\mathbb{R}^N)$, and the question is whether or not (10.1.1) gives rise to a solution which is regular for $t > 0$ (i.e. whether or not g has the same effect on the regularity as $-\Delta$). It has been proved in [36] that this indeed happens: there exists a unique bounded solution to (10.1.1), in a suitable sense, and this solution belongs to $C^\infty(]0,\infty[\times\mathbb{R}^N)$. It is constructed via a splitting method, and inherits thus all the properties that are common to both the conservation law and the equation $\partial_t v + \varepsilon g[v] = 0$, such as essential bounds, comparison and contraction principles, etc...; its regularity is proved using the Banach fixed point theorem on Duhamel's formula for $\partial_t u^\varepsilon + \varepsilon g[u^\varepsilon] = -\mathrm{div}(f(u^\varepsilon))$.
Once it has been established that (10.1.1) has the same regularizing effect, with respect to (10.1.3), as the parabolic equation, the next question is to know if, as in the parabolic case, the solution to (10.1.1) remains close to the solution of (10.1.3) for small ε. This is the aim of the present work, and our main results are the following.

Theorem 10.1.1 *Let $u_0 \in L^\infty(\mathbb{R}^N)$. As $\varepsilon \to 0$, the solution to (10.1.1) converges to the entropy solution of (10.1.3), in $C([0,T];L^1_{\mathrm{loc}}(\mathbb{R}^N))$ for all $T > 0$.*

Remark 10.1.1 *This theorem (as well as the results of [36]) is valid for more general g (roughly speaking, the methods work for operators whose kernels are approximate units — see subsection 10.2.1). For example, sums of operators of the kind (10.1.2) (or more general Lévy operators) can be considered, with, as a special case, the equation $\partial_t u^\varepsilon + \mathrm{div}(f(u^\varepsilon)) + \varepsilon g[u^\varepsilon] - \varepsilon\Delta u^\varepsilon = 0$ (as in [8], [9]).*

As a by-product of the proof of Theorem 10.1.1, we also obtain the following error estimate.

Theorem 10.1.2 *Let $u_0 \in L^\infty(\mathbb{R}^N) \cap L^1(\mathbb{R}^N) \cap BV(\mathbb{R}^N)$, u^ε be the solution to (10.1.1) and u be the entropy solution to (10.1.3). Then, for all $T > 0$, $||u^\varepsilon - u||_{C([0,T];L^1(\mathbb{R}^N))} = \mathcal{O}(\varepsilon^{1/\lambda})$.*

Remark 10.1.2 *This result in the case of parabolic regularization ($\lambda = 2$) has already been proved in [61]. The special feature of Theorem 10.1.2 is that it establishes an elegant relationship between the rate of convergence and the order of the operator chosen for the regularization of (10.1.3); in fact, this error estimate is optimal (see Remark 10.2.1).*

Remark 10.1.3 *Note that, since u^ε and u are bounded (by $||u_0||_\infty$), a convergence in $L^1(\mathbb{R}^N)$ (respectively in $L^1_{\mathrm{loc}}(\mathbb{R}^N)$) implies, by interpolation, a convergence in $L^p(\mathbb{R}^N)$ (respectively in $L^p_{\mathrm{loc}}(\mathbb{R}^N)$) for all finite p. For example, under the hypotheses of Theorem 10.1.2, we have, for all $p \in [1,\infty[$ and all $T > 0$, $||u^\varepsilon - u||_{C([0,T];L^p(\mathbb{R}^N))} = \mathcal{O}(\varepsilon^{\frac{1}{p\lambda}})$.*

The paper is organized as follows. In the next section, we prove approximate entropy inequalities for u^ε; this function has been obtained in [36], using a splitting method, as a limit of explicit functions: we first prove approximate entropy inequalities on those explicit functions, and then deduce the corresponding inequalities for u^ε; it is not clear that these estimates could be inferred from the methods of [7]. In Section 10.3, we use Krushkov's classical doubling variable technique to combine the approximate entropy inequalities on u^ε and the entropy inequalities on u, which gives an estimate on $|u^\varepsilon - u|$ and proves Theorem 10.1.1. Section 10.4 is devoted to the proof of Theorem 10.1.2, which is an easy consequence of the estimates obtained in Section 10.3. We have gathered, in Section 10.5, some results concerning g and its kernel, which we use in the rest of the work.

10.2 Approximate entropy inequalities for (10.1.1)

In order to prove approximate entropy inequalities for u^ε, we need to recall the construction of this function (see [36]).

10.2.1 Construction of u^ε

The solution to $\partial_t v + \varepsilon g[v] = 0$ with initial condition $v(0,\cdot) = v_0$ is (at least formally) given by $v(t,\cdot) = K_\varepsilon(t,\cdot) * v_0$, where

$$K_\varepsilon(t,x) = \mathcal{F}^{-1}(e^{-\varepsilon t|\cdot|^\lambda})(x).$$

The main property of this kernel is that $(K_\varepsilon(t,\cdot))_{t\to 0}$ is an approximate unit. This means that $K_\varepsilon(t,\cdot)$ is non-negative (see [65]), has integral equal to 1 and that, for all $\nu > 0$, $\int_{|y|\geq\nu} K_\varepsilon(t,y)\,dy \to 0$ as $t \to 0$ ([1]).

We assume here that $u_0 \in C_c^\infty(\mathbb{R}^N)$ (though $u_0 \in L^\infty(\mathbb{R}^N) \cap L^1(\mathbb{R}^N) \cap BV(\mathbb{R}^N)$ would be enough). Let $\delta > 0$ and $u^{\varepsilon,\delta} : [0,\infty[\times\mathbb{R}^N \to \mathbb{R}$ be defined by $u^{\varepsilon,\delta}(0,\cdot) = u_0$ and

- for all even p, $u^{\varepsilon,\delta}$ is, on $]p\delta,(p+1)\delta]\times\mathbb{R}^N$, the solution to $\partial_t u^{\varepsilon,\delta} + 2\varepsilon g[u^{\varepsilon,\delta}] = 0$ with initial condition $u^{\varepsilon,\delta}(p\delta,\cdot)$, that is to say $u^{\varepsilon,\delta}(t,x) = K_\varepsilon(2(t-p\delta),\cdot) * u^{\varepsilon,\delta}(p\delta,\cdot)(x)$ for $(t,x) \in]p\delta,(p+1)\delta] \times \mathbb{R}^N$.
- for all odd p, $u^{\varepsilon,\delta}$ is, on $]p\delta,(p+1)\delta] \times \mathbb{R}^N$, the entropy solution to $\partial_t u^{\varepsilon,\delta} + 2\mathrm{div}(f(u^{\varepsilon,\delta})) = 0$ with initial condition $u^{\varepsilon,\delta}(p\delta,\cdot)$.

We have then $u^{\varepsilon,\delta} \in C([0,\infty[;L^1(\mathbb{R}^N))$ with $u^{\varepsilon,\delta}(0,\cdot) = u_0$ and

$$\begin{aligned}\forall t \geq 0\,,\ ||u^{\varepsilon,\delta}(t,\cdot)||_{L^\infty(\mathbb{R}^N)} \leq ||u_0||_{L^\infty(\mathbb{R}^N)}\,,\ ||u^{\varepsilon,\delta}(t,\cdot)||_{L^1(\mathbb{R}^N)} \leq ||u_0||_{L^1(\mathbb{R}^N)} \\ \text{and } |u^{\varepsilon,\delta}(t,\cdot)|_{BV(\mathbb{R}^N)} \leq ||\nabla u_0||_{L^1(\mathbb{R}^N)}.\end{aligned} \tag{10.2.1}$$

It has been proved in [36] that $u^{\varepsilon,\delta}$ converges, as $\delta \to 0$ and in $C([0,T];L^1_{\mathrm{loc}}(\mathbb{R}^N))$ for all $T > 0$, to the solution u^ε of (10.1.1). It has also been noticed that, for δ small enough, $u^{\varepsilon,\delta}$ is in fact, on $[p\delta,(p+1)\delta]\times\mathbb{R}^N$ for all odd p, a regular solution to $\partial_t u^{\varepsilon,\delta} + 2\mathrm{div}(f(u^{\varepsilon,\delta})) = 0$; moreover, for such δ and all $t \geq 0$, $u^{\varepsilon,\delta}(t,\cdot)$ is regular.

The results of [36] are stated in dimension $N = 1$ and with g instead of εg (with K_1 instead of K_ε) but, as indicated in this reference, they are valid in any dimension $N \geq 1$ and substituting K_ε for K_1 (K_ε has the same properties, for a fixed $\varepsilon > 0$, as K_1: in fact, $K_\varepsilon(t,x) = K_1(\varepsilon t,x)$), they also hold with εg.

Remark 10.2.1 *If $f = 0$, the solution to (10.1.1) is $u^\varepsilon(t,x) = K_\varepsilon(t,\cdot) * u_0(x)$ and the solution to (10.1.3) is $u(t,x) = u_0(x)$. Taking, for example, u_0 the characteristic function of $[-1,1]^N$, some easy computations and the homogeneity property $K_\varepsilon(1,x) = K_1(\varepsilon,x) = \varepsilon^{-N/\lambda}K_1(1,\varepsilon^{-1/\lambda}x)$ (see footnote 1 on page 184) show that $||u^\varepsilon(1,\cdot)||_{L^1(\mathbb{R}^N\setminus[-1,1]^N)} \geq c\varepsilon^{1/\lambda}$ for some $c > 0$. Hence, the estimate of Theorem 10.1.2 is optimal.*

10.2.2 The approximate entropy inequalities

We now establish the following approximate entropy inequalities for the solution to (10.1.1).

Proposition 10.2.1 *Assume that $u_0 \in L^\infty(\mathbb{R}^N)$ and let u^ε be the solution to (10.1.1). Let $\eta : \mathbb{R} \to \mathbb{R}$ be a regular convex function and $\phi = (\phi_1,\ldots,\phi_N)$ such that $\phi_i' = \eta' f_i'$. Then, for all non-negative*

[1] This comes from $K_\varepsilon(t,x) = t^{-N/\lambda}K_\varepsilon(1,t^{-1/\lambda}x)$ (change of variable in the definition of K_ε) and from $K_\varepsilon(1,\cdot) \in L^1(\mathbb{R}^N)$ (because the $N+1$ first derivatives of $\xi \to e^{-\varepsilon|\xi|^\lambda}$ are integrable on $\mathbb{R}^N$).

$\varphi \in C_c^\infty([0,\infty[\times\mathbb{R}^N)$, *we have*

$$\int_0^\infty \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x))\partial_t\varphi(t,x) + \phi(u^\varepsilon(t,x))\cdot\nabla\varphi(t,x)\,dtdx + \int_{\mathbb{R}^N} \eta(u_0(x))\varphi(0,x)\,dx$$
$$\geq \quad \varepsilon \int_0^\infty \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x)) g[\varphi(t,\cdot)](x)\,dtdx. \quad (10.2.2)$$

Remark 10.2.2 *If $\varphi \in C_c^\infty([0,\infty[\times\mathbb{R}^N)$, then $t \in [0,\infty[\to \nabla\varphi(t,\cdot) \in L^1(\mathbb{R}^N)^N$ and $t \in [0,\infty[\to \Delta\varphi(t,\cdot) \in L^1(\mathbb{R}^N)$ are continuous; hence, by Lemma 10.5.1 and the linearity of g, the function $t \in [0,\infty[\to g[\varphi(t,\cdot)] \in L^1(\mathbb{R}^N)$ is continuous. In particular, since $\varphi(t,\cdot) = 0$ for t large enough, $(t,x) \to g[\varphi(t,\cdot)](x)$ is integrable on $]0,\infty[\times\mathbb{R}^N$.*

Proof of Proposition 10.2.1
Note that (10.2.2) with η or $\eta - \eta(0)$ are the same inequalities. Indeed, the entropy fluxes ϕ associated to η and $\eta - \eta(0)$ are identical,

$$\int_0^\infty \int_{\mathbb{R}^N} \partial_t\varphi(t,x)\,dtdx + \int_{\mathbb{R}^N} \varphi(0,x)\,dx = 0$$

and, since $g[\varphi(t,\cdot)] \in L^1(\mathbb{R}^N)$ for all $t \geq 0$ (see Lemma 10.5.1),

$$\int_{\mathbb{R}^N} g[\varphi(t,\cdot)](x)\,dx = \mathcal{F}(g[\varphi(t,\cdot)])(0) = (|\cdot|^\lambda \mathcal{F}(\varphi(t,\cdot)))(0) = 0.$$

Hence, there is no loss in generality if we assume that $\eta(0) = 0$, which we do from now on.
The proof is done in two steps. We first suppose that the initial condition is regular, in which case we establish approximate entropy inequalities for the functions $u^{\varepsilon,\delta}$ constructed in subsection 10.2.1, and we deduce the result of the proposition by letting $\delta \to 0$. We then prove the proposition for general initial conditions.

Step 1: assume that $u_0 \in C_c^\infty(\mathbb{R}^N)$.
We take δ small enough so that $u^{\varepsilon,\delta}$ is, on $[p\delta,(p+1)\delta] \times \mathbb{R}^N$ for all odd p, a regular solution to $\partial_t u^{\varepsilon,\delta} + 2\mathrm{div}(f(u^{\varepsilon,\delta})) = 0$. For odd p, we therefore have

$$\int_{p\delta}^{(p+1)\delta} \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x) + 2\phi(u^{\varepsilon,\delta}(t,x))\cdot\nabla\varphi(t,x)\,dtdx$$
$$= \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}((p+1)\delta,x))\varphi((p+1)\delta,x)\,dx - \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(p\delta,x))\varphi(p\delta,x)\,dx.$$

Summing on odd p's (note that, since the support of φ is compact, this sum is finite), and defining χ_δ as the characteristic function of $\cup_{\text{odd } p}]p\delta,(p+1)\delta]$, we find

$$\int_0^\infty \int_{\mathbb{R}^N} \left(\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x) + 2\phi(u^{\varepsilon,\delta}(t,x))\cdot\nabla\varphi(t,x)\right)\chi_\delta(t)\,dtdx = \sum_{\text{odd } p} (a_{p+1} - a_p)$$

where

$$a_p = \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(p\delta,x))\varphi(p\delta,x)\,dx.$$

Since

$$\sum_{\text{odd } p} (a_{p+1} - a_p) = \sum_{\text{even } p\,,\ p\geq 2} a_p - \sum_{\text{odd } p} a_p = \sum_{\text{even } p} (a_p - a_{p+1}) - a_0\,,$$

we deduce that

$$\int_0^\infty \int_{\mathbb{R}^N} \left(\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x) + 2\phi(u^{\varepsilon,\delta}(t,x))\cdot\nabla\varphi(t,x)\right)\chi_\delta(t)\,dtdx + \int_{\mathbb{R}^N}\eta(u_0(x))\varphi(0,x)\,dx$$
$$= \sum_{\text{even } p}(a_p - a_{p+1}). \quad (10.2.3)$$

If p is even, we have, by definition, $u^{\varepsilon,\delta}((p+1)\delta) = K_\varepsilon(2\delta) * u^{\varepsilon,\delta}(p\delta)$ (it is convenient, because of the convolution product, to omit the space variable). Since η is convex and $K_\varepsilon(2\delta)$ is positive with integral equal to 1, Jensen's inequality gives then $\eta(u^{\varepsilon,\delta}((p+1)\delta)) \le K_\varepsilon(2\delta) * \eta(u^{\varepsilon,\delta}(p\delta))$. The function φ being non-negative, we deduce that $\eta(u^{\varepsilon,\delta}((p+1)\delta))\varphi((p+1)\delta) \le K_\varepsilon(2\delta) * \eta(u^{\varepsilon,\delta}(p\delta))\varphi((p+1)\delta)$ and thus

$$\begin{aligned} a_{p+1} - a_p &\le \int_{\mathbb{R}^N} K_\varepsilon(2\delta) * \eta(u^{\varepsilon,\delta}(p\delta))\varphi((p+1)\delta) - \eta(u^{\varepsilon,\delta}(p\delta))\varphi(p\delta) \\ &= \int_{\mathbb{R}^N} F((p+1)\delta)\varphi((p+1)\delta) - F(p\delta)\varphi(p\delta) \end{aligned}$$

where $F(p\delta) = \eta(u^{\varepsilon,\delta}(p\delta))$ and, for $t \in]p\delta,(p+1)\delta]$, $F(t) = K_\varepsilon(2(t-p\delta)) * F(p\delta)$ (i.e. F satisfies $\partial_t F + 2\varepsilon g[F] = 0$ on $]p\delta,(p+1)\delta]$). We have $\eta(0) = 0$, so that, letting C_0 be the Lipschitz constant of η on $[-||u_0||_\infty, ||u_0||_\infty]$ and using (10.2.1),

$$||F(p\delta)||_{L^1(\mathbb{R}^N)} \le C_0||u^{\varepsilon,\delta}(p\delta)||_{L^1(\mathbb{R}^N)} \le C_0||u_0||_{L^1(\mathbb{R}^N)} \quad (10.2.4)$$
$$||\nabla F(p\delta)||_{L^1(\mathbb{R}^N)} \le C_0||\nabla u^{\varepsilon,\delta}(p\delta)||_{L^1(\mathbb{R}^N)} \le C_0||\nabla u_0||_{L^1(\mathbb{R}^N)} \quad (10.2.5)$$

(recall that δ is small enough so that $u^{\varepsilon,\delta}(t,\cdot)$ is regular for all $t \ge 0$). Lemma 10.5.2 in the appendix enables then to write

$$\begin{aligned} a_{p+1} - a_p &\le \int_{\mathbb{R}^N} F((p+1)\delta)\varphi((p+1)\delta) - F(p\delta)\varphi(p\delta) \\ &= \int_{p\delta}^{(p+1)\delta}\int_{\mathbb{R}^N} F(t,x)\partial_t\varphi(t,x) - 2\varepsilon F(t,x)g[\varphi(t,\cdot)](x)\,dtdx. \end{aligned} \quad (10.2.6)$$

We have, by Lemma 10.5.3 in the appendix, for all $\nu > 0$ and all $t \in]p\delta,(p+1)\delta]$,

$$\begin{aligned} ||F(t) - \eta(u^{\varepsilon,\delta}(p\delta))||_{L^1(\mathbb{R}^N)} &= ||K_\varepsilon(2(t-p\delta)) * F(p\delta) - F(p\delta)||_{L^1(\mathbb{R}^N)} \\ &\le 2||F(p\delta)||_{L^1(\mathbb{R}^N)}\int_{|y|\ge\nu} K_\varepsilon(2(t-p\delta),y)\,dy + \nu||\nabla F(p\delta)||_{L^1(\mathbb{R}^N)}. \end{aligned}$$

Using (10.2.4) and (10.2.5), we deduce

$$||F(t) - \eta(u^{\varepsilon,\delta}(p\delta))||_{L^1(\mathbb{R}^N)} \le 2C_0||u_0||_{L^1(\mathbb{R}^N)}\sup_{0<s\le 2\delta}\int_{|y|\ge\nu} K_\varepsilon(s,y)\,dy + C_0\nu||\nabla u_0||_{L^1(\mathbb{R}^N)}. \quad (10.2.7)$$

We have $|\eta(u^{\varepsilon,\delta}(t)) - \eta(u^{\varepsilon,\delta}(p\delta))| \le C_0|u^{\varepsilon,\delta}(t) - u^{\varepsilon,\delta}(p\delta)|$ and, for $t \in]p\delta,(p+1)\delta]$, $u^{\varepsilon,\delta}(t) = K_\varepsilon(2(t-p\delta)) * u^{\varepsilon,\delta}(p\delta)$. Hence, Lemma 10.5.3 and (10.2.1) give, for all $t \in]p\delta,(p+1)\delta]$ and all $\nu > 0$,

$$||\eta(u^{\varepsilon,\delta}(t)) - \eta(u^{\varepsilon,\delta}(p\delta))||_{L^1(\mathbb{R}^N)} \le 2C_0||u_0||_{L^1(\mathbb{R}^N)}\sup_{0<s\le 2\delta}\int_{|y|\ge\nu} K_\varepsilon(s,y)\,dy + C_0\nu||\nabla u_0||_{L^1(\mathbb{R}^N)}. \quad (10.2.8)$$

Gathering (10.2.7) and (10.2.8), we find, for all $t \in]p\delta,(p+1)\delta]$ and all $\nu > 0$,

$$||F(t) - \eta(u^{\varepsilon,\delta}(t))||_{L^1(\mathbb{R}^N)} \le 4C_0||u_0||_{L^1(\mathbb{R}^N)}\sup_{0<s\le 2\delta}\int_{|y|\ge\nu} K_\varepsilon(s,y)\,dy + 2C_0\nu||\nabla u_0||_{L^1(\mathbb{R}^N)} = \omega_\varepsilon(\delta,\nu)$$

with $\lim_{\nu\to 0}(\lim_{\delta\to 0}\omega_\varepsilon(\delta,\nu))=0$ (because $(K_\varepsilon(t,\cdot))_{t\to 0}$ is an approximate unit). Using this inequality in (10.2.6), we obtain, for all $\nu>0$,

$$\begin{aligned} a_{p+1}-a_p &\leq \int_{p\delta}^{(p+1)\delta}\int_{\mathbb{R}^N}\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x)-2\varepsilon\eta(u^{\varepsilon,\delta}(t,x))g[\varphi(t,\cdot)](x)\,dtdx \\ &\quad+\omega_\varepsilon(\delta,\nu)\int_{p\delta}^{(p+1)\delta}\left(||\partial_t\varphi(t,\cdot)||_{L^\infty(\mathbb{R}^N)}+2\varepsilon||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)}\right)\,dt. \end{aligned} \tag{10.2.9}$$

Note that, by definition of g, we can write

$$\begin{aligned} ||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)} &\leq \left|\left|\,|\cdot|^\lambda\mathcal{F}(\varphi(t,\cdot))\right|\right|_{L^1(\mathbb{R}^N)} \\ &= \left|\left|\frac{|\cdot|^\lambda}{1+(2\pi|\cdot|)^{2(N+1)}}\mathcal{F}\left(\varphi(t,\cdot)+(-\Delta)^{N+1}\varphi(t,\cdot)\right)\right|\right|_{L^1(\mathbb{R}^N)} \\ &\leq \left|\left|\frac{|\cdot|^\lambda}{1+(2\pi|\cdot|)^{2(N+1)}}\right|\right|_{L^1(\mathbb{R}^N)}||\mathcal{F}\left(\varphi(t,\cdot)+(-\Delta)^{N+1}\varphi(t,\cdot)\right)||_{L^\infty(\mathbb{R}^N)} \\ &\leq \left|\left|\frac{|\cdot|^\lambda}{1+(2\pi|\cdot|)^{2(N+1)}}\right|\right|_{L^1(\mathbb{R}^N)}||\varphi(t,\cdot)+(-\Delta)^{N+1}\varphi(t,\cdot)||_{L^1(\mathbb{R}^N)} \end{aligned} \tag{10.2.10}$$

with $\lambda-2(N+1)\leq -2N<-N$. Hence, $t\to||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)}$ is integrable on $[0,\infty[$ (in fact, this function is continuous and null for t large).

Summing (10.2.9) on even p's and coming back to (10.2.3), we deduce

$$\begin{aligned} &\int_0^\infty\int_{\mathbb{R}^N}\left(\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x)+2\phi(u^{\varepsilon,\delta}(t,x))\cdot\nabla\varphi(t,x)\right)\chi_\delta(t)\,dtdx+\int_{\mathbb{R}^N}\eta(u_0(x))\varphi(0,x)\,dx \\ &\quad\geq -\int_0^\infty\int_{\mathbb{R}^N}\left(\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x)-2\varepsilon\eta(u^{\varepsilon,\delta}(t,x))g[\varphi(t,\cdot)](x)\right)(1-\chi_\delta(t))\,dtdx \\ &\qquad-\omega_\varepsilon(\delta,\nu)\int_0^\infty\left(||\partial_t\varphi(t,\cdot)||_{L^\infty(\mathbb{R}^N)}+2\varepsilon||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)}\right)\,dt \end{aligned}$$

(note that $1-\chi_\delta$ is the characteristic function of $\cup_{\text{even }p}]p\delta,(p+1)\delta]$), that is to say

$$\begin{aligned} &\int_0^\infty\int_{\mathbb{R}^N}\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x)+2\phi(u^{\varepsilon,\delta}(t,x))\cdot\nabla\varphi(t,x)\chi_\delta(t)\,dtdx+\int_{\mathbb{R}^N}\eta(u_0(x))\varphi(0,x)\,dx \\ &\qquad\geq 2\varepsilon\int_0^\infty\int_{\mathbb{R}^N}\eta(u^{\varepsilon,\delta}(t,x))g[\varphi(t,\cdot)](x)(1-\chi_\delta(t))\,dtdx \\ &\qquad\quad-\omega_\varepsilon(\delta,\nu)\int_0^\infty\left(||\partial_t\varphi(t,\cdot)||_{L^\infty(\mathbb{R}^N)}+2\varepsilon||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)}\right)\,dt. \end{aligned} \tag{10.2.11}$$

As $\delta\to 0$, we have $u^{\varepsilon,\delta}\to u^\varepsilon$ in $C([0,T];L^1_{\text{loc}}(\mathbb{R}^N))$ for all $T>0$; hence, η and ϕ being Lipschitz-continuous on $[-||u_0||_\infty,||u_0||_\infty]$ and $u^{\varepsilon,\delta}$ taking its values in this interval, we deduce that $\eta(u^{\varepsilon,\delta})\to\eta(u^\varepsilon)$ and $\phi(u^{\varepsilon,\delta})\to\phi(u^\varepsilon)$, as $\delta\to 0$ and in $C([0,T];L^1_{\text{loc}}(\mathbb{R}^N))$ for all $T>0$. This allows to see that, as $\delta\to 0$,

$$\int_0^\infty\int_{\mathbb{R}^N}\eta(u^{\varepsilon,\delta}(t,x))\partial_t\varphi(t,x)\,dtdx\to\int_0^\infty\int_{\mathbb{R}^N}\eta(u^\varepsilon(t,x))\partial_t\varphi(t,x)\,dtdx. \tag{10.2.12}$$

We also deduce that

$$\int_{\mathbb{R}^N}\phi(u^{\varepsilon,\delta}(\cdot,x))\cdot\nabla\varphi(\cdot,x)\,dx\to\int_{\mathbb{R}^N}\phi(u^\varepsilon(\cdot,x))\cdot\nabla\varphi(\cdot,x)\,dx\quad\text{ in }L^\infty_{\text{loc}}([0,\infty[),$$

and thus in $L^1(]0,\infty[)$ (these functions are null for t large). Since $\chi_\delta \to 1/2$ in $L^\infty(]0,\infty[)$ weak-$*$, this implies

$$\int_0^\infty \int_{\mathbb{R}^N} 2\phi(u^{\varepsilon,\delta}(t,x)) \cdot \nabla\varphi(t,x)\chi_\delta(t)\,dtdx \to \int_0^\infty \int_{\mathbb{R}^N} \phi(u^\varepsilon(t,x)) \cdot \nabla\varphi(t,x)\,dtdx. \tag{10.2.13}$$

For all $M \geq 0$, we have

$$\begin{aligned}&\left| \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(t,x))g[\varphi(t,\cdot)](x)\,dx - \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x))g[\varphi(t,\cdot)](x)\,dx \right| \\ &\quad \leq ||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)} \int_{|x|\leq M} |\eta(u^{\varepsilon,\delta}(t,x)) - \eta(u^\varepsilon(t,x))|\,dx + 2C_1 \int_{|x|\geq M} |g[\varphi(t,\cdot)](x)|\,dx \end{aligned} \tag{10.2.14}$$

with $C_1 = \sup\{|\eta(z)|\,,\ |z| \leq ||u_0||_\infty\}$. By Remark 10.2.2, the function $t \in [0,\infty[\to g[\varphi(t,\cdot)] \in L^1(\mathbb{R}^N)$ is continuous and null for t large enough; this implies that $\{g[\varphi(t,\cdot)]\,,\ t \geq 0\}$ is compact in $L^1(\mathbb{R}^N)$, and thus, by Vitali's theorem, that

$$\lim_{M\to\infty} \int_{|x|\geq M} |g[\varphi(t,\cdot)](x)|\,dx = 0 \quad \text{uniformly with respect to } t \geq 0.$$

For a fixed M, we have $\int_{|x|\leq M} |\eta(u^{\varepsilon,\delta}(t,x)) - \eta(u^\varepsilon(t,x))|\,dx \to 0$ as $\delta \to 0$, locally uniformly with respect to $t \geq 0$; since $\sup_{t\geq 0} ||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)} < \infty$ (see (10.2.10)), these considerations and (10.2.14) show that, as $\delta \to 0$,

$$\int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(\cdot,x))g[\varphi(\cdot,\cdot)](x)dx \to \int_{\mathbb{R}^N} \eta(u^\varepsilon(\cdot,x))g[\varphi(\cdot,\cdot)](x)\,dx \quad \text{ in } L^\infty_{\rm loc}([0,\infty[),$$

and thus also in $L^1(]0,\infty[)$ (because $\varphi(t,\cdot) = 0$ for t large). We have $1 - \chi_\delta \to 1/2$ in $L^\infty(]0,\infty[)$ weak-$*$, which implies

$$2\varepsilon \int_0^\infty \int_{\mathbb{R}^N} \eta(u^{\varepsilon,\delta}(t,x))g[\varphi(t,\cdot)](x)(1-\chi_\delta(t))\,dtdx \to \varepsilon \int_0^\infty \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x))g[\varphi(t,\cdot)](x)\,dtdx. \tag{10.2.15}$$

Passing to the limit $\delta \to 0$ in (10.2.11), thanks to (10.2.12), (10.2.13) and (10.2.15), we deduce

$$\begin{aligned}&\int_0^\infty \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x))\partial_t\varphi(t,x) + \phi(u^\varepsilon(t,x)) \cdot \nabla\varphi(t,x)\,dtdx + \int_{\mathbb{R}^N} \eta(u_0(x))\varphi(0,x)\,dx \\ &\qquad \geq \quad \varepsilon \int_0^\infty \int_{\mathbb{R}^N} \eta(u^\varepsilon(t,x))g[\varphi(t,\cdot)](x)\,dtdx \\ &\qquad\qquad - \left(\lim_{\delta\to 0} \omega_\varepsilon(\delta,\nu)\right) \int_0^\infty \left(||\partial_t\varphi(t,\cdot)||_{L^\infty(\mathbb{R}^N)} + 2\varepsilon||g[\varphi(t,\cdot)]||_{L^\infty(\mathbb{R}^N)}\right)\,dt.\end{aligned}$$

Since this is satisfied for all $\nu > 0$, we can let $\nu \to 0$ and use the property of $\omega_\varepsilon(\delta,\nu)$ to see that (10.2.2) holds.

Step 2: we now only assume that $u_0 \in L^\infty(\mathbb{R}^N)$.
Let $u_{0,n} \in C_c^\infty(\mathbb{R}^N)$ which converges a.e. to u_0 and is bounded by $||u_0||_\infty$; we define u_n^ε as the solution to (10.1.1) with $u_{0,n}$ as initial datum. As in Section 6.4 of [36], we can see that $(u_n^\varepsilon)_{n\geq 1}$ is bounded ([2]) and converges pointwise to u^ε as $n \to \infty$ ([3]).
u_n^ε satisfies (10.2.2), with $u_{0,n}$ instead of u_0. Hence, using the dominated convergence theorem, we let $n \to \infty$ in this inequality to see that it is also satisfied by u^ε, and the proof is complete. ■

[2] This can be deduced from (10.2.1) by letting $\delta \to 0$, and using $||u_{0,n}||_\infty \leq ||u_0||_\infty$.

[3] This is a consequence of estimates in [36] which show that all the derivatives of u_n^ε are bounded on $]t_0,\infty[\times\mathbb{R}^N$, for all $t_0 > 0$, uniformly with respect to n; there is thus a subsequence of $(u_n^\varepsilon)_{n\geq 1}$ which converges pointwise and, to prove that the limit is a solution to (10.1.1), we let $n \to \infty$ in Duhamel's formula which defines these solutions.

10.3 Proof of the convergence

Proposition 10.3.1 *Let $u_0 \in L^\infty(\mathbb{R}^N)$, u^ε be the solution to (10.1.1) and u be the entropy solution to (10.1.3). Let L be a Lipschitz constant of f on $[-||u_0||_\infty, ||u_0||_\infty]$ and $T > 0$. If B is a subset of $\mathbb{R}^N$, we define $\widetilde{B} = \{x \in \mathbb{R}^N \mid \mathrm{dist}(x, B) \leq 1\}$ and, for $(\mu, \nu) \in]0,1[^2$,*

$$\omega_1^B(\mu,\nu) = \sup_{0<t<T}\left(\sup_{0<r<\mu\,,|z|<\nu}\int_B |u(t,x)-u(t+r,x+z)|\,dx\right) \tag{10.3.1}$$

$$\omega_2^B(\mu,\nu) = \sup_{|z|<\nu}\int_B |u_0(x)-u_0(x+z)|\,dx + \sup_{0<s<\mu}\int_{\widetilde{B}} |u_0(x)-u(s,x)|\,dx. \tag{10.3.2}$$

$B(R)$ denotes the ball in $\mathbb{R}^N$ of center 0 and radius R. Then, for all $M > LT$, there exists $C_1 > 0$ such that, for all $t_0 \in [0,T]$, for all $\varepsilon > 0$, for all $\mu \in]0,1[$ and for all $\nu \in]0,1[$,

$$\begin{aligned}\int_{B(M-LT)} |u^\varepsilon(t_0,x)-u(t_0,x)|\,dx &\leq C_1\omega_1^{B(M+1)}(\mu,\nu)+\omega_2^{B(M+1)}(\mu,\nu)\\ &\quad +2\varepsilon||u_0||_\infty\int_{\mathbb{R}^N}\int_0^T\int_{\mathbb{R}^N}|g[h_{\nu,M}(y,t,\cdot)](x)|\,dydtdx\end{aligned} \tag{10.3.3}$$

for some $h_{\nu,M} \in C_c^\infty(\mathbb{R}^N \times [0,T] \times \mathbb{R}^N)$ only depending on ν and M.

Remark 10.3.1 *As in Remark 10.2.2, the regularity of $h_{\nu,M}$ enables us to see that $(y,t) \in \mathbb{R}^N \times [0,T] \to g[h_{\nu,M}(y,t,\cdot)] \in L^1(\mathbb{R}^N)$ is continuous and that $(y,t,x) \to g[h_{\nu,M}(y,t,\cdot)](x)$ is integrable on $\mathbb{R}^N \times [0,T] \times \mathbb{R}^N$.*

Proof of Proposition 10.3.1
We use the famous doubling variable technique of Krushkov (see [60]).
(10.2.2) has been obtained for regular convex η but it is easy, thanks to an approximation technique, to see that it also holds with the entropy $\eta_\kappa(z) = |z-\kappa|$, associated to the flux $\phi_\kappa(z) = f(z\top\kappa) - f(z\bot\kappa)$ (where $z\top\kappa = \max(z,\kappa)$ and $z\bot\kappa = \min(z,\kappa)$).
Let $\varphi \in C_c^\infty([0,\infty[\times\mathbb{R}^N \times [0,\infty[\times\mathbb{R}^N)$ be non-negative. Applying, for fixed $(s,y) \in]0,\infty[\times\mathbb{R}^N$, (10.2.2) to $\eta_{u(s,y)}$ and $\varphi(\cdot,\cdot,s,y)$, and integrating on $(s,y) \in]0,\infty[\times\mathbb{R}^N$, we find

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N}|u^\varepsilon(t,x)-u(s,y)|\partial_t\varphi(t,x,s,y)+F(u^\varepsilon(t,x),u(s,y))\cdot\nabla_x\varphi(t,x,s,y)\,dsdydtdx\\ +\int_0^\infty\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}|u_0(x)-u(s,y)|\varphi(0,x,s,y)\,dsdydx\\ \geq \varepsilon\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N}|u^\varepsilon(t,x)-u(s,y)|\,g[\varphi(t,\cdot,s,y)](x)\,dsdydtdx\end{aligned} \tag{10.3.4}$$

where $F(z,w) = f(z\top w) - f(z\bot w)$ is symmetric. We can see, as in Remarks 10.2.2 and 10.3.1, that $(t,x,s,y) \to g[\varphi(t,\cdot,s,y)](x)$ is integrable on $]0,\infty[\times\mathbb{R}^N\times]0,\infty[\times\mathbb{R}^N$, so that all the integral signs in the right-hand side can be manipulated at wish, using Fubini's theorem.
Since u is the entropy solution to (10.1.3), we have, for all $\kappa \in \mathbb{R}$ and all non-negative $\psi \in C_c^\infty([0,\infty[\times\mathbb{R}^N)$,

$$\int_0^\infty\int_{\mathbb{R}^N}\eta_\kappa(u(s,y))\partial_s\psi(s,y)+\phi_\kappa(u(s,y))\cdot\nabla_y\psi(s,y)\,dsdy+\int_{\mathbb{R}^N}\eta_\kappa(u_0(y))\psi(0,y)\,dy \geq 0.$$

Applying this inequality to $\kappa = u^\varepsilon(t,x)$ and $\psi = \varphi(t,x,\cdot,\cdot)$, and integrating on $(t,x) \in]0,\infty[\times\mathbb{R}^N$, we obtain

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N}|u(s,y)-u^\varepsilon(t,x)|\partial_s\varphi(t,x,s,y)+F(u(s,y),u^\varepsilon(t,x))\cdot\nabla_y\varphi(t,x,s,y)\,dsdydtdx\\ +\int_0^\infty\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}|u_0(y)-u^\varepsilon(t,x)|\varphi(t,x,0,y)\,dtdxdy \geq 0.\end{aligned} \tag{10.3.5}$$

Summing (10.3.4) and (10.3.5), we see that

$$\begin{aligned}
&\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(s,y)|(\partial_t\varphi(t,x,s,y)+\partial_s\varphi(t,x,s,y))\,dsdydtdx\\
&+\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} F(u^\varepsilon(t,x),u(s,y))\cdot(\nabla_x\varphi(t,x,s,y)+\nabla_y\varphi(t,x,s,y))\,dsdydtdx\\
&+\int_0^\infty\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |u_0(x)-u(s,y)|\varphi(0,x,s,y)\,dsdydx\\
&+\int_0^\infty\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |u_0(y)-u^\varepsilon(t,x)|\varphi(t,x,0,y)\,dtdxdy\\
&\geq \varepsilon\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(s,y)|\,g[\varphi(t,\cdot,s,y)](x)\,dsdydtdx. \qquad (10.3.6)
\end{aligned}$$

Let $\rho_\nu\in C_c^\infty(\mathbb{R}^N)$ and $\theta_\mu\in C_c^\infty(\mathbb{R})$ be smoothing kernels such that $\mathrm{supp}(\rho_\nu)\subset\{x\in\mathbb{R}^N \mid |x|<\nu\}$ and $\mathrm{supp}(\theta_\mu)\subset]0,\mu[$. We take $\psi\in C_c^\infty([0,\infty[\times\mathbb{R}^N)$ a non-negative function and we let $\varphi(t,x,s,y)=\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)$; we have $\partial_t\varphi(t,x,s,y)+\partial_s\varphi(t,x,s,y)=\partial_t\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)$, $\nabla_x\varphi(t,x,s,y)+\nabla_y\varphi(t,x,s,y)=\nabla_x\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)$ and $\varphi(t,x,0,y)=0$ (for $t\geq 0$). Hence, (10.3.6) gives

$$\begin{aligned}
&\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(s,y)|\partial_t\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx\\
&+\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} F(u^\varepsilon(t,x),u(s,y))\cdot\nabla_x\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx\\
&+\int_0^\infty\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |u_0(x)-u(s,y)|\psi(0,x)\rho_\nu(y-x)\theta_\mu(s)\,dsdydx\\
&\geq \varepsilon\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} \theta_\mu(s-t)|u^\varepsilon(t,x)-u(s,y)|\,g[\rho_\nu(y-\cdot)\psi(t,\cdot)](x)\,dsdydtdx. \qquad (10.3.7)
\end{aligned}$$

Let A_1, A_2 and A_3 be the first three lines of this inequality ([4]). We take $T>0$ and B a bounded set in $\mathbb{R}^N$, and we suppose that $\mathrm{supp}(\psi)\subset[0,T]\times B$. Then

$$\begin{aligned}
&\left|A_1-\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(t,x)|\partial_t\psi(t,x)\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx\right|\\
&\leq \int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_B |u(t,x)-u(s,y)|\,|\partial_t\psi(t,x)|\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx\\
&\leq ||\partial_t\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\sup_{0<t<T}\left(\int_0^\infty\int_{\mathbb{R}^N}\int_B |u(t,x)-u(s,y)|\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydx\right)\\
&\leq ||\partial_t\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\omega_1^B(\mu,\nu).
\end{aligned}$$

Since $\int_0^\infty\theta_\mu(s-t)\,ds=1$ for all $t>0$ and $\int_{\mathbb{R}^N}\rho_\nu(y-x)\,dy=1$ for all $x\in\mathbb{R}^N$, this gives

$$A_1\leq\int_0^\infty\int_{\mathbb{R}^N}|u^\varepsilon(t,x)-u(t,x)|\partial_t\psi(t,x)\,dtdx+||\partial_t\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\omega_1^B(\mu,\nu). \qquad (10.3.8)$$

[4] We keep the precise expression of the fourth term up to the end, since it will be useful in the proof of Theorem 10.1.2.

We have $|F(u^\varepsilon(t,x),u(s,y))| \le L|u^\varepsilon(t,x)-u(s,y)|$ (because both functions u^ε and u take their values in $[-||u_0||_\infty, ||u_0||_\infty]$) and therefore

$$\begin{aligned}
|A_2| &\le L\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(s,y)|\,|\nabla_x\psi(t,x)|\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx \\
&\le L\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(t,x)|\,|\nabla_x\psi(t,x)|\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx \\
&\quad + L\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_B |u(t,x)-u(s,y)|\,|\nabla_x\psi(t,x)|\rho_\nu(y-x)\theta_\mu(s-t)\,dsdydtdx \\
&\le L\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(t,x)|\,|\nabla_x\psi(t,x)|\,dtdx + L||\nabla_x\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\omega_1^B(\mu,\nu). \qquad (10.3.9)
\end{aligned}$$

Note that if $x\in B$ and $\rho_\nu(y-x)\neq 0$, then $\mathrm{dist}(y,B)\le\nu\le 1$. Therefore, for $\nu\le 1$,

$$\begin{aligned}
|A_3| &\le ||\psi(0,\cdot)||_{L^\infty(\mathbb{R}^N)}\int_0^\infty\int_{\mathbb{R}^N}\int_B |u_0(x)-u_0(y)|\rho_\nu(y-x)\theta_\mu(s)\,dsdydx \\
&\quad + ||\psi(0,\cdot)||_{L^\infty(\mathbb{R}^N)}\int_0^\infty\int_{\mathbb{R}^N}\int_B |u_0(y)-u(s,y)|\rho_\nu(y-x)\theta_\mu(s)\,dsdydx \\
&\le ||\psi(0,\cdot)||_{L^\infty(\mathbb{R}^N)}\omega_2^B(\mu,\nu). \qquad (10.3.10)
\end{aligned}$$

Gathering (10.3.8), (10.3.9) and (10.3.10) in (10.3.7), we deduce

$$\begin{aligned}
&\int_0^\infty\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(t,x)|(\partial_t\psi(t,x)+L|\nabla_x\psi(t,x)|)\,dtdx \\
&\quad + \left(||\partial_t\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))} + L||\nabla_x\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\right)\omega_1^B(\mu,\nu) + ||\psi(0,\cdot)||_{L^\infty(\mathbb{R}^N)}\omega_2^B(\mu,\nu) \\
&\quad \ge \varepsilon\int_0^\infty\int_{\mathbb{R}^N}\int_0^\infty\int_{\mathbb{R}^N} \theta_\mu(s-t)|u^\varepsilon(t,x)-u(s,y)|\,g[\rho_\nu(y-\cdot)\psi(t,\cdot)](x)\,dsdydtdx. \qquad (10.3.11)
\end{aligned}$$

Let $M>LT$ and $w_M\in C_c^\infty([0,\infty[)$ be non-increasing, with values in $[0,1]$, such that $w_M\equiv 1$ on $[0,M]$ and $\mathrm{supp}(w_M)\subset[0,M+1]$. Let $\Theta\in C_c^\infty([0,T[)$ with values in $[0,1]$. Then $\psi(t,x)=w_M(|x|+Lt)\Theta(t)$ is non-negative, belongs to $C_c^\infty([0,\infty[\times\mathbb{R}^N)$ (the function Θ has its support in $[0,T[$ and $(t,x)\to w_M(|x|+Lt)$ is regular on $[0,T]\times\mathbb{R}^N$ since, in the neighborhood of $[0,T]\times\{0\}$, we have $w_M(|x|+Lt)=1$) and $\mathrm{supp}(\psi)\subset[0,T[\times B(M+1)$. We have

$$\begin{aligned}
\partial_t\psi(t,x) &= Lw_M'(|x|+Lt)\Theta(t)+w_M(|x|+Lt)\Theta'(t) \\
|\nabla_x\psi(t,x)| &= \left|w_M'(|x|+Lt)\Theta(t)\frac{x}{|x|}\right| = (-w_M'(|x|+Lt))\Theta(t)
\end{aligned}$$

(recall that w_M is non-increasing). Hence, $\partial_t\psi(t,x)+L|\nabla_x\psi(t,x)| = w_M(|x|+Lt)\Theta'(t)$. Moreover,

$$||\partial_t\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\le LT||w_M'||_\infty + ||\Theta'||_{L^1(0,T)}, \quad ||\nabla_x\psi||_{L^1(0,T;L^\infty(\mathbb{R}^N))}\le T||w_M'||_\infty.$$

Therefore, (10.3.11) gives

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}^N} |u^\varepsilon(t,x)-u(t,x)|w_M(|x|+Lt)\Theta'(t)\,dtdx \\
&+(2LT||w_M'||_\infty + ||\Theta'||_{L^1(0,T)})\omega_1^{B(M+1)}(\mu,\nu)+\omega_2^{B(M+1)}(\mu,\nu) \\
&-\varepsilon\int_0^\infty\int_{\mathbb{R}^N}\int_0^T\int_{\mathbb{R}^N} \Theta(t)\theta_\mu(s-t)|u^\varepsilon(t,x)-u(s,y)|\,g[\rho_\nu(y-\cdot)w_M(|\cdot|+Lt)](x)\,dsdydtdx \\
&\ge 0. \qquad (10.3.12)
\end{aligned}$$

Let $t_0 \in [0,T[$ and take $\Theta(t) = \Theta_\beta(t) = \int_t^\infty \theta_\beta(s-t_0)\,ds$. Then, for β small enough, $\Theta_\beta \in C_c^\infty([0,T[)$, has its values in $[0,1]$ and $||\Theta'_\beta||_{L^1(0,T)} \leq 1$. Since, for all $t \in [0,T]$, $w_M(|\cdot| + Lt) \equiv 1$ on $B(M-LT)$ and $\Theta'_\beta(t) = -\theta_\beta(t-t_0)$, we deduce from (10.3.12) that

$$\begin{aligned}
&\int_0^T \int_{B(M-LT)} |u^\varepsilon(t,x) - u(t,x)|\theta_\beta(t-t_0)\,dtdx \\
&\quad \leq \quad (2LT||w'_M||_\infty + 1)\omega_1^{B(M+1)}(\mu,\nu) + \omega_2^{B(M+1)}(\mu,\nu) \\
&\qquad -\varepsilon \int_0^\infty \int_{\mathbb{R}^N} \int_0^T \int_{\mathbb{R}^N} \Theta_\beta(t)\theta_\mu(s-t)|u^\varepsilon(t,x) - u(s,y)|\, g[\rho_\nu(y-\cdot)w_M(|\cdot| + Lt)](x)\,dsdydtdx.
\end{aligned}$$

For all $t_0 \in [0,T[$, $\theta_\beta(\cdot - t_0)$ converges, as $\beta \to 0$ and in the weak-$*$ sense of the measures on $[0,T]$, to the Dirac mass at t_0; as $\beta \to 0$, we also have $\Theta_\beta \to \mathbf{1}_{[0,t_0]}$ everywhere and $|\Theta_\beta| \leq 1$. Since both u and u^ε are continuous $[0,T] \to L^1_{\rm loc}(\mathbb{R}^N)$ and

$$t \to \int_0^\infty \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \theta_\mu(s-t)|u^\varepsilon(t,x) - u(s,y)|\, g[\rho_\nu(y-\cdot)w_M(|\cdot| + Lt)](x)\,dsdydx$$

is integrable on $[0,T]$ (see Remark 10.3.1), we can let $\beta \to 0$ to find

$$\begin{aligned}
\int_{B(M-LT)} |u^\varepsilon(t_0,x) - u(t_0,x)|\,dx \quad &\leq \quad (2LT||w'_M||_\infty + 1)\omega_1^{B(M+1)}(\mu,\nu) + \omega_2^{B(M+1)}(\mu,\nu) \\
&\quad +\varepsilon T_{\varepsilon,\mu,\nu,M}(t_0) \qquad (10.3.13)
\end{aligned}$$

where

$$T_{\varepsilon,\mu,\nu,M}(t_0) = -\int_0^\infty \int_{\mathbb{R}^N} \int_0^{t_0} \int_{\mathbb{R}^N} \theta_\mu(s-t)|u^\varepsilon(t,x) - u(s,y)|\, g[\rho_\nu(y-\cdot)w_M(|\cdot| + Lt)](x)\,dsdydtdx \qquad (10.3.14)$$

satisfies

$$|T_{\varepsilon,\mu,\nu,M}(t_0)| \leq 2||u_0||_\infty \int_{\mathbb{R}^N} \int_0^T \int_{\mathbb{R}^N} |g[h_{\nu,M}(y,t,\cdot)](x)|\,dydtdx$$

with $h_{\nu,M}(y,t,x) = \rho_\nu(y-x)w_M(|x| + Lt) \in C_c^\infty(\mathbb{R}^N \times [0,T] \times \mathbb{R}^N)$. This concludes the proof of the proposition for $t_0 < T$, and the estimate for $t_0 = T$ is obtained by letting $t_0 \to T$ in (10.3.3). ■

The result in Theorem 10.1.1 is then an easy consequence of the following lemma.

Lemma 10.3.1 *Let $u \in C([0,\infty[; L^1_{\rm loc}(\mathbb{R}^N))$ and $T > 0$. If B is a bounded subset of $\mathbb{R}^N$, we define $\omega_1^B(\mu,\nu)$ and $\omega_2^B(\mu,\nu)$ from u by (10.3.1) and (10.3.2), with $u_0 = u(0,\cdot)$. Then, as $(\mu,\nu) \to (0,0)$, $\omega_1^B(\mu,\nu)$ and $\omega_2^B(\mu,\nu)$ go to 0.*

Proof of Theorem 10.1.1
Let $T > 0$ and $M > LT$, with L a Lipschitz constant of f on $[-||u_0||_\infty, ||u_0||_\infty]$. Let C_1 and $h_{\nu,M}$ be given by Proposition 10.3.1.
Take $\alpha > 0$. Since u is the entropy solution to (10.1.3), it is in $C([0,\infty[; L^1_{\rm loc}(\mathbb{R}^N))$. Hence, applying Lemma 10.3.1, we fix $\mu \in]0,1[$ and $\nu \in]0,1[$ small enough so that

$$C_1\omega_1^{B(M+1)}(\mu,\nu) + \omega_2^{B(M+1)}(\mu,\nu) \leq \alpha.$$

By Remark 10.3.1, we can choose $\varepsilon_0 > 0$ (depending on ν and M) such that

$$2\varepsilon_0||u_0||_\infty \int_{\mathbb{R}^N} \int_0^T \int_{\mathbb{R}^N} |g[h_{\nu,M}(y,t,\cdot)](x)|\,dydtdx \leq \alpha.$$

Proposition 10.3.1 then shows that, for all $\varepsilon \leq \varepsilon_0$,

$$\sup_{t\in[0,T]} \int_{B(M-LT)} |u^\varepsilon(t,x) - u(t,x)|\, dx \leq 2\alpha.$$

This reasoning can be made for all $T > 0$ and all $M > LT$, which proves that $u^\varepsilon \to u$ in $C([0,T]; L^1_{\rm loc}(\mathbb{R}^N))$ for all $T > 0$. ■

Proof of Lemma 10.3.1
The convergence of $\omega_2^B(\mu,\nu)$ is quite easy. Indeed, since $u_0 = u(0,\cdot) \in L^1_{\rm loc}(\mathbb{R}^N)$ and B is bounded, we know that

$$\int_B |u_0(x) - u_0(x+z)|\, dx \to 0 \quad \text{ as } z \to 0.$$

By continuity of $u : [0,\infty[\to L^1_{\rm loc}(\mathbb{R}^N)$ and since $\widetilde{B}$ is bounded, we also have $||u(s,\cdot) - u_0||_{L^1(\widetilde{B})} \to 0$ as $s \to 0$. Hence, this proves that $\omega_2^B(\mu,\nu) \to 0$ as $(\mu,\nu) \to 0$.
The convergence of $\omega_1^B(\mu,\nu)$ is a bit more tricky. We split it in two parts:

$$\begin{aligned}
\omega_1^B(\mu,\nu) \;\leq\; & \sup_{0<t<T} \left(\sup_{|z|<\nu} \int_B |u(t,x) - u(t,x+z)|\, dx \right) \\
& + \sup_{0<t<T} \left(\sup_{0<r<\mu\,,|z|<\nu} \int_B |u(t,x+z) - u(t+r,x+z)|\, dx \right) \\
\leq\; & \sup_{0<t<T} \left(\sup_{|z|<\nu} \int_B |u(t,x) - u(t,x+z)|\, dx \right) \qquad (10.3.15) \\
& + \sup_{0<t<T} \left(\sup_{0<r<\mu} \int_{\widetilde{B}} |u(t,y) - u(t+r,y)|\, dy \right). \qquad (10.3.16)
\end{aligned}$$

By hypothesis, $u \in C([0,T+1]; L^1(\widetilde{B}))$; hence, u is uniformly continuous $[0,T+1] \to L^1(\widetilde{B})$ and

$$\sup_{0<t<T} \left(\sup_{0<r<\mu} \int_{\widetilde{B}} |u(t,y) - u(t+r,y)|\, dy \right) \leq \sup_{(t,s)\in[0,T+1]^2\,,\,0<s-t<\mu} ||u(t,\cdot) - u(s,\cdot)||_{L^1(\widetilde{B})} \to 0 \quad (10.3.17)$$

as $\mu \to 0$. Moreover, since $u \in C([0,T]; L^1(\widetilde{B}))$, the set $\mathcal{K} = \{u(t,\cdot)\,,\ 0 \leq t \leq T\}$ is compact in $L^1(\widetilde{B})$; therefore, by Kolmogorov's compactness theorem, the translations are equicontinuous on $\mathcal{K}$, that is to say

$$\sup_{v\in\mathcal{K}} \left(\sup_{|z|<\nu} \int_B |v(x) - v(x+z)|\, dx \right) \to 0$$

as $\nu \to 0$. This quantity bounds (10.3.15), which proves, together with (10.3.17), that $\omega_1^B(\mu,\nu) \to 0$ as $(\mu,\nu) \to 0$. ■

10.4 Proof of the error estimate

We prove here Theorem 10.1.2, beginning with a stronger version of Lemma 10.3.1 in the case of more regular functions.

Lemma 10.4.1 *Let $u \in \mathrm{Lip}([0,\infty[; L^1(\mathbb{R}^N))$ such that $\sup_{t\geq 0} |u(t,\cdot)|_{BV(\mathbb{R}^N)} < \infty$. We define $\omega_1^{\mathbb{R}^N}(\mu,\nu)$ and $\omega_2^{\mathbb{R}^N}(\mu,\nu)$ from u by (10.3.1) and (10.3.2), with $T = \infty$, $u_0 = u(0,\cdot)$ and $B = \mathbb{R}^N$. Then $\omega_1^{\mathbb{R}^N}(\mu,\nu) = \mathcal{O}(\mu+\nu)$ and $\omega_2^{\mathbb{R}^N}(\mu,\nu) = \mathcal{O}(\mu+\nu)$.*

Proof of Lemma 10.4.1
It is classical (see e.g. [38] or (10.5.7)) that, if $v \in BV(\mathbb{R}^N)$ then,

$$\int_{\mathbb{R}^N} |v(x+h) - v(x)|\, dx \leq |h|\, |v|_{BV(\mathbb{R}^N)}. \tag{10.4.1}$$

Thus,

$$\sup_{|z|<\nu} \int_{\mathbb{R}^N} |u_0(x) - u_0(x+z)|\, dx = \mathcal{O}(\nu)$$

and, since $u : [0,\infty[\to L^1(\mathbb{R}^N)$ is Lipschitz continuous, we deduce that $\omega_2^{\mathbb{R}^N}(\mu,\nu) = \mathcal{O}(\mu+\nu)$.
We split $\omega_1^{\mathbb{R}^N}(\mu,\nu)$ as in the proof of Lemma 10.3.1 (with $\widetilde{B} = \mathbb{R}^N$ here). By the Lipschitz continuity of u, (10.3.16) is a $\mathcal{O}(\mu)$; by (10.4.1) and the bound on $|u(t,\cdot)|_{BV(\mathbb{R}^N)}$, (10.3.15) is a $\mathcal{O}(\nu)$. This concludes the proof of the lemma. ■

Proof of Theorem 10.1.2
Since $u_0 \in L^\infty(\mathbb{R}^N) \cap L^1(\mathbb{R}^N) \cap BV(\mathbb{R}^N)$, it is classical that $|u(t,\cdot)|_{BV(\mathbb{R}^N)} \leq |u_0|_{BV(\mathbb{R}^N)}$. The function f being regular, the BV semi-norm of $f(u(t,\cdot))$ is also bounded and, thanks to $\partial_t u + \mathrm{div}(f(u)) = 0$, we see that u is Lipschitz continuous $[0,\infty[\to L^1(\mathbb{R}^N)$. Hence, Lemma 10.4.1 and (10.3.13) show that, for all $T > 0$, for all $M > LT$ and all $t_0 \in [0,T]$, if $\mu \in]0,1[$ and $\nu \in]0,1[$,

$$\int_{B(M-LT)} |u^\varepsilon(t_0,x) - u(t_0,x)|\, dx \leq C_0(2LT||w_M'||_\infty + 2)(\mu+\nu) + \varepsilon T_{\varepsilon,\mu,\nu,M}(t_0), \tag{10.4.2}$$

where we recall that $T_{\varepsilon,\mu,\nu,M}(t_0)$ is defined by (10.3.14).
To bound $T_{\varepsilon,\mu,\nu,M}(t_0)$, we use (10.5.1). We handle the case $\lambda \in]1,2[$, the other one being easier (and, anyway, well-known). We define $\beta = -N - (\lambda - 2)$. It is not hard to check, differentiating under the integral sign, that

$$g[h_{\nu,M}(y,t,\cdot)](x) = E_\lambda |\cdot|^\beta * (\Delta_x h_{\nu,M}(y,t,\cdot))(x) = E_\lambda \mathrm{div}_x \left(|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot) \right)(x)$$

(recall that $h_{\nu,M}(y,t,x) = \rho_\nu(y-x) w_M(|x|+Lt) \in C_c^\infty(\mathbb{R}^N \times [0,T] \times \mathbb{R}^N)$). Let A be such that the support of $h_{\nu,M}(y,t,\cdot)$ is contained in the ball of center 0 and radius A. From the definition of the convolution product, we see that, for $|x| > A$, $||\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot)(x)| \leq \Lambda(|x|-A)^\beta$; hence, $|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot)(x)$ goes to 0, as $|x| \to \infty$, quicker than $|x|^{-N+1}$ (because $\beta = -N - (\lambda-2) < -N+1$). We know that $u^\varepsilon(t,\cdot)$ is regular for all $t > 0$ (see [36]), and that $||\nabla u^\varepsilon(t,\cdot)||_{L^1(\mathbb{R}^N)} \leq |u_0|_{BV(\mathbb{R}^N)}$ (this can be easily seen letting $\delta \to 0$ in (10.2.1) — we have noticed that the construction of u^ε in subsection 10.2.1 is valid for initial data in $L^\infty(\mathbb{R}^N) \cap L^1(\mathbb{R}^N) \cap BV(\mathbb{R}^N)$). We can therefore use Stokes formula on a ball of radius R and let $R \to \infty$ to find

$$\begin{aligned}
&\int_{\mathbb{R}^N} |u^\varepsilon(t,x) - u(s,y)|\, g[h_{\nu,M}(y,t,\cdot)](x)\, dx \\
&= -E_\lambda \int_{\mathbb{R}^N} \nabla_x(|u^\varepsilon(t,\cdot) - u(s,y)|)(x) \cdot (|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x)\, dx \\
&= -E_\lambda \int_{\mathbb{R}^N} \mathrm{sgn}(u^\varepsilon(t,x) - u(s,y)) \nabla u^\varepsilon(t,x) \cdot (|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x)\, dx\,,
\end{aligned}$$

which implies

$$\left| \int_{\mathbb{R}^N} |u^\varepsilon(t,x) - u(s,y)|\, g[h_{\nu,M}(y,t,\cdot)](x)\, dx \right| \leq |E_\lambda| \int_{\mathbb{R}^N} |\nabla u^\varepsilon(t,x)|\, |(|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x)|\, dx.$$

Therefore, by (10.3.14),

$$|T_{\varepsilon,\mu,\nu,M}(t_0)| \leq |E_\lambda| \int_{\mathbb{R}^N} \int_0^{t_0} \int_{\mathbb{R}^N} |\nabla u^\varepsilon(t,x)|\, |(|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x)|\, dy dt dx. \tag{10.4.3}$$

We choose w_M such that $(w'_M)_{M\geq 1}$ is bounded by C_1. Let $\delta \in]0,1[$ and B_δ be the ball of center 0 and radius δ; cutting as in the proof of Lemma 10.5.1 and using Stokes formula, we have

$$\begin{aligned}(|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x) &= \int_{B_\delta} |z|^\beta \nabla_x h_{\nu,M}(y,t,x-z)\,dz - \int_{B_\delta^c} |z|^\beta \nabla_z(h_{\nu,M}(y,t,x-z))\,dz \\ &= \int_{B_\delta} |z|^\beta \nabla_x h_{\nu,M}(y,t,x-z)\,dz - \delta^\beta \int_{\partial B_\delta^c} h_{\nu,M}(y,t,x-z)\mathbf{n}(z)\,d\sigma_\delta(z) \\ &\quad +\beta \int_{B_\delta^c} h_{\nu,M}(y,t,x-z)|z|^{\beta-1}\frac{z}{|z|}\,dz \qquad (10.4.4)\end{aligned}$$

(σ_δ is the $(N-1)$-dimensional measure on ∂B_δ^c and $\mathbf{n}$ is the unit normal to ∂B_δ^c outward to B_δ^c). Since

$$|h_{\nu,M}(y,t,x)| = |\rho_\nu(y-x)w_M(|x|+Lt)| \leq \rho_\nu(y-x)$$

and

$$\begin{aligned}|\nabla_x h_{\nu,M}(y,t,x)| &= \left|-\nabla\rho_\nu(y-x)w_M(|x|+Lt) + \rho_\nu(y-x)w'_M(|x|+Lt)\frac{x}{|x|}\right| \\ &\leq |\nabla\rho_\nu(y-x)| + C_1\rho_\nu(y-x),\end{aligned}$$

(10.4.4) shows that

$$\begin{aligned}|(|\cdot|^\beta * \nabla_x h_{\nu,M}(y,t,\cdot))(x)| &\leq \int_{B_\delta} |z|^\beta \left(|\nabla\rho_\nu(y-x+z)| + C_1\rho_\nu(y-x+z)\right)\,dz \\ &\quad +\delta^\beta \int_{\partial B_\delta^c} \rho_\nu(y-x+z)\,d\sigma_\delta(z) + |\beta| \int_{B_\delta^c} \rho_\nu(y-x+z)|z|^{\beta-1}\,dz.\end{aligned}$$

By Fubini-Tonelli's theorem and (10.4.3), we obtain

$$\begin{aligned}&|T_{\varepsilon,\mu,\nu,M}(t_0)| \\ &\leq |E_\lambda| \int_{\mathbb{R}^N}\int_0^{t_0}\int_{\mathbb{R}^N} |\nabla u^\varepsilon(t,x)| \left(\int_{B_\delta} |z|^\beta \left(|\nabla\rho_\nu(y-x+z)| + C_1\rho_\nu(y-x+z)\right)\,dz\right. \\ &\qquad \left. +\delta^\beta \int_{\partial B_\delta^c} \rho_\nu(y-x+z)\,d\sigma_\delta(z) + |\beta| \int_{B_\delta^c} \rho_\nu(y-x+z)|z|^{\beta-1}\,dz\right)\,dydtdx \\ &\leq |E_\lambda| \left(||\nabla\rho_\nu||_{L^1(\mathbb{R}^N)} + C_1\right) |||\cdot|^\beta||_{L^1(B_\delta)} ||\nabla u^\varepsilon||_{L^1(]0,t_0[\times\mathbb{R}^N)} + |E_\lambda|\delta^\beta\sigma_\delta(\partial B_\delta^c)||\nabla u^\varepsilon||_{L^1(]0,t_0[\times\mathbb{R}^N)} \\ &\quad +|E_\lambda|\,|\beta|\,|||\cdot|^{\beta-1}||_{L^1(B_\delta^c)}||\nabla u^\varepsilon||_{L^1(]0,t_0[\times\mathbb{R}^N)}.\end{aligned}$$

By change of variable, $|||\cdot|^\beta||_{L^1(B_\delta)} = C_2\delta^{N+\beta}$, $|||\cdot|^{\beta-1}||_{L^1(B_\delta^c)} = C_3\delta^{N+\beta-1}$ and $\sigma_\delta(\partial B_\delta^c) = C_4\delta^{N-1}$, where C_2, C_3 and C_4 do not depend on δ (recall that $\beta - 1 < -N < \beta$). Since $||\nabla u^\varepsilon(t,\cdot)||_{L^1(\mathbb{R}^N)} \leq |u_0|_{BV(\mathbb{R}^N)}$, we deduce that

$$|T_{\varepsilon,\mu,\nu,M}(t_0)| \leq C_5T(||\nabla\rho_\nu||_{L^1(\mathbb{R}^N)} + 1)\delta^{N+\beta} + C_5T\delta^{N+\beta-1} \qquad (10.4.5)$$

where C_5 does not depend on t_0, ε, μ, ν, M or δ.

Choosing smoothing kernels $(\rho_\nu)_{\nu>0}$ of the kind $\rho_\nu(x) = \nu^{-N}\rho(\nu^{-1}x)$, we have $||\nabla\rho_\nu||_{L^1(\mathbb{R}^N)} = C_6\nu^{-1}$. Since $(w'_M)_{M\geq 1}$ is bounded by C_1, (10.4.2) and (10.4.5) give, for all $T > 0$, for all $M > LT$ and all $t_0 \in [0,T]$,

$$\int_{B(M-LT)} |u^\varepsilon(t_0,x) - u(t_0,x)|\,dx \leq C_0(2LTC_1+2)(\mu+\nu) + \varepsilon\left(\frac{C_5C_6T\delta^{2-\lambda}}{\nu} + C_5T\delta^{2-\lambda} + C_5T\delta^{1-\lambda}\right)$$

(we have $N+\beta = 2-\lambda$). We let $M \to \infty$ and $\mu \to 0$; since $\nu < 1$, this gives

$$||u^\varepsilon(t_0,\cdot) - u(t_0,\cdot)||_{L^1(\mathbb{R}^N)} = \mathcal{O}\left(\nu + \varepsilon\left(\frac{\delta^{2-\lambda}}{\nu} + \delta^{1-\lambda}\right)\right).$$

Minimizing on δ and then on ν, we see that the best choices are (up to multiplicative constants) $\delta = \nu$ and $\nu = \varepsilon^{1/\lambda}$, which proves Theorem 10.1.2. ■

10.5 Appendix

10.5.1 An expression and an estimate of $g[\varphi]$

Lemma 10.5.1 *Let $\lambda \in]1,2]$. There exist $E_\lambda \in \mathbb{R}$ and $C_\lambda > 0$ such that, for all $\varphi \in \mathcal{S}(\mathbb{R}^N)$,*

$$\begin{cases} g[\varphi] = E_\lambda |\cdot|^{-N-(\lambda-2)} * \Delta\varphi & \text{for } \lambda \in]1,2[\\ g[\varphi] = E_\lambda \Delta\varphi & \text{for } \lambda = 2 \end{cases} \tag{10.5.1}$$

and

$$||g[\varphi]||_{L^1(\mathbb{R}^N)} \le C_\lambda \left(||\nabla\varphi||_{L^1(\mathbb{R}^N)} + ||\Delta\varphi||_{L^1(\mathbb{R}^N)}\right).$$

Proof of Lemma 10.5.1
If $\lambda = 2$, the result is obvious since, up to a multiplicative constant, $g[\varphi]$ is $\Delta\varphi$. We thus assume that $\lambda \in]1,2[$ and we have

$$g[\varphi] = \mathcal{F}^{-1}(|\cdot|^\lambda \mathcal{F}(\varphi)) = (2i\pi)^{-2}\mathcal{F}^{-1}(|\cdot|^{\lambda-2}\mathcal{F}(\Delta\varphi))$$

(note that $|\cdot|^{\lambda-2} \in L^1_{\rm loc}(\mathbb{R}^N)$, as $\lambda - 2 > -N$, and that $\mathcal{F}(\Delta\varphi) \in \mathcal{S}(\mathbb{R}^N)$, so that $|\cdot|^{\lambda-2}\mathcal{F}(\Delta\varphi)$ is integrable on $\mathbb{R}^N$). Since $\lambda - 2 \in]-N, 0[$, it is classical that $\mathcal{F}^{-1}(|\cdot|^{\lambda-2}) = C_1|\cdot|^{-N-(\lambda-2)}$ in $\mathcal{S}'(\mathbb{R}^N)$, for some $C_1 \in \mathbb{R}$. We can then check, using the definition (by duality) of $\mathcal{F}^{-1}$ on $\mathcal{S}'(\mathbb{R}^N)$, that $\mathcal{F}^{-1}(|\cdot|^{\lambda-2}\mathcal{F}(\Delta\varphi)) = C_1|\cdot|^{-N-(\lambda-2)} * \Delta\varphi$, which proves (10.5.1).
Let $\beta = -N-(\lambda-2) \in]-N, 0[$. We now estimate $||\,|\cdot|^\beta * \Delta\varphi||_{L^1(\mathbb{R}^N)}$, which will conclude the proof (note that this estimate is not a straightforward consequence of Young's inequalities for convolution, since $|\cdot|^\beta$ is not integrable on $\mathbb{R}^N$). We have, if $\mathbf{1}_B$ is the characteristic function of the ball B of center 0 and radius 1 and $\mathbf{1}_{B^c}$ is the characteristic function of $B^c = \mathbb{R}^N \setminus B$,

$$|\cdot|^\beta * \Delta\varphi = (\mathbf{1}_B|\cdot|^\beta) * \Delta\varphi + (\mathbf{1}_{B^c}|\cdot|^\beta) * \Delta\varphi. \tag{10.5.2}$$

But $\mathbf{1}_B|\cdot|^\beta \in L^1(\mathbb{R}^N)$ (because $\beta > -N$), and thus

$$||(\mathbf{1}_B|\cdot|^\beta) * \Delta\varphi||_{L^1(\mathbb{R}^N)} \le ||\,\mathbf{1}_B|\cdot|^\beta\,||_{L^1(\mathbb{R}^N)}||\Delta\varphi||_{L^1(\mathbb{R}^N)}. \tag{10.5.3}$$

By Stokes formula, we write

$$\begin{aligned} (\mathbf{1}_{B^c}|\cdot|^\beta) * \Delta\varphi(x) &= \int_{B^c} |y|^\beta \Delta\varphi(x-y)\,dy \\ &= -\int_{\partial B^c} \nabla\varphi(x-y)\cdot\mathbf{n}(y)\,d\sigma(y) + \beta\int_{B^c} \nabla\varphi(x-y)\cdot\left(|y|^{\beta-1}\frac{y}{|y|}\right)dy \end{aligned}$$

where $\mathbf{n}$ is the outward unit normal to B^c and σ is the measure on ∂B^c. We deduce that

$$|(\mathbf{1}_{B^c}|\cdot|^\beta) * \Delta\varphi(x)| \le \int_{\partial B^c} |\nabla\varphi(x-y)|\,d\sigma(y) + |\beta|\int_{B^c} |\nabla\varphi(x-y)|\,|y|^{\beta-1}\,dy$$

and, integrating this thanks to Fubini-Tonelli's theorem,

$$\begin{aligned}\int_{\mathbb{R}^N} |(\mathbf{1}_{B^c}|\cdot|^\beta) * \Delta\varphi(x)|\,dx &\leq \int_{\partial B^c}\int_{\mathbb{R}^N} |\nabla\varphi(x-y)|\,dx d\sigma(y) + |\beta| \int_{B^c}\int_{\mathbb{R}^N} |\nabla\varphi(x-y)|\,dx\,|y|^{\beta-1}\,dy \\ &= \left(\sigma(\partial B^c) + |\beta| \int_{B^c} |y|^{\beta-1}\,dy\right) \int_{\mathbb{R}^N} |\nabla\varphi(z)|\,dz. \qquad (10.5.4)\end{aligned}$$

Since $\beta - 1 = -N - (\lambda - 2) - 1 = -N - \lambda + 1 < -N$, $\int_{B^c} |y|^{\beta-1}\,dy$ is finite. Gathering (10.5.3) and (10.5.4) in (10.5.2), we deduce that $||\,|\cdot|^\beta * \Delta\varphi||_{L^1(\mathbb{R}^N)} \leq C(||\Delta\varphi||_{L^1(\mathbb{R}^N)} + ||\nabla\varphi||_{L^1(\mathbb{R}^N)})$ for some C not depending on φ, and the proof is complete. ■

10.5.2 Technical lemmas on the kernel of g

The results in the following two lemmas have already been used in [36], but their proofs were left to the reader. We include them here for sake of completeness.

Lemma 10.5.2 *Let $r > 0$, $w_0 \in L^1(\mathbb{R}^N)$ and, for $t > 0$, $w(t,\cdot) = K_r(t,\cdot) * w_0$. Then, for all $\varphi \in C_c^\infty([0,\infty[\times\mathbb{R}^N)$ and all $t_0 > 0$,*

$$\int_0^{t_0}\int_{\mathbb{R}^N} w(t,x)\partial_t\varphi(t,x) - rw(t,x)g[\varphi(t,\cdot)](x)\,dtdx = \int_{\mathbb{R}^N} w(t_0,x)\varphi(t_0,x)\,dx - \int_{\mathbb{R}^N} w_0(x)\varphi(0,x)\,dx.$$

Proof of Lemma 10.5.2
We ignore, as in the proof of Proposition 10.2.1, the space variable. Since $K_r(t)$ and w_0 are integrable, $w(t)$ is integrable and we have

$$\mathcal{F}^{-1}(w(t)) = \mathcal{F}^{-1}(K_r(t))\mathcal{F}(w_0) = e^{-rt|\cdot|^\lambda}\mathcal{F}^{-1}(w_0) \qquad (10.5.5)$$

(note that, since $K_r(t)$ is even, $\mathcal{F}^{-1}(K_r(t)) = \mathcal{F}(K_r(t))$). By Fubini's theorem, for all $(a,b) \in L^1(\mathbb{R}^N)$,

$$\int_{\mathbb{R}^N} a\mathcal{F}^{-1}(b) = \int_{\mathbb{R}^N} \mathcal{F}^{-1}(a)b. \qquad (10.5.6)$$

Thus, writing $g[\varphi(t)] = \mathcal{F}^{-1}(|\cdot|^\lambda\mathcal{F}(\varphi(t)))$ and $\partial_t\varphi(t) = \mathcal{F}^{-1}(\mathcal{F}(\partial_t\varphi(t)))$, we have, thanks to (10.5.5) and (10.5.6), for all $t > 0$,

$$\begin{aligned}&\int_{\mathbb{R}^N} w(t)\partial_t\varphi(t) - rw(t)g[\varphi(t)] \\ &= \int_{\mathbb{R}^N} e^{-rt|\xi|^\lambda}\mathcal{F}^{-1}(w_0)(\xi)\mathcal{F}(\partial_t\varphi(t))(\xi) - r|\xi|^\lambda e^{-rt|\xi|^\lambda}\mathcal{F}^{-1}(w_0)(\xi)\mathcal{F}(\varphi(t))(\xi)\,d\xi \\ &= \int_{\mathbb{R}^N} \partial_t\left(e^{-rt|\xi|^\lambda}\mathcal{F}(\varphi(t))(\xi)\right)\mathcal{F}^{-1}(w_0)(\xi)\,d\xi.\end{aligned}$$

$(t,\xi) \to e^{-rt|\xi|^\lambda}\mathcal{F}^{-1}(w_0)(\xi)$ is bounded on $]0,t_0[\times\mathbb{R}^N$ and, by regularity of φ, $(t,\xi) \to |\xi|^\lambda\mathcal{F}(\varphi(t))(\xi)$ and $(t,\xi) \to \mathcal{F}(\partial_t\varphi(t))(\xi)$ are integrable on $]0,t_0[\times\mathbb{R}^N$ (see e.g. (10.2.10)); hence, integrating the preceding equality on $]0,t_0[$ and using Fubini's theorem, we find

$$\int_0^{t_0}\int_{\mathbb{R}^N} w(t)\partial_t\varphi(t) - rw(t)g[\varphi(t)]\,dt = \int_{\mathbb{R}^N}\left(e^{-rt_0|\xi|^\lambda}\mathcal{F}(\varphi(t_0))(\xi) - \mathcal{F}(\varphi(0))(\xi)\right)\mathcal{F}^{-1}(w_0)(\xi)\,d\xi.$$

Using once again (10.5.5) and (10.5.6), we get

$$\begin{aligned}\int_0^{t_0}\int_{\mathbb{R}^N} w(t)\partial_t\varphi(t) - rw(t)g[\varphi(t)]\,dt &= \int_{\mathbb{R}^N} \mathcal{F}^{-1}(w(t_0))\mathcal{F}(\varphi(t_0)) - \mathcal{F}^{-1}(w_0)\mathcal{F}(\varphi(0)) \\ &= \int_{\mathbb{R}^N} w(t_0)\varphi(t_0) - w_0\varphi(0)\end{aligned}$$

which concludes the proof. ■

Lemma 10.5.3 *Let $r > 0$ and $w_0 \in W^{1,1}(\mathbb{R}^N) \cap C^1(\mathbb{R}^N)$. We define, for $t > 0$, $w(t,\cdot) = K_r(t,\cdot) * w_0$. Then, for all $\nu > 0$ and all $t > 0$,*

$$||w(t,\cdot) - w_0||_{L^1(\mathbb{R}^N)} \leq 2||w_0||_{L^1(\mathbb{R}^N)} \int_{|y|\geq\nu} K_r(t,y)\, dy + \nu||\nabla w_0||_{L^1(\mathbb{R}^N)}.$$

Proof of Lemma 10.5.3
The proof relies on classical cuttings of integration domain when approximate units are involved. Since $K_r(t,\cdot)$ is non-negative with integral equal to 1, we can write

$$|w(t,x) - w_0(x)| = \left| \int_{\mathbb{R}^N} K_r(t,y)(w_0(x-y) - w_0(x))\, dy \right| \leq \int_{\mathbb{R}^N} K_r(t,y)|w_0(x-y) - w_0(x)|\, dy.$$

Now,

$$\begin{aligned} &||w(t,\cdot) - w_0||_{L^1(\mathbb{R}^N)} \\ &\leq \int_{|y|\geq\nu} K_r(t,y) \int_{\mathbb{R}^N} |w_0(x-y) - w_0(x)|\, dxdy + \int_{|y|<\nu} K_r(t,y) \int_{\mathbb{R}^N} |w_0(x-y) - w_0(x)|\, dxdy \\ &\leq 2||w_0||_{L^1(\mathbb{R}^N)} \int_{|y|\geq\nu} K_r(t,y)\, dy + \sup_{|z|<\nu} \int_{\mathbb{R}^N} |w_0(x+z) - w_0(x)|\, dx. \end{aligned}$$

We then write, using Fubini-Tonelli's theorem and a change of variable,

$$\int_{\mathbb{R}^N} |w_0(x+z) - w_0(x)|\, dx \leq \int_{\mathbb{R}^N} \int_0^1 |\nabla w_0(x+\zeta z)|\, |z|\, d\zeta\, dx \leq |z| \int_{\mathbb{R}^N} |\nabla w_0(y)|\, dy\,, \tag{10.5.7}$$

and the proof is complete. ■

Bibliographie

[1] Adams R.A., Sobolev Spaces. Academic Press (1975).

[2] Bacuta C., Bramble J., Pasciak J., *New interpolation results and applications to finite element methods for elliptic boundary value problems.* East-West J. Numer. Math. **9** (2001), no. 3, 179-198.

[3] Bardos C., le Roux A. Y., Nédélec J.-C., *First order quasilinear equations with boundary conditions.* Comm. Partial Differential Equations, **4** (1979), 1017–1034.

[4] Benilan P., Boccardo L., Gallouët T., Gariepy R., Pierre M., Vazquez J. L., *An L^1-Theory of Existence and Uniqueness of Solutions of Nonlinear Elliptic Equations.* Ann. Scuola Norm. Sup. Pisa Sci. Fis. Mat., IV, **22** (1995), no. 2, 241-273.

[5] Bénilan P., Brezis H., Crandall M.G., *A semilinear equation in L^1.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. **2** (1975), 523-555.

[6] Bergh J., Löfström J., Interpolation Spaces. Springer-Verlag (1976).

[7] Biler P., Funaki T., Woyczynski W. A., *Fractal Burgers Equations.* J. Diff. Eq. **148** (1998), 9-46.

[8] Biler P., Karch G., Woyczynski W. A., *Asymptotics for conservation laws involving Lévy diffusion generators.* Studia Math., **148** (2001), no. 2, 171–192.

[9] Biler P., Karch G., Woyczynski W. A., *Asymptotics for multifractal conservation laws.* Studia Math. **135** (1999), no. 3, 231–252.

[10] Boccardo L., *Problemi differenziali ellittici e parabolici con dati misure.* Boll. Un. Mat. Ital. A (7) **11** (1997), no. 2, 439-461.

[11] Boccardo L., Gallouët T., *Nonlinear elliptic and parabolic equations involving measure data.* J. Funct. Anal. **87** (1989), 241-273.

[12] Boccardo L., Gallouët T., *Nonlinear elliptic equations with right-hand side measures.* Comm. Partial Differential Equations **17** (1992), 641-655.

[13] Boccardo L., Gallouët T., Orsina L., *Existence and uniqueness of entropy solutions for nonlinear elliptic equations with measure data.* Ann. Inst. H. Poincaré Anal. Non Linéaire, **13** (1996), 539-551.

[14] Boccardo L., Gallouët T., Vazquez J-L., *Nonlinear elliptic equations in $\mathbb{R}^N$ without growth restrictions on the data.* Jour. Diff. Eqns. **105** (1993), 334-363.

[15] Boivin S., Cayré F., Hérard J.M., *A finite volume method to solve the Navier-Stokes equations for incompressible flows on unstructured meshes.* Int. J. Therm. Sci. **39** (2000), 806-825.

[16] Bony J-M., Courrège P., Priouret P., *Semi-groupes de Feller sur une variété à bord compacte et problèmes aux limites intégro-différentiels du second ordre donnant lieu au principe du maximum.* Ann. Inst. Fourier (Grenoble) **18** (1968), no. 2, 369-521.

[17] Bramble J., *Interpolation between Sobolev Spaces in Lipschitz domains with an application to multigrid theory.* Math. Comp. **64** (1995), 1359-1365.

[18] Bramble J., Hilbert S., *Bounds for a class of linear functionals with applications to Hermite interpolation.* Numer. Math. **16** (1974), 362-369.

[19] Brezis H., Strauss W., *Semilinear elliptic equations in L^1.* Jour. Math. Soc. Japan **25** (1973), 565-590.

[20] Chainais-Hillairet C., Grenier E., *Numerical boundary layers for hyperbolic systems in 1-D.* M2AN Math. Model. Numer. Anal. **35** (2001), 91-106.

[21] Chatzipantelidis P., Lazarov R.D., *The finite volume element method in nonconvex polygonal domains.* Finite Volumes for Complex Applications III, Hermes Penton Science, 171-178. Porquerolles, France, 2002.

[22] Chou S. H., Vassilevski P.S. , *A general mixed covolume framework for constructing conservative schemes for elliptic problems.* Math. Comp. **68** (1999), 991-1011.

[23] Clavin P., Denet B., *Theory of cellular detonations in gases. part 2. mach-stem formation at strong overdrive*, C. R. Acad. Sci. Paris **329** (2001), no. Série II b, 489-496.

[24] Clavin P., Denet B., *Diamond patterns in the cellular front of an overdriven detonation.* Phys. Rev. Letters **88** (2002).

[25] Clavin P., He L., *Theory of cellular detonations in gases. part 1. stability limits at strong overdrive.* C. R. Acad. Sci. Paris **329** (2001), no. Série II b, 463–471.

[26] Coudière Y., Gallouët T., Herbin R., *Discrete Sobolev Inequalities and L^p error estimates for approximate finite volume solutions of convection diffusion equations.* M2AN Math. Model. Numer. Anal. **35** (2001), no. 4, 767-778.

[27] Coudière Y., Vila J.P., Villedieu P., *Convergence Rate of a Finite Volume Scheme for a Two Dimensional Convection Diffusion Problem.* M2AN Math. Model. Numer. Anal. **33** (1999), no. 3, 493-516.

[28] Dal Maso G., Murat F., Orsina L., Prignet A., *Renormalized solutions for elliptic equations with general measure data.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. **28** (1999), no. 4, 741-808.

[29] Dall'Aglio A., *Approximated solutions of equations with L^1 data. Application to the H-convergence of quasi-linear parabolic equations.* Ann. Mat. Pura Appl. (4) **170** (1996), 207-240.

[30] DiPerna R. J., *Measure-valued solutions to conservation laws.* Arch. Rational Mech. Anal. **88** (1985), no. 3, 223-270.

[31] Droniou J., *Non Coercive Linear Elliptic Problems.* Potential Anal. **17** (2002), no. 2, 181-203.

[32] Droniou J., *PhD Thesis*, CMI, Université de Provence.
Available at `http://www-gm3.univ-mrs.fr/~droniou/these/index-en.html`.

[33] Droniou J., Gallouët T., *Finite volume methods for convection-diffusion equations with right-hand side in H^{-1}.* M2AN Math. Model. Numer. Anal. **36** (2002), no. 4, 705-724.

[34] Droniou J., Gallouët T, *A finite volume scheme for noncoercive Dirichlet problems with right-hand side in H^{-1}*. Finite volume for complex applications III, R. Herbin and D. Kröner eds, Hermes Penton Science (2002), 195-202.

[35] Droniou J., Gallouët T., *A uniqueness result for quasilinear elliptic equations with measures as data*. Rend. Mat. Appl. (7) **21** (2001), no. 1-4, 57-86.

[36] Droniou J., Gallouët T., Vovelle J., *Global solution and smoothing effect for a non-local regularization of an hyperbolic equation*. J. Evol. Equ. **3** (2003), no. 3, 499-521.

[37] Droniou J., Gallouët T. and Herbin R., *A finite volume scheme for a noncoercive elliptic equation with measure data*. SIAM J. Numer. Anal. **41** (2003), no. 6, 1997-2031.

[38] Eymard R., Gallouët T., Herbin R., Finite Volume Methods. Handbook of Numerical Analysis, Vol. VII, pp. 713-1020. Edited by P.G. Ciarlet and J.L. Lions (North Holland).

[39] Eymard R., Gallouët T., Herbin R., *Convergence of finite volume approximations to the solutions of semilinear convection diffusion reaction equations*. Numer. Math. **82** (1999), 91-116.

[40] Eymard R., Gallouët T., Herbin R., *Finite volume approximation of elliptic problems and convergence of an approximate gradient*. Appl. Numer. Math. **37** (2001), no. 1-2, 31-53.

[41] Eymard R., Gallouët T., Herbin R., Michel A., *Convergence of a finite volume scheme for nonlinear degenerate parabolic equations*. Numer. Math. **92** (2002), no. 1, 41-82.

[42] Fabrie P., Gallouët T., *Modeling wells in porous media flow*. Math. Models Methods Appl. Sci. **10** (2000), no. 5, 673-709.

[43] Farkas W., Jacob N., Schilling R. L., *Feller semigroups, L^p-sub-Markovian semigroups, and applications to pseudo-differential operators with negative definite symbols*. Forum Math. **13** (2001), no. 1, 51-90.

[44] Fiard J.M., Herbin R., *Comparison between finite volume finite element methods for the numerical simulation of an elliptic problem arising in electrochemical engineering*. Comput. Meth. Appl. Mech. Engin. **115** (1994), 315-338.

[45] Forsyth P.A., Sammon P.H., *Quadratic Convergence for Cell-Centered Grids*. Appl. Num. Math. **4** (1988), 377-394.

[46] Gallouët T., Herbin R., *Finite volume methods for diffusion problems and irregular data*. Finite volumes for complex applications, Problems and Perspectives, II, F. Benkhaldoun, M. Hänel and R. Vilsmeier eds, Hermes, 155-162 (1999).

[47] Gallouët T., Herbin R., *Convergence of linear finite elements for diffusion equations with measure data*. C. R. Acad. Sci. Paris, **338**, issue 1, 2004, 81-84.

[48] Gallouët T., Herbin R., Vignal M.H., *Error estimate for the approximate finite volume solutions of convection diffusion equations with Dirichlet, Neumann or Fourier boundary conditions*. SIAM J. Numer. Anal. **37** (2000), no. 6, 1935-1972.

[49] Gallouët T., Monier A., *On the regularity of solutions to elliptic equations*. Rend. Mat., VII **19** (1999), 471-488.

[50] Gisclon M., Serre D., *Conditions aux limites pour un système strictement hyperbolique fournies par le schéma de Godunov*. RAIRO Modél. Math. Anal. Numér. **31** (1997), 359-380.

[51] Grenier E., Guès O., *Boundary layers for viscous perturbations of noncharacteristic quasilinear hyperbolic problems*. J. Differential Equations **143** (1998), 110-146.

[52] GRISVARD P., Elliptic problems in nonsmooth domains. Pitman 1985.

[53] GUÈS O., *Perturbations visqueuses de problèmes mixtes hyperboliques et couches limites.* Ann. Inst. Fourier (Grenoble) **45** (1995), 973-1006.

[54] GUIBE O., BEN CHEIKH M., *Résultats d'existence et d'unicité pour une classe de problèmes non linéaires et non coercitifs.* C. R. Acad. Sci. Paris Sér. I Math., **329** (1999), 11, 967-972.

[55] HERBIN R., *An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh.* Num. Meth. P.D.E. **11** (1995), 165-173.

[56] HERBIN R., *Finite volume approximation of elliptic problems with irregular data.* Finite volumes for complex applications, Problems and Perspectives, F. Benkhaldoun, D. Hanel and R. Vilsmeier eds, Hermes, 1999, 153-160.

[57] HOH W., *Pseudodifferential operators with negative definite symbols and the martingale problem.* Stochastics Stochastics Rep. **55** (1995), no. 3-4, 225-252.

[58] IMBERT C., VOVELLE J., *A kinetic formulation for multidimensional scalar conservation laws with boundary conditions and applications.* To appear in "SIAM - Mathematical Analysis".

[59] JOSEPH K. T., LEFLOCH P. G., *Boundary layers in weak solutions of hyperbolic conservation laws.* Arch. Ration. Mech. Anal. **147** (1999), 47-88.

[60] KRUŽKOV S. N., *First order quasilinear equations with several independent variables.* Mat. Sb. (N.S.) **81 (123)** (1970), 228-255.

[61] KUZNECOV N. N., *The accuracy of certain approximate methods for the computation of weak solutions of a first order quasilinear equation.* Ž. Vyčisl. Mat. i Mat. Fiz. **16** (1976), 1489-1502, 1627.

[62] LADYZENSKAJA O. A., SOLONNIKOV V. A., URALČEVA N. N., *Linear and quasilinear equations of parabolic type.* Translations of Mathematical Monographs, Vol. **23**, American Mathematical Society, Providence R.I., 1967.

[63] LAZAROV R.D., MISHEV I.D., VASSILEVSKI P.S., *Finite volume methods for convection-diffusion problems.* SIAM J. Numer. Anal. **33** (1996), 31-55.

[64] LERAY J., LIONS J.L., *Quelques résultats de Višik sur les problèmes elliptiques semi-linéaires par les méthodes de Minty et Browder.* Bull. Soc. Math. France **93** (1965), 97-107.

[65] LÉVY P.. Calcul des Probabilités, 1925.

[66] LIONS J-L., MAGENES E., Non-Homogeneous Boundary Value Problems and Applications. Springer-Verlag (1972).

[67] LIONS P.-L., PERTHAME B., TADMOR E., *A kinetic formulation of multidimensional scalar conservation laws and related equations.* J. Amer. Math. Soc. **7** (1994), 169-191.

[68] MANTEUFEL T.A., WHITE A.B., *The numerical solution of second order boundary value problem on non uniform meshes.* Math. Comp. **47** (1986), no. 176, 511-535.

[69] MEYERS N.G., *An L^p estimate for the gradient of solutions of second order divergence equations.* Ann. Sc. Norm. Sup. Pisa **17** (1963), 189-206.

[70] MISHEV I., *Finite volume element methods for non-definite problems.* Numer. Math. **83** (1999), 161-175.

[71] MISHEV I.D., *Finite volume methods on Voronoï meshes.* Numer. Methods Partial Differential Equations **14** (1998), no. 2, 193-212.

[72] Nečas J., Les méthodes directes en théorie des équations elliptiques. Masson, 1967.

[73] Otto F., *Initial-boundary value problem for a scalar conservation law.* C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), 729-734.

[74] Perthame B., *Uniqueness and error estimates in first order quasilinear conservation laws via the kinetic entropy defect measure.* J. Math. Pures Appl. (9) **77** (1998), no. 10, 1055–1064.

[75] Pierre M., *Parabolic capacity and Sobolev spaces.* Siam J. Math. Anal. **14** (1983), 522-533.

[76] Prignet A., *Remarks on existence and uniqueness of solutions of elliptic problems with right-hand side measures.* Rend. Mat. Appl. (7) **15** (1995), no. 3, 321-337.

[77] Stampacchia G., *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus.* Ann. Inst. Fourier, Grenoble, **15** (1965), 189-258.

[78] Tadmor E., Tang T., *Pointwise error estimates for relaxation approximations to conservation laws.* SIAM J. Math. Anal. **32** (2000), 870-886.

[79] Tang T., *Error estimates of approximate solutions for nonlinear scalar conservation laws.* Hyperbolic problems: theory, numerics, applications, Vol. I, II (Magdeburg, 2000), vol. 141 of Internat. Ser. Numer. Math., 140, Birkhäuser, Basel, 2001, pp. 873–882.

[80] Vol'pert A. I., *Spaces* bv *and quasilinear equations.* Mat. Sb. (N.S.) **73 (115)** (1967), 255-302.

[81] Woyczyński W., *Lévy processes in the physical sciences.* Lévy processes, 241-266, Birkhuser Boston, Boston, MA, 2001.