Optimisation d'interfaces

Edouard Oudet

UNIVERSITÉ DE SAVOIE

Demande d'habilitation à diriger des recherches

Spécialité : Mathématiques appliquées

Présentée par

Édouard Oudet

Optimisation d'interfaces

Soutenue le premier décembre 2009 devant le jury composé de Messieurs

G. ALLAIRE	Rapporteur externe
E. BONNETIER	
Y. BRENIER	Rapporteur externe
D. BUCUR	
A. HENROT	
V. KOMORNIK	
B. MAURY	Rapporteur externe

À Thomas

Au moment de mettre un point final à ce manuscrit, ma première pensée va à mes collaborateurs, une grande partie du mérite de ces travaux leur revient. Qu'ils trouvent ici l'expression de ma plus sincère reconnaissance.

Merci aux membres du LAMA qui contribuent à ce que ce laboratoire soit un cadre de travail dynamique et agréable comme il en existe peu.

Je tiens à remercier Grégoire Allaire, Yann Brenier et Bertrand Maury d'avoir accepté de rapporter mon habilitation. Merci à Eric Bonnetier, Dorin Bucur, Antoine Henrot et Vilmos Komornik de participer à mon jury. J'en suis très honoré.

Mes parents ont leur part de responsabilités dans mon égarement universitaire. Pour cette liberté et pour tant d'autres choses, une nouvelle fois merci.

Merci à Aude, Capucine et Simon, vous rendez tout possible. Je vous aime.

Pour conclure, merci à toi Thomas qui a été le maître d'œuvre du début de cette histoire. Je suis tellement désolé que tu n'ais pu en connaître la suite. Tu me manques.

Table des matières

2 Optimisation de forme à plusieurs phases

3 Transport optimal et irrigation optimale

Préface

Ce mémoire, relatif à l'optimisation des interfaces, est composé de trois parties subdivisées en neuf chapitres. L'état en cours des publications scientifiques associées à chacun de ces chapitres est indiqué page xi.

La première partie (chapitres I à IV) porte sur les problèmes d'optimisation sous contraintes de convexité et de largeur constante. Nous y présentons de nouveaux résultats relatifs au problème de Newton et à la conjecture de Meissner.

La seconde partie (chapitres V à VII) s'intéresse à des problèmes d'optimisation dont la variable est un ensemble de domaines du plan ou de l'espace. Des méthodes numériques adaptées à l'approximation de pavages optimaux pour des problèmes de nature géométrique (le problème de Kelvin) et spectrale (une conjecture de L. A. Caffarelli & F. H. Lin) y sont développées.

La dernière partie (chapitres VIII et IX) examine des questions liées à la théorie du transport optimal. Le chapitre VIII propose une modélisation de l'effet de congestion alors que le dernier chapitre est dédié à la présentation de travaux en cours sur l'approximation de réseaux optimaux d'irrigation.

Première partie

Corps convexes et corps de largeur constante

A — Minimisation numérique sous contrainte de convexité

À mon arrivée au LAMA il y a maintenant 6 ans, Thomas Lachand-Robert m'a proposé de travailler numériquement sur un sujet qu'il avait beaucoup étudié : le problème de la minimisation de la résistance d'un corps à la pénétration dans l'air, qui avait été proposé par Isaac Newton. Sous des hypothèses très simplificatrices (milieu faible en particules, unicité d'impact des particules sur le corps, déplacement unidirectionnel), Newton modélisa un profil optimal comme le graphe d'une fonction solution de

$$\min_{u \text{ convexe}} \int_\Omega \frac{dx}{1+|\nabla u|^2}, \tag{1}$$

où $u : \Omega \to [0, M]$, $M > 0$ est un paramètre positif et Ω le disque unité de $\mathbb{R}^2$. Thomas souhaitait à l'époque poursuivre l'étude des propriétés qualitatives des formes optimales qui avait connu d'importants progrès ces dernières années. Rappelons ici quelques unes des grandes étapes qui ont jalonnées l'histoire du problème de Newton :

En 1687, Isaac Newton publie *Philosophiæ Naturalis Principia Mathematica* où il décrit la solution optimale radiale du problème. La contrainte de convexité est utilisée de manière implicite dans la modélisation. La solution de Newton, en raison de sa « partie plate » (voir le profil le plus à droite de la figure 1), ne laisse pas la communauté scientifique de l'époque indifférente .

En 1786, A.-M. Legendre critique la solution de Newton lorsqu'il dérive ses conditions d'optimalité du second ordre. En l'absence de contrainte de convexité, le profil de Newton n'est en aucun cas optimal.

À la fin du $XIX^{\text{ième}}$, F. August et E. Armanini complètent la démonstration de Newton.

En 1993, G. Buttazzo, V. Ferone et B. Kawohl démontrent l'existence d'une solution dans la classe générale des fonctions convexes.

En 1996, F.Brock, V. Ferone et B. Kawohl démontrent que la solution ne peut être radialement symétrique.

En 2000, M. Pelletier et T. Lachand-Robert montrent qu'un minimiseur de la fonctionnelle de Newton n'est nulle part strictement convexe.

En 2001, Après ce résultat de non stricte convexité, M. Pelletier et T. Lachand-Robert identifient les profils optimaux dans une classe naturelle de fonctions non strictement convexes : la classe des fonctions développables. Les graphes optimaux s'obtiennent comme l'enveloppe convexe d'un polygone régulier de $\Omega \times \{M\}$ et du cercle $\partial\Omega$. Le nombre de côtés du polygone étant déterminé par la valeur de M (voir la figure 1).

La question que souhaitait aborder Thomas était précisément l'étude du caractère développable des minimiseurs. S'il était possible de démontrer à priori une telle propriété, ses travaux antérieurs apportaient une réponse complète au problème de Newton. C'est ce point précis qui a motivé le travail d'approximation numérique développé dans ce premier chapitre.

Cette problématique, issue du calcul des variations, se distingue de la formulation classique par le fait que la contrainte de convexité porte ici, non sur la fonction coût, mais sur l'état que l'on cherche à optimiser. Si l'on associe à chaque état le corps convexe que délimite son graphe, ce type de problème peut s'écrire de manière équivalente sous la forme intégrale

$$\inf_{A \in \mathcal{A}} \mathcal{F}(A), \text{ avec } \mathcal{F}(A) := \int_{\partial A} f(x, \nu_A(x), \varphi_A(x))\, d\mathcal{H}^2(x), \tag{2}$$

où $\mathcal{A}$ désigne un ensemble de corps convexes adéquat, ∂A est le bord du corps convexe A, ν_A son champ normal extérieur, φ_A sa fonction support et où f est une fonction régulière de ses arguments.

D'autres problèmes que celui de la résistance minimale de Newton entrent dans la catégorie ci-dessus. Lorsque nous avons commencé à étudier cette thématique, nous étions plus partic-

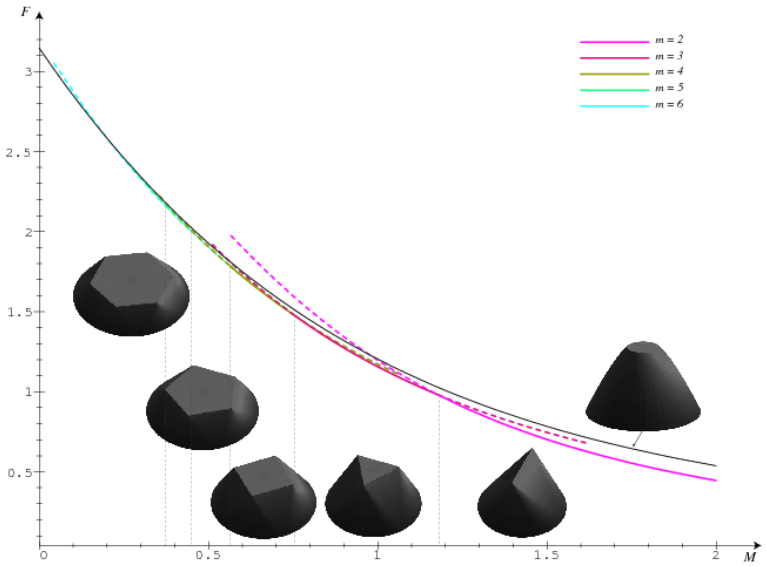

FIG. 1 – Formes optimales développables pour le problème de Newton (T. Lachand-Robert & M. Pelletier)

ulièrement intéressés par trois d'entre eux :

- le problème de Newton,
- le problème de Cheeger, qui pour un domaine borné de l'espace, consiste à chercher un sous ensemble mesurable qui minimise le quotient de la surface du bord par le volume de l'ensemble,
- le problème d'Alexandrov qui, étant donné une famille finie de couples (a_i, ν_i) de $\mathbb{R}_+ \times \mathbb{S}^2$ cherche à reconstruire (quand il existe) un polytope dont l'ensemble des vecteurs normaux est exactement la famille (ν_i) et dont chacune des faces de vecteur normal ν_i a pour aire a_i.

Une difficulté fondamentale associée à l'approximation numérique d'un corps sous la contrainte de convexité vient du fait qu'il n'est pas possible, pour ce type de problème, de dériver une équation d'optimalité d'Euler Lagrange sans information à priori, sur la régularité d'un corps optimal et sur sa stricte convexité. Nous verrons dans les résultats que nous présentons plus loin que ces situations défavorables apparaissent dans les problèmes qui nous intéressent.

Rappelons que d'autres auteurs s'étaient intéressés à la gestion numérique de cette contrainte. En particulier, P. Choné et H. Le Meur ont montré qu'une approche classique basée sur une discrétisation par éléments finis de type $\mathbb{P}_1$ peut conduire à des phénomènes de non convergence. Plus précisément, ils ont établi que l'ensemble des fonctions $\mathbb{P}_1$ convexes associées à une suite de raffinements d'un maillage peut ne pas être dense dans H^1 si ce maillage initial comporte certaines anisotropies. Afin de contourner cette difficulté, T. Lachand-Robert, G. Carlier et B. Maury ont proposé un algorithme d'approximation extérieur, en ce sens que l'espace d'approximation n'est pas inclus dans l'ensemble des fonctions convexes, ce qui amène à la résolution d'un problème d'optimisation sous un grand nombre (relativement au nombre de noeuds du maillage) de contraintes d'inégalité linéaires.

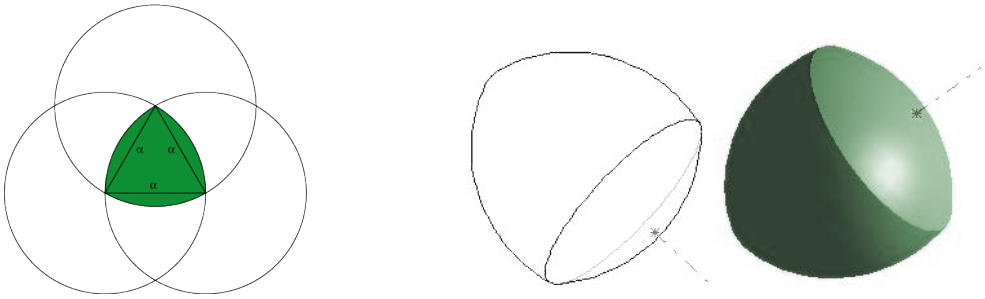

FIG. 2 – Le triangle de Reuleaux et le triangle de Reuleaux tourné

L'approche que nous proposons dans le chapitre I repose sur une représentation implicite des corps convexes basée sur leur fonction support couplée à une méthode classique de gradient projeté. De manière plus standard en géométrie algorithmique, notre approche revient tout simplement à chercher à approcher un convexe optimal comme une intersection finie de demi-espaces (qui correspondent à la discrétisation de la fonction support comme une somme de mesures de Dirac). Par la définition de notre paramétrisation, notre méthode est de type intérieur : tout au long du processus d'optimisation l'ensemble des paramètres discrets décrit un polytope convexe admissible pour notre problème. En revanche, l'évaluation de certaines quantités associées au corps convexe comme son volume, sa surface ou toute autre quantité intégrale du type (2) peuvent nécessiter la reconstruction du polytope associé à une fonction support. Cette étape, qui a été une des questions les plus étudiées en géométrie algorithmique, nécessite en dimension 2 et 3 au plus $n \log n$ opérations où n représente le nombre de demi-espaces. Nous décrivons dans ce premier chapitre les détails de la mise en oeuvre d'une telle approche ainsi que le calcul de la variation de fonctions coût de type intégrale par rapport à la fonction support. À la fin du chapitre nous illustrons l'effectivité de notre méthode sur les trois problèmes pré-cités.

Le résultat le plus marquant que notre méthode ait pu obtenir est une réponse numérique satisfaisante à la question de la développabilité des formes optimales du problème de Newton. Contrairement à ce que pouvait laisser présager le résultat de non stricte convexité, ces formes optimales ne sont pas développables (voir figures I.5 et I.7) en raison de lignes de singularités liant le polygone sommital et l'arc de cercle du support.

B — Corps de largeur constante et la conjecture de Meissner

Les chapitresII, III et IV sont dédiés à l'étude des objets de largeur constante. Un corps convexe de $\mathbb{R}^N$ est dit de *largeur constante* si sa projection sur toute droite est un segment de même longueur. Les boules sont bien sûr de largeur constante mais ce ne sont pas les seules ; des continuums d'autres corps convexes vérifient aussi cette propriété.

De nombreux travaux du $XIX^{\text{ième}}$ et de la première moitié du $XX^{\text{ième}}$ siècle ont porté sur l'étude des propriétés géométriques en dimension 2 de ces objets très particuliers. En particulier Frank Reuleaux, dont le nom est aujourd'hui associé au convexe obtenu comme intersection de trois disques de mêmes rayons R dont les centres sont les sommets d'un triangle équilatéral de côté R (voir figure 2), ainsi que H. Lebesgue et W. Blaschke apportèrent les contributions les plus

significatives de cette époque. Ces deux derniers auteurs démontrèrent que le triangle de Reuleaux est l'unique objet qui minimise l'aire parmi les objets de largeur constante fixée.

Au vu de ce résultat, il est naturel de conjecturer que la « pyramide de Reuleaux », intersection de quatre boules de même rayon R dont les centres sont les sommets d'un simplexe régulier de côté R, minimise le volume parmi les objets de largeur constante R. Il n'en est rien ! Bien qu'il soit facile d'obtenir des objets de largeur constante en dimension 2, comme intersection de disques, une telle approche n'est plus possible en dimension supérieure où l'intersection d'un nombre fini de boules n'est jamais de largeur constante. Une manière élémentaire de générer de tels objets en dimension 3 est de considérer les solides de révolution obtenus par la rotation autour d'un axe de symétrie d'un objet de largeur constante du plan. F. Meissner démontra que le solide obtenu à partir du triangle de Reuleaux, appelé « triangle de Reuleaux tourné » (voir la figure 2) minimise le volume parmi les corps de révolution de largeur constante fixée. Quelques années plus tard, F. Meissner proposa une construction d'un objet de largeur constante ayant un volume plus petit que le « triangle de Reuleaux tourné ». Ce solide qui est obtenu en lissant trois des arêtes de la « pyramide de Reuleaux » est appelé le « tétraèdre de Meissner » (voir la figure III.1).

La question de l'optimalité du tétraèdre de Meissner est, encore aujourd'hui, un problème ouvert. C'est cette question qui est à l'origine des travaux présentés dans ces trois chapitres.

Nous présentons dans le chapitre II plusieurs caractérisations géométriques des ensembles de largeur constante en dimension quelconque. Ce travail constitue une première étape essentielle dans la paramétrisation de tels objets. Ces différentes caractérisations nous fournissent un procédé de construction canonique pour générer un ensemble de largeur constante de dimension N à partir d'une de ses projections (de largeur constante) en dimension $N-1$. De plus, il est à noter que ce procédé est, à notre connaissance, la première description « canonique » du « tétraèdre de Meissner ». Plus précisément, en l'appliquant à un segment (qui est le seul objet de largeur constante en dimension 1), on retrouve le « triangle de Reuleaux ». Partant du triangle de Reuleaux, on obtient l'un des « tétraèdre de Meissner ». Nous proposons une illustration du corps de largeur constante de dimension 4 obtenu à partir du « tétraèdre de Meissner » (voir la planche II.4).

Le chapitre III porte sur la paramétrisation des objets de largeur constante de l'espace. On introduit une bijection entre de tels objets géométriques et un espace fonctionnel lié à la paramétrisation du bord de l'objet par ses normales. Les contraintes qui définissent cet espace fonctionnel sont de deux natures : des propriétés d'anti-symétrie liées à la périodicité de la sphère ainsi que des contraintes de positivité d'ordre deux (i.e. faisant intervenir des combinaisons linéaires et quadratiques des dérivées secondes). Certaines quantités géométriques simples comme l'aire ou le volume sont calculées en fonction des représentants de cet espace fonctionnel. Cette nouvelle approche nous permet de proposer un nouvelle démonstration de l'identité de W. Blaschke liant l'aire et le volume des objets de largeur constante en dimension 3. À notre connaissance, cette démonstration est la première de nature purement algébrique. Pour finir nous en déduisons une condition d'optimalité faible pour les corps de largeur constante minimisant le volume en dimension 3. Cette condition est une conséquence de la nature concave de la surface pour notre paramétrisation. Il est à noter que le « tétraèdre de Meissner » satisfait cette condition d'optimalité : pour tout sous ensemble ω de la sphère assez petit, les deux morceaux de surfaces d'un corps optimal dont les normales sont ω et $-\omega$ ne peuvent être tous les deux des ensembles réguliers. Dans le cadre du solide de Meissner, les sommets singuliers du simplexe correspondent aux morceaux sphériques de la frontière alors que les arcs de cercles, eux aussi singuliers, correspondent aux arêtes lissées. Ce dernier résultat est, de

notre point de vue, un premier pas significatif dans la démonstration de la conjecture de Meissner.

Le chapitre IV porte sur l'étude numérique des problèmes d'optimisation sous contraintes de largeur. Nous nous intéressons à la fois à des contraintes de type largeur constante, comme aux chapitres précédents, et à des contraintes de type inégalité. Ces deux questions concernent des objets de nature profondément différente. Alors que des polytopes peuvent vérifier des contraintes d'inégalité liée à leurs largeurs, ils ne sont jamais de largeur constante : tout corps de largeur constante est toujours strictement convexe. Pour chacune de ces deux situations nous introduisons des méthodes de type interne afin de pouvoir confronter nos résultats numériques aux conjectures du domaine.

Pour travailler avec des objets de largeur constante, nous utilisons une discrétisation de la paramétrisation introduite au chapitre III par des splines cubiques. Le point clé de ce travail réside en l'introduction d'une linéarisation interne des contraintes amenant à la résolution d'un problème d'optimisation quadratique standard. Par cette approche, nous vérifions numériquement l'optimalité locale du solide de Meissner. De plus, nous illustrons la non locale optimalité du « triangle de Reuleaux tourné » parmi les ensembles de largeur constante. Partant du « triangle de Reuleaux tourné » nous obtenons une forme discrète singulière (voir la figure **??**) ayant une surface moindre et vérifiant les conditions d'optimalité faible décrites précédemment.

Pour traiter numériquement des problèmes d'inégalité, nous décrivons des ensembles convexes comme des combinaisons convexes de Minkowski de polytopes élémentaires. Grâce à la théorie de Brunn-Minkowski, il est possible de calculer des quantités géométriques associées à ces combinaisons convexes sans pour autant être capable de décrire complètement (par ses sommets par exemple) un tel polytope. Nous illustrons l'effectivité de notre approche algébrique en l'utilisant pour étudier une conjecture due à E. Heil. Le polytope que nous obtenons nous autorise à infirmer cette conjecture.

Deuxième partie

Optimisation de forme à plusieurs phases

C — Minimisation à volume d'ensemble de niveaux fixés

Le chapitre V porte sur l'analyse de problèmes variationnels où le volume d'ensembles de niveaux de l'état est fixé. Ce type de problèmes apparaît dans la modélisation de fluides non miscibles et de mélanges de matériaux micro-magnétiques. Lorsque la contrainte sur l'état porte sur plus d'un ensemble de niveaux, l'existence d'un état d'équilibre n'a été établie que pour des énergies très spécifiques (voir en particulier les travaux antérieurs de L. Ambrosio, P. Marcellini, I. Foncesca et L. Tartar).

L'étude des méthodes à frontière libre de type « multi-level set » est un sujet de recherche très actif. La gestion numérique de la non intersection des ensembles de niveaux est un problème crucial et délicat dans bon nombre d'applications. Nous sommes ici dans un cadre favorable où nous démontrons et illustrons qu'une situation d'auto-intersection ne peut se développer dans notre

contexte. Notre approche a permis d'illustrer aussi bien des phénomènes d'absence d'existence que la non unicité de points critiques pour l'énergie de Dirichlet.

Les deux chapitres suivants portent sur la recherche de partitions optimales. Dans ce contexte, les méthodes « multi-level set » sont délicates à mettre en oeuvre en raison du grand nombre de fonctions niveaux à faire évoluer et de la non convexité des fonctionnelles étudiées. Nous proposons deux alternatives basées sur une description des partitions en termes de densités autorisant une convexification (au moins partielle) des fonctionnelles étudiées.

D — Une conjecture due à L. A. Caffarelli & F. H. Lin

L'étude de la dépendance des modes propres de Dirichlet par rapport au domaine sur lequel est défini l'opérateur Laplacien, remonte aux travaux de Lord Rayleigh sur l'analogue spectral de l'inégalité isopérimétrique. Un des objectifs de ces travaux est de lier des quantités différentielles, que sont les modes d'un domaine, à ses caractéristiques géométriques. En 1923, E. Krahn et G. Faber démontrent que la boule minimise la première valeur propre du Laplacien-Dirichlet sous contrainte de volume. Ainsi, dans ce contexte, la première valeur propre joue le rôle du périmètre. Quelques années plus tard, ils établissent que l'union de deux boules disjointes et de même rayon minimise la seconde valeur propre sous cette même contrainte de volume.

Un problème bien plus récent, dû à L. A. Caffarelli & F. H. Lin, propose d'étudier le comportement asymptotique de partitionnement optimaux. Plus précisément, étant donné un domaine borné du plan, on s'intéresse aux partitions de ce domaine en n ensembles telles que la somme des premiers modes de Dirichlet soit minimale. La conjecture de L. A. Caffarelli & F. H. Lin porte sur le comportement asymptotique de telles partitions optimales lorsque n tend vers l'infini : est-il vrai que les partitions optimales « tendent » (au sens de la valeur de la somme de leurs modes) vers un partitionnement hexagonal régulier ?

Une telle conjecture peut, à première vue, paraître surprenante (voir artificielle…). Pour en comprendre l'origine, il convient de la mettre en parallèle avec la fameuse « conjecture des nids d'abeilles » démontrée rigoureusement en 1999 par T. C. Hales. Il s'agit dans cette dernière d'identifier le pavage du plan en des cellules de même aire de telle manière à minimiser le périmètre (en un sens asymptotique) de ce pavage : le pavage du plan par des hexagones réguliers est la solution de ce problème. À la lumière de ce résultat, la conjecture de L. A. Caffarelli & F. H. Lin est très naturelle : si la première valeur propre joue un rôle analogue au périmètre comme dans l'inégalité spéctrale de Lord Rayleigh, il est raisonnable d'envisager qu'asymptotiquement, les hexagones seront aussi optimaux dans cette situation.

Le chapitre VI porte sur l'étude théorique et numérique des partitionnements optimaux d'un domaine, associés à la somme des valeurs propres du Laplacien-Dirichlet. L'originalité de ce travail réside dans la complémentarité des résultats théoriques et numériques qu'il propose : une analyse fine, basée sur une formulation à l'aide de mesures du problème de partitionnement optimal, nous a permis de traiter numériquement ce problème d'optimisation de très grande taille (voir la figure VI.6). Bien que ce problème soit concave, l'approche récursive et parallèle que nous avons mise en oeuvre a permis d'exhiber des structures asymptotiquement en accord avec la conjecture.

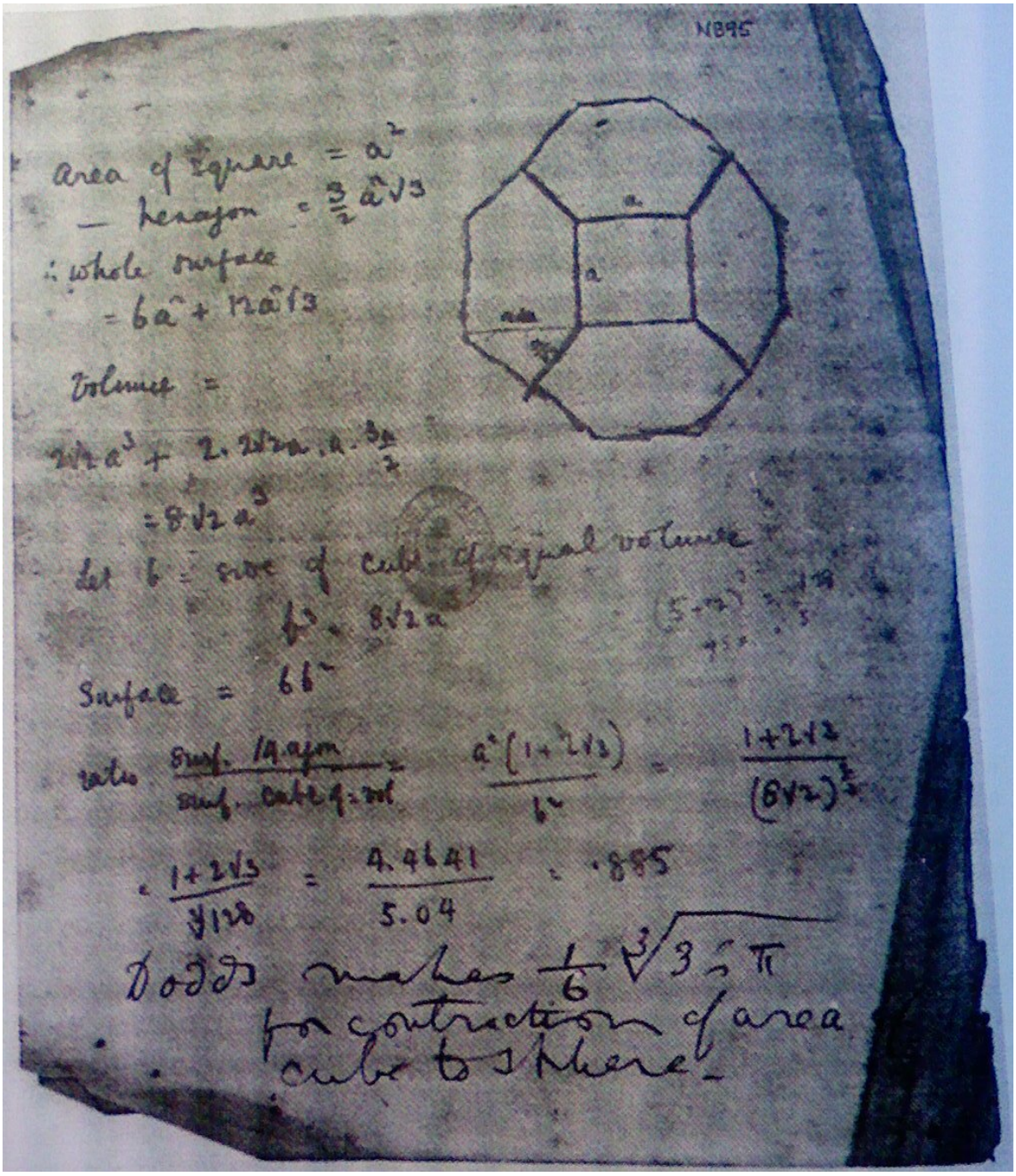

FIG. 3 – L'octaèdre tronqué dans des notes de Lord Kelvin (d'après « The physics of foams » de D. Weaire

E — Le problème de Kelvin

En 1894, Lord Kelvin se propose d'étudier l'analogue de la « conjecture des nids d'abeilles » en dimension 3 : quel est le découpage de l'espace en cellules de même volume qui minimise la mesure surfacique totale ? Il conjectura qu'un pavage proche de celui obtenu par des octaèdres tronqués est solution de ce problème (voir la figure 3). Il est à noter que, contrairement au cas de la dimension 2, la condition d'optimalité ne garantit pas que le bord du pavage soit affine par morceaux. Cette condition du premier ordre (sous réserve de régularité...) affirme que le bord est localement à courbure moyenne constante ce qui était le cas de la construction de Kelvin. Cette dernière avait de plus l'avantage de satisfaire aux conditions angulaires d'optimalité déduite par Joseph Antoine Ferdinand Plateau.

En 1996, deux physiciens D. Weaire et P. Phelan identifient, par de l'optimisation locale de maillages, un pavage constitué de deux types de cellules qui améliore le coût du pavage de Kelvin de 0.02%. Ces deux types de cellules ont respectivement 12 et 14 faces qui sont de nature pentagonales et hexagonales (voir la figure 4).

Nous exposons dans le chapitre VII une approche numérique de cette question basée sur les résultats de Γ-convergence de L. Modica, S. Mortola et S. Baldo. Nous proposons une extension au cadre périodique, des résultats de Γ-convergence vers la mesure surfacique. La difficulté de cette généralisation provient du fait que les morceaux de frontière qui intersectent le bord du domaine, contrairement au cadre classique, doivent être comptabilisés dans le coût limite. Notre formulation sous contraintes linéaires permet d'autre part de travailler avec des potentiels dont le degré reste faible. Cette modeste originalité par rapport au résultat de S. Baldo est d'une importance capitale dans notre approche numérique. Basée sur cette relaxation, nous illustrons l'efficacité de cette « convexification » par Γ-convergence : bien que des branchements, lors du processus d'optimisation, puissent amener à des minima locaux, nos résultats sont d'une surprenante régularité. Partant de 16 cellules paramétrées par des densités initialement aléatoires, notre algorithme fait converger ces densités vers le pavage de Kelvin. Partant de 8 cellules, elles aussi initialisées aléatoirement (contrairement aux travaux numériques de D. Weaire et P. Phelan), nous retrouvons le pavage bi-cellulaire de D. Weaire et P. Phelan (voir la figure VII.4). D'autres expériences numériques visant à améliorer le précédent pavage ont été menées à bien, malheureusement sans succès, laissant penser que cette structure bi-cellulaire est globalement optimale (voir la figure VII.5).

TROISIÈME PARTIE

Transport optimal et irrigation optimale

F — Effet de congestion en transport optimal

La théorie du transport de masse, introduite par Gaspard Monge en 1781 dans son « Mémoire sur la théorie des déblais et des remblais », a connu d'importants développements ces dernières

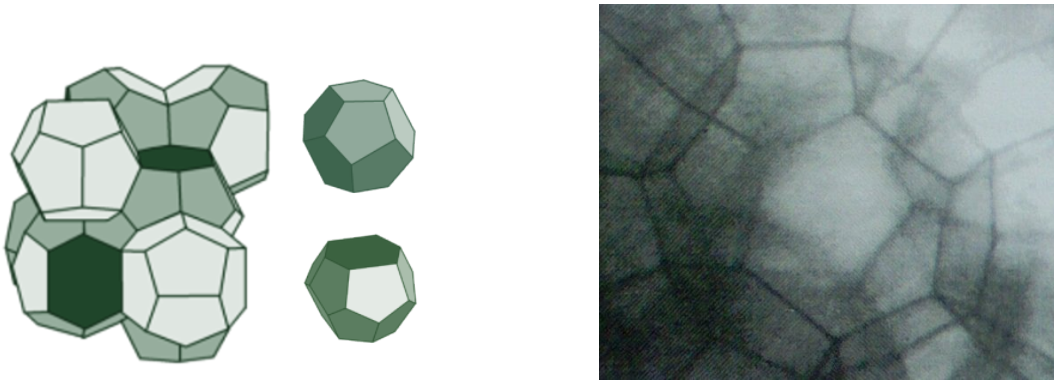

FIG. 4 – Le pavage de D. Weaire et R. Phelan : observation expérimentale de films minces présentant cette structure à l'équilibre (d'après « The physics of foams » de D. Weaire)

années. La modélisation adoptée par G. Monge est par définition de nature statique : la variable du problème est une application de transport T qui associe à une densité initiale une densité cible ne tenant pas compte des étapes transitoires du transport. Dans un tel contexte, des éventuels effets de congestion ne peuvent être pris en compte. L'objet du chapitre VIII est de proposer un cadre permettant de modéliser ces phénomènes.

La première formulation dynamique équivalente a été proposée par J.-D. Benamou et Y. Brenier en 2000 dans le cadre de la distance de Wasserstein 2. Dans ce chapitre, nous proposons une formalisation synthétique des problèmes de transport sous une forme pouvant décrire l'approche de type dynamique des fluides introduite par J.-D. Benamou et Y. Brenier. Plus précisément, introduisant la nouvelle variable (ρ, E), où E désigne le champ des vitesses multiplié par la densité, notre formalisme décrit un transport optimal comme la solution d'un problème de minimisation en cette variable sous une contrainte de divergence. Dans ce contexte, nous proposons différentes adaptations de la fonctionnelle coût permettant de tenir compte de l'effet de congestion dans une dynamique de transport. Suivant une approche de type Lagrangien augmenté introduite par J.-D. Benamou et Y. Brenier, nous décrivons une méthode numérique d'approximation des champs optimaux (ρ, E) permettant d'apprécier qualitativement les effets qu'induisent ces modifications de la fonctionnelle. Dans nos expérimentations numériques, nous nous intéressons en particulier aux effets de l'ajout des termes $\int \rho^2$ et $\chi_{\rho \leq M}$ à la distance classique de Wasserstein 2 pour prendre en compte les fortes concentrations qui peuvent apparaître lors d'un transport en présence d'obstacles.

G — Quelques perspectives : approximation de réseaux d'irrigation

Ce dernier chapitre constitue une brève description de travaux en cours concernant l'approximation numérique des réseaux d'irrigation optimaux. L'objectif est, à terme, de pouvoir proposer des méthodes performantes pour modéliser et simuler la croissance des structures minces en dimension 2 et 3 par des mouvements minimisants. Ce travail a toute sa place dans ce mémoire car il est l'aboutissement de travaux de plusieurs chapitres précédents. Le point de départ est un résultat de Γ-convergence obtenu récemment par F. Santambrogio relatif aux réseaux d'irrigation. Comme dans le chapitre VII, où le calcul de mesures surfaciques est remplacé par l'évaluation de l'énergie

relaxée de L. Modica et S. Mortola, il est possible d'associer une énergie à une densité vectorielle jouant le rôle asymptotiquement d'un coût d'irrigation. Suivant la démarche introduite au chapitre VII, nous avons mis en oeuvre une approche de type gradient conjugué projeté afin d'approximer séquentiellement les minima des fonctionnelles relaxées. Notre approche a toutefois nécessité des adaptations importantes en raison de la contrainte de divergence ainsi que de la nature singulière de l'objet recherché. Ce dernier point nécessitant l'utilisation de grilles fines, la plus grande attention a dû être portée à l'opérateur de projection associé à la contrainte de divergence. De manière analogue à l'algorithme de J.-D. Benamou et Y. Brenier que nous avons évoqué au chapitre précédent, l'étape de projection nécessite la résolution d'un problème de Poisson de grande taille. Afin de rendre notre démarche effective sur des grilles fines en dimension 2 et 3 nous avons mis en oeuvre une méthode de type Fourier à même de résoudre de manière efficace l'inversion du système linéaire de Poisson discret.

Nous présentons les premiers résultats d'approximation que nous avons obtenus en faisant varier le paramètre de concavité associé au coût du transport. On retrouve dans ces simulations la corrélation attendue entre ce paramètre et le nombre de branchements. Nous concluons notre présentation par les perspectives d'applications de ce travail.

H — Situation des publications de ce mémoire

Chapitre I : T. Lachand-Robert & É. Oudet, *Minimizing within convex bodies using a convex hull method*, SIAM Journal on Optimization, **16.2** (2006), pp. 368–379.

Chapitre II : T. Lachand-Robert & É. Oudet, *Bodies of constant width in arbitrary dimension*, Math. Nachrichten, **280** (2007), pp. 740–750.

Chapitre III : T. Bayen, T. Lachand-Robert & É. Oudet, *Analytic parametrization and volume minimization of three dimensional bodies of constant width*, Archive for Rational Mechanics and Analysis, **186** (2007), pp. 225–249.

Chapitre IV : É. Oudet, *Numerical shape optimisation under width constraint*, actuellement soumis.

Chapitre V : É. Oudet & M. O. Rieger, *Local minimizers of functionals with multiple volume constraints*, ESAIM COCV, **14** (2008), pp. 780–794.

Chapitre VI : B. Bourdin, D. Bucur & É. Oudet, *Optimal partition for eigenvalues*, accepté pour publication dans SIAM Journal on Scientific Computing.

Chapitre VII : É. Oudet, *Approximation of partitions of least perimeter by* Γ*-convergence : around Kelvin's conjecture*, actuellement soumis.

Chapitre VIII : G. Buttazzo, C. Jimenez & É. Oudet, *An optimization problem for mass transportation with congested dynamics*, SIAM Journal on Control and Optimization, **48** (2009), pp. 1961–1976.

I — Autres publications postérieures à la soutenance de thèse

– M. Belloni & É. Oudet, *Minimal gap between* $\Lambda_2(\Omega)$ *and* $\Lambda_\infty(\Omega)$ *in a class of convex domain*, Journal of convex Analysis, **15** (2008), pp. 507–521.

– D. Bucur , I .Durus & É. Oudet, *The conductivity eigenvalue problem*, Control and Cybernetics, (2008).
– R. Hassani, I. R. Ionescu & É. Oudet, *Critical friction and wedged configurations : A genetic algorithm approach*, Internat. J. Solids Structures, **44** (2007), pp. 6187–6200.
– E. Sonnendruc, F. Filbet, A. Friedman, É. Oudet & J.-L. Vay, *Vlasov simulations of beams with a moving grid*, Computer Physics Communications, **164** (2004), pp. 390–395.

Première partie

Corps convexes et corps de largeur constante

Minimizing within convex bodies using a convex hull method

Thomas Lachand-Robert & Édouard Oudet

I.1 Introduction

In this paper, we present numerical methods to solve optimization problems among convex bodies or convex functions. Several problems of this kind appear in geometry, calculus, applied mathematics, etc. As applications, we present some of them together with our corresponding numerical results.

Dealing with convex bodies or convex functions is usually considered easier in optimization theory. Unfortunately, this is not true when *the optimization space itself* is (a subset of) the set of convex functions or bodies. As an example, consider the following minimization problem, where $M > 0$ is a given parameter, Ω a regular bounded convex subset of $\mathbb{R}^n$ and g a continuous function on $\Omega \times \mathbb{R} \times \mathbb{R}^n$):

$$\inf_{u \in C_M} \int_\Omega g(x, u(x), \nabla u(x))\, dx, \tag{1}$$
$$\text{where } C_M = \{u : \Omega \to [-M, 0], \quad u \text{ convex } \}.$$

I.1.1 Convexity constraint

Without the convexity constraint, this problem is usually handled in a numerical way by considering the associated Euler equation $g_2'(x, u(x), \nabla u(x)) = \operatorname{div} g_3'(x, u(x), \nabla u(x)))$. Such an equation is discretized and solved on a mesh defined on Ω (or more precisely, a sequence of meshes, in order to achieve a given precision), using for instance finite element methods.

These classical numerical methods do not work at all with our problem:

1. The convexity constraint prevents us to use an Euler equation. In fact, just stating a correct Euler equation for this sort of problem is a difficult task [12, 20, 8]. Discretizing the corresponding equation is rather difficult, then.

2. The set C_M of admissible functions, considered as a subset of a Sobolev space like $H^1_{\text{loc}}(\Omega)$, is compact [5]. This makes it easy to prove the existence of a solution to (1) without any other assumption on g. But this also implies that C_M is a very small subset of the functions space, with empty interior. Therefore most numerical approximations of a candidate function u are not convex. Evaluating the functional on those approximations is likely to yield a value much smaller than the sought minimum.

3. The natural way to evade the previous difficulty is to use only convex approximations. For instance, on a triangular mesh of Ω, it is rather easy to characterize those P1-functions (that is, continuous and affine by parts functions) which are convex. Unfortunately, such an approximation introduces a geometric bias from the mesh. The set of convex functions that are limits of this sort of approximation is much smaller than C_M [13].

4. Penalization processes are other ways to deal with this difficulty. But finding a good penalization is not easy, and this usually yields very slow algorithms, which in this particular case are not very convincing. This yields approximation difficulties similar to those given in 2 above.

A first solution for this kind of numerical problems was presented in [10], and an improved version is given in [9]. However the algorithms given in these references are not very fast, since they deal with a large number of constraints, and do not apply for those problems where local minimizers exist. The latter are common in the applications since there is not need for the functional itself to be convex to prove the existence of solution of (1): the mere compacity of C, together with the continuity of the functional on an appropriate function space, is enough.

I.1.2 A mixed-type algorithm

Our main idea to handle numerically (1) is to mix geometrical and numerical algorithms. It is standard that any convex body (or equivalently, the graph of any convex function) can be described as an intersection of half-spaces or as a convex hull of points. Our discretization consists in considering only a finite number of half-spaces, or a finite number of points (this is not equivalent, and choosing either mode is part of the method). Reconstructing the convex body is a standard algorithm, and computing the value of the functional is straightforward then. Obviously the convex hull algorithm used implies an additionnal cost that can not be neglected. On the other hand, this method makes it easy to deal with additionnal constraints like the fact that functions get values in $[0, M]$, for instance. We also show that it is possible to compute the derivative of the functional. Hence we may use gradient methods for minimization.

Note that since this always deals with convex bodies, we are guaranteed that the evaluations of the functional are not smaller than the sought minimum, up to numerical errors. Because the approximation process is valid for any convex body, we can ensure that all minimizers can be approximated arbitrary closely.

The detailed presentation of the method requires to explain how the half-spaces or points are moved, whether or not their number is increased, and which information on the specific problem is useful for this. We present quite different examples in our applications, in order to pinpoint the corresponding difficulties. Whenever the minimizer of the functional is not unique, gradient methods may get stuck in local minima. We present a "genetic algorithm" to deal with these, too.

In this paper, we concentrate on the three-dimensional settings. The two-dimensional case is much easier, and convex sets in the plane can be parametrized in a number of very simple ways. Even though our methods could be applied to dimensions $n \geq 4$, the convex hull computation may become too expensive.

I.1.3 Generalized problem

This algorithm's design does not involve any mesh or interpolation process. As an important consequence, we are not limited to convex functions but may also consider convex bodies. This allows us to study problems like

$$\inf_{A\in\mathcal{A}} \mathcal{F}(A), \quad \text{where } \mathcal{F}(A) := \int_{\partial A} f(x, \nu_A(x), \varphi_A(x))\, d\mathcal{H}^2(x), \tag{2}$$

and $\mathcal{A}$ is a subset of the class of closed convex bodies of $\mathbb{R}^3$. We make use of the notations:

- ∂A is the boundary of a convex body A;
- ν_A is the almost everywhere defined outer normal vector field on ∂A, with values on the sphere $\mathbf{S}^2$;
- $\varphi_A(x)$ is the signed distance from the supporting plane at x to the origin of coordinates;
- f is a continuous function $\mathbb{R}^3 \times \mathbf{S}^2 \times \mathbb{R} \to \mathbb{R}$.

As reported in [7], the problem (1) can be reformulated in terms of (2) whenever g depends only on its third variable. In this formulation $\mathcal{A}$ stands for the set of convex subsets of $Q_M := \Omega \times [0, M]$ containing $Q_0 = \Omega \times \{0\}$. Any convex body $A \in \mathcal{A}$ has the form

$$\mathcal{A} = \{(x', x_3) \in \Omega \times \mathbb{R}, 0 \le x_3 \le -u(x')\}, \text{ with } u \in C_M.$$

Therefore any $x \in \partial A \setminus Q_0$ has the form $x = (x', -u(x'))$, with $x' \in \Omega$. Then

$$\nu_A(x) = (\nabla u(x'), 1)/\sqrt{1 + |\nabla u(x')|^2},$$

and the function f is deduced from g by the relation $f(\nu) = \nu_3 g(\frac{1}{\nu_3}\nu')$, for every $\nu = (\nu', \nu_3) \in \mathbf{S}^2$. Several other problems with a geometrical background may also be formulated in a similar way.

Actually the formulation (2) allows us to study any problem of the form (1). It is enough to define $f(x, \nu, \varphi) = \nu_3 g(x', -x_3, \frac{1}{\nu_3}\nu')$, taking into account that $x = (x', -u(x'))$.

On the other hand, it is much more practical in the numerical implementation to consider functions f depending only on ν, φ. This avoids numerical integration on surfaces altogether, as explained in section I.2, hence reducing greatly the computation time. With such a restriction, only some problems of the form (1) can be considered. Since

$$\varphi_A(x) = \frac{1}{\sqrt{1 + |\nabla u(x')|^2}} (x' \cdot \nabla u(x') + u(x')),$$

we can handle functions g depending on $\nabla u(x')$ and the aggregate $x' \cdot \nabla u(x') + u(x')$.

I.2 Half-spaces and discretization

For every $\nu \in \mathbf{S}^2$ and every $\varphi \ge 0$, let us define the half-space of $\mathbb{R}^3$:

$$[\![\nu, \varphi]\!] = \left\{x \in \mathbb{R}^3,\ x \cdot \nu \le \varphi\right\}.$$

Lemma 1 *Let A be a convex body of $\mathbb{R}^3$. Then $\forall \varepsilon > 0$, there exists a convex polytope $P \supset A$ such that:*

$$|\mathcal{F}(P) - \mathcal{F}(A)| \leq \varepsilon.$$

Proof. Let us note

$$\partial^* A := \{a \in \partial A; \nu_A(a) \text{ exists}\}.$$

Let $(X_j)_{j\in\mathbb{N}}$ be a dense sequence of points in $\partial^* A$ and consider the sequence of convex polytopes $(P_j)_{j\in\mathbb{N}}$ defined by:

$$P_j := \bigcup_{k=1}^{j} [\![\nu_A(X_k), \varphi_A(X_k)]\!].$$

Clearly $P_j \supset A$ and $\lim_{j\to\infty} P_j = A$ for the Hausdorff distance. From a classical theorem of Rockafellar [22], for any $a \in \partial^* A$, and any sequence (p_j), converging to a, with $p_j \in \partial^* P_j$ for all j, we have $\nu_{P_j}(p_j)$ converges to $\nu_A(a)$. Since $\partial A \setminus \partial^* A$ is $\mathcal{H}^2$-negligible, we get $\mathcal{F}(P_j) \to \mathcal{F}(A)$. □

As every convex polytope is the finite intersection of half-spaces, the natural discretization of (2) is the finite dimensional problem:

$$\min_{N,\Phi} G(N, \Phi) \tag{3}$$
$$\text{where } N := (\nu_1, \ldots, \nu_k) \in (\mathbf{S}^2)^k, \Phi := (\varphi_1, \ldots, \varphi_k) \in \mathbb{R}^k,$$
$$G(N, \Phi) := \int_{\partial P} f(x, \nu_P(x), \varphi_P(x))\, d\mathcal{H}^2(x),$$
$$\text{and } P := P(N, \Phi) := \bigcap_{i=1}^{k} [\![\nu_i, \varphi_i]\!].$$

Notice that, whenever f does not depend explicitely on x, $G(N, \Phi)$ can be computed as a finite sum, namely

$$G(N, \Phi) = \sum_{i=1}^{k} f(\nu_i, \varphi_i)\mathcal{H}^2(F_i), \text{ where } F_i := [\![\nu_i, \varphi_i]\!] \cap \partial P.$$

This is of primary importance in the numerical algorithms. More general functions f require the computation of integrals like $\int_{F_i} f(x, \nu_i, \varphi_i)\, d\mathcal{H}^2(x)$, which are computationaly expensive.

I.2.1 Computation of the derivatives

In this paragraph we compute the derivatives of G, in order to use the results in a gradient-like method. We focus on the case where f depends only on ν, φ, since this is the special case used in our actual programs. Straightforward modifications can be done to handle the general case. It suffices to change the term $\frac{\partial f}{\partial \varphi_i}(\nu_i, \varphi_i)\ \mathcal{H}^2(F_i)$ by the integral $\int_{F_i} \frac{\partial f}{\partial \varphi_i}(x, \nu_i, \varphi_i)\, d\mathcal{H}^2(x)$, and similarly with the $\mathcal{H}^1$ term.

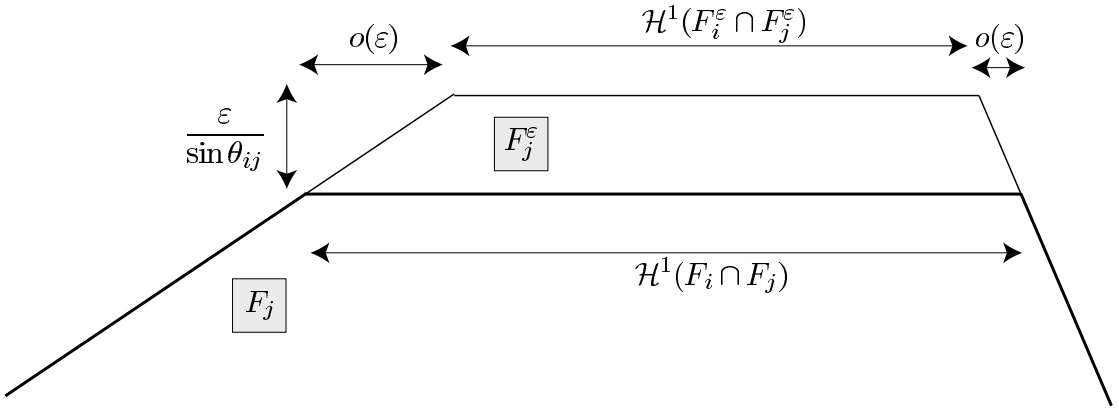

Figure I.1: Variation of the surface area of F_j (pictured in the plane of F_j), for the variation $\varphi_i \to \varphi_i + \varepsilon$.

Theorem 1 *Let $P := P(N, \Phi)$ be a convex polytope and $F_i = [\![\nu_i, \varphi_i]\!] \cap \partial P$. Then for almost every value of φ_i we have:*

$$\frac{\partial G}{\partial \varphi_i}(N, \Phi) = \frac{\partial f}{\partial \varphi_i}(\nu_i, \varphi_i)\, \mathcal{H}^2(F_i) + \sum_{\substack{j \neq i \\ \mathcal{H}^1(F_i \cap F_j) \neq 0}} \mathcal{H}^1(F_i \cap F_j) \left(\frac{f(\nu_j, \varphi_j) - \cos\theta_{ij} f(\nu_i, \varphi_i)}{\sin\theta_{ij}} \right), \quad (4)$$

where $\theta_{ij} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is defined by $\cos\theta_{ij} = |\nu_i \cdot \nu_j|$ and $\sin\theta_{ij}(\nu_i \cdot \nu_j) \geq 0$.

Proof. Let $\varepsilon > 0$ and consider the difference

$$G(\dots, \varphi_i + \varepsilon, \dots) - G(\dots, \varphi_i, \dots) = f(\nu_i, \varphi_i + \varepsilon)\, \mathcal{H}^2(F_i^\varepsilon) - f(\nu_i, \varphi_i)\, \mathcal{H}^2(F_i^\varepsilon) + \sum_j f(\nu_j, \varphi_j)\, (\mathcal{H}^2(F_j^\varepsilon) - \mathcal{H}^2(F_j))$$

where

$$F_j^\varepsilon = [\![\nu_j, \varphi_j]\!] \cap \partial P(\dots, \varphi_i + \varepsilon, \dots).$$

The first difference $f(\nu_i, \varphi_i + \varepsilon)\, \mathcal{H}^2(F_i^\varepsilon) - f(\nu_i, \varphi_i)\, \mathcal{H}^2(F_i^\varepsilon)$ has the form $\varepsilon \frac{\partial f}{\partial \varphi_i}(\nu_i, \varphi_i)\, \mathcal{H}^2(F_i) + o(\varepsilon)$.

To evaluate the remainding sum asymptotically we have to assume that the value of φ_i is such that there is no topological change in the polytope whenever φ_i becomes $\varphi_i + \varepsilon$. This is obviously true for all except a finite number of values of φ_i. We then distinguish two cases:

- $j \neq i$: $\mathcal{H}^2(F_j^\varepsilon) - \mathcal{H}^2(F_j) = \varepsilon \dfrac{\mathcal{H}^1(F_i \cap F_j)}{\sin\theta_{ij}} + o(\varepsilon)$ (see Figure I.1);
- $j = i$: $\mathcal{H}^2(F_i^\varepsilon) - \mathcal{H}^2(F_i) = -\varepsilon \displaystyle\sum_{\substack{j \neq i \\ \mathcal{H}^1(F_i \cap F_j) \neq 0}} \mathcal{H}^1(F_i \cap F_j) \cot\theta_{ij} + o(\varepsilon)$ (see Figure I.2).

This completes the proof of the theorem. □

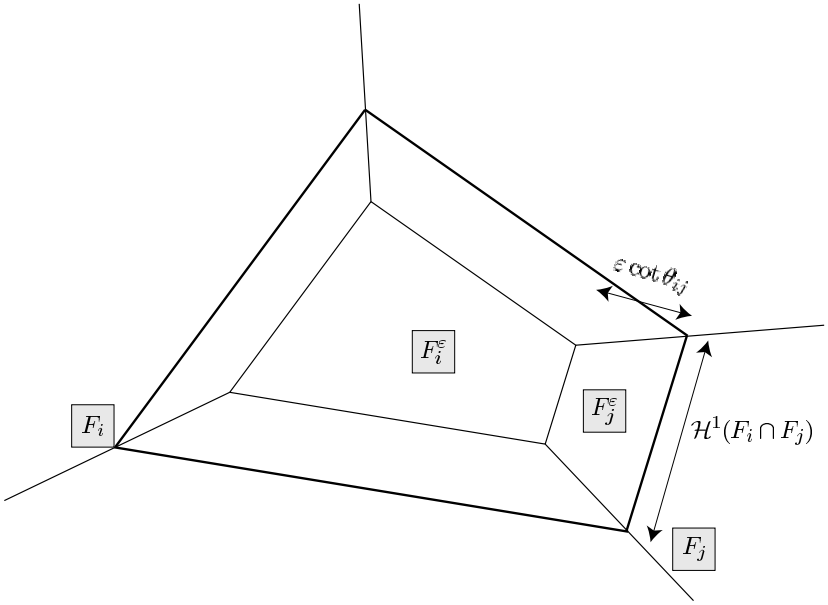

Figure I.2: Variation of the surface area of F_i (pictured in the plane of F_i), for the variation $\varphi_i \to \varphi_i + \varepsilon$.

Remark 2.A. The polyhedral representation used here, as an intersection of half-planes, yields a technical difficulty that should not be underestimated: some of the boundary planes $\partial[\![\nu_i, \varphi_i]\!]$ are "dormant", meaning the polytope is actually included in the interior of $[\![\nu_i, \varphi_i]\!]$.

In such a situation, formula (4) effectively yields zero, since $\mathcal{H}^2(F_i) = 0 = \mathcal{H}^1(F_i \cap F_j)$.

A similar computation can be achieved for derivatives of G with respect to ν_i, with another algebraic formula as a result. However numerical evidence proves that using a "full" gradient method is of little advantage.

It turns out that it is faster and accurate enough to use only the derivatives with respect to φ_i (as detailed in the next section), and to increase if necessary the number of hyperplanes by considering additional half-spaces. We can make profit of the "dormant" property by introducing these new ones in a tangent dormant position, letting the minimization method changing their position after that. This can be done in different ways, depending on the actual problem considered.

I.2.2 Summary of the algorithm

Thanks to Theorem 1, it is possible to apply a classical gradient algorithm to the problem (3). Let us summarize the different steps:

0. Choose one admissible polytope $P([\![\nu_1, \varphi_1^0]\!], \dots, [\![\nu_k, \varphi_k^0]\!])$, set $n = 0$.

1. Compute the geometry (vertexes, faces ...) of the polytope

$$P([\![\nu_1, \varphi_1^n]\!], \dots, [\![\nu_k, \varphi_k^n]\!]).$$

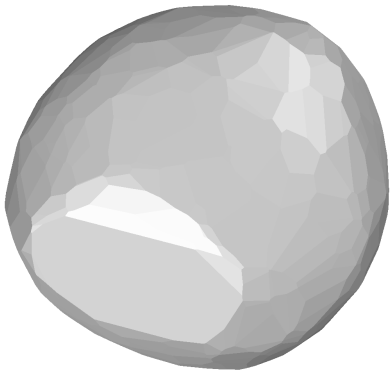

Figure I.3: A 1000 faces polyhedron of given faces areas and normals reconstructed.

2. Evaluate the gradient of G with respect to the φ_j using (4). If the euclidian norm of the gradient is small, then stop here.

3. Project the gradient into the set of admissible directions.

4. Set $\rho_n = \arg\min_{\rho>0} G(\nu_1, \ldots, \nu_k, \varphi_1^n - \rho\frac{\partial G}{\partial \varphi_1}, \ldots, \varphi_k^n - \rho\frac{\partial G}{\partial \varphi_k})$.

5. Define the new variables $\varphi_1^{n+1} = \varphi_1^n - \rho_n\frac{\partial G}{\partial \varphi_1}, \ldots, \varphi_k^{n+1} = \varphi_k^n - \rho_n\frac{\partial G}{\partial \varphi_k}$, $n \leftarrow n+1$ and go to step 1.

Step 3 in particular depends on the set of admissible bodies. So additional details are given in the examples hereafter.

I.2.3 Application to Alexandrov's Theorem

It is a classical result from Minkowski [21], that given n different vectors $\nu_1, \ldots, \nu_n$ on $\mathbf{S}^2$ such that the dimension of $\mathrm{Span}\{\nu_1, \ldots, \nu_n\}$ is equal to 3, and n positive real numbers $a_1, \ldots, a_n$ such that $\sum_{i=1}^n a_i\nu_i = 0$, then there exists a three-dimensional convex polytope having n faces $F_1, \ldots, F_n$ such that the outward normal vector to F_i equals ν_i and $\mathcal{H}^2(F_i) = a_i$. Moreover this polytope is unique up to translations.

This result has been extended by Alexandrov [1] to arbitrary convex bodies as follows: given a positive measure μ on $\mathbf{S}^2$ satisfying $\int_{\mathbf{S}^2} y\, d\mu(y) = 0$ and $\mathrm{Span}(\mathrm{supp}\,\mu) = \mathbb{R}^3$, then there exists a unique body A, up to translations, whose surface function measure is equal to μ.

G. Carlier proved recently [6] that this body is the unique (up to translations) solution of the variational problem

$$\sup_{\varphi\in\Sigma} \left|A_\varphi\right|, \tag{5}$$
$$\text{with } \Sigma := \left\{\varphi \in C^0(\mathbf{S}^2, \mathbb{R}_+);\ \int_{\mathbf{S}^2} \varphi\, d\mu = 1\right\} \text{ and } A_\varphi := \bigcap_{\nu\in\mathbf{S}^2} [\![\nu, \varphi(\nu)]\!],$$

where $\left|A_\varphi\right|$ is the volume of A_φ. Whenever A_φ is optimal, its support function equals φ on the support of μ [6].

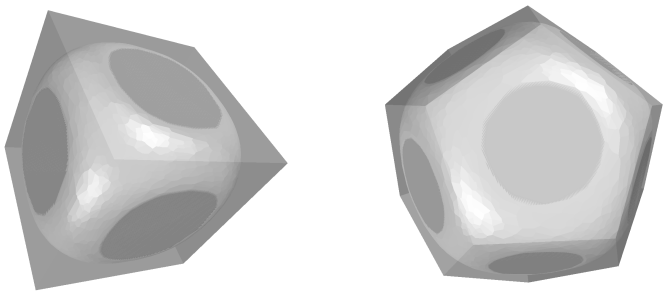

Figure I.4: Computed solutions for the Cheeger problem in the cube and the dodecahedron.

Now we recall that the volume of a convex body can be expressed as a boundary integral of its support function, that is:

$$|A| = \frac{1}{3}\int_{\partial A} \varphi_A(x)\, d\mathcal{H}^2(x).$$

Consequently Alexandrov's problem can be formulated in the form (2) with $f(x, \nu, \varphi) = -\varphi$ and

$$\mathcal{A} = \left\{A \subset \mathbb{R}^3, A \text{ convex } ; \varphi_A \geq 0, \int_{\mathbf{S}^2} \varphi_A\, d\mu = 1\right\}.$$

(The sign condition on φ_A is only a normalization expressing the fact that $0 \in A$.)

Whenever μ has a discrete support, namely $\mu = \sum a_i \delta_{\nu_i}$, then (5) solves Minkowski's problem for polytopes. In particular, the value of φ outside the support of μ does not matter for the maximization, hence only the numbers $\varphi_i := \varphi(\nu_i)$ have to be considered.

Replacing an arbitrary measure μ on $\mathbf{S}^2$ by a sum of Dirac masses is also the more natural discretization of this problem. For polytopes, the set of admissibles bodies has the form

$$\mathcal{A} = \left\{P = P(N, \Phi); \varphi_i \geq 0, \sum_{i=1}^{n} \varphi_i a_i = 1\right\}.$$

(Again the conditions $\varphi_i \geq 0$ are only here to limit translations ensuring that $0 \in A$. This is essential in the numerical method.) These are very simple constraints on the admissible values, so step 3 in the algorithm is an elementary projection onto $\mathbb{R}^n_+$ and a hyperplane. Hence the given algorithm can be implemented in a straightforward way.

We present an example result on figure I.3. Here we chosed at random 999 vectors ν_i on $\mathbf{S}^2$, and 999 numbers a_i in $[0, 1]$ uniformly; ν_{1000} and a_{1000} are determined such that the existence condition $\sum_{i=1}^{1000} a_i \nu_i = 0$ is satisfied.

I.2.4 Application: Cheeger sets

Let us now present a more involved application. In 1970, Jeff Cheeger [11] proposed to study the problem

$$\inf_{X \subset M} \frac{\mathcal{H}^{n-1}(\partial X)}{\mathcal{H}^n(X)} \tag{6}$$

where M is an n-dimensional manifold with boundary. The resulting optimal value, known as the *Cheeger constant*, can be used to give bounds for the first eigenvalue of the Laplace-Beltrami operator on M, and even more general operators [14]. There is a number of variations and applications of this problem, see for example [2, 16].

The theorical results on the problem (6) are rather sparse. It is easy to show that the infimum is usually not attained in this general formulation. On the other hand it can be proved that minimizers exist whenever $M = \overline{\Omega}$, where $\Omega \subset \mathbb{R}^n$ is a nonempty open set. Moreover, if Ω is convex and $n = 2$, there is a unique convex optimum X which can be computed by algebraic algorithms [18]. On the other hand, if $n \geq 3$, it is not known whether the optimum set is unique or convex, even with Ω convex. However Ω convex implies that there exists at least one convex optimum [17]. But this optimum is not known for any particular Ω except balls.

Our algorithm allows us to compute an approximation of a convex optimum when $\Omega \subset \mathbb{R}^3$ is convex. Indeed (6) can be reformulated as follows:

$$\min_{A \in \mathcal{A}} \frac{3 \int_{\partial A} d\mathcal{H}^2(x)}{\int_{\partial A} \varphi_A(x)\, d\mathcal{H}^2(x)}, \text{ with } \mathcal{A} = \{A \subset \overline{\Omega}, A \text{ convex and 3-dimensional}\}.$$

So the numerator and denominator here have the form $\int_{\partial A} f(\nu_A, \varphi_A)$, and the algorithm can be used with straightforward modifications.

A key difference with respect to our previous application is the management of the constraint $A \subset \overline{\Omega}$. The set Ω itself is approximated by a polytope (whenever necessary). The corresponding enclosing half-spaces are kept in the algorithm in order to ensure that the approximating polytopes belong to $\mathcal{A}$. For example, if Ω is a unit cube, we fix $\nu_1 = (1, 0, 0), \ldots, \nu_6 = (0, 0, -1)$ and $\varphi_1 = \cdots = \varphi_6 = 1$.

This approach allows to handle any problem with constraints of the form

$$Q_0 \subset A \subset Q_1, \tag{7}$$

assuming that Q_1 is convex. (For Q_0 it is not a restriction to assume it is convex.) Other examples of problems of this kind come from mathematical economy, see references in [9], and also [4].

I.3 Newton's problem of the body of minimal resistance

The problem of the body of minimal resistance has been settled by I. Newton in its *Principia*: given a body progressing at constant speed in a fluid, what shape should it be given in order to minimize its resistance? Expressed in its more classical way, this can be formulated as the following optimization problem:

$$\min_{\substack{u:\Omega \to [0,M] \\ u \text{ convex}}} \int_\Omega \frac{dx}{1 + |\nabla u|^2}, \tag{8}$$

where $M > 0$ is a given parameter and Ω is the unit disk of $\mathbb{R}^2$. There is a lot of variants from this formulation and a huge litterature on this problem, see [5, 19] and their references.

I. Newton considered only radial solutions of this problem, and his solution was already considered surprising. But it has been proved in [3] that the solutions of (8) are not radially symmetric. Unfortunately it has been impossible until now to describe more precisely the minimizers. Some theorical results suggests that they should be developable in a sense given in [19].

So in this application, we are considering a problem of the form (1), with $g(x, u, p) = 1/(1 + |p|^2)$. As explained in Section I.1.3, this can be reformulated as (2) with $f(x, \nu, \varphi) = (\nu_3)_+^3$, where $t_+ := \max(t, 0)$ for any $t \in \mathbb{R}$. The set $\mathcal{A}$ is the set of convex bodies with a constraint of the kind (7), with $Q_0 := \Omega \times \{0\}$ and $Q_1 := \Omega \times [0, M]$.

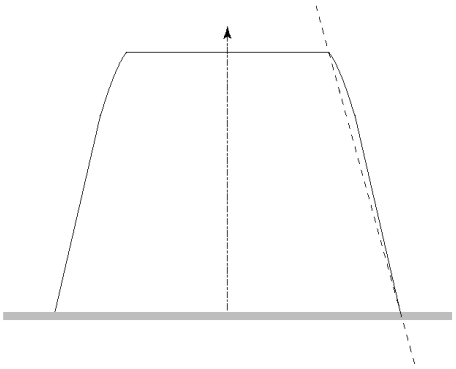

Figure I.5: Profile of computed optimal shape ($M = 3/2$): the solution is not developable.

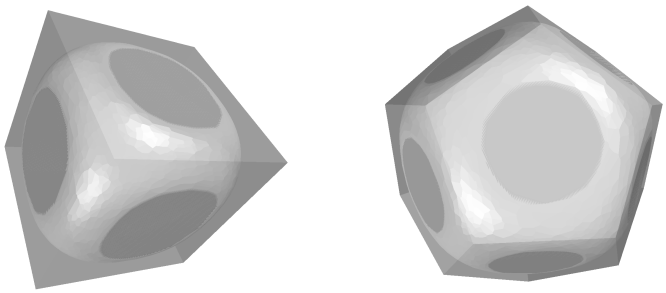

Figure I.6: Computed solutions for the Cheeger problem in the cube and the dodecahedron.

In the classical application, Ω is a disk. So we discretize these constraints by replacing the disk by a regular polygon Ω_ℓ, with ℓ sides. (In practice we used $\ell = 300$.) In this particular problem, this yields an overestimated value of the functional. Indeed if $A \subset \Omega_\ell \times [0, M]$ is convex, then $\tilde{A} := A \cap Q_1$ belongs to $\mathcal{A}$, and $\mathcal{F}(\tilde{A}) \leq \mathcal{F}(A)$ since $f \geq 0$ and vanishes on $\partial\tilde{A} \setminus \partial A$, where the normal vectors belong to $\{e_3\}^\perp$. Obviously for a minimization problem, this is not a predicament to overestimate the functional.

Using our gradient method on this problem yields different results starting with different initial shapes. This is likely the consequence of the existence of local minima. (Note that no theoretical result is known on the number or on the kind of critical points in this problem.) So our method needs to be preprocessed to start closer from a global minimum.

We use a genetic algorithm for this task. It is inspired from the ideas developed by J. Holland [15].

Our tests exhibit a behavior corresponding to the theoretical results given in [19]. Even for local minimizers, the image set of ν_A is sparse in $\mathbf{S}^2$. This suggests that optimal sets could be described with a lot fewer parameters as convex hulls of points instead of as an intersection of half-spaces. Therefore, we use the information given in the stochastic step (from the genetic algorithm) in two ways: as an initial set for the gradient method, and as an initial guess of the appropriate set of normal vectors to use. But the stochastic step itself represents the convex bodies as convex hull of points in $\Omega_\ell \times [0, M]$, together with the vertices $\Omega_\ell \times \{0\}$. The genetic algorithm optimizes the position of these points.

With these improvements, we get similar shapes for any run of the algorithm. Some of them are

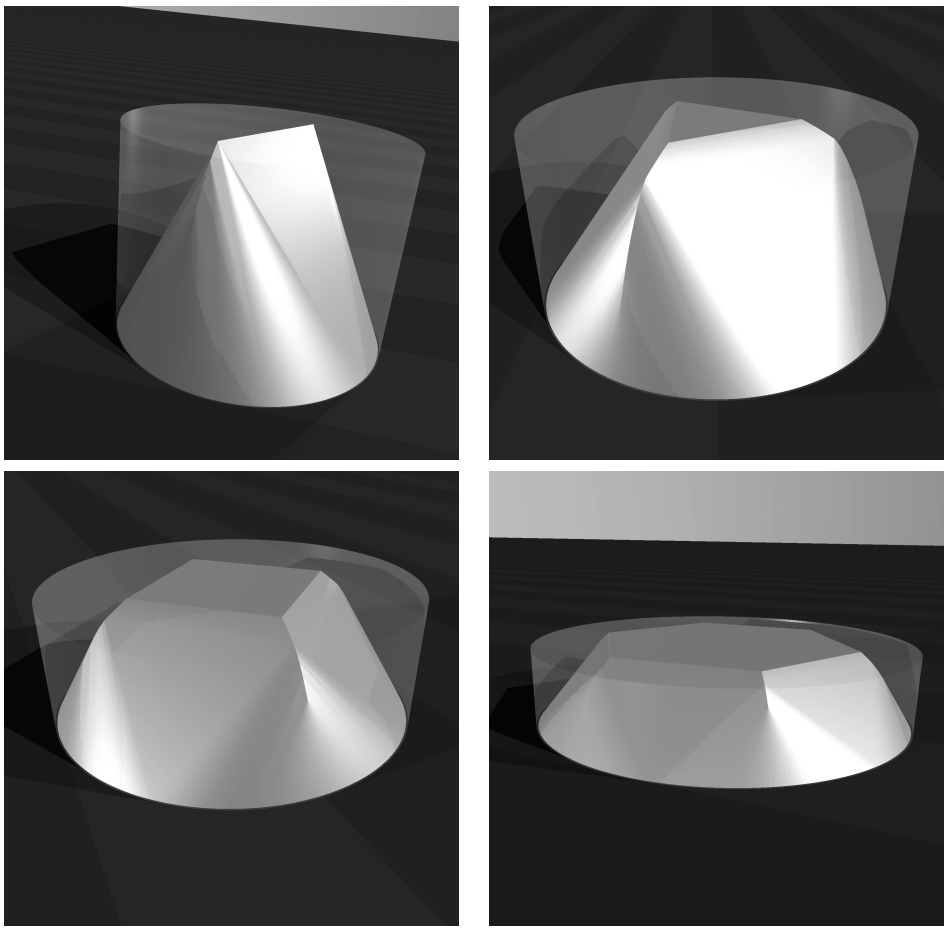

Figure I.7: Computed solutions of Newton's problem of the body of minimal resistance.

M	Newton's radial value	best theoretical values	numerical values
3/2	0.7526	0.7019	0.7012
1	1.1775	1.1561	1.1379
7/10	1.5685	1.5566	1.5457
4/10	2.1074	2.1034	2.1006

Table I.1: Minimal values of the Newton's resistance.

pictured in Figure I.7, for different values of the parameter M. These solutions are not developable in the sense of [19]. This can be seen more precisely on Figure I.5, where only the profile of the body is pictured.

Note that the corresponding values obtained by our method are smaller than the best theoretical values given in [19], even though they are slightly overestimated as explained before: see Table I.3.

It is a common conjecture on this problem that the solution is smooth except on the top and bottom parts, that is on $u^{(-1)}(0, M)$. However C^2-regularity would imply the developability property given in [19, Conjecture 2]. Our results demonstrate the non optimality of the best previously known profiles, and consequently the non regularity of the minimizers.

Bibliography

[1] A. D. Alexandrov, *Theory of mixed vloumes for convex bodies*, Mathem. Sb. USSR **2** (1937), pp. 947–972.

[2] G.Bellettini, V. Caselles & M. Novaga, *The total variation flow in* $\mathbb{R}^N$, J. Differential Equations **184** (2002), pp. 475–525.

[3] F. Brock, V. Ferone, B. Kawohl, *A Symmetry Problem inthe Calculus of Variations*, Calc. Var. Partial Differential Equations,**4** (1996), pp. 593–599.

[4] G. Buttazzo, P. Guasoni, *Shape optimization problems over classes of convex domains*, J. Convex Anal. **4** (1997), no. 2, pp. 343–351.

[5] G. Buttazzo, V. Ferone & B. Kawohl, *Minimum Problems over Sets of Concave Functions and Related Questions*, Math. Nachrichten, **173** (1993), pp. 71–89.

[6] G. Carlier, *On a theorem of Alexandrov*, to appear.

[7] G. Carlier & T. Lachand-Robert, *Convex bodies of optimal shape.* J. Convex Anal. **10** (2003) pp. 265-273.

[8] G. Carlier & T. Lachand-Robert, *Regularity of solutions for some variational problems subject to convexity constraint*, Comm. Pure Appl. Math. **54** (2001), pp. 583–594.

[9] G. Carlier, T. Lachand-Robert & B. Maury, *H^1-projection into sets of convex functions: a saddle point formulation*, Proceedings ESAIM (2001).

[10] G. Carlier, T. Lachand-Robert & B. Maury, *A numerical approach to variational problems subject to convexity constraint*, Numerische Math. **88** (2001), pp. 299–318.

[11] Cheeger, J., *A lower bound for the smalles eigenvalue of the Laplacian*, in: *Problems in Analysis, A Symposium in Honor of Salomon Bochner, Ed.: R.C.Gunning*, Princeton Univ. Press (1970) pp. 195–199.

[12] P. Choné & J.-C. Rochet, *Ironing, Sweeping and Multidimensional screening*, Econometrica, vol. 66 (1998), pp. 783–826.

[13] P. Choné & H. Le Meur, *Non-convergence result for conformal approximation of variational problems subject to a convexity constraint*, Numer. Funct. Anal. Optim. 22 (2001), no. 5-6, pp. 529–547.

[14] V. Fridman & B. Kawohl, *Isoperimetric estimates for the first eigenvalue of the p-Laplace operator and the Cheeger constant*, Comment. Math. Univ. Carol. **44** (2003) pp. 659–667.

[15] J. Holland, *Adaptation in natural and artificial systems*, Univ. Michigan Press (1975).

[16] I. Ionescu & T. Lachand-Robert, *Generalized Cheeger sets related to landslides*, submitted.

[17] B. Kawohl, *On a family of torsional creep problems*, J. reine angew. Math. **410** (1990) pp. 1–22.

[18] B. Kawohl & T. Lachand-Robert, *Characterization of Cheeger sets for convex subsets of the plane*, to appear.

[19] T. Lachand-Robert & M. A. Peletier, *Newton's problem of the body of minimal resistance in the class of convex developable functions*, Math. Nachrichten **226** (2001), pp. 153–176.

[20] P.-L. Lions, *Identification du cône dual des fonctions convexes et applications*, C. R. Acad. Sci. Paris (1998).

[21] H. Minkowski, *Allgemeine Lehrsätze über die Konvexen Polyeder*, Nach. Ges. Wiss., Göttingen (1897), pp. 198–219.

[22] R. T. Rockafellar, *Convex Analysis*, PrincetonUniversity Press (1970).

Bodies of constant width in arbitrary dimension

Thomas Lachand-Robert & Édouard Oudet

II.1 Introduction

A body (that is, a compact connected subset K of $\mathbb{R}^n$) is said to be of *constant width* α if its projection on any straight line is a segment of length $\alpha \in \mathbb{R}_+$, the same value for all lines. This can also be expressed by saying that the *width map*

$$w_K : \nu \in \mathbb{S}^{n-1} \longmapsto \max_{x\in K} \nu \cdot x - \min_{x\in K} \nu \cdot x \tag{1}$$

has constant value α. This is also equivalent to the geometrical fact that two parallel support hyperplanes on K are always separated by a distance α, independent of their direction. For an extended survey of properties and references about these bodies, see [3].

Note that the width of a body K and of its convex hull are the same. So, as many authors do, we will focus here on convex bodies of constant width.

Obvious bodies of constant width are the balls; but they are many others. These bodies, also called *orbiforms* in dimension two, or *spheroforms* in dimension three (as in [2]), have many interesting properties and applications. Orbiforms in particular have been studied a lot during the nineteenth century and later, particularly by Frank Reuleaux, whose name is now attached to those orbiforms you get by intersecting a finite number of disks of equal radii α, whose center are vertices of a regular polygon of diameter α. In particular the *Reuleaux triangle* is the intersection of three discs of radius α, centered on vertices of an equilateral triangle with side length α.

The mere existence of non trivial three-dimensional bodies of constant width is not so easy to establish. In particular, no finite intersection of balls has constant width (except balls themselves, see Corollary 3 in this article), a striking difference with the two-dimensional case.

To obtain spheroforms, a simple construction is to consider a two dimensional body of constant width having an axis of symmetry (like the Reuleaux triangle for instance): the corresponding body of revolution obtained by rotation around this axis is a spheroform. F. Meissner was able to construct another spheroform (usually called "Meissner's tetrahedron") which does not have the symmetry of revolution [5]. We give in Section II.4 of this paper an alternate construction of this body. Let us just say for the moment that it looks like an intersection of four balls centered on the vertices of a regular tetrahedron, but some of the edges are smoothed; in particular, it doesn't have all the symmetries of a regular tetrahedron.

Although many properties of bodies of constant width are known, very few characterizations exist in the literature, except basic ones. The aim of this paper is to provide a number of them,

that is, properties of bodies (in arbitrary dimension) that ensure they have constant width. This is important in particular for variational studies, like the still open problem of finding the orbiform of minimal volume. We will use these characterizations in a forthcoming article on this problem [1].

As a side effect, we give here a different application of these characterizations: namely we give a way (Theorem 7) to construct a body of constant width in dimension n, starting from a given projection in dimension $n-1$. This yields some sort of canonical construction of Meisner's body, starting from the Reuleaux triangle.

In Section II.2, we recall some properties of convex bodies and give a few notations. In Section II.3, we state the promised characterizations. In Section II.4, we give the raising dimension Theorem, together with a number of examples and numerical simulations. For instance we show a four-dimensional body of constant width whose projection is Meissner's body.

II.2 Constant width bodies

We recall here the main properties of bodies of constant width for convenience. Most of these are easy to prove, so we give only insight of the proofs. Details can be found in [2].

Let us first recall the definition of the support function for a body $K \in \mathbb{R}^n$: it is the map $h_K : \mathbb{S}^{n-1} \to \mathbb{R}$ defined by $h_K(v) := \max_{x\in K} x \cdot v$. It is related to the width function by the identity:

$$\forall v \in \mathbb{S}^{n-1}, \qquad w_K(v) = h_K(v) + h_K(-v). \tag{2}$$

Given two bodies K and L, their *Minkowski sum* is $K + L := \{x + y \,;\, x \in K, y \in L\}$. More generally, we define for any $\lambda, \mu \in \mathbb{R}$, the *Minkowski combination*

$$\lambda K + \mu L := \{\lambda x + \mu y; x \in K, y \in L\}.$$

It follows easily from the definitions that, for $\lambda \geq 0$ and $\mu \geq 0$, we have $h_{\lambda K+\mu L} = \lambda h_K + \mu h_L$, and then $w_{\lambda K+\mu L} = \lambda w_K + \mu w_L$ as well. We also have $h_{-K}(v) = h_K(-v)$ for all v, so $w_{-K} = w_K$. We deduce generally that

$$\forall \lambda, \mu \in \mathbb{R}, \qquad w_{\lambda K+\mu L} = |\lambda|\, w_K + |\mu|\, w_L. \tag{3}$$

As a consequence, the fact that K has constant width can easily be expressed with a Minkowski difference, namely

$$K - K = \alpha B_1 \tag{4}$$

where B_1 is the unit ball of $\mathbb{R}^n$. We see also that if K has constant width α, then $K + \beta B_1$ has constant width $\alpha + \beta$ for any $\beta \geq 0$.

A simple consequence of this property is that no body of constant width has a center of symmetry, in the sense that after a suitable translation $K = -K$, unless it is a ball. Indeed if $-K$ is (a translate of) K, (4) proves that K is a ball of radius $\alpha/2$. This explains why it is not possible to find an orbiform based on a square, or any even-sided polygon, as the Reuleaux polygons are based on odd-sided polygons. Similarly in dimension 3, there are no spheroforms having the same group of symmetries than the cube, octahedron, dodecahedron or icosahedron, except the ball. On the other hand there are some bodies of constant width whose group of symmetries is the same than the tetrahedron's. In order to construct one, it suffices to start from one of the Meissner's bodies K_1, that have all the required symmetries except one. So let K_2 be the symmetrical of K_1 with respect to the missing plane of symmetry. Now $K := \frac{1}{2}(K_1 + K_2)$ has constant width α, and has all the symmetries of the tetrahedron.

For a convex body K, we say that a hyperplane H is a *hyperplane of support for K at x*, if $x \in K \cap H$ and K is included in one of the half-spaces limited by H. If $\nu \in \mathbb{S}^{n-1}$ is a normal vector to H, pointing outside the half space containing K, we say that ν is *an outward support vector at x*. Obviously if K is smooth (that is, has a differentiable boundary), then ν is just the outward unit normal at x. In this particular case, there is a map $x \mapsto \nu$ which is usually called *the Gauss map*.

Note that a body of constant width is not always smooth, as the Reuleaux triangle shows. It turns out that for our purpose, we are more interested in the *reverse Gauss map*: for a **strictly convex** body K, and for any given $\nu \in \mathbb{S}^{n-1}$, the linear map $x \in K \mapsto x \cdot \nu$ attains its maximum at a unique point $x := R_K(\nu)$ (and the corresponding value is $h_K(\nu)$). The map

$$R_K \left| \begin{array}{l} \mathbb{S}^{n-1} \longrightarrow \partial K \\ \quad \nu \longmapsto x \text{ such that } x \cdot \nu = \max\limits_{y \in K} y \cdot \nu \end{array} \right.$$

so defined is surjective; it is a bijection if and only K is smooth.

Property 1 *Let $K \subset \mathbb{R}^n$ be a convex body of constant width α. Then the following properties hold:*

1. *the diameter of K is α;*
2. *K is strictly convex;*
3. *$K = \bigcap_{x \in \partial K} \overline{\mathrm{B}}(x, \alpha)$, where $\overline{\mathrm{B}}(x, \alpha)$ is the closed ball of center x and radius α;*
4. *if K has a C^2 boundary, then the radii of curvature at any point $x \in \partial K$ are all smaller than α.*

Property 3 is also called the "spherical intersection property" after Eggleston, see [3]. The proof of the proposition follows easily from a more technical lemma expressing how "farthest" points on ∂K are related to the reverse Gauss map:

Lemma 2 *Let K be a convex body of constant width α; for any $x \in \partial K$, consider F_x the set of points in K which are as far as possible from x:*

$$F_x := \{x' \in K; |x - x'| = \max_{y \in K} |x - y|\}.$$

Then for any $x' \in F_x$, we have $|x - x'| = \alpha$ and $x = R_K(\frac{1}{\alpha}(x - x'))$.

This expresses the fact that if x and x' are as far as possible, then they are at distance α and the unit vectors parallel to $x - x'$ are support vector at x or x'.

Proof of the lemma and of the proposition. Let $x \in \partial K$ be given. The set F_x is nonempty since K is compact. Let $x' \in F_x$, and $\delta := |x - x'| = \max_{y \in K} |x - y|$.

We have $\delta \leq \alpha$ since otherwise the projection of K on the line joining x and x' would have a length at least $\delta > \alpha$. So the diameter of K is smaller than α.

Let us show that K is strictly convex. Assume by contradiction that there exists $y \neq x$ such that the segment $[x, y]$ is contained on ∂K. Let us denote by $x_t := tx + (1 - t)y$, with $t \in (0, 1)$, the intermediate points in the segment. If H is any support hyperplane at some x_t, it contains the whole segment. Given an outward unit normal vector ν to H, we have $x_t \cdot \nu = h_K(\nu)$ for all $t \in (0, 1)$.

Using again the compactness of K, there exists some $z \in \partial K$ such that $z \cdot (-\nu) = h_K(-\nu)$. Now K has constant width, so

$$\alpha = w_K(\nu) = x_t \cdot \nu - z \cdot \nu \leq |x_t - z| \,.$$

The latter inequality must be an equality, since K has diameter smaller than α. This implies $x_t - z = \alpha\nu$, which is not possible for all $t \in (0,1)$ since $x \neq y$.

Hence we have proved that K is strictly convex. Repeating the argument with x instead of x_t, and H any support hyperplane at x, we deduce again $x - z = \alpha\nu$. This implies $\delta = \alpha$, and also $z \in F_x$. Thus $K \subset \overline{\mathrm{B}}(x,\alpha)$; since this property holds for any $x \in \partial K$, we deduce that $K \subset \bigcap_{x\in\partial K} \overline{\mathrm{B}}(x,\alpha)$. Also if K is C^2 near z, this implies that the curvature radii at this point are smaller than α.

Now for any given $x' \in F_x$, let us define $\nu := \frac{1}{\alpha}(x - x')$. Since $K \subset \overline{\mathrm{B}}(x',\alpha)$, and since ν is the outward unit normal vector at x on this ball, the hyperplane containing x and orthogonal to ν is a support plane to K. This implies $x = R_K(\nu)$ from the definition of the reverse Gauss map.

We have finally to prove $K \supset \bigcap_{x\in\partial K} \overline{\mathrm{B}}(x,\alpha)$, since the reverse inclusion was proved herebefore. Assume by contradiction that there exists some $z \in \bigcap_{x\in\partial K} \overline{\mathrm{B}}(x,\alpha)$, such that $z \notin K$. Since K is convex, it follows from the Hahn-Banach theorem that there exists $\nu \in \mathbb{S}^{n-1}$ such that $z \cdot \nu < x \cdot \nu$ for all $x \in K$. Consider $x := R_K(\nu)$ and $x' := R_K(-\nu)$. From the previous study $x - x' = \alpha\nu$ so $x \cdot \nu = \alpha + x' \cdot \nu > \alpha + z \cdot \nu$. This contradicts $z \in \overline{\mathrm{B}}(x,\alpha)$. □

For strictly convex bodies, we also have the following classical property:

Lemma 3 *Let K be a strictly convex body. Its support function h_K is a C^1 function on $\mathbb{S}^{n-1}$. Its reverse Gauss map R_K is continuous.*

Proof. Let $(\nu_i) \subset \mathbb{S}^{n-1}$ be any converging sequence, with limit ν. Define $x_i := R_K(\nu_i)$. Since K is compact, we may extract a subsequence (with no change of notation) in order to ensure that (x_i) converges. Let x be its limit. For all $y \in K$, and all i, we have $y \cdot \nu_i \leq x_i \cdot \nu_i$ from the definition of R_K. Passing to the limit yields $y \cdot \nu \leq x \cdot \nu$ for all $y \in K$, so $x = R_K(\nu)$ since this is the only maximizer of $y \mapsto y \cdot \nu$ from the strict convexity of K. This proves that R_K is continuous.

Note that this implies that h_K is continuous since $h_K(\nu) = \nu \cdot R_K(\nu)$. It is well-known that h_K can be extended to a convex 1-homogeneous function $\bar{h}_K : \mathbb{R}^n \to \mathbb{R}$, also called the support function of K, and defined by $\bar{h}_K(d) = |d|\, h_K\left(\frac{d}{|d|}\right)$. The subdifferential of $\bar{h}_K(d)$ at some $d \neq 0$ is the face of K associated with d, that is $\{x \in K \,;\, x \cdot d = \bar{h}_K(d)\}$ [4, Section D.3.1]. If $d = \nu \in \mathbb{S}^{n-1}$ and K is strictly convex, this reduces to $\{R_K(\nu)\}$. Hence $\bar{h}_K$ is C^1 on $\mathbb{S}^{n-1}$, and so is h_K. □

II.3 Characterizations of bodies of constant width

As explained before, finding bodies of constant width is not so easy. One simple way to construct a spheroform for instance, is to start with an orbiform having an axis of symetry, and then to consider the body of revolution generated by its rotation around the axis. However this process usually yields bodies with large volume.

We describe in the next section a new process that allows us to construct a body of constant width in dimension $n \geq 2$ from any body of constant width in dimension $n-1$. In order to do that, we need some characterizations of bodies of constant width.

Property 2 *A strictly convex body K has constant width α if and only if its reverse Gauss map satisfies:*

$$\forall \nu \in \mathbb{S}^{n-1}, \qquad R_K(-\nu) = R_K(\nu) - \alpha\nu. \tag{5}$$

Proof. If K has constant width, we have from the definition of R_K and taking into account the fact that the diameter of K is smaller than α:

$$\alpha = h_K(\nu) + h_K(-\nu) = (R_K(\nu) - R_K(-\nu)) \cdot \nu \leq |R_K(\nu) - R_K(-\nu)| \leq \alpha.$$

Therefore we have equality in the above inequality, and this implies (5).

Let us assume that K is strictly convex and satisfies (5). Then

$$w_K(\nu) = h_K(\nu) + h_K(-\nu) = (R_K(\nu) - R_K(-\nu)) \cdot \nu = \alpha.$$

So K has constant width α. □

Let us draw a number of consequences of Proposition 2. Here and in the rest of the paper, a *singular point* x on the boundary of some convex K is a point where more than one unit outward support vector exists.

Corollary 2 *Let x be a singular point on the boundary of some convex body K of constant width α. Then there exists a nontrivial arc of circle of radius α with center x on ∂K.*

A circle denotes as usual the intersection of some ball with a plane (dimension 2), and the center is in this plane. By "nontrivial" we mean that the arc must have more than one point.

Proof. The corollary follows from the proposition by noticing that for a convex set K, a point $x \in \partial K$ is singular if and only if $R_K^{(-1)}(x)$ contains more than one vector. By convexity it contains a spherical arc $\widehat{\nu_0\nu_1}$. From the proposition, ∂K contains $x - \alpha\nu$ for all $\nu \in \widehat{\nu_0\nu_1}$, which is an arc of circle of radius α. □

Corollary 3 *In dimension $n \geq 3$, no finite intersection of balls have constant width, unless it reduces to a single ball.*

Proof. Indeed consider $K = \bigcap_{i=1}^m \overline{\mathrm{B}}(x_i, r_i)$. We assume that this intersection is reduced, that is, no smaller intersection among the same balls yields the same set K. In particular, for each i, there is a relatively open part Q_i of the boundary of $\overline{\mathrm{B}}(x_i, r_i)$ which is contained in ∂K. Any point $x \in Q_i$ is a differentiability point for ∂K, and $x = R_K(\nu)$ where $\nu = (x - x_i)/r_i$ is the common outward unit normal at x to $\overline{\mathrm{B}}(x_i, r_i)$ and to K. If we denote by $\Sigma_i \subset \mathbb{S}^{n-1}$ the corresponding subset of unit vectors, we have $Q_i = R_K(\Sigma_i)$. According to Proposition 2, $R_K(\nu) - \alpha\nu \in \partial K$ for all ν, in particular $\nu \in \Sigma_i$. This implies $r_i \leq \alpha$, for otherwise the corresponding image $\nu \mapsto R_K(\nu) - \alpha\nu$ yields a concave surface on ∂K, which is impossible.

So $r_i \leq \alpha$, and in particular, any nontrivial arc of circle of radius α on ∂K has its center at some x_i. So the family of allowed centers for arcs of circle of radius α on ∂K is finite. From the previous corollary, we deduce that ∂K has only a finite number of singular points. This is clearly not possible in dimension $n \geq 3$, unless ∂K is just a sphere. □

Theorem 4 *Let K be a convex body. Then K has constant width α if and only if it satisfies both conditions:*

$$\operatorname{diam} K \leq \alpha \tag{6}$$

$$\forall x \in \partial K,\ \exists x' \in K,\ |x - x'| = \alpha. \tag{7}$$

Proof. We already know from Lemma 2 that a body of constant width satisfies these properties, so we just have to prove the reciprocal. Moreover the property is obvious in dimension $n = 1$, so we assume that $n \geq 2$ in the following.

So assume that K satisfies (6) and (7). Let us first prove that K is strictly convex. Indeed if ∂K contains a segment $[x_0, x_1]$ with nonempty (relative) interior, choose any x in this interior, say $x = (x_0 + x_1)/2$. From condition (7), there exists $x' \in K$ such that $|x - x'| = \alpha$. This implies that $|x_i - x'| > \alpha$ for $i = 0$ or $i = 1$, since a ball is strictly convex. So we get a contradiction with condition (6).

Now that we know that K is strictly convex, we can use its reverse Gauss map and Proposition 2. Observe first that the strict convexity implies in particular that K is n-dimensional (not contained in a strict affine subspace).

Let $\nu \in \mathbb{S}^{n-1}$ be given, and $x := R_K(\nu)$. Let us first assume that K is smooth at x, so that ν is the outward unit normal vector at x. From (7), there exists $x' \in K$ such that $|x - x'| = \alpha$. (Actually $x' \in \partial K$ since $\operatorname{diam} K \leq \alpha$.)

Since the ball $\overline{\mathrm{B}}(x', \alpha)$ contains K from (6), the tangent hyperplanes to the ball and to K at x coincide. Therefore ν is equal to the unit outward normal vector to the ball $\overline{\mathrm{B}}(x', \alpha)$ at x, which is $(x - x')/\alpha$. Hence $x' = x - \alpha\nu$, and in particular, for all $y \in K$, we have with (6):

$$y \cdot (-\nu) = (x - y) \cdot \nu - x \cdot \nu \leq \alpha - x \cdot \nu = x' \cdot (-\nu).$$

Taking the supremum on all y yields $x' = R_K(-\nu)$. So we have proved (5) for any ν such that ∂K is smooth at $R_K(\nu)$. Let us recall that the subset of points where the boundary is smooth is dense in ∂K.

So consider an arbitrary ν, $x := R_K(\nu)$ and $x' := R_K(-\nu)$. Notice that $x' \neq x$ since K is n-dimensional. The sets $U := R_K^{(-1)}(x)$ and $V := -R_K^{(-1)}(x')$ are closed in the sphere $\mathbb{S}^{n-1}$ and have a common element, namely ν. However none of them contains $-\nu$ (since $x' \neq x$), so they are not equal to the whole sphere. Since it is not possible to have each one included in the interior of the other, one of $\partial U \cap V$ or $U \cap \partial V$ must be nonempty (the boundary here is considered with respect to the natural topology of the sphere $\mathbb{S}^{n-1}$). Say for instance $\partial U \cap V \neq \emptyset$, so that there exists $\nu' \in \partial U \cap V$. In particular we can find a sequence $(\nu_k)_{k\geq 1}$ converging to ν' as $k \to \infty$ such that $x_k := R_K(\nu_k) \neq x$ for all k. (Note that R_K is continuous according to Lemma 3, so x_k converges to x.) We may even assume, with no loss of generality, that ∂K is smooth at x_k for all $k \geq 1$, since the set of such points is dense.

In particular $x_k - \alpha\nu_k = R_K(-\nu_k)$ from our previous study. Letting k going to infinity yields $x - \alpha\nu' = R_K(-\nu')$. But the latter is x' since we assumed $\nu' \in V = -R_K^{(-1)}(x')$. So we have proved $\nu' = \nu_0$ where $\nu_0 := (x - x')/\alpha$. This shows that $\partial U \cap V = \{\nu_0\}$ whenever this set is nonempty. In general, we have $\partial U \cap V \subset \{\nu_0\}$. A symetrical argument shows that $U \cap \partial V \subset \{\nu_0\}$ also. This means that one of the two sets U or V has empty interior. In particular, since $\nu \in U \cap V$, we get $\nu = \nu_0$. This proves (5). □

Later on we will need a slightly different version of this theorem:

Theorem 5 *Let K be a closed subset of $\mathbb{R}^n$. Then K is a convex body of constant width α if and only if it satisfies* (6) *and*

$$\forall x \in \partial K,\ \exists x',\ [x, x'] \subset K \quad \textit{and} \quad |x - x'| = \alpha. \tag{8}$$

The difference is that we do not assume K convex here, but require instead that the whole segment $[x, x']$ is contained in K.

Proof. It is clear that (8) implies (7). So we just have to prove that K is convex, and use Theorem 4 to conclude the proof.

So assume by contradiction that K is not convex. Hence there exists $x_0, x_1 \in K$ such that $[x_0, x_1] \not\subset K$. Let x be some point in $[x_0, x_1] \cap \partial K$. We can find $x' \in K$ with $|x - x'| = \alpha$ using (8). Since $x \in [x_0, x_1]$, we have $|x - x_0| + |x - x_1| = |x_1 - x_0|$. From Ptolemy's inequality we get, taking into account that $\operatorname{diam} K \leq \alpha$:

$$\begin{aligned} \alpha\,|x_1 - x_0| &\leq |x - x_0|\,|x_1 - x'| + |x - x_1|\,|x_0 - x'| \\ &\leq \alpha(|x - x_0| + |x - x_1|) = \alpha\,|x_1 - x_0|\,. \end{aligned}$$

So there must be equality everywhere, which means that the four points are cocyclic, and that $|x_i - x'| = \alpha$ for $i = 0, 1$. Note that this implies in particular that the four points are not aligned. Therefore the assumption $x \in [x_0, x_1] \cap \partial K$ implies $x = x_0$ or $x = x_1$.

So we proved that for any $(x, y) \in K^2$, the whole interior of the segment $[x, y]$ is outside K or is interior to K. In particular $x_0 \in \partial K$, so there exists x_0' such that $\left|x_0 - x_0'\right| = \alpha$ and $[x_0, x_0'] \subset K$. More precisely the interior of this segment is included in the interior of K. In particular, for any $x \in (x_0, x_0')$, we have $[x, x_1] \subset K$ since $x \notin \partial K$. Passing to the limit $x \to x_0$, this implies that $[x_0, x_1] \subset \partial K$, a contradiction. □

The preceding characterizations are useful, but the diameter condition is difficult to handle in the variational context we consider in [1]. So let us give a slightly different characterization of bodies of constant width.

Theorem 6 *Let K be closed subset of $\mathbb{R}^n$. Then K has constant width α if and only if it satisfies:*

$$\forall \nu \in \mathbb{S}^{n-1},\ \exists x_\nu \in K,$$
$$x_\nu + \alpha\nu \in K \quad \textit{and} \quad \forall y \in K,\ (y - x_\nu) \cdot \nu \in [0, \alpha]. \tag{9}$$

So (4) expresses that the projection of K on the line $\mathbb{R}\nu$ is included in an interval of length α (condition $\forall y \dots$) and that the corresponding extremal points x_ν and $x_\nu' := x_\nu + \alpha\nu$ do exist in K.

Proof. If K has constant width α, it satisfies (4) with $x_\nu = R_K(-\nu)$. Indeed we know from Proposition 2 that $x_\nu + \alpha\nu = R_K(\nu)$ in that case, and the remaining part follows from the very definition of R_K.

So let us prove the converse, starting with some closed set satisfying (4). Let us first prove that $\operatorname{diam} K \leq \alpha$. Let $x, y \in K$ be given, with $x \neq y$, and consider $\nu := (y - x)/|y - x|$. From (4), there exists $x_\nu \in K$ such that $(y - x_\nu) \cdot \nu$ and $(x - x_\nu) \cdot \nu$ both belong to $[0, \alpha]$. Since

$$(y - x_\nu) \cdot \nu = |y - x| + (x - x_\nu) \cdot \nu$$

according to the definition of ν, this implies in particular that $|y - x| \leq \alpha$. So $\operatorname{diam} K \leq \alpha$.

Let $\hat{K}$ be the closed convex hull of K. Notice that $\hat{K}$ also satisfies (4) since $K \subset \hat{K}$ and $y \in \hat{K}$ implies that y is a finite convex combination of elements of K. In particular, $\hat{K}$ has diameter $\leq \alpha$, too.

Consider some $x \in \partial\hat{K}$. Since $\hat{K}$ is convex, there exists some outward support vector $\nu \in \mathbb{S}^{n-1}$. So we have $x \cdot \nu \geq y \cdot \nu$ for all $y \in \hat{K}$. Let x_ν be given by (4), and $x'_\nu := x_\nu + \alpha\nu \in K$. Since $(x - x'_\nu) \cdot \nu \leq 0$ from (4) and $(x - x'_\nu) \cdot \nu \geq 0$ from the definition of ν, we deduce that $(x - x'_\nu) \cdot \nu = 0$. Therefore

$$|x - x_\nu|^2 = \left|x - x'_\nu\right|^2 + \left|x_\nu - x'_\nu\right|^2 = \left|x - x'_\nu\right|^2 + \alpha^2.$$

Since this is also less than α^2 from the diameter property, we have proved that $x = x'_\nu$. Therefore the point $x' := x_\nu$ satisfies $|x - x'| = \alpha$ and $x' \in \hat{K}$.

We have proved that $\hat{K}$ satisfies (6) and (8), so $\hat{K}$ has constant width α according to Theorem 5. In particular, $\hat{K}$ is strictly convex. Therefore any $x \in \partial\hat{K}$ is exposed. Since such an x can be expressed as a convex combination of $(n + 1)$ points of K according to Caratheodory's Theorem [8, Theorem 1.1.4], and since the exposure property implies that all these points coincide with x, we deduce that $x \in K$. Hence $K = \hat{K}$ and this concludes the proof. □

II.4 Raising dimensions

Now we have the required elements to exhibit the raising dimensions process:

Theorem 7 *Let $H \subset \mathbb{R}^n$ be an affine hyperplane, E_+ and E_- the two open half-spaces separated by H, and $K_0 \subset H$ be an $(n-1)$-dimensional convex body of constant width α. Let Q be any set satisfying*

$$K_0 \subset Q \subset \overline{E}_- \cap \bigcap_{x \in K_0} \overline{\mathrm{B}}(x, \alpha). \tag{10}$$

Consider the set K defined as follows:

$$K \cap \overline{E}_+ = K_+ := \overline{E}_+ \cap \bigcap_{x \in Q} \overline{\mathrm{B}}(x, \alpha) \tag{11}$$

$$K \cap E_- = K_- := E_- \cap \bigcap_{x \in K_+} \overline{\mathrm{B}}(x, \alpha). \tag{12}$$

Then K is a n-dimensional convex body of constant width α, and $K \cap H = K_0$.

The property $K_0 = K \cap H$ shows that this process raises dimension. Note that since $K_0 \subset H$ and has diameter α, it is possible to find sets Q satisfying (10).

The simplest choice for Q is just $Q = K_0$. With this choice, and starting from a one-dimensional convex body (a segment of width α), we get a two-dimensional Reuleaux triangle. Starting from a two-dimensional disk, we get a rotated Reuleaux triangle. Starting from a two-dimensional Reuleaux triangle, we get Meissner's body.

Let us explain that in more details. Figure II.1 shows an example that makes the construction of the Theorem 7 easier to understand. We start from a 1-dimensional convex body of constant width, that is a segment K_0 of length α (bold on the left and middle parts of the figure). In the figure, H is the horizontal line, E_+ and E_- are the upper and lower part of the plane. All the construction takes

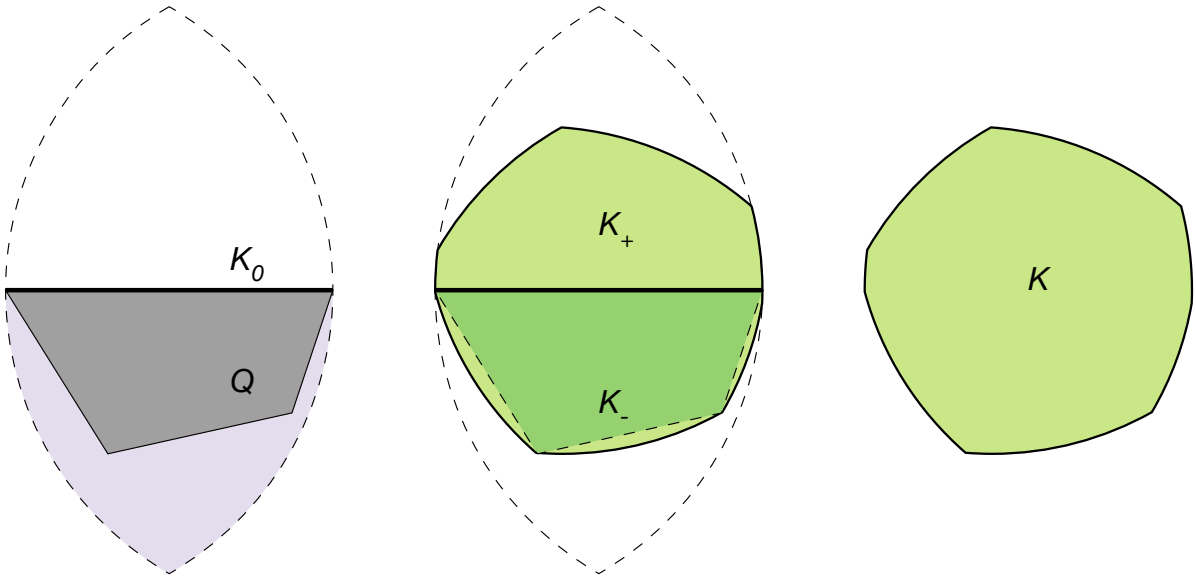

Figure II.1: Illustration of the raising dimension process, 1D to 2D.

place in the intersection of balls of radius α on K, which reduces here to the intersection of two balls (dashed lines). We choose a set Q satisfying (10), that is in the lower part of this intersection (shown in light gray). In our example, Q is a quadrilateral (gray on the left of the figure). The set K_+ is the intersection of balls centered on Q (only four balls really since Q is a polygon) and the upper part of the plane (middle part of the figure). Then K_- is the intersection of balls centered on K_+ (again a finite number of balls is enough here), and it contains Q (dashed lines). The resulting set K is their union, and is shown again on the right. It is now easier to understand that if we had chosen $Q = K_0$, then K_+ would have been the upper part in the intersection of the two disks, and then K_- would have been the intersection of the disk centered on the upper point of K_+ with the lower half plane. So we would get a Reuleaux triangle as claimed.

Figure II.2 shows an example starting from a two-dimensional body K_0 of constant width shown on the left. This body is the intersection of a rather large number of disks. We chose $Q = K_0$ in the construction. The resulting three-dimensional body is shown on the right (with a different scale). Notice in particular on the lower left part, the body is shown from above, so the projection is just K_0. (This figure and the following ones have been obtained using softwares [7] and [6].)

Should we have started from a Reuleaux triangle K_0, then K_+ would have been the intersection of the upper-space and the three balls centered on vertices of K_0. Then K_- would have been determined from K_+ as defined in the theorem, but would have not been a finite intersection of disks, according to Corollary 3. The resulting spheroform has an upper part that is identical to a spherical tetrahedron: that is exactly one of the variant of Meissner's body, shown in Figure II.3. (The original construction by Meissner of this body is quite different and can be found in [3]: the original body is the one obtained by flipping couples of edges).

There is no limit on the dimensions that can be reached by the process given by Theorem 7. Starting from Meissner's body K_0, and using $Q = K_0$, we get a four dimensional body shown in

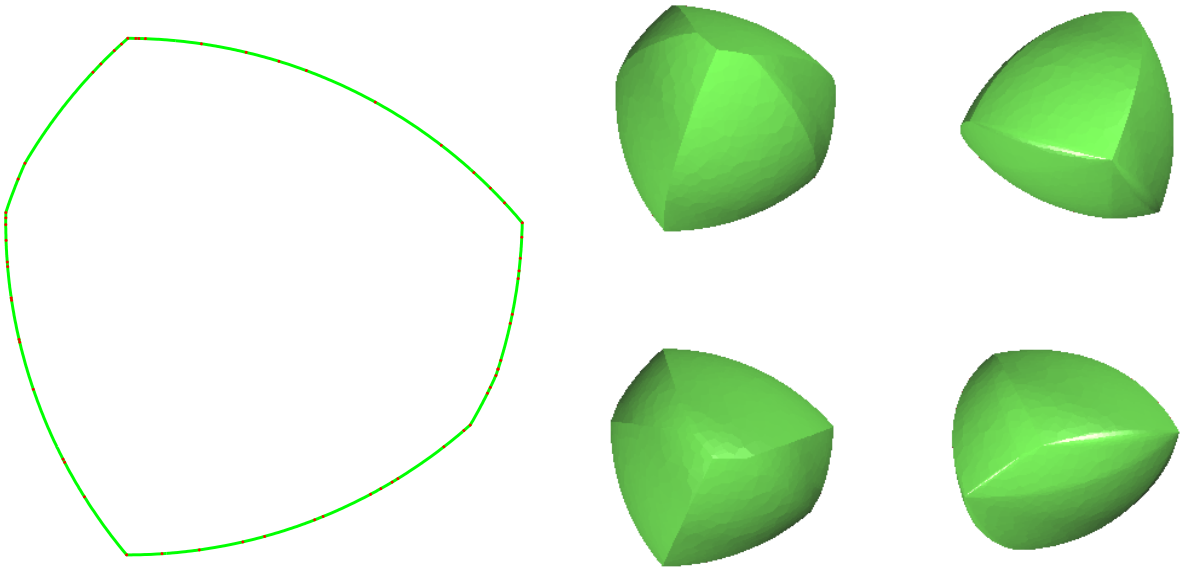

Figure II.2: Illustration of the raising dimension process, 2D to 3D.

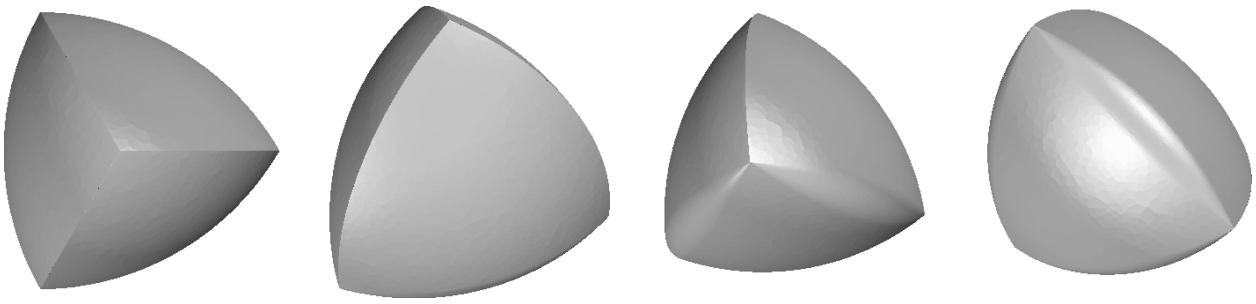

Figure II.3: One of the two forms of Meissner's body. We can get this using Theorem 7 with $K_0 = Q$ a Reuleaux triangle.

Figure II.4 (Displayed are parallel cross sections). A numerical computation of the (four dimensional) volume and surface area of this body shows the following values (with $\alpha = 1$):

$$\text{Volume} \simeq 0.223 \pm 10^{-3}, \qquad \text{Surface} \simeq 2.12 \pm 10^{-2}.$$

For the proof of Theorem 7, we need a small geometrical lemma:

Lemma 4 *Let a, b, x, y be four points in a plane. Assume that*

$$\max\Big(|a-x|, |a-y|, |b-x|, |b-y|\Big) \leq |b-a|$$

and that the segment $[x, y]$ *does not intersect the line generated by a and b. Then* $|x-y| \leq |b-a|$.

Proof. Let $\alpha := |b-a|$ and c be the point at distance α from a and b, on the same side of the line generated by a and b than x and y. So a, b, c forms an equilateral triangle. Let T be the Reuleaux triangle supported by this triangle. From the assumptions on x, y, we see that these two points belong to T. Since T has constant width α, its diameter is α also from Lemma 2. □

Proof of Theorem 7. Observe first that K_+ is closed and $K_+ \cap H = K_0$ from its definition and (10). Also $\overline{K}_- \cap H = K_0$, and $K_- \cap K_+ = \emptyset$. The set K_- is not closed, but $\overline{K}_- = K_- \cup K_0$. So in particular $K = K_+ \cup K_-$ is closed, and $K \cap H = K_0$ as claimed.

We will prove that K is convex and satisfies both conditions of Theorem 4. (Note that it is obvious that K_+ and K_- are convex, but it is not so clear that $K = K_+ \cup K_-$ is.)

Let x, y be any two different points in K, and define $\nu := (y-x)/|y-x| \in \mathbb{S}^{n-1}$. We will prove that $[x, y] \subset K$ and $|x-y| \leq \alpha$. Note that we don't need to prove that for any pair (x, y) in order to prove that K is convex and has diameter α. It is enough to consider a dense subset, since K is closed, so we may assume that the straight line δ joining x and y (that is $x + \mathbb{R}\nu$) is not parallel to H, and that $\delta \cap H \not\subset \partial K_0$.

From these assumptions, δ intersects H at a point $z \notin \partial K_0$. Let $p \in \partial K_0$ be defined as follows:

1. if $z \notin K_0$, let p be the orthogonal projection of z onto K_0 (in H);
2. otherwise z is in the relative interior of K_0; let p be the farthest point from z in K_0 (so that $|p-z| \geq |m-z|$ for all $m \in K$).

We define also $\nu_0 \in \mathbb{S}^{n-1}$ by $\nu_0 := (z-p)/|p-z|$ in the first case, and $\nu_0 := (p-z)/|p-z|$ otherwise. Note that this vector belongs to the vector space directing H. We claim that $p = R_{K_0}(\nu_0)$ in all cases. Indeed let us just check the two different cases in the same order:

1. if $p \neq p_0 := R_{K_0}(\nu_0)$, then $p \cdot \nu_0 < p_0 \cdot \nu_0$. Since $z - p \in \mathbb{R}_+^* \nu_0$ by definition, this implies $0 < (p_0 - p) \cdot (z - p)$ in contradiction to the fact that p is the projection of z onto the convex set K_0;
2. again if $p \neq p_0 := R_{K_0}(\nu_0)$, then $p \cdot \nu_0 < p_0 \cdot \nu_0$. Since $p - z \in \mathbb{R}_+^* \nu_0$ by definition we get

$$0 > 2(p_0 - p) \cdot (z - p) = |z-p|^2 - |z-p_0|^2 + |p-p_0|^2 .$$

Hence we have $|z - p_0| > |z - p|$ in contradiction to the fact that p is the farthest point from z in K_0.

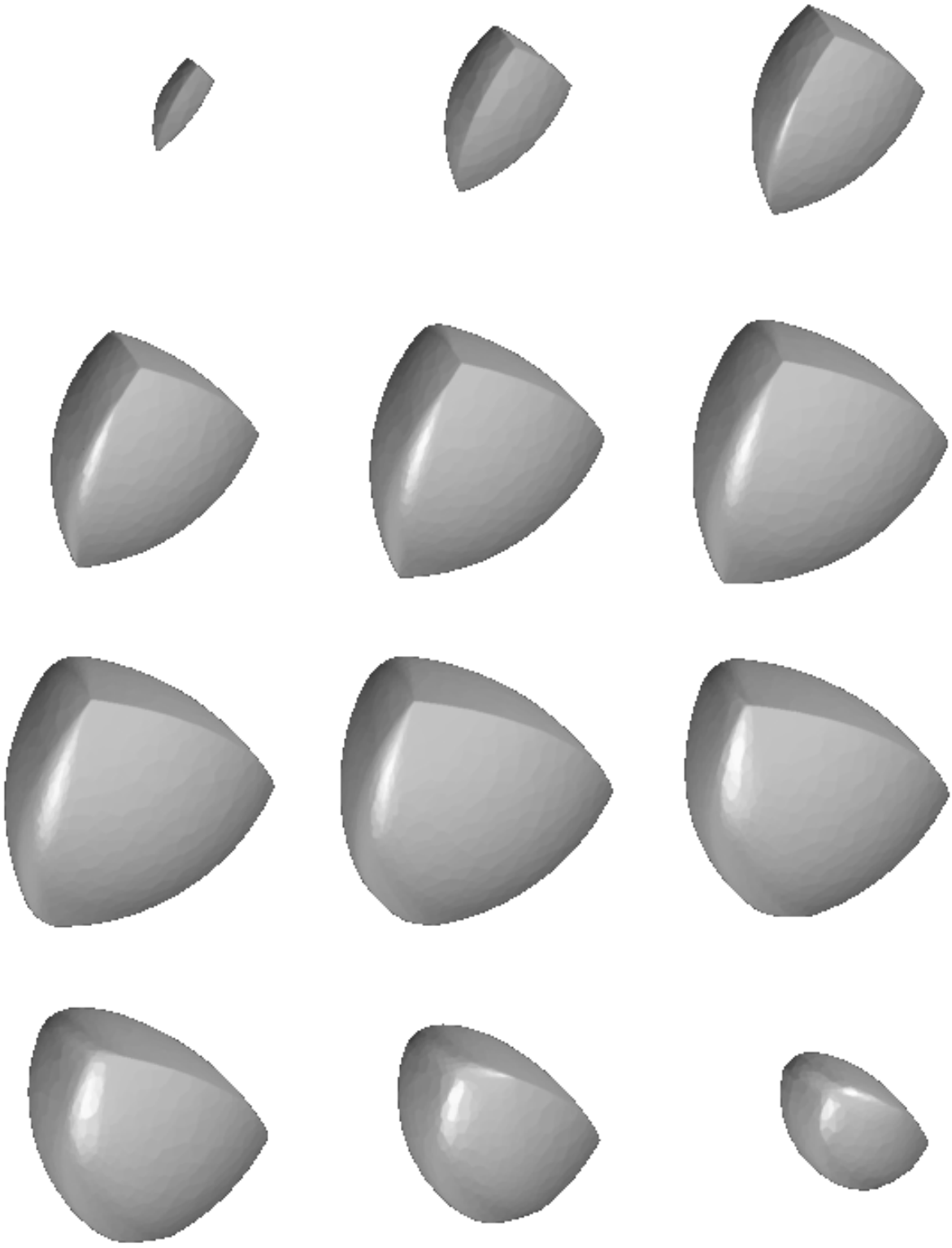

Figure II.4: A four dimensional body of constant width (parallel cross-sections), obtained using Theorem 7 with $K_0 = Q$ a Meissner's body.

Now let us consider $p' := p - \alpha\nu_0$. We know from Lemma 2 that $p' = R_{K_0}(-\nu_0) \in \partial K_0$ since K_0 has constant width.

Let P be the two dimensional plane $p + \mathbb{R}\nu + \mathbb{R}\nu_0$. It contains z, and also the four points x, y, p, p'. So in particular these four points are coplanar.

We now discuss the different cases:

- if $x \in K_+$ and $y \in K_+$, then $[x, y] \subset K_+$ since K_+ is convex, and we have from its definition that $K_+ \subset \overline{\mathrm{B}}(p, \alpha) \cap \overline{\mathrm{B}}(p', \alpha)$. In particular, we get:

$$|x - p| \leq \alpha, |x - p'| \leq \alpha, |y - p| \leq \alpha, |y - p'| \leq \alpha. \tag{13}$$

  Since $|p - p'| = \alpha$, we get $|x - y| \leq \alpha$ using Lemma 4 (with $a = p$, $b = p'$).

- if $x \in K_-$ and $y \in K_-$, we get the same results using similar arguments (note that $K_0 \subset K_+$, so (13) holds true again).

- if one of the point is in K_+, say x, and the other in K_-, we get $|x - y| \leq \alpha$ from the definition of K_-. Here x and y are on opposite sides of the hyperplane H, so in particular z lies on the segment $[x, y]$, that is $z = [x, y] \cap H$. Remember also that z is contained in the straight line generated by p and p'. Since K_+ and K_- are contained in $\overline{\mathrm{B}}(p, \alpha) \cap \overline{\mathrm{B}}(p', \alpha)$ by definition, z lies on the projection of $\overline{\mathrm{B}}(p, \alpha) \cap \overline{\mathrm{B}}(p', \alpha)$ onto the straight line generated by p and p'. But such a projection is just the segment $[p, p']$, so we have $z \in [p, p'] \subset K_0$. In particular $z \in K_+$, so $[x, z] \subset K_+$, and $z \in \overline{K}_-$, so $[z, y] \subset \overline{K}_-$. This implies $[x, y] \subset K$.

So we proved that K is convex and $\operatorname{diam} K \leq \alpha$. To conclude the proof we need to prove (7). So let us consider some $x \in \partial K$. Note that the property is obvious if $x \in K_0$, since K_0 has constant width and $K_0 = K \cap H$.

If $x \in \partial K_- \setminus K_0$, let us define $\delta := \sup_{y \in K_+} |x - y|$. Note that $\delta \leq \alpha$ from the definition of K_-, and that there exists $x' \in K_+$ such that $|x - x'| = \delta$ since K_+ is compact. So if $\delta = \alpha$, (7) is satisfied. If on the contrary $\delta < \alpha$, then there exists a (small) open neighborhood U of x such that $U \subset \overline{\mathrm{B}}(y, \alpha)$ for all $y \in K_+$. Since $x \in E_-$, an open set, we may assume that $U \subset E_-$, reducing the neighborhood if necessary. So $U \subset K_-$, in contradiction to the assumption $x \in \partial K_-$.

If $x \in \partial K_+ \setminus K_0$, let us define similarly $\delta := \sup_{y \in \overline{Q}} |x - y|$. Again there is some $x' \in \overline{Q}$ such that $|x - x'| = \delta$ since $\overline{Q}$ is compact. If $\delta < \alpha$ there is some open neighborhood U of x such that $U \subset E_+$ and $U \subset \overline{\mathrm{B}}(y, \alpha)$ for all $y \in \overline{Q}$. So we get also a contradiction. Hence $\delta = \alpha$ and $x' \in \partial Q$. This latter relation implies in particular $x' \in K$. Indeed, from the definition of K_+ we have

$$\forall y \in \overline{Q},\ \forall z \in K_+, \quad |y - z| \leq \alpha.$$

Since $Q \subset \overline{E}_-$ from (10), this implies $\overline{Q} \subset \overline{K}_- \subset K$. □

Acknowledgements *The authors express their gratitude to C. Raffalli, whose software* `GlSurf` *has permitted the creation of the figures of this paper [7].*

Bibliography

[1] T. Bayen, T. Lachand-Robert & É. Oudet, *Analytic parametrization and volume minimization of three dimensional bodies of constant width*, to appear. See also http://www.lama.univ-savoie.fr/~lachand/index.php?page=Publis .

[2] T. Bonnesen, W. Fenchel, *Theory of convex bodies*, BCS Associates, pp. 135-149 (1987).

[3] G. D. Chakerian, H. Groemer, *Convex bodies of constant width*, Convexity and its Applications, Birkhäuser, pp. 49-96 (1983).

[4] J. B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of convex analysis*, Springer-Verlag, (2001).

[5] F. Meissner, *Über Punktmengen konstanter Breite.* Vierteljahresschr. naturforsch. Ges. Zürich **56** , pp. 42-50 (1911).

[6] E. Oudet, *The Convex Geometry toolbox*, http://www.lama.univ-savoie.fr/~oudet/ConvGeomToolbox/ConvGeomToolbox.html .

[7] C. Raffalli, *GlSurf: OpenGl implicit surface drawing*, http://www.lama.univ-savoie.fr/~raffalli/glsurf.html .

[8] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Cambridge Univ. Press, (1993).

Analytic parametrization and volume minimization of three dimensional bodies of constant width

T. Bayen, T. Lachand-Robert & É. Oudet

III.1 Introduction

A body (that is, a compact connected subset K of $\mathbb{R}^n$) is said to be of *constant width* α if its projection on any straight line is a segment of length $\alpha \in \mathbb{R}_+$, the same value for all lines. This can also be expressed by saying that the *width map*

$$w_K : \nu \in \mathbb{S}^{n-1} \longmapsto \max_{x\in K} \nu \cdot x - \min_{x\in K} \nu \cdot x \tag{1}$$

has constant value α. This is also equivalent to the geometrical fact that two parallel support hyperplanes on K are always separated by a distance α, independent of their direction.

Obvious bodies of constant width are the balls; but they are many others. These bodies, also called *orbiforms* in dimension two, or *spheroforms* in dimension three (as in [2]), have many interesting properties and applications. Orbiforms in particular have been studied a lot during the nineteenth century and later, particularly by Frank Reuleaux, whose name is now attached to those orbiforms you get by intesecting a finite number of disks of equal radii α, whose center are vertices of a regular polygon of diameter α.

Among the oldest problems related to these bodies of constant width are the question of which are those with maximal or minimal volume, for a given value of the width α. It is not difficult to prove that the ball (of radius $\alpha/2$) has maximal volume: this follows from the isoperimetric inequality.

On the other hand, the question of which body of constant width α has minimal volume proved to be much more difficult. First notice that this problem is not correctly stated: indeed, one can remove the interior of a body to decrease its volume, without changing its constant width property. Therefore, we need to add an additional requirement for the problem to make sense (even though this is not needed for the maximization problem). The problem is well-posed if we consider only *convex* bodies, and this is the usual statement considered.

So let us define formally the following class:

$$\mathcal{W}_\alpha := \{K \subset \mathbb{R}^n \,;\, K \text{ compact convex and } \forall \nu \in \mathbb{S}^{n-1},\ w_K(\nu) = \alpha\}. \tag{2}$$

The problem of interest is now to minimize the n-dimensional volume, denoted by $|K|$ hereafter:

$$\text{Find } K^* \in \mathcal{W}_\alpha \text{ such that } |K^*| = \min_{K\in\mathcal{W}_\alpha} |K|\,. \tag{3}$$

Note that the existence of K^* is easy to establish. Indeed $\mathcal{W}_\alpha$ is a compact class of sets for most reasonable topologies (for instance the Haussdorff topology), and the volume is a continuous function.

In dimension two, the problem was solved by Lebesgue and Blaschke: the solution turns out to be a *Reuleaux triangle*.

In dimension three, the problem is still open. Indeed the mere existence of non trivial three-dimensional bodies of constant width is not so easy to establish. In particular, no finite intersection of balls has constant width (except balls themselves), a striking difference with the two-dimensional case.

A simple construction is to consider a two dimensional body of constant width having an axis of symmetry (like the Reuleaux triangle for instance): the corresponding body of revolution obtained by rotation around this axis is a spheroform. F. Meissner proved that the *rotated Reuleaux triangle* has the smaller volume among bodies of revolution in $\mathcal{W}_\alpha$.

Later on he was able to construct another spheroform (usually called "Meissner's tetrahedron") which does not have the symmetry of revolution. The volume of this body is smaller than any other known of constant width, so it is a good candidate as a solution to the problem (3). We describe this body in more details later on in this paper. Let us just say for the moment that it looks like an intersection of four balls centered on the vertices of a regular tetrahedron, but some of the edges are smoothed; in particular, it doesn't have all the symmetries of a regular tetrahedron.

In this paper we first present a complete analytic parametrization of constant width bodies in dimension 3 based on the median surface. More precisely, we define a bijection between the space of functions $\mathrm{C}^{1,1}_\sigma(\Omega)$ and constant width bodies. Then, we compute simple geometrical quantities like the volume and the surface area in terms of those functions. As a corollary we give a new algebraic proof of Blaschke's formula and compute the surface and the volume of Meissner's tetrahedron. Finally, we derive weak optimality conditions for the problem (3).

III.2 The Median Surface

In this section, we introduce a geometrical tool, which we call the *median surface*.

III.2.1 Definition and basics

For a convex body K, we say that a hyperplane H is a *hyperplane of support for K at x*, if $x \in K \cap H$ and K is included in one of the half-spaces limited by H. If $\nu \in \mathbb{S}^{n-1}$ is a normal vector to H, pointing outside the half space containing K, we say that ν is *an outward support vector at x*. Obviously if K is smooth (that is, has a differentiable boundary), then ν is just the outward unit normal at x. In this particular case, there is a map $x \mapsto \nu$ which is usually called *the Gauss map*.

The reverse Gauss map (which is well defined for a body of constant see for instance [11]), satisfies $R_K(\nu) - R_K(-\nu) = \alpha\nu$ for all ν. We may now introduce a parallel surface to ∂K. Consider, for all $\nu \in \mathbb{S}^{n-1}$ the point,

$$M_K(\nu) := R_K(\nu) - \frac{\alpha}{2}\nu = R_K(-\nu) + \frac{\alpha}{2}\nu.$$

Notice that $M_K(-\nu) = M_K(\nu)$. The set of points $M_K(\nu)$ is called the *median surface* of the body K.

Let us recall from [11] one geometrical characterization of constant width bodies:

Theorem 8 *Let K be closed subset of $\mathbb{R}^n$. Then K has constant width α if and only if it satisfies:*

$$\forall \nu \in \mathbb{S}^{n-1},\ \exists x_\nu \in K,$$
$$x_\nu + \alpha\nu \in K \quad \textit{and} \quad \forall y \in K,\ (y - x_\nu)\cdot \nu \in [0, \alpha]. \quad (4)$$

III.2.2 Construction of constant width sets

We present in this section a construction process of constant width bodies starting from an appropriate surface, which will be their median surface. More precisely:

Theorem 9 *Let $\alpha > 0$ be given and $M : \mathbb{S}^{n-1} \to \mathbb{R}^n$ be a continuous application satisfying*

$$\forall \nu \in \mathbb{S}^{n-1}, \qquad M(-\nu) = M(\nu); \quad (5)$$
$$\forall \nu_0, \nu_1 \in \mathbb{S}^{n-1}, \qquad (M(\nu_1) - M(\nu_0))\cdot \nu_0 \le \frac{\alpha}{4}\,|\nu_1 - \nu_0|^2\,. \quad (6)$$

Define a subset $K \subset \mathbb{R}^n$ as follows:

$$K := \left\{ M(\nu) + t\nu\ ;\ \nu \in \mathbb{S}^{n-1}, t \in \left[0, \frac{\alpha}{2}\right]\right\}. \quad (7)$$

Then K is a convex body of constant width α, and $M_K \equiv M$.

Conversely, any convex body of constant width α can be described by (7)*, where $M = M_K$.*

Notice that we could have defined K by

$$K := \left\{ M(\nu) + t\nu\ ;\ \nu \in \mathbb{S}^{n-1}, t \in \left[-\frac{\alpha}{2}, \frac{\alpha}{2}\right]\right\}. \quad (8)$$

This is equivalent to (7), due to (5). Similarly, taking (5) into consideration, we can rewrite (6) with $-\nu_0, -\nu_1$. We deduce that for an application M satisfying (5), (6) is equivalent to:

$$\forall \nu_0, \nu_1 \in \mathbb{S}^{n-1}, \qquad \left|(M(\nu_1) - M(\nu_0))\cdot \nu_0\right| \le \frac{\alpha}{4}\,|\nu_1 - \nu_0|^2\,. \quad (9)$$

In order to prove this theorem, we make use of a lemma:

Lemma 5 *Under the assumptions of Theorem 9, let K be defined by* (7)*. Then $\mathbb{R}^n = \{M(\nu) + t\nu\,;\nu \in \mathbb{S}^{n-1}, t \in \mathbb{R}_+\}$, K is compact, and*

$$\partial K \subset \left\{ M(\nu) + \frac{\alpha}{2}\nu\ ;\ \nu \in \mathbb{S}^{n-1}\right\}. \quad (10)$$

(It will come from Theorem 9 that there is actually equality for the sets in (10).)

Proof. Consider the map $Q : \mathbb{S}^{n-1} \times \mathbb{R} \mapsto M(\nu) + t\nu$ where M satisfies (5). Since M is continuous, $K = Q\left(\mathbb{S}^{n-1} \times [0, \frac{\alpha}{2}]\right)$ is a compact set.

Let us first prove that $Q(\mathbb{S}^{n-1} \times \mathbb{R}_+) = \mathbb{R}^n$. Note that $Q(\mathbb{S}^{n-1} \times \mathbb{R}_+) = Q(\mathbb{S}^{n-1} \times \mathbb{R})$ from (5). We consider some $x \in \mathbb{R}^n$, and assume by contradiction that $x \notin Q(\mathbb{S}^{n-1} \times \mathbb{R})$. For each ν, define x_ν as the projection of x onto the straight line $M(\nu) + \mathbb{R}\nu$. Our assumption implies $x \neq x_\nu$. Moreover

$$x_\nu = M(\nu) + t_\nu \nu \quad \text{where} \quad t_\nu := \nu \cdot (x - M(\nu))$$

as a classical property of the projection.

Since in particular $x \neq M(\nu)$ for all ν, we can define a map $f : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ by $f(\nu) := (x - M(\nu))/|x - M(\nu)|$. Note that f is continuous, and $f(-\nu) = f(\nu)$. Such a map has an even topological degree, and in particular has a fixed point [5]. Therefore there exists some ν such that $f(\nu) = \nu$. For such a ν, we get $x_\nu = x$, a contradiction.

We now turn on the proof of (10). Consider some $x \in \partial K$. In particular, $x \in K$, so $x = M(\nu_0) + t_0\nu_0$ for some ν_0 and $t_0 \in [-\frac{\alpha}{2}, \frac{\alpha}{2}]$. There exists a sequence $(x_n) \subset \mathbb{R}^n \setminus K$ with limit x. From our previous study, we known that $x_n = M(\nu_n) + t_n\nu_n$ for some $\nu_n \in \mathbb{S}^{n-1}$ and $t_n \in \mathbb{R}_+$. The assumption $x_n \notin K$ implies $t_n > \alpha/2$, but on the other hand the sequence (t_n) is bounded since (x_n) is bounded and $M(\mathbb{S}^{n-1})$ is compact. Therefore we may assume that the sequences (ν_n) and (t_n) are convergent. Let us denote by ν_∞ and $t_\infty \geq \frac{\alpha}{2}$ their limits. Since M is continuous, we have $x = M(\nu_\infty) + t_\infty\nu_\infty$.

In particular, $M(\nu_0) = x - t_0\nu_0 = M(\nu_\infty) + t_\infty\nu_\infty - t_0\nu_0$. Let us assume with no loss of generality that $\nu_0 \cdot \nu_\infty \geq 0$ (otherwise we just have to change ν_0 to $-\nu_0$ and t_0 to $-t_0$). We write (6) for ν_∞, ν_0, so

$$\begin{aligned} & (M(\nu_0) - M(\nu_\infty)) \cdot \nu_\infty \leq \frac{\alpha}{4} |\nu_\infty - \nu_0|^2 = \frac{\alpha}{2}(1 - \nu_0 \cdot \nu_\infty) \\ \iff \quad & t_\infty - t_0\, \nu_0 \cdot \nu_\infty \leq \frac{\alpha}{2}(1 - \nu_0 \cdot \nu_\infty) \\ \iff \quad & t_\infty \leq \frac{\alpha}{2} - \left(\frac{\alpha}{2} - t_0\right)\nu_0 \cdot \nu_\infty \leq \frac{\alpha}{2} \end{aligned}$$

since $t_0 \in \left[-\frac{\alpha}{2}, \frac{\alpha}{2}\right]$. This proves that $t_\infty = \frac{\alpha}{2}$. Hence $x \in Q(\mathbb{S}^{n-1}, \frac{\alpha}{2})$. □

Proof of Theorem 9. We begin with the proof of the reciprocal statement in the Theorem. Let K be a body of constant width. We already know that its median surface $M = M_K$ is continuous and satisfies (5). Since $M_K(\nu) = R_K(\nu) - \frac{\alpha}{2}\nu$, and $R_K(\nu_1) \cdot \nu_0 \leq R_K(\nu_0) \cdot \nu_0$ from the definition of R_K, we have

$$(M(\nu_1) - M(\nu_0)) \cdot \nu_0 \leq \frac{\alpha}{2}(1 - \nu_0 \cdot \nu_1) = \frac{\alpha}{4} |\nu_1 - \nu_0|^2 .$$

This proves (6).

Since K is convex and $M_K(\nu) + \frac{\alpha}{2}\nu \in K$, $M_K(\nu) - \frac{\alpha}{2}\nu \in K$, we see that K contains the right hand side of (8). Now let $x \in K$ be given, and let y be the farthest point from x in K. Define $\nu := (y - x)/|y - x|$. For any $z \in K$, we have

$$y \cdot \nu = |y - x| + x \cdot \nu \geq |z - x| + x \cdot \nu \geq (z - x) \cdot \nu + x \cdot \nu = z \cdot \nu$$

so $y = R_K(\nu) = M_K(\nu) + \frac{\alpha}{2}\nu$. Hence $x = M_K(\nu) + t\nu$ with $t = \frac{\alpha}{2} - |x - y|$. Since $|x - y| \leq \alpha$, we have $|t| \leq \frac{\alpha}{2}$, which concludes the proof of (8).

We now prove the direct statement in the Theorem. So consider a map M satisfying (5) and (6), and K be defined by (7) (or (8) equivalently). In view of Theorem 8, we need to prove (4).

Let $\nu \in \mathbb{S}^{n-1}$ be given. Consider $x_\nu := M(\nu) - \frac{\alpha}{2}\nu$, so that $x_\nu + \alpha\nu = M(\nu) + \frac{\alpha}{2}\nu \in K$ from its definition.

Consider any $y \in K$, so that $y = M(\hat{\nu}) + t\hat{\nu}$. Changing $\hat{\nu}$ and t to their opposite, if necessary, we may assume that $\nu \cdot \hat{\nu} \geq 0$. Note that

$$(y - x_\nu) \cdot \nu = (M(\hat{\nu}) - M(\nu)) \cdot \nu + t\nu \cdot \hat{\nu} + \frac{\alpha}{2}.$$

Using (9) with $\nu_0 = \nu, \nu_1 = \hat{\nu}$, we get

$$-\frac{\alpha}{2}(1 - \nu \cdot \hat{\nu}) \leq \big(M(\hat{\nu}) - M(\nu)\big) \cdot \nu \leq \frac{\alpha}{2}(1 - \nu \cdot \hat{\nu}).$$

Hence, since $t \in \left[-\frac{\alpha}{2}, \frac{\alpha}{2}\right]$:

$$0 \leq (t + \frac{\alpha}{2})\ \nu \cdot \hat{\nu} \leq (y - x_\nu) \cdot \nu \leq \alpha + (t - \frac{\alpha}{2})\ \nu \cdot \hat{\nu} \leq \alpha.$$

This concludes the proof of the theorem. □

Applications M satisfying (5) and (6) will play an important role in the remaining of this paper. So let us give a few additionnal properties on them. We start here with simple inequalities, and will consider what happens on a differential level in the next section. Note that all these results apply in particular to the median surface of any convex body of constant width according to Theorem 9.

Lemma 6 *Let M be a continuous application satisfying* (5) *and* (6)*. Then M is $\frac{\alpha}{2}$-lipschitzian:*

$$\forall \nu_0, \nu_1 \in \mathbb{S}^{n-1}, \qquad |M(\nu_1) - M(\nu_0)| \leq \frac{\alpha}{2} |\nu_1 - \nu_0| . \tag{11}$$

and satisfies:

$$\forall \nu_0, \nu_1 \in \mathbb{S}^{n-1}, \qquad \left| M(\nu_1) + \frac{\alpha}{2}\nu_1 - M(\nu_0) - \frac{\alpha}{2}\nu_0 \right| \leq \alpha. \tag{12}$$

Proof. According to Theorem 9, M is the median surface of some $K \in \mathcal{W}_\alpha$ defined by (7). Since K contains $M(\nu) + \frac{\alpha}{2}\nu$ for any ν, and has diameter α, we get (12).

Squaring the left hand side of (12) and expanding it, we get

$$|M(\nu_1) - M(\nu_0)|^2 - \alpha(M(\nu_1) - M(\nu_0)) \cdot (\nu_1 - \nu_0) \leq \frac{\alpha^2}{4} |\nu_1 + \nu_0|^2 \tag{13}$$

since $|\nu_1 - \nu_0|^2 + |\nu_1 + \nu_0|^2 = 4$. The above relation is true for any pair of unit vectors, so we can write it for $(\nu_1, -\nu_0)$ and $(-\nu_1, \nu_0)$. We get, taking (5) into account:

$$|M(\nu_1) - M(\nu_0)|^2 - \alpha(M(\nu_1) - M(\nu_0)) \cdot (\nu_1 + \nu_0) \leq \frac{\alpha^2}{4} |\nu_1 - \nu_0|^2$$
$$|M(\nu_1) - M(\nu_0)|^2 + \alpha(M(\nu_1) - M(\nu_0)) \cdot (\nu_1 + \nu_0) \leq \frac{\alpha^2}{4} |\nu_1 - \nu_0|^2 .$$

Summing these relations yields (11). □

III.2.3 Smooth median surface

In this section we reduce (6) to local differential properties. This is easy whenever M is differentiable, but requires more involved statements in the general case. Note that M will always be defined on the sphere $\mathbb{S}^{n-1}$, and if differentiable, its derivative $DM(\nu)$ is defined on the tangent space to the sphere at ν, which is simply $\nu^\perp := \{w \in \mathbb{R}^n\ ;\ w \cdot \nu = 0\}$. In the following proposition, we consider C^2 maps $\widetilde{\nu} : [0, 1] \to \mathbb{S}^{n-1}$, and $\dot{\widetilde{\nu}}$ is the derivative of $\widetilde{\nu}$. Notice that only the end point $\widetilde{\nu}(0)$ and the corresponding derivatives do matter.

Property 3 *Let $M : \mathbb{S}^{n-1} \to \mathbb{R}^n$ be given. Then M satisfies* (9) *if and only if it satisfies*

$$\forall \widetilde{\nu} \in \mathrm{C}^2([0,1];\mathbb{S}^{n-1}),$$
$$\limsup_{t \overset{>}{\to} 0} \frac{1}{t^2} \left| \left(M(\widetilde{\nu}(t)) - M(\widetilde{\nu}(0)) \right) \cdot \widetilde{\nu}(0) \right| \leq \frac{\alpha}{4} \left| \dot{\widetilde{\nu}}(0) \right|^2 . \tag{14}$$

If M is differentiable, then (14) *is equivalent to*

$$\forall \nu_0 \in \mathbb{S}^{n-1}, \forall w \in \nu_0^\perp,$$
$$\nu_0 \cdot DM(\nu_0)w = 0 \quad \textit{and} \quad |w \cdot DM(\nu_0)w| \leq \frac{\alpha}{2} |w|^2 . \tag{15}$$

We will shorten (15) in the following by writing it $\nu_0 \cdot DM(\nu_0) = 0$ (as vectors) and $\pm DM(\nu_0) \leq \frac{\alpha}{2}\,\mathrm{Id}$ (as matrices). This expresses the fact that ν_0 is the normal vector to the surface $\nu_0 \mapsto M(\nu_0)$ at $M(\nu_0)$, and that the absolute values of the curvature radii does not exceed $\frac{\alpha}{2}$. (See also the parametric equivalent in the next section.)

Proof. Assume first that M satisfies (9). Let $\nu \in \mathrm{C}^2([0,1];\mathbb{S}^{n-1})$, and define $\nu_0 := \widetilde{\nu}(0)$ for short. Note that $|\widetilde{\nu}(t)|^2 = 1$ for all t, so

$$\forall t, \qquad \widetilde{\nu}(t) \cdot \dot{\widetilde{\nu}}(t) = 0 \quad \text{and} \quad \widetilde{\nu}(t) \cdot \ddot{\widetilde{\nu}}(t) = -\left| \dot{\widetilde{\nu}}(t) \right|^2 . \tag{16}$$

In particular a Taylor expansion near $t = 0$ yields

$$\widetilde{\nu}(t) \cdot \nu_0 = (\nu_0 + t\dot{\widetilde{\nu}}(0) + \frac{t^2}{2}\ddot{\widetilde{\nu}}(0) + o(t^2)) \cdot \nu_0 = 1 - \frac{t^2}{2} \left| \dot{\widetilde{\nu}}(0) \right|^2 + o(t^2).$$

Using (9) with $\widetilde{\nu}(t)$ and ν_0, we get:

$$\left| \left(M(\widetilde{\nu}(t)) - M(\nu_0) \right) \cdot \nu_0 \right| \leq \frac{\alpha}{2}(1 - \widetilde{\nu}(t) \cdot \nu_0) \leq \frac{\alpha}{4} t^2 \left| \dot{\widetilde{\nu}}(0) \right|^2 + o(t^2).$$

Dividing by t^2, we get (14).

If M is differentiable and $w \in \nu_0^\perp$, consider $\widetilde{\nu}(t) := p_{\mathbb{S}^{n-1}}(\nu_0 + tw)$ where $p_{\mathbb{S}^{n-1}} : x \mapsto x/|x|$ is the projection on the sphere. So $\widetilde{\nu}(0) = \nu_0$ and $\dot{\widetilde{\nu}}(0) = w$. Hence we have

$$M(\widetilde{\nu}(t)) \cdot \nu_0 = M(\nu_0) \cdot \nu_0 + t\nu_0 \cdot DM(\nu_0)w + o(t)$$

so (14) clearly implies $\nu_0 \cdot DM(\nu_0)w = 0$.

Assume for a moment that M is twice differentiable and satisfies (14). We already know that $\nu_0 \cdot DM(\nu_0)w = 0$ for any ν_0 and any $w \in \nu_0^\perp$. Therefore we have, for any $\nu \in \mathrm{C}^2([0,1];\mathbb{S}^{n-1})$:

$$\forall t, \qquad 0 = \widetilde{\nu}(t) \cdot DM(\widetilde{\nu}(t))\dot{\widetilde{\nu}}(t).$$

Differentiating this relation with respect to t, we get

$$0 = \dot{\widetilde{\nu}}(t) \cdot DM(\widetilde{\nu}(t))\dot{\widetilde{\nu}}(t) + \widetilde{\nu}(t) \cdot D^2M(\widetilde{\nu}(t))(\dot{\widetilde{\nu}}(t), \dot{\widetilde{\nu}}(t))$$

since $\widetilde{\nu}(t) \cdot DM(\widetilde{\nu}(t)) = 0$. Considering $t = 0$ and $w := \dot{\widetilde{\nu}}(0) \in \nu_0^\perp$ yields

$$\forall w \in \nu_0^\perp, \qquad w \cdot DM(\nu_0)w = -\nu_0 \cdot D^2M(\nu_0)(w, w). \tag{17}$$

Therefore a Taylor expansion yields

$$\begin{aligned} M(\widetilde{\nu}(t)) \cdot \nu_0 &= M(\nu_0) \cdot \nu_0 + \frac{1}{2} t^2 \ \nu_0 \cdot D^2 M(\nu_0)(w, w) + o(t^2) \\ &= M(\nu_0) \cdot \nu_0 - \frac{1}{2} t^2 \ w \cdot DM(\nu_0) w + o(t^2). \end{aligned} \tag{18}$$

It is now clear that (14) implies (15).

If M is not twice differentiable, we use an approximation argument as follows. For any $\beta > \alpha$ and any $\varepsilon > 0$, there exists an approximating map $M_\varepsilon \in \mathrm{C}^2(\mathbb{S}^{n-1}, \mathbb{R}^n)$, such that

$$\|M - M_\varepsilon\|_{W^{1,\infty}(\mathbb{S}^{n-1};\mathbb{R}^n)} \leq \varepsilon \tag{19}$$

and M_ε satisfies (14) with α replaced by β. Hence M_ε satisfies (15), also with α replaced by β. Letting ε go to zero and using (19), we deduce that M satisfies (15) with α replaced by β. Since this holds for any $\beta > \alpha$, it holds for α as well.

Conversely, if M is differentiable and satisfies (15), let us prove that it satisfies (14). Using exactly the same approximation, we see that we just have to prove that for M twice differentiable. In such a case, (15) implies (17). Hence the Taylor expansion (18) holds true. This yields (14).

Let us now prove the reverse statement of the proposition, that is, a map M satisfying (14) also satisfies (9). Again it is enough to prove it for a twice differentiable map, for (19) implies in particular uniform convergence of M_ε to M.

So let us consider two vectors ν_0, ν_1 in $\mathbb{S}^{n-1}$ and prove (9). We consider a geodesic path $\widetilde{\nu} \in \mathrm{C}^2([0,1]; \mathbb{S}^{n-1})$ such that $\widetilde{\nu}(0) = \nu_0$ and $\widetilde{\nu}(1) = \nu_1$. Such a path satisfies $\widetilde{\nu}(t) \in (\mathbb{R}\nu_0 + \mathbb{R}\nu_1)$, and $\nu_0 \cdot \dot{\widetilde{\nu}}(t) \leq 0$ for all t.

The function $f : t \mapsto \nu_0 \cdot M(\widetilde{\nu}(t))$ has derivative $f'(t) = \nu_0 \cdot DM(\widetilde{\nu}(t))\dot{\widetilde{\nu}}(t)$. Since $\widetilde{\nu}(t) \in (\mathbb{R}\nu_0 + \mathbb{R}\nu_1)$, we have

$$\nu_0 = (\nu_0 \cdot \widetilde{\nu}(t))\, \widetilde{\nu}(t) + \frac{(\nu_0 \cdot \dot{\widetilde{\nu}}(t))}{\left|\dot{\widetilde{\nu}}(t)\right|^2}\, \dot{\widetilde{\nu}}(t).$$

Taking (15) and $\nu_0 \cdot \dot{\widetilde{\nu}}(t) \leq 0$ into account, we get

$$|f'(t)| = \frac{\left|\nu_0 \cdot \dot{\widetilde{\nu}}(t)\right|}{\left|\dot{\widetilde{\nu}}(t)\right|^2}\ \left|\dot{\widetilde{\nu}}(t) \cdot DM(\widetilde{\nu}(t))\dot{\widetilde{\nu}}(t)\right| \leq -\frac{\alpha}{2}(\nu_0 \cdot \dot{\widetilde{\nu}}(t)).$$

Therefore

$$\begin{aligned} \left|(M(\nu_1) - M(\nu_0)) \cdot \nu_0\right| &= |f(1) - f(0)| \\ &\leq -\frac{\alpha}{2} \int_0^1 (\nu_0 \cdot \dot{\widetilde{\nu}}(t))\, \partial t = \frac{\alpha}{2}(1 - \nu_0 \cdot \nu_1). \end{aligned}$$

This completes the proof of the proposition. □

Remark 2.B. Observe that (14) is equivalent to

$$\forall \widetilde{\nu} \in \mathrm{C}^2([0,1]; \mathbb{S}^{n-1}),$$
$$\limsup_{t \xrightarrow{>} 0} \frac{1}{t^2} \left|\Big(M(\widetilde{\nu}(t)) - M(\widetilde{\nu}(0))\Big) \cdot \widetilde{\nu}(t)\right| \leq \frac{\alpha}{4} \left|\dot{\widetilde{\nu}}(0)\right|^2. \tag{20}$$

Indeed we just have to prove that for a smooth M again. Then we may rewrite (18) with v_0 and $\widetilde{v}(t)$ reversed:

$$M(v_0)\cdot\widetilde{v}(t) = M(\widetilde{v}(t))\cdot\widetilde{v}(t) - \frac{1}{2}t^2\,\dot{\widetilde{v}}(t)\cdot DM(\widetilde{v}(t))\dot{\widetilde{v}}(t) + o(t^2).$$

Since $w = \dot{\widetilde{v}}(0) = \dot{\widetilde{v}}(t) + O(t)$ and DM is continuous, we get by substracting (18):

$$\Big|\Big(M(\widetilde{v}(t)) - M(\widetilde{v}(0))\Big)\cdot(\widetilde{v}(t) - \widetilde{v}(0))\Big| = o(t^2)$$

as $t \to 0$. This proves that the limits on the right hand sides in (14) and (20) are equal.

Let us recall a classical geometrical definition: two smooth oriented surfaces S and S' are said to be *parallel at distance* δ if S' is the image of S through the map $x \mapsto x + \delta\vec{n}_S(x)$, where $\vec{n}_S$ is the normal vector field on S. It is classical that the normal vector on S' at $x + \delta\vec{n}_S(x)$ is actually $\vec{n}_S(x)$. (We will give a proof of this result in the next section.) In particular, S is also a surface parallel to S', at distance $-\delta$. Moreover, if S have well defined radii of curvature $\rho_i(x)$ $(i = 1, 2)$, then S' also have radii of curvature at $x + \delta\vec{n}_S(x)$, equal to $\rho_i(x) + \delta$.

So we see that for a body K of constant width α with median surface M_K, the median surface and the boundary ∂K are parallel at distance $\pm\alpha$, whenever they are smooth. In general, these surfaces are not smooth, but only have Lipschitz regularity, though.

III.3 Parametrizations

In this section, we give a parametrization of the median surface of a body K of constant width. This provides a simple parametrization of the boundary of K, and gives a simple formula to compute the volume and surface area of K.

From now on we focus on the three-dimensional setting. A similar work can easily be done in dimension two, but the properties of orbiforms are already quite well known.

III.3.1 Isothermal parametrization of the sphere

Let us start with a parametrization of the unit sphere $\mathbb{S}^2$ in the form $(u, v) \in \Omega \mapsto \nu(u, v)$, where Ω is some subset of $\mathbb{R}^2$. We assume that this parametrization is *isothermal*, that is, satisfies for all $(u, v) \in \Omega$:

$$\partial_u\nu(u,v)\cdot\partial_v\nu(u,v) = 0 \qquad \text{and} \qquad |\partial_u\nu(u,v)| = |\partial_v\nu(u,v)| =: \frac{1}{\lambda(u,v)}. \tag{21}$$

We also assume that the map $\nu : \Omega \to \mathbb{S}^2$ is injective and almost surjective, that is, its image set is equal to $\mathbb{S}^2$ except possibly a finite number of points.

An example of such a parametrization is

$$(u, v) \in (\mathbb{R}/2\pi\mathbb{Z}) \times \mathbb{R} \longmapsto \left(\frac{\cos u}{\cosh v}, \frac{\sin u}{\cosh v}, \tanh v\right) \tag{22}$$

and in such a case $\lambda(u, v) = \cosh v$, and $\nu(\Omega) = \mathbb{S}^2 \setminus \{(0, 0, \pm 1)\}$. However we do not rely on this particular form in the following.

For technical reasons, we will also assume that λ satisfies, for all values of (u, v), the identity

$$\lambda^2 \nabla \cdot \left(\lambda^{-1} \nabla \lambda\right) = \lambda \Delta \lambda - |\nabla \lambda|^2 = 1. \tag{23}$$

(Gradient and Laplacian taken relative to (u, v).) This is clearly true for the particular parametrization given above.

Let us shorten the notations by not writing the dependencies on the parameters (u, v). We introduce the unit vectors $\nu_u := \lambda \partial_u \nu$, $\nu_v := \lambda \partial_v \nu$. Since ν is also a unit vector, we have $\nu \cdot \partial_u \nu = 0$, so $\nu \cdot \nu_u = 0$; and similarly $\nu \cdot \nu_v = 0$. Hence the family (ν, ν_u, ν_v) is an orthonormal basis of $\mathbb{R}^3$, taking (21) into account.

Lemma 7 *For such an isothermal parametrization of the unit sphere, we have*

$$\partial_u \nu_u = -\lambda^{-1} \nu + \lambda^{-1} \partial_v \lambda \; \nu_v \tag{24}$$

$$\partial_v \nu_u = -\lambda^{-1} \partial_u \lambda \; \nu_v \tag{25}$$

$$\partial_u \nu_v = -\lambda^{-1} \partial_v \lambda \; \nu_u \tag{26}$$

$$\partial_v \nu_v = -\lambda^{-1} \nu + \lambda^{-1} \partial_u \lambda \; \nu_u \tag{27}$$

Proof.

Since $\nu \cdot \partial_u \nu = 0$, we get by differentiating $\nu \cdot \partial^2_{uv} \nu = -\partial_u \nu \cdot \partial_v \nu = 0$, so $\partial^2_{uv} \nu$ has the form $\alpha \nu_u + \beta \nu_v$. On the other hand

$$\partial^2_{uv} \nu = \partial_u (\partial_v \nu) = \partial_u (\lambda^{-1} \nu_v) = \lambda^{-1} \partial_u \nu_v - \lambda^{-2} \partial_u \lambda \; \nu_v.$$

Since $|\nu_v| = 1$ implies $\nu_v \cdot \partial_u \nu_v = 0$, we get $\beta = \partial^2_{uv} \nu \cdot \nu_v = -\lambda^{-2} \partial_u \lambda$. Similarly $\alpha = -\lambda^{-2} \partial_v \lambda$. Putting this relation in the value of $\partial^2_{uv} \nu$ above, we deduce (26). We get (25) using $\partial^2_{uv} \nu = \partial_v (\partial_u \nu)$ in the same way.

Differentiating the three relations $|\nu_u|^2 = 1$, $\nu \cdot \nu_u = 0$ and $\nu_u \cdot \nu_v$ with respect to u, we get $\nu_u \cdot \partial_u \nu_u = 0$,

$$\nu \cdot \partial_u \nu_u = -\nu_u \cdot \partial_u \nu = -\lambda^{-1} |\nu_u|^2 = -\lambda^{-1}$$

and

$$\nu_v \cdot \partial_u \nu_u = -\nu_u \cdot \partial_u \nu_v = \lambda^{-1} \partial_v \lambda.$$

This gives (24). The proof of (27) is similar. □

Let us finish this section with a note about the antipodal symmetry on $\mathbb{S}^2$ that we will use in the following sections. There must be some involutive map $\sigma : \Omega \to \Omega$ such that $\nu \circ \sigma(u, v) = -\nu(u, v)$ for all $(u, v) \in \Omega$. For instance with the parametrization (22) we have

$$\nu(u + \pi, -v) = -\nu(u, v) \tag{28}$$

so $\sigma : (u, v) \mapsto (u + \pi, -v)$. We will call this map the *antipodal symmetry of the parametrization*. In the following, we will always assume that σ is C^1 and is *consistent* with the isothermal parametrization, that is satisfies:

$$\lambda \circ \sigma = \lambda. \tag{29}$$

Since $\nu = -\nu \circ \sigma$, we have $\partial_u \nu = -\partial_u\sigma\partial_u\nu \circ \sigma$. Considering the norm of both sides, we deduce with (29) that $|\partial_u\sigma| = 1$. Similarly we have $|\partial_v\sigma| = 1$. Since $\sigma \in \mathrm{C}^1$, we see that

$$\partial_u\sigma = \text{const.} = \pm 1 \quad \text{and} \quad \partial_u\sigma = \text{const.} = \pm 1. \tag{30}$$

These relations, together with the definition of ν_u, ν_v and (29) imply

$$\nu_u \circ \sigma = -\partial_u\sigma\nu_u, \qquad \nu_v \circ \sigma = -\partial_v\sigma\nu_v. \tag{31}$$

III.3.2 Parametrization of the median surface

Since the three vectors (ν, ν_u, ν_v) are independent, any point $P \in \mathbb{R}^3$ can be written in the form $P = h\nu + h_1\nu_u + h_2\nu_v$. If h, h_1, h_2 are actually some smooth functions of (u, v), P depends on (u, v) and describe a surface. In this section we investigate the conditions on h, h_1, h_2 ensuring that such a surface is the median surface of a spheroform, with support vector $\nu(u, v)$ at $P(u, v)$.

Property 4 *Given an isothermal parametrization $\nu : \Omega \to \mathbb{S}^2$ of the sphere, let K be a strictly convex body. There exists a C^1 map $h : \Omega \to \mathbb{R}$ such that $R_K(\nu) = \mathcal{M}(h)(u, v)$ for all $\nu = \nu(u, v)$, where*

$$\mathcal{M}(h) : \left| \begin{array}{rcl} \Omega & \longrightarrow & \mathbb{R}^3 \\ (u, v) & \longmapsto & h\,\nu + \lambda\partial_u h\,\nu_u + \lambda\partial_v h\,\nu_v. \end{array} \right. \tag{32}$$

Proof. For any given $\nu = \nu(u, v)$, consider $P(u, v) := R_K(\nu(u, v))$. Since the three vectors (ν, ν_u, ν_v) are independent, $P(u, v)$ can be written in the form $P = h\nu + h_1\nu_u + h_2\nu_v$, for some functions h, h_1, h_2 of (u, v). These functions are continuous since R_K is continuous.

Note that $h_K(\nu(u, v)) = \nu(u, v) \cdot P(u, v) = h(u, v)$. So h is just the support function of K, and in particular is of class C^1. Moreover we have from the definition of R_K:

$$\forall (u_1, v_1) \in \Omega, \qquad P(u_1, v_1) \cdot \nu \le P \cdot \nu.$$

(All values of the functions are at (u, v), unless otherwise specified.) Let us write this relation with $u_1 = u + t, v_1 = v$. For small values of t, we have from (24–27):

$$\begin{aligned} \nu(u + t, v) &= \nu + t\lambda^{-1}\nu_u + o(t), \\ \nu_u(u + t, v) &= \nu_u - t\lambda^{-1}(\nu - \partial_v\lambda\nu_v) + o(t), \\ \nu_v(u + t, v) &= \nu_v - t\lambda^{-1}\partial_v\lambda\nu_u + o(t). \end{aligned}$$

Also since h is of class C^1, we have $h(u + t, v) = h + t\partial_u h + o(t)$. Hence

$$\begin{aligned} 0 = P \cdot \nu - h &\ge P(u_1, v_1) \cdot \nu - h \\ &\ge t\Big(\partial_u h - \lambda^{-1}h_1(u + t, v)\Big) + o(t). \end{aligned}$$

Passing to the limit $t = 0$ with either $t > 0$ or $t < 0$, we deduce that $h_1(u, v) = \lambda\partial_u h$. Similarly $h_2 = \lambda\partial_v h$. □

Remark 3.C. Notice that $\mathcal{M}(h)$ is obviously linear with respect to h. Since in the previous proposition, $h = \mathcal{M}(h) \cdot \nu = R_K \cdot \nu$ is the support function of K, then the mapping from K to h is additive (with respect to the Minkowski addition). However, not any h yields an interesting body K. In particular, if $h(u, v) = \vec{w} \cdot \nu(u, v)$ for some fixed vector $\vec{w} \in \mathbb{R}^3$, then $\partial_u h = \vec{w} \cdot \partial_u \nu = \lambda^{-1} \vec{w} \cdot \nu_u$, so $\mathcal{M}(h) = \vec{w}$ is constant, and the corresponding body K reduces to a point. Due to the additivity property, we see that adding $\vec{w} \cdot \nu$ to some given h is equivalent to a translation of the corresponding body K by the vector $\vec{w}$.

We prove in the next theorem that for a constant width body, the corresponding function h is actually $\mathrm{C}^{1,1}$ (the derivatives are lipschitzian). Here and in the following, differential operators like ∇ (gradient) or Δ (laplacian) are taken relative to the variables (u, v). We denote by $\nabla^\perp$ the operator $(-\partial_v, \partial_u)$. Whenever h is twice differentiable, we denote by $D^2 h$ the 2×2 matrix of its second-order derivatives (hessian matrix).

An inequality like $D^2 h(u, v) \leq A$, where A is also a 2×2 symmetrical matrix, means that the difference $A - D^2 h(u, v)$ is nonnegative definite. For $h \in \mathrm{C}^{1,1}$ only, the second-order derivatives do not necessarily exists, but the Taylor expansion

$$T[h](u, v; \xi, \eta) := h(u + \xi, v + \eta) - h(u, v) - \xi \partial_u h(u, v) - \eta \partial_v h(u, v)$$

is of order $O(\xi^2 + \eta^2)$ for (ξ, η) small.

Definition 1 *We shall say that* $D^2 h(u, v) \leq A = (a_{i,j})$ in a generalized sense*, if the following occurs:*

$$\limsup_{(\xi,\eta)\to(0,0)} \frac{T[h](u, v; \xi, \eta) - \frac{1}{2}(a_{11}\xi^2 + 2a_{12}\xi\eta + a_{22}\eta^2)}{\xi^2 + \eta^2} \leq 0. \tag{33}$$

Similarly we say that $D^2 h(u, v) \geq A$ in a generalized sense*, if a similar property holds with a limit-inf* ≥ 0 *instead.*

Clearly this is the same as the usual meaning for a twice-differentiable function h, since

$$T[h](u, v; \xi, \eta) = \frac{1}{2}\xi^2 \partial^2_{uu} h(u, v) + \xi\eta \partial^2_{uv} h(u, v) + \frac{1}{2}\eta^2 \partial^2_{vv} h(u, v) + o(\xi^2 + \eta^2)$$

in that case.

Definition 2 *Given an isothermal parametrization* $\nu : \Omega \to \mathbb{S}^2$ *of the sphere, let* σ *be its antipodal symmetry. Let* $\mathrm{C}^{1,1}_\sigma(\Omega)$ *be the set of all* $\mathrm{C}^{1,1}$ *maps* $h : \Omega \to \mathbb{R}$ *such that*

$$h \circ \sigma = -h. \tag{34}$$

Let $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ *be the subset of functions* $h \in \mathrm{C}^{1,1}_\sigma(\Omega)$ *satisfying everywhere on* Ω *in a generalized sense (see Definition 1 above):*

$$-\frac{\alpha}{2\lambda^2}\,\mathrm{Id} \leq U[h] \leq \frac{\alpha}{2\lambda^2}\,\mathrm{Id} \tag{35}$$

where

$$U[h] := D^2 h + \lambda^{-2} h\,\mathrm{Id} + \lambda^{-1} \nabla\lambda \otimes \nabla h - \lambda^{-1} \nabla^\perp \lambda \otimes \nabla^\perp h. \tag{36}$$

Theorem 10 *Given an isothermal parametrization of the sphere, let* $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ *be given by the Definition 2 above.*

Then an application $M : \mathbb{S}^2 \to \mathbb{R}^3$ *is the median surface of a spheroform if and only if there exists* $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ *such that* $M(\nu) = \mathcal{M}(h)(u, v)$ *for all* $\nu = \nu(u, v)$*, where the map* $\mathcal{M}(h) : \Omega \to \mathbb{R}^3$ *is defined by* (32)*. In this case, the map* $\mathcal{M}(h + \frac{\alpha}{2}) : \Omega \to \mathbb{R}^3$ *describes all but a finite number of the points on* ∂K*.*

The restriction about exceptional points on ∂K comes from the fact that $\nu(\Omega)$ equals $\mathbb{S}^2$, excepts some exceptional points. (The points $(0, 0, \pm 1)$ with the parametrization (22).)

Proof. Given $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$, define $M : \mathbb{S}^2 \to \mathbb{R}^3$ by $M(\nu) = \mathcal{M}(h)(u, v)$ for all $\nu = \nu(u, v)$. Let us prove that M is the median surface of some spheroform. In view of Theorem 9, Proposition 3 and Remark 2.B, we just have to prove (5) and (20). From (29–31) we get

$$M(-\nu) = \mathcal{M}(h) \circ \sigma(u, v) = -h \circ \sigma\nu + \lambda\partial_u h \circ \sigma\partial_u\sigma\nu_u + \lambda\partial_v h \circ \sigma\partial_v\sigma\nu_v.$$

But (34) implies in particular $\partial_u h \circ \sigma = \partial_u h/\partial_u\sigma$ and a similar relation for v. So $\mathcal{M}(h) \circ \sigma = \mathcal{M}(h)$ and M satisfies (5).

Let us now prove (20). Any $\widetilde{\nu} \in \mathrm{C}^2([0, 1]; \mathbb{S}^2)$ can be written in the parametrization as $\widetilde{\nu}(t) = \nu(u(t), v(t))$ where $u(t), v(t) \in \mathrm{C}^2([0, 1])$. If $\widetilde{\nu}(0) = \nu(u_0, v_0) =: \nu_0$, we also have $u(0) = u_0$, $v(0) = v_0$. Let us consider $\xi := u(t) - u_0$, $\eta := v(t) - v_0$. Since $\partial_u\nu = \lambda^{-1}\nu_u$, and

$$\partial^2_{uu}\nu = \partial_u(\lambda^{-1}\nu_u) = \lambda^{-2}(-\nu + \partial_v\lambda\nu_v - \partial_u\lambda\nu_u)$$

with the help of (24). With similar relations for the other derivatives, we get the Taylor expansion of $\widetilde{\nu}$ near $t = 0$:

$$\begin{aligned}\widetilde{\nu}(t) = \nu(u_0 + \xi, v_0 + \eta) &= \nu_0 + \xi\lambda^{-1}\nu_u + \eta\lambda^{-1}\nu_v \\ &+ \frac{1}{2\lambda^2}\Big[\xi^2(-\nu + \partial_v\lambda\nu_v - \partial_u\lambda\nu_u) - 2\xi\eta(\partial_u\lambda\nu_v + \partial_v\lambda\nu_u) + \eta^2(-\nu + \partial_u\lambda\nu_u - \partial_v\lambda\nu_v)\Big] \\ &+ o(\xi^2 + \eta^2), \quad (37)\end{aligned}$$

where all functions on the right hand side are computed at (u_0, v_0).

In particular we get using $\xi = u(t) - u_0 = t\dot{u}(0) + o(t)$ and $\eta = v(t) - v_0 = t\dot{v}(0) + o(t)$,

$$\dot{\widetilde{\nu}}(0) = \lim_{t\to 0}\frac{1}{t}(\widetilde{\nu}(t) - \nu_0) = \lambda^{-1}(\dot{u}(0)\nu_u + \dot{v}(0)\nu_v).$$

This implies

$$\frac{1}{\lambda^2}(\xi^2 + \eta^2) = t^2\left|\dot{\widetilde{\nu}}(0)\right|^2 + o(t^2). \quad (38)$$

Similarly since $M(\nu(u,v)) \cdot \nu(u,v) = h(u,v)$ from the definition of h, we have:

$$\begin{aligned}
&(M(\widetilde{\nu}(t)) - M(\nu_0)) \cdot \widetilde{\nu}(t) \\
&\qquad = h(u(t), v(t)) - M(\nu_0) \cdot \nu(u(t), v(t)) \\
&\qquad = h(u_0 + \xi, v_0 + \eta) - (h\nu_0 + \lambda\partial_u h \nu_u + \lambda\partial_v h \nu_v) \cdot \nu(u_0 + \xi, v_0 + \eta) \\
&\qquad = h(u_0 + \xi, v_0 + \eta) - h - \xi\partial_u h - \eta\partial_v h + \frac{h}{2\lambda^2}(\xi^2 + \eta^2) \\
&\qquad\qquad + \frac{1}{2\lambda}\Big(\xi^2(-\partial_v\lambda\partial_v h + \partial_u\lambda\partial_u h) + 2\xi\eta(\partial_u\lambda\partial_v h + \partial_v\lambda\partial_u h) \\
&\qquad\qquad\qquad + \eta^2(-\partial_u\lambda\partial_u h + \partial_v\lambda\partial_v h)\Big) \\
&\qquad\qquad + o(\xi^2 + \eta^2) \\
&\qquad = T[h](u_0, v_0; \xi, \eta) + \frac{1}{2}\begin{pmatrix}\xi \\ \eta\end{pmatrix} A \begin{pmatrix}\xi & \eta\end{pmatrix} + o(t^2)
\end{aligned}$$

where $A := \lambda^{-2} h \,\mathrm{Id} + \lambda^{-1}\nabla\lambda \otimes \nabla h - \lambda^{-1}\nabla^\perp\lambda \otimes \nabla^\perp h$.

Using the right inequality in (35), and the definition of the corresponding generalized sense, we deduce with (38):

$$\limsup_{t\to 0} \frac{1}{t^2}(M(\widetilde{\nu}(t)) - M(\nu_0)) \cdot \widetilde{\nu}(t) \le \limsup_{t\to 0} \frac{\alpha}{4\lambda^2 t^2}(\xi^2 + \eta^2) = \frac{\alpha}{4}\left|\dot{\widetilde{\nu}}(0)\right|.$$

Similarly the left inequality in (35) yields the reverse inequality, which achieves the proof of (20).

Conversely, let K be a spheroform. We know from Proposition 4 that there exists some function $\tilde{h} \in \mathrm{C}^1(\Omega)$ such that $R_K(\nu) = \mathcal{M}(\tilde{h})(u,v)$ for all $\nu = \nu(u,v)$.

Consider now the function $h := \tilde{h} - \frac{\alpha}{2}$. From the definition of $\mathcal{M}$, it is clear that $\mathcal{M}(h)(u,v) = \mathcal{M}(\tilde{h}) - \frac{\alpha}{2}\nu(u,v)$, so for any $\nu = \nu(u,v)$ we have

$$M_K(\nu(u,v)) = R_K(\nu(u,v)) - \frac{\alpha}{2}\nu(u,v) = \mathcal{M}(h)(u,v).$$

Moreover the map $\nu \mapsto M_K(\nu)$ is lipschitzian from Lemma 6. Hence $\partial_u h(u,v) = M_K(\nu)\cdot\nu_u(u,v)$ is lipschitzian, too. And similarly for $\partial_v h$. So $h \in \mathrm{C}^{1,1}$. Additionally $h(u,v) = M_K(\nu(u,v)) \cdot \nu(u,v)$ implies $h \circ \sigma = M_K(-\nu)\cdot(-\nu) = -h$, so h satisfies (34). Hence $h \in \mathrm{C}^{1,1}_\sigma(\Omega)$.

We know that M_K satisfies (20). If we consider the special path $\tilde{\nu} : t \mapsto \nu(u_0 + t\xi, v_0 + t\eta)$, we can expand $\widetilde{\nu}(t)$ near $t = 0$ as before, obtaining something similar to (37). This implies with a similar computation:

$$(M(\widetilde{\nu}(t) - M(\nu_0)) \cdot \widetilde{\nu}(t) = t^2 T[h](u_0, v_0; \xi, \eta) + \frac{t^2}{2}\begin{pmatrix}\xi \\ \eta\end{pmatrix} A \begin{pmatrix}\xi & \eta\end{pmatrix} + o(t^2).$$

Therefore (20) implies (35) in the generalized sense. This completes the proof that $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. □

III.3.3 Regularity of the parametrization

In this section, we investigate the consequences of (35) on h, whenever h is regular enough.

Property 5 *Let h be C^2 on some open set $\omega \subset \Omega$. Then h satisfies* (35) *on ω if and only if it satisfies*

$$|R(h)| \leq \min\left(\frac{\alpha}{\lambda}, \frac{\alpha}{2\lambda} + \frac{2\lambda}{\alpha} J(h)\right) \tag{39}$$

on ω, where $R(h)$ and $J(h)$ are the trace and determinant of the matrix $\lambda^{-1}hI + \nabla\lambda \otimes \nabla h - \nabla^\perp\lambda \otimes \nabla^\perp h + \lambda D^2 h$, that is

$$R(h) := \frac{2h}{\lambda} + \lambda\Delta h \tag{40}$$

$$J(h) := \lambda^{-2}h^2 + h\Delta h + \lambda^2 \det D^2 h + \lambda\nabla^\perp\lambda \cdot D^2 h \cdot \nabla^\perp h - \lambda\nabla\lambda \cdot D^2 h \cdot \nabla h - |\nabla\lambda|^2 \ |\nabla h|^2 . \tag{41}$$

Proof. For a C^2 function, the generalized sense for (35) is just the common pointwise sense. We can multiply by λ and get (39) since a 2×2 matrix is nonnegative definite, if, and only if, its trace and determinant are nonnegative. □

Let us note for further references that $R(h)$ and $J(h)$ are the trace and determinant of a symmetric matrix. Therefore it has real eigenvalues, and in particular the discriminant of its characteristic polynomial is nonnegative:

$$R(h)^2 \geq 4J(h). \tag{42}$$

This holds for any C^2 function h.

Notice that for any $\delta \in \mathbb{R}$,

$$R(h + \delta) = R(h) + \frac{2\delta}{\lambda}, \tag{43}$$

$$J(h + \delta) = J(h) + \delta\lambda^{-1}R(h) + \lambda^{-2}\delta^2. \tag{44}$$

Therefore (39) may be equivalently written

$$R(h + \frac{\alpha}{2}) \geq 0, \quad R(h - \frac{\alpha}{2}) \leq 0, \quad J(h + \frac{\alpha}{2}) \geq 0 \quad \text{and} \quad J(h - \frac{\alpha}{2}) \geq 0. \tag{45}$$

Remark 3.D. The appearance of the matrix in the previous proposition seems quite odd at first. Here is another way to obtain it, which is easier to understand, but requires again $h \in \mathrm{C}^2$, so we can compute the derivatives of $M := \mathcal{M}(h)$. We get using (24–27):

$$\partial_u M = a\nu_u + b\nu_v \quad \text{and} \quad \partial_v M = c\nu_u + d\nu_v \tag{46}$$

where

$$\begin{aligned} a &:= \lambda^{-1}h + \partial_u\lambda\partial_u h - \partial_v\lambda\partial_v h + \lambda\partial^2_{uu}h \\ b = c &:= \partial_v\lambda\partial_u h + \partial_u\lambda\partial_v h + \lambda\partial^2_{uv}h \\ d &:= \lambda^{-1}h + \partial_v\lambda\partial_v h - \partial_u\lambda\partial_u h + \lambda\partial^2_{vv}h. \end{aligned}$$

So we find $DM\,\nu = 0$ in agreement to Proposition 3. We also see from their definition that $R(h) = a + d$ and $J(h) = ad - bc$.

Since $M = \mathcal{M}(h) \in \mathrm{C}^1$, (14) is equivalent to (15) according to Proposition 3. Since for $\nu_0 = \nu(u_0, v_0)$, we have $\nu_0^\perp = \mathrm{Span}(\nu_u(u_0, v_0), \nu_v(u_0, v_0))$, we just have to check (15) for $w = \xi\nu_u + \eta\nu_v$, with arbitrary (ξ, η). This inequality becomes then, using (24–27):

$$\forall (\xi, \eta) \in \mathbb{R}^2, \quad \left|a\xi^2 + 2b\xi\eta + d\eta^2\right| \leq \frac{\alpha}{2\lambda}(\xi^2 + \eta^2).$$

This means

$$-\frac{\alpha}{2\lambda}\,\mathrm{Id} \leq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \leq \frac{\alpha}{2\lambda}\,\mathrm{Id}$$

in the sense of matrices, which is (35).

Remark 3.E. The previous proposition has also a geometrical meaning and can be proved using corresponding considerations. Indeed, in matrix notations, we have $\nabla M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \nu_u \\ \nu_v \end{pmatrix}$, using again the notations of the previous remark. Consequently we get:

$$\nabla\nu = \begin{pmatrix} \partial_u\nu \\ \partial_v\nu \end{pmatrix} = \lambda^{-1}\begin{pmatrix} \nu_u \\ \nu_v \end{pmatrix} = \lambda^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}\nabla M.$$

By definition, the curvatures of the surface $(u, v) \mapsto M(u, v)$ are the eigenvalues of the matrix A such that $\nabla\nu = A\nabla M$, since ν is normal to the surface. And the curvature radii, their inverse, are the eigenvalues of A^{-1}. We see that in our case $A = \lambda^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$. So the curvature radii are the solutions ρ_i $(i = 1, 2)$ of the equation

$$\rho^2 - \lambda^2 R(h)\rho + \lambda^2 J(h) = 0. \tag{47}$$

Therefore, if we change h to $\tilde{h} := h + \delta$ in order to consider a parallel surface, we see that the curvature radii $\tilde{\rho}_i$ on this new surface are solutions of the equation

$$\begin{aligned} 0 &= \tilde{\rho}^2 - (2\delta + 2h + \lambda^2\Delta h)\tilde{\rho} + \lambda^2 J(h) + \delta(2h + \lambda^2\Delta h) + \delta^2 \\ &= (\tilde{\rho} + \delta)^2 - (2h + \lambda^2\Delta h)(\tilde{\rho} + \delta) + \lambda^2 J(h). \end{aligned}$$

Hence $\tilde{\rho}_i = \rho_i + \delta$ as claimed before.

For a body of constant width α, the parallel to the median surface at distance $\pm\frac{\alpha}{2}$ are part of the boundary of K. So they are convex, with opposite directions (the outward normal vector on $\mathcal{M}(h - \frac{\alpha}{2})(u, v)$ is $-\nu(u, v)$). Hence we must have $\rho_i \in [-\frac{\alpha}{2}, +\frac{\alpha}{2}]$. This is equivalent to saying that the left hand side of (47) is nonnegative whenever $\rho = \pm\frac{\alpha}{2}$, and that the sum of the roots belongs to $[-\alpha, \alpha]$. This in turn is equivalent to (39). In other words, (39) expresses the fact that the radii of curvature on the median surface are in $[-\frac{\alpha}{2}, +\frac{\alpha}{2}]$, whenever they are defined.

The equivalent formula (45) expresses the fact that the Gaussian curvatures $J(h \pm \frac{\alpha}{2})$ are nonnegative, while the mean curvatures $R(h \pm \frac{\alpha}{2})$ have opposite signs, since the convex surfaces are opposite.

III.3.4 Surface area and volume

According to Theorem 10, there is a one to one correspondence between $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ and $\mathcal{W}_\alpha$. We investigate now the way to compute the volume and surface area of some $K \in \mathcal{W}_\alpha$ through the corresponding function h.

Property 6 *Let $\omega \subset \Omega$ be a symmetrical subset of the parametrization space, that is $\sigma(\omega) = \omega$. Let $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ be C^2 on $\partial\omega$, and let K be the corresponding spheroform.*

The set $R_K(\omega) \subset \partial K$ has surface area:

$$\begin{aligned} |R_K(\omega)| = \frac{\alpha^2}{4}\,|\nu(\omega)| + \int_\omega \left(\lambda^{-2}h^2 - \frac{1}{2}\,|\nabla h|^2\right) \\ + \int_{\partial\omega} \left(h + \frac{1}{2}\lambda^2\Delta h\right)\nabla h\cdot\vec{n} - \frac{1}{4}\int_{\partial\omega} \nabla\left(\lambda^2\,|\nabla h|^2\right)\cdot\vec{n}. \end{aligned} \tag{48}$$

(Here $|\nu(\omega)|$ stands for the surface area of the subset $\nu(\omega)$ of $\mathbb{S}^2$.)

Proof. It is enough to prove the proposition for $h \in \mathrm{C}^2(\overline{\omega})$. Indeed an approximation argument allows to generalize to others h, since the right hand side of (48) involves second-order derivatives only on the boundary of ω.

We make use of the notations of Proposition 5 and Remark 3.D. We have

$$\partial_u M \times \partial_v M = (a\nu_u + b\nu_v) \times (c\nu_u + d\nu_v) = J(h)\nu.$$

Hence the area of the surface $\mathcal{M}(h)(\omega)$ is $\int_\omega |J(h)|$. Since $\mathcal{M}(h)$ is the median surface of $K \in \mathcal{W}_\alpha$, $\mathcal{M}(h+\frac{\alpha}{2})$ and $\mathcal{M}(h-\frac{\alpha}{2})$ both describe the boundary of K. If we restrict the parameters to the subset ω, they both describe $R_K(\omega)$ since we assumed $\omega = \sigma(\omega)$. So the surface area of $R_K(\omega)$ is equal to $\int_\omega \left|J(h \pm \frac{\alpha}{2})\right|$. This implies, using (44) and (45):

$$\begin{aligned} |R_K(\omega)| &= \frac{1}{2}\int_\omega J(h+\frac{\alpha}{2})(u,v)\partial u\partial v + \frac{1}{2}\int_\omega J(h-\frac{\alpha}{2})(u,v)\partial u\partial v \\ &= \int_\omega \left(\frac{\alpha^2}{4\lambda^2(u,v)} + J(h)(u,v)\right)\partial u\partial v. \\ &= \frac{\alpha^2}{4}\,|\nu(\omega)| + \int_\omega J(h)\partial u\partial v. \end{aligned}$$

(The latter equality follows from $\partial_u\nu \times \partial_v\nu = \lambda^{-2}\nu$.) To complete the proof of the proposition, we have now to prove:

$$\begin{aligned} \int_\omega J(h)(u,v)\partial u\partial v = \int_\omega \left(\lambda^{-2}h^2 - \frac{1}{2}\,|\nabla h|^2\right) \\ + \int_{\partial\omega} \left(h + \frac{1}{2}\lambda^2\Delta h\right)\nabla h\cdot\vec{n} - \frac{1}{4}\int_{\partial\omega} \nabla\left(\lambda^2\,|\nabla h|^2\right)\cdot\vec{n}. \end{aligned} \tag{49}$$

By expanding products in (41), we get

$$J(h) = \lambda^{-2}h^2 + h\Delta h + J_1(h) - J_2(h), \tag{50}$$

where

$$J_1(h) := \lambda^2 \left(\partial^2_{uu} h \partial^2_{vv} h - (\partial^2_{uv} h)^2\right) + \lambda \left(\partial_u \lambda \partial_u h \partial^2_{vv} h + \partial_v \lambda \partial_v h \partial^2_{uu} h - \partial_u \lambda \partial_v h \partial^2_{uv} h - \partial_v \lambda \partial_u h \partial^2_{uv} h\right)$$

and

$$J_2(h) := \left((\partial_u \lambda)^2 + (\partial_v \lambda)^2\right)\left((\partial_u h)^2 + (\partial_v h)^2\right) + \lambda \left(\partial_u \lambda \partial_u h \partial^2_{uu} h + \partial_v \lambda \partial_v h \partial^2_{vv} h + \partial_v \lambda \partial_u h \partial^2_{uv} h + \partial_u \lambda \partial_v h \partial^2_{uv} h\right).$$

Let us define $w_1 := \partial_u h \partial^2_{uv} h - \partial_v h \partial^2_{uu} h$, $w_2 := \partial_u h \partial^2_{vv} h - \partial_v h \partial^2_{uv} h$ and $\vec{w} := (w_1, w_2)$. We have $\partial_u w_2 - \partial_v w_1 = 2(\partial^2_{uu} h \partial^2_{vv} h - (\partial^2_{uv} h)^2)$. Therefore

$$2J_1(h) = \partial_u \left(\lambda^2 w_2\right) - \partial_v \left(\lambda^2 w_1\right).$$

This implies using Green's formula

$$\int_\omega J_1(h) = \frac{1}{2} \int_{\partial\omega} \lambda^2 \vec{w} \cdot \vec{\ell}.$$

Now let us denote by H the scalar function $|\nabla h|^2$. Since $\vec{w} = \Delta h \nabla^\perp h - \frac{1}{2}\nabla^\perp H$ (where $\nabla^\perp = (-\partial_v, \partial_u)$), we also have

$$\int_\omega J_1(h) = \frac{1}{2} \int_{\partial\omega} \lambda^2 \left(\Delta h \nabla h - \frac{1}{2}\nabla H\right) \cdot \vec{n} \partial s.$$

Considering now J_2, we can check easily that $J_2(h) = |\nabla\lambda|^2 H + \frac{1}{2}\lambda \nabla\lambda \cdot \nabla H$. Integrating by parts we get

$$\begin{aligned}
\int_\omega J_2(h) &= \int_\omega H \left(|\nabla\lambda|^2 - \frac{1}{2}\nabla \cdot (\lambda \nabla\lambda)\right) + \frac{1}{2} \int_{\partial\omega} H \lambda \nabla\lambda \cdot \vec{n} \partial s \\
&= \frac{1}{2} \int_\omega H \left(|\nabla\lambda|^2 - \lambda\Delta\lambda\right) + \frac{1}{2} \int_{\partial\omega} H \lambda \nabla\lambda \cdot \vec{n} \partial s \\
&= -\frac{1}{2} \int_\omega H + \frac{1}{2} \int_{\partial\omega} H \lambda \nabla\lambda \cdot \vec{n} \partial s
\end{aligned}$$

using (23).

Finally we have

$$\int_\omega h\Delta h = \int_{\partial\omega} h\, |\nabla h| \cdot \vec{n} \partial s - \int_\omega |\nabla h|^2 = \int_{\partial\omega} h \nabla h \cdot \vec{n} \partial s - \int_\omega H.$$

So integrating (50) yields (49). □

We are now in position to compute the volume and surface area of any spheroform, expressed as integrals of the corresponding function h:

Theorem 11 *Let $h \in C^{1,1}_{\sigma,\alpha}(\Omega)$ be given, and $K \in \mathcal{W}_\alpha$ the corresponding spheroform. The surface area $|\partial K|$ and the volume $|K|$ are given by*

$$|\partial K| = \int_\Omega \left(\lambda^{-2}h^2 - \frac{1}{2}|\nabla h|^2\right) + \pi\alpha^2, \tag{51}$$

$$|K| = \frac{\alpha}{2}\int_\Omega \left(\lambda^{-2}h^2 - \frac{1}{2}|\nabla h|^2\right) + \frac{\pi\alpha^3}{6}. \tag{52}$$

Corollary 12 (Blaschke) *Let K be any convex body of constant width α in dimension 3. Then the volume and surface area of K satisfy:*

$$|K| = \frac{\alpha}{2}|\partial K| - \frac{\pi\alpha^3}{3}. \tag{53}$$

We refer the reader to [1] for the original proof of this property. Here it follows directly from Theorem 11.

Proof. The parametrization domain $\Omega = \mathbb{S}^1 \times \mathbb{R}$ has no boundary. So if we apply (48) with $\omega = \Omega$, we get (51) since $|\nu(\Omega)| = \left|\mathbb{S}^2\right| = 4\pi$.

The volume of K can be expressed as $|K| = \frac{1}{3}\int_{\partial K} \overrightarrow{OM}\cdot\vec{n}\partial_\sigma$, using Stokes' formula. We can choose $M = \mathcal{M}(h + \frac{\alpha}{2})(u,v)$ as a parametrization, and then $\vec{n} = \nu(u,v)$ and $\partial_\sigma = J(h+\frac{\alpha}{2})\partial_u\partial_v$. But we may also choose $M = \mathcal{M}(h - \frac{\alpha}{2})$, and in such a case $\vec{n} = -\nu(u,v)$ since $\vec{n}$ is the outward normal in Stokes' formula, and $\partial_\sigma = J(h-\frac{\alpha}{2})\partial_u\partial_v$. So we have

$$|K| = \frac{1}{3}\int_\Omega (h+\frac{\alpha}{2})J(h+\frac{\alpha}{2}) = -\frac{1}{3}\int_\Omega (h-\frac{\alpha}{2})J(h-\frac{\alpha}{2}).$$

In particular this implies, using (44) and an integration by parts:

$$\begin{aligned}
|K| &= \frac{1}{6}\int_\Omega \left\{(h+\frac{\alpha}{2})J(h+\frac{\alpha}{2}) - (h-\frac{\alpha}{2})J(h-\frac{\alpha}{2})\right\}. \\
&= \frac{\alpha}{6}\int_\Omega J(h) + \frac{\alpha^3}{24}\int_\Omega \lambda^{-2} + \frac{\alpha}{6}\int_\Omega h(2\lambda^{-2}h + \Delta h) \\
&= \frac{\alpha}{6}\int_\Omega J(h) + \frac{\pi\alpha^3}{6} + \frac{\alpha}{6}\int_\Omega \left(2\lambda^{-2}h^2 - |\nabla h|^2\right).
\end{aligned}$$

This proves (52) using (49) with $\omega = \Omega$. □

III.3.5 Description of Meissner's tetrahedron

A description of this volume can be found in [3],[15] and [8]. We shall give a brief definition of this volume and describe its parametrization.

Meissner's tetrahedron is geometrically defined in the following way: consider a body K_t obtained as the intersection of four balls of radius α which centers are the vertices of a regular tetrahedron (of edge lengths α). Thus, the boundary of K_t is composed of four pieces of balls connected

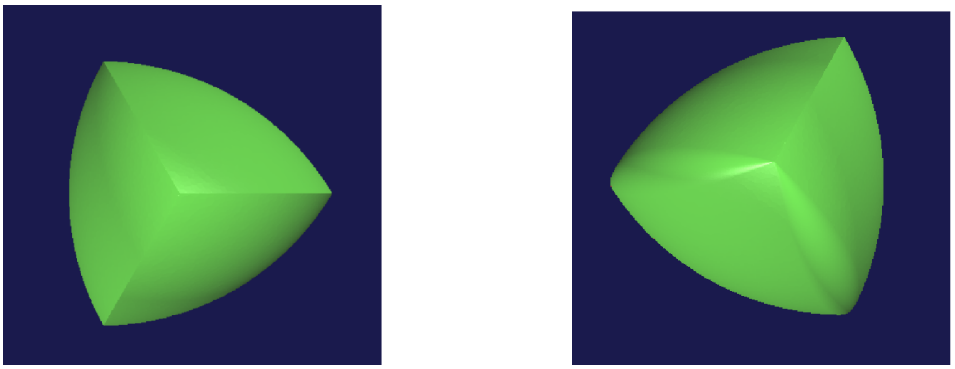

Figure III.1: Two views of one Meissner's tetrahedron

by six arc of circles. Surprisingly, this set K_t is not of constant width: geometrical considerations show that opposite circular edges are too far away. Meissner proposed to smooth three edges of K_t in order to get a constant width body. Consider E the union of three circular edges which share a common vertex S. Then, the body K defined as

$$K = \bigcap_{x \in E} B(x, \alpha) \cap K_t$$

is a body of constant width called Meissner's tetrahedron (see figure III.1). Notice that it is possible to build an other constant width body based on the regular etrahedron by smoothing a different set of edges.

We give below an analytical representation in terms of its h function based on the parametrization of the sphere described by (22). In order to take benefit of the invariance of the previous body K by rotations of angles $\pm 2\pi/3$, we consider a body K built on a regular tetrahedron which has its vertex S on the z-axes and the others on the plane $z = 0$. Moreover, we assume that the equilateral triangle formed by other vertices on $z = 0$ is symmetric with respect to the y-axes. It is straightforward to check that such a Meissner's tetrahedron is invariant with respect to the rotations about the z-axes of angles $\pm 2\pi/3$ and also invariant by orthogonal symmetry with respect to the plane $x = 0$. Then, the function h is completely defined if we give an analytical representation of h on $\omega = [0, \frac{\pi}{3}] \times [0, +\infty[$ since relations (34) and $h(u, v) = h(-u, v)$ define h on all $\Omega = [-\pi, \pi] \times]-\infty, +\infty[$ (see figure III.3.5). On ω, for $\alpha = 1$, the function h may be described in the following way:

$$\begin{cases} \sqrt{2/3}\tanh v - 1/2, & \text{if } \sinh v > 2\sqrt{2}\cos u, \\ -1/2 - (1/2\sqrt{3})(\cos u/\cosh v) + \dots & \\ (\sqrt{3}/2)(\cosh v^2 - \sin u^2)/\cosh v, & \text{if } \sinh v \leq 2\sqrt{2}\cos u, \cosh v \geq 2\sin u, \\ 1/2 + (1/\sqrt{3})\cos(u + 2\pi/3)/\cosh v & \text{if } \sinh v \leq 2\sqrt{2}\cos u, \cosh v > 2\sin u. \end{cases}$$

Notice that it is possible to compute the volume and the surface area of Meissner's tetrahedron thanks to equations (51) and (53). After some symbolic computations, we get the formulas

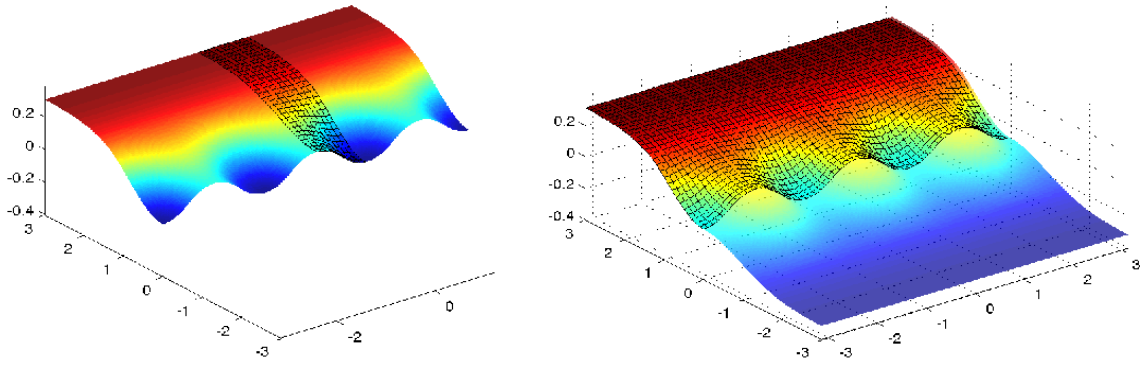

Figure III.2: Construction of Meissner's h function

presented in [7]:

$$|K| = \frac{2\pi}{3} - \frac{\pi\sqrt{3}}{4}\arccos\frac{1}{3}$$
$$|\partial K| = 2\pi - \frac{\pi\sqrt{3}}{2}\arccos\frac{1}{3}$$

III.3.6 Local optimality

We now come back to the volume functional $K \mapsto |K|$ in order to investigate the properties of its minimizers. A striking consequence of Theorem 11 is that minimizing the volume in $\mathcal{W}_\alpha$ is equivalent to minimizing the surface area. More precisely, the volume minimization problem is equivalent to

$$\min_{h\in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)} L(h) \qquad \text{where} \qquad L(h) = \int_\Omega \left(\lambda^{-2}h^2 - \frac{1}{2}|\nabla h|^2\right). \tag{54}$$

Let us first observe that the maximum value of L is zero:

Lemma 8 *For any $h \in \mathrm{C}^{1,1}_\sigma(\Omega)$, we have $L(h) \le 0$.*

In particular, a maximizer of L in $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ is always $h = 0$, which corresponds to a ball of radius $\alpha/2$. Hence such a ball has maximal volume among all spheroforms, a well-known result.

Proof. Let W_σ be the space of all functions $h \in W^{1,2}(\Omega)$ satisfying (34). This is a closed subspace of the Sobolev space $W^{1,2}(\Omega)$, so it is a Banach space. Let us define $s \in \mathbb{R}$ as follows:

$$s = \inf_{h\in W_\sigma} \frac{\displaystyle\int_\Omega |\nabla h|^2}{\displaystyle\int_\Omega \lambda^{-2}h^2}.$$

This is a "weighted Sobolev constant", and it is classical in PDE theory that the infimum is actually attained by a smooth function $\varphi \in W_\sigma$ satisfying the corresponding Euler equation

$$\int_\Omega \nabla h\cdot\nabla\varphi = s\int_\Omega \lambda^{-2}h\varphi, \qquad \forall h \in W_\sigma.$$

In other words, φ is an eigenfunction of the operator $-\lambda^2\Delta$, with the symmetry condition $\varphi\circ\sigma = -\varphi$. Additionally, if we choose two open sets $\Omega_+ \subset \Omega$ and $\Omega_- = \sigma(\Omega_+)$ such that $\Omega_+ \cap \Omega_- = \emptyset$ and $\Omega = \overline{\Omega_+} \cup \Omega_-$, then it follows from Krein-Rutman's theorem that there exists an eigenfunction φ satisfying $\varphi > 0$ on Ω_+. One way to choose such a set Ω_+ is to consider some fixed vector $\vec{w} \in \mathbb{R}^3$, and to set

$$\Omega_+ := \{(u,v) \in \Omega \,;\, \vec{w}\cdot\nu(u,v) > 0\}.$$

Given such a $\vec{w}$, define $g := (u,v) \mapsto \vec{w}\cdot\nu(u,v)$. As explained in Remark 3.C, $\mathcal{M}(g) = \vec{w}$ for all (u,v), and the body corresponding to $h+g$, for any $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$, is just a translation of the body corresponding to h. In particular, they have the same volume, so $L(h+g) = L(h)$. Since L is quadratic, this means that

$$0 = L(g) + \int_\Omega \lambda^{-2} hg - \nabla h\cdot\nabla g$$

for all $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. In particular, $L(g) = 0$ since we can take $h = 0$, and $\Delta g + 2\lambda^{-2} g$ is orthogonal (for the L^2 scalar product) to all $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. The latter implies it is orthogonal to W_σ, since $\bigcup_{\alpha>0} \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ contains $\mathrm{C}^2_\sigma(\Omega)$. Hence it is orthogonal to φ, so we get

$$2\int_\Omega \lambda^{-2} g\varphi = \int_\Omega \nabla g\cdot\nabla\varphi = s\int_\Omega \lambda^{-2} g\varphi.$$

Now both functions g and φ are positive on Ω_+ and odd with respect to σ, so

$$\int_\Omega \lambda^{-2} g\varphi = 2\int_{\Omega_+} \lambda^{-2} g\varphi > 0$$

and therefore $s = 2$. This implies $L(h) \le 0$ for any $h \in W_\sigma$, and in particular in $\mathrm{C}^{1,1}_\sigma(\Omega)$. □

Remark 3.F. Since balls are the unique maximizers of the volume among spheroforms of given width, it follows that for $h \in \mathrm{C}^{1,1}_\sigma(\Omega)$:

$$L(h) = 0 \iff \exists \vec{w} \in \mathbb{R}^3,\ h(u,v) = \vec{w}\cdot\nu(u,v),$$

for all $(u,v) \in \Omega$.

An interesting consequence of the previous lemma is that the functional L is actually strictly concave with respect to h (when considered on the quotient of $\mathrm{C}^{1,1}_\sigma(\Omega)$ by the smallest subspace containing all the functions $\vec{w}\cdot\nu(u,v)$ for $\vec{w} \in \mathbb{R}^3$).

Indeed L is quadratic, so for any $h, g \in \mathrm{C}^{1,1}_\sigma(\Omega)$ and for all $t \in [0,1]$:

$$L(th + (1-t)g) - tL(h) - (1-t)L(g) = -t(1-t)L(h-g) \ge 0.$$

From the remark 3.F, the equality holds if and only if it exists $\vec{w} \in \mathbb{R}^3$ such that $h = g + \vec{w}\cdot\nu(u,v)$.

The following weak optimality result applies not only to global minimizers, but also to local ones. Notice that this condition is very close from the one established in (**??**) for a relaxed problem of (3).

Theorem 13 *Let K be a body of constant width, and a local minimizer of the volume functional. Then K is everywhere irregular in the following sense: for any $A \subset \mathbb{S}^{n-1}$, one of the two subsets $R_K(A)$ or $R_K(-A)$ of ∂K is not a smooth surface.*

In this context, a "smooth surface" means that the set of points can be described as the graph of a regular function. Observe that this result is obvious in dimension two for global minimizers, since these are Reuleaux triangles.

Proof. Let K be a local minimizer of the volume and $A \subset \mathbb{S}^2$ with $R_K(A)$ a smooth surface. Let h be the function of $C^{1,1}_\sigma(\Omega)$ associated to K by the proposition 4. Since every constant width bodies are strictly convex, we can assume without loss of generality that $R_K(A)$ is the graph of a strictly convex function. In this context, it is standard that the reverse Gauss map is a smooth diffeomorphism. Moreover, the function h is also locally smooth on the points of $\omega \subset \Omega$ corresponding to A since:

$$h(u,v) = R_K(\nu(u,v)) \cdot \nu(u,v).$$

Let us first establish that h saturates the pointwise constraint (35) on a subset of ω. By reducing ω to a smaller set if necessary, we suppose that $\omega \cap \sigma(\omega) = \emptyset$. Assume by contradiction that the four inequalities are strict. Let $g \in C^2(\omega)$ with compact support. We extend it to $\sigma(\omega)$ by symmetry, defining $g(\sigma(u,v)) = -g(u,v)$, so that the new function, still denoted g, belongs to $C^{1,1}_\sigma(\Omega)$. Due to the non-saturation property, the functions $f_+ := h + tg$ and $f_- := h - tg$ belong to $C^{1,1}_\sigma(\Omega)$ for $|t|$ small enough. Now L is strictly concave so we have:

$$L(h) = L(\frac{1}{2}f_+ + \frac{1}{2}f_-) \geq \min(L(f_+), L(f_-)). \tag{55}$$

for all g. Since an equality in (55) is not possible because of the remark 3.F (none of the function $\vec{w} \cdot \nu(u,v)$ has a compact support), we have that $L(h) > \min(L(f_+), L(f_-))$. This contradicts the local minimality of h.

We established in subsection III.3.3, that the saturation of the constraints for the regular function h is equivalent to the fact that one or both of the radii of curvature on $R_K(A)$ are equal to α or 0. Since $R_K(A)$ is a strictly convex regular surface, its curvature radii are not zero. As a consequence, on all points of $R_K(A)$ at least one of the curvature radii is equal to α. Consider now the surjective application from ω to $R_K(-A)$ given by

$$(u,v) \mapsto R_K(\nu(u,v)) - \alpha\nu(u,v).$$

If this application is not injective, $R_K(-A)$ is not smooth since at least one point of this surface has a non empty subdiferrential. We conclude that the previous application is an admissible parametrization of $R_K(-A)$. It is now straightforward to compute that on all points of $R_K(-A)$, at least one of the curvature radii is equal to 0. Again, this fact contradicts the regularity of $R_K(-A)$ which concludes the proof. □

Remark 3.G. If we assume additionally that the lines of curvature on $R_K(A)$ of the body K have no torsion, it is possible to show that $R_K(-A)$ is a convex curve. In this situation we would conclude that one of the two pieces of the boundary $R_K(A)$ or $R_K(-A)$ has measure 0. Notice that Meissner's tetrahedron satisfies the previous assumption.

Bibliography

[1] W. Blaschke, *Konvexe Bereiche gegebener konstanter Breite und kleinsten Inhalts*, Math. Ann., 76, 1915, pp. 504–513.

[2] T. Bonnesen, W. Fenchel, *Theory of convex bodies*, BCS Associates, 1987, pp. 135-149.

[3] G. D. Chakerian, H. Groemer, *Convexity and its Applications*, Chap. Convex bodies of constant width, ed P. M. Gruber and J. M. Wills, Birkhauser, 1983, pp. 49-96.

[4] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, 1976.

[5] J. Dugundji & A. Granas, *Fixed point theory*, Monog. Mat. Polska Ak., 61, 1982.

[6] E. M. Harrell, *A direct proof ot a theorem of Blaschke and Lebesgue*, Journal of Geometric Analysis, 12, 2002, pp. 81 88.

[7] E. M. Harrell, *Calculations for Convex Bodies. Example: The rotated Reuleaux triangle*, http://www.mathphysics.com/convex/Links/Convexnb_lnk_67.html.

[8] D. Hilbert, *Elf Eigenschaften der Kugel*.

[9] J. B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of convex analysis*, Springer-Verlag, 2001.

[10] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Series of Comprehensive Studies in Mathematics, 224, 2. Edition, Springer 1977, 1983.

[11] T. Lachand-Robert, É. Oudet, *Bodies of constant width in arbitrary dimension*, to appear in Math. Nachrichten.

[12] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[13] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Cambridge Univ. Press, 1993.

[14] K. Toyama, *Self-Parallel Constant Mean Curvature Surfaces*

[15] I.M. Yaglom, V.G. Boltyanskii, *Convex Figures*, Holt, Rinehart and Winston New York.

Shape optimisation under width constraint

Édouard Oudet

IV.1 Introduction

This article deals with numerical shape optimisation problems involving convex shapes under width constraints in $\mathbb{R}^3$. Throughout the article, we make use of the following notations:

- K is a convex body of $\mathbb{R}^3$ with nonempty interior which contains the origin,
- ∂K denotes its boundary,
- ν_K is the almost everywhere defined outer normal vector field on ∂K, with values on the sphere $\mathbb{S}^2$,
- for $\nu \in \mathbb{S}^2$, $\varphi_K(\nu)$ is the distance to the origin of the supporting plane to K of exterior normal ν. More explicitly,
$$\varphi_K(\nu) = \sup_{x \in K} x.\nu$$
where $x.\nu$ stands for the usual scalar product of $\mathbb{R}^3$. φ_K is called the support function of K,
- $w_K(\nu) = \varphi_K(\nu) + \varphi_K(-\nu)$ for $\nu \in \mathbb{S}^2$ is called the *width* in the direction ν.

The two kinds of optimisation problem that we will study are :

$$\min_{K \in \mathcal{K}} F(K)$$

where

$$\mathcal{K} = \{K \text{ convex}, \ w_K(\nu) = 1, \ \forall \nu \in \mathbb{S}^2\}. \tag{1}$$

or

$$\mathcal{K} = \{K \text{ convex}, \ w_K(\nu) \geq 1, \ \forall \nu \in \mathbb{S}^2\} \tag{2}$$

In particular, we focus our work on the numerical study of the previous problems when $F(K)$ has a geometrical meaning. More precisely, we restrict our study to $F(K)$ equal to the volume of K denoted by $|K|$ or the surface area of ∂K denoted by S_K.

Taking $F(K)$ equal to $|K|$ (or equivalently to S_K by Blaschke's formula), problem (1) is a well known question called Meissner's conjecture. In dimension two, this problem was solved by Lebesgue and Blaschke: the solution turns out to be a *Reuleaux triangle*. In dimension three, this problem is still open. Indeed the mere existence of non trivial three-dimensional bodies of constant

width is not so easy to establish. In particular, no finite intersection of balls has constant width (except balls themselves), a striking difference with the two-dimensional case. A simple construction, to obtain constant width bodies in dimension 3, is to consider a two dimensional body of constant width having an axis of symmetry (like the Reuleaux triangle for instance): the corresponding body of revolution obtained by rotation around this axis is of constant width. F. Meissner proved that the *rotated Reuleaux triangle* has the smaller volume among constant width bodies of revolution. Later on he was able to construct another spheroform (usually called "Meissner's tetrahedron") which does not have the symmetry of revolution. The volume of this body is smaller than any other known of constant width, so it is a good candidate as a solution to the problem (1) (see [9], [10], [11], [12]).

In a first part of this article, we study constant width constraints of type (1) using an analytical parametrisation introduced in [1]. We discuss a cubic spline method based on [6] which approximates problem (1) by a standard quadratic programming problem under equalities and inequalities constraints. The main interest of the method is that it gives a discrete way to parameterise (and not to approximate) constant with bodies. This point is of dramatic importance to study Meissner's conjecture in a numerical way. Based on that method, we perform numerical experiments to study the local optimality of Meissner's body and of the *rotated Reuleaux triangle*. Our numerical results satisfy the weak optimality condition which has been described in [1]. More precisely, any constant width body K^* which minimises the area is irregular in the sense that for any $\omega \in \mathbb{S}^2$ small enough, the part of ∂K^* whose normals are ω and the other part whose normals are $-\omega$ are not both regular.

In a second part we study the relaxed problem (2). The question to minimise the surface area among convex bodies of prescribed minimal width was first addressed in [5]. One convex body based on a regular simplex, whose precise description is recalled in the following, has been conjectured by E. Heil to be optimal in 1978. The previous analytical parametrisation is not relevant in that context. Thus, we give a new algebraic discretisation of convex bodies based on Minkowski's sums. To illustrate the efficiency of our method for inequality constraints, we solve numerically Heil's problem. Our numerical optimisation gives a polytope which is admissible (in the sense that it satisfies exactly the constraints up to round off errors) and has a surface area smaller than Heil's polytope. This result disprove Heil's conjecture.

IV.2 A geometrical approach and its difficulties

For every $\nu \in \mathbb{S}^2$ and every $\varphi \geq 0$, let us define the half-space of $\mathbb{R}^3$:

$$[\![\nu, \varphi]\!] = \left\{ x \in \mathbb{R}^3,\ x \cdot \nu \leq \varphi \right\}.$$

In a previous article [7] the authors present a discretisation of convex bodies based on half spaces. A convex set K is approached by a polytope P which is defined in the following way. Let $n \in \mathbb{N}^*$, choose randomly and uniformly n vectors ν_i of $\mathbb{S}^2$ and define

$$P_n = \bigcap_{i=1}^{n} [\![\nu_i, \varphi_i]\!],$$

where $\varphi_i = \varphi_K(\nu_i)$. It is straightforward to show that when n tends to infinity this outer approximation converges with respect to the Hausdorff distance to the set K. This discretisation has been

used in [7] to solve numerically different optimisation's problems where convex bodies are involved. The key idea is to start with a given convex polytope and to adjust the parameters φ_i in order to minimise the cost functional.

As it has been noticed in the introduction, a width constraint can be written in terms of the support function. Namely, $w_K(\nu) = 1$ is equivalent by definition to $\varphi_K(\nu) + \varphi_K(-\nu) = 1$. A simple idea would be to reproduce the method of [7] adding linear constraints to the parameters φ_i such that

$$\varphi_i^+ + \varphi_i^- = 1,$$

where $\varphi_i^\pm$ are the parameters associated to the normal vectors $\pm\nu_i$. Here is the crucial difficulty: the latter statement on the parameters $\varphi_i^\pm$ is not equivalent to $\varphi_K(\nu_i) + \varphi_K(-\nu_i) = 1$. It may happen for instance that

$$\bigcap_{i=1}^{n-1} [\![\nu_i, \varphi_i]\!] \subset \{x \cdot \nu_n \leq \varphi_n - \varepsilon\}$$

with $\varepsilon > 0$. In this case the hyperplane

$$\left\{x \in \mathbb{R}^3,\ x \cdot \nu_n = \varphi_n\right\}$$

is not anymore in a tangent position since it has an empty intersection with the body P_n. This difficulty turns the previous algorithm inefficient for this kind of constraint.

We present in this article two alternative methods to handle width constraints in geometrical optimisation. Those two discretisations of problems (1) and (2) leads to standard non-convex quadratic programming problems which are solved by classical solvers (see section IV.3.3).

IV.3 Minimisation among sets of constant width

In this section we are interested in the numerical study of Meissner's conjecture. Does Meissner's tetrahedron minimise the volume (or equivalently in dimension 3, the surface area) among sets of constant and fixed width (see [1] for a complete description of this convex body) ? In order to be able to eventually contradict the conjecture we have to propose a discrete description of constant width bodies which is an exact sub-problem of (1). More precisely, we would like to restrict our optimisation procedure to a subset of $\mathcal{K}$. Moreover, we would like to be able to evaluate exactly (up to round off errors) its surface area in order to compare our results and Meissner's conjecture.

We first recall a functional parametrisation result of constant width bodies obtained in [1]. Based on this parametrisation, problem (1) becomes a more classical optimisation problem on some convex space functional. Then, in order to approximate an optimal function we follow an approach introduced in [6] based on tensor-product splines. We stress the point on the fact that our method gives at the end of the process a discrete description (based on the cubic splines parametrisation) of some real constant width body of $\mathcal{K}$. Based on the previous formulation, our optimisation problem becomes a large scale quadratic optimisation problem. Finally, some numerical results are presented.

IV.3.1 Parametrisation by the median surface

A major difficulty to handle the constant width constraint is the potential irregularity of those bodies. As it is suggested by the 2 dimensional case, we have to consider shapes which may have

singularities (consider for instance Reuleaux's triangle which solves the question we are interested in, in dimension 2).

A framework designed to parametrise those kind of potentially irregular shapes is presented in [1]. We recall here the main results related to this parametrisation which will be useful to describe our optimisation approach.

First, we recall from [1] that constant width sets can all be described by vector fields on the sphere which satisfy the following global conditions :

Theorem 14 *Let $\alpha > 0$ be given and $M : \mathbb{S}^2 \to \mathbb{R}^3$ be a continuous application satisfying*

$$\forall \nu \in \mathbb{S}^2, \qquad M(-\nu) = M(\nu); \tag{3}$$

$$\forall \nu_0, \nu_1 \in \mathbb{S}^2, \qquad (M(\nu_1) - M(\nu_0)) \cdot \nu_0 \leq \frac{\alpha}{4} |\nu_1 - \nu_0|^2 . \tag{4}$$

Define a subset $K \subset \mathbb{R}^n$ as follows:

$$K := \left\{ M(\nu) + t\nu \, ; \, \nu \in \mathbb{S}^2, t \in \left[0, \frac{\alpha}{2}\right] \right\}. \tag{5}$$

Then K is a convex body of constant width α.

Conversely, any convex body of constant width α can be described by (5) *with some vector field M satisfying* (3) *and* (4)*.*

Next, we recall that the previous vector fields M on $\mathbb{S}^2$ can be parametrised by some smooth scalar functions satisfying second order differential conditions.

To this purpose, consider a parametrisation of the sphere $(u, v) \in \Omega \mapsto \nu(u, v) \in \mathbb{S}^2$, where Ω is some subset of $\mathbb{R}^2$. We assume that this parametrisation is *isothermal*, that is, satisfies for all $(u, v) \in \Omega$:

$$\partial_u \nu(u, v) \cdot \partial_v \nu(u, v) = 0 \quad \text{and} \quad |\partial_u \nu(u, v)| = |\partial_v \nu(u, v)| =: \frac{1}{\lambda(u, v)}. \tag{6}$$

Let K a body of constant width, then there exists a C^1 map $h : \Omega \to \mathbb{R}$ such that

$$M(\nu) = \mathcal{M}(\nu(u, v)) = h\, \nu + \lambda \partial_u h\, \nu_u + \lambda \partial_v h\, \nu_v \tag{7}$$

for all $(u, v) \in \Omega$, where M is a vector field associated to K defined by theorem 14.

Conversely, sets of constant width are all described analytically with additional constraints on the previous function h. In order to present those conditions, we recall the two definitions:

Definition 3 *We shall say that $D^2 h(u, v) \leq A = (a_{i,j})$* in a generalised sense*, if the following occurs:*

$$\limsup_{(\xi,\eta)\to(0,0)} \frac{T[h](u, v; \xi, \eta) - \frac{1}{2}(a_{11}\xi^2 + 2a_{12}\xi\eta + a_{22}\eta^2)}{\xi^2 + \eta^2} \leq 0. \tag{8}$$

where

$$T[h](u, v; \xi, \eta) := h(u + \xi, v + \eta) - h(u, v) - \xi \partial_u h(u, v) - \eta \partial_v h(u, v)$$

Similarly we say that $D^2 h(u, v) \geq A$ in a generalised sense*, if a similar property holds with a limit-inf ≥ 0 instead.*

Notice that in the regular case (that is h of class C^2), the inequality $D^2h(u,v) \leq A$ is equivalent to the standard positiveness of the matrix $A - D^2h(u,v)$.

Definition 4 *Given an isothermal parametrisation $\nu : \Omega \to \mathbb{S}^2$ of the sphere, let σ be its antipodal symmetry. Let $\mathrm{C}^{1,1}_{\sigma}(\Omega)$ be the set of all $\mathrm{C}^{1,1}$ maps $h : \Omega \to \mathbb{R}$ such that*

$$h \circ \sigma = -h. \tag{9}$$

Let $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ be the subset of functions $h \in \mathrm{C}^{1,1}_{\sigma}(\Omega)$ satisfying everywhere on Ω in a generalised sense (see Definition 3 above):

$$-\frac{\alpha}{2\lambda^2}\,\mathrm{Id} \leq U[h] \leq \frac{\alpha}{2\lambda^2}\,\mathrm{Id} \tag{10}$$

where

$$U[h] := D^2h + \lambda^{-2}h\,\mathrm{Id} + \lambda^{-1}\nabla\lambda \otimes \nabla h - \lambda^{-1}\nabla^{\perp}\lambda \otimes \nabla^{\perp}h. \tag{11}$$

We can now recall the main result obtain in [1] to describe constant width bodies. Then we have the characterisation of constant width set in terms of their support function:

Theorem 15 *Given an isothermal parametrisation of the sphere, let $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ be given by the Definition 4 above. Then an application $M : \mathbb{S}^2 \to \mathbb{R}^3$ is the median surface of a constant width body (that is corresponds to a constant width body by 5) if and only if there exists $h \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$ such that $M(\nu) = \mathcal{M}(h)(u,v)$ for all $\nu = \nu(u,v)$, where the map $\mathcal{M}(h) : \Omega \to \mathbb{R}^3$ is defined by (7). In this case, the map $\mathcal{M}(h + \frac{\alpha}{2}) : \Omega \to \mathbb{R}^3$ describes all but a finite number of the points on ∂K.*

IV.3.2 Discretisation of $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$

Based on theorem 15, the discretisation of our optimisation problem can be reduced to the discretisation of the space functional $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. We follow an approach introduced in [6] based on tensor-product splines to obtain splines which satisfy exactly (and not approximately) the differential constraints (10).

In the following we will use the standard isothermal parametrisation of the sphere ν:

$$(u,v) \in \Omega \longmapsto \left(\frac{\cos u}{\cosh v}, \frac{\sin u}{\cosh v}, \tanh v\right) \tag{12}$$

where $\Omega = [-\pi,\pi] \times \mathbb{R}$, $\lambda(u,v) = \cosh v$, and $\nu(\Omega) = \mathbb{S}^2 \setminus \{(0,0,\pm 1)\}$.

The starting point of our approach is to discretise the space of parameters $[-\pi,\pi] \times [0, v_{max}]$ by a bounded regular orthogonal grid where v_{max} is a parameter of the method. In order to satisfy exactly the antipodal symmetry constraint (9), we impose to the grid to contain the origin. Consider now a tensor-product spline h_d defined on that grid. We want to find sufficient conditions on the coefficients of h_d which ensure that $h_d \in \mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. Since the final goal of the discretisation is to achieve an optimisation procedure, we want the constraints on the coefficients of h_d to be linear.

Notice first that the periodicity and the antipodal symmetry constraint (9) are equivalent to linear equality constraints on those coefficients. The most challenging problem is to manage the constraints (10) in a linear way. We will describe how to deduce a set of linear inequality constraints

which is asymptotically equivalent to those conditions. For simplicity we restrict our description to the differential inequality

$$0 \leq U[h_d] + \frac{\alpha}{2\lambda^2}\,\mathrm{Id}. \tag{13}$$

In [6], the author describes how to obtain a set of linear inequality which ensures that the tensor-product spline to be a convex function. In that sense it is an interior approximation of the convexity constraint. Moreover, it is also proved that any strictly convex patch satisfies this kind of constraints for a suitable choice of the set of constraints. Due to the weight λ which appears in (10), we need to adapt the method to the space $\mathrm{C}^{1,1}_{\sigma,\alpha}(\Omega)$. Let us first describe more precisely the constraint (13) on a patch of the tensor-product spline assuming for simplicity $\alpha = 1$:

$$\left(\begin{array}{c|c} \partial_{uu}h_d + \dfrac{h_d + 1/2}{\lambda^2} - \dfrac{\sinh(v)\partial_v h_d}{\lambda} & \partial_{uv}h_d + \dfrac{\sinh(v)\partial_u h_d}{\lambda} \\ \hline \partial_{uv}h_d + \dfrac{\sinh(v)\partial_u h_d}{\lambda} & \partial_{vv}h_d + \dfrac{h_d + 1/2}{\lambda^2} + \dfrac{\sinh(v)\partial_v h_d}{\lambda} \end{array}\right) \geq 0 \tag{14}$$

The key point is to remark that the previous matrix may be rewritten only in terms of $\tanh(v)$ and $\tanh^2(v)$ by the standard formula $1/\lambda^2 = 1 - \tanh^2(v)$. Regarding $Y := \tanh(v)$ as a new parameter and using the approach of [6], we can force the differential constraints by a set of linear inequalities imposed on the coefficients of the cubic spline. We do not recall here all the technical description of those inequalities but we illustrate the principle of the method in the following to avoid a continuous set of linear constraints depending on Y. If one consider one of the inequalities provided by [6] regarding Y as a parameter, it is straightforward to observe that is has the form

$$Y^2 l_1 + Y l_2 + l_3 \leq 0 \tag{15}$$

where l_1, l_2 and l_3 are affine forms of the Bernstein/Bezier coefficients of the cubic polynomial h_d on the patch which is considered. Notice that we want (15) to be satisfied for all $Y \in [\tanh(v_1), \tanh(v_2)]$ for some $v_1 < v_2$ depending on the patch. In order to reduce this set of constraints to a finite number of inequalities we use the same strategy as in [6]. Consider the polynomial of two variables

$$p(x, y) = xy l_1 + \frac{x + y}{2} l_2 + l_3. \tag{16}$$

Let $\Sigma = (\sigma_0 \ldots, \sigma_Q)$ be a strictly increasing sequence satisfying

$$\sigma_0 = v_1 < \cdots < \sigma_Q = v_2$$

for some integer $Q > 1$. We define a new set of inequalities

$$\mathcal{I}(l_1, l_2, l_3) = \{\, p(v_1, v_1) \leq 0,\ \ p(v_2, v_2) \leq 0, \text{ and } p(\sigma_i, \sigma_{i+1}) \leq 0\ \ \forall i = 0 \ldots Q - 1\}\,. \tag{17}$$

Following the proof of the Lemma 1 of [6] we obtain:

Lemma 9 *Let $(l_1, l_2, l_3) \in \mathbb{R}^3$, $v_{max} > 0$ and $Q \in \mathbb{N}^*$. Suppose that (l_1, l_2, l_3) satisfies a set of constraints of type (17) for some increasing sequence*

$$\sigma_0 = v_1 < \cdots < \sigma_Q = v_2.$$

Then (l_1, l_2, l_3) satisfies (15) for all $Y \in [v_1, v_2]$.

Proof. First observe that $p(\sigma_i, \sigma_i) \leq 0, \ \forall i = 0 \dots Q$. If $i = 0, Q$ this is a consequence of the definition of $\mathcal{I}(l_1, l_2, l_3)$. For $0 < i < Q$, we have

$$p(\sigma_i, \sigma_i) = \frac{\sigma_{i+1} - \sigma_i}{\sigma_{i+1} - \sigma_{i-1}} p(\sigma_i, \sigma_{i-1}) + \frac{\sigma_i - \sigma_{i-1}}{\sigma_{i+1} - \sigma_{i-1}} p(\sigma_i, \sigma_{i+1}) \leq 0$$

since both coefficients are positive. Now let $Y \in [v_1, v_2]$, it exists i such that $Y \in [\sigma_i, \sigma_{i+1}]$. In the same way as before we have

$$\begin{aligned} p(Y, Y) &= p(\frac{Y - \sigma_i}{\sigma_{i+1} - \sigma_i} \sigma_{i+1} + \frac{\sigma_{i+1} - Y}{\sigma_{i+1} - \sigma_i} \sigma_i, Y) \\ &= \frac{Y - \sigma_i}{\sigma_{i+1} - \sigma_i} p(\sigma_{i+1}, Y) + \frac{\sigma_{i+1} - Y}{\sigma_{i+1} - \sigma_i} p(\sigma_i, Y) \\ &= \left(\frac{Y - \sigma_i}{\sigma_{i+1} - \sigma_i} \right)^2 p(\sigma_{i+1}, \sigma_{i+1}) + \left(\frac{\sigma_{i+1} - Y}{\sigma_{i+1} - \sigma_i} \right)^2 p(\sigma_i, \sigma_i) \\ &\quad + 2 \frac{(Y - \sigma_i)(\sigma_{i+1} - Y)}{(\sigma_{i+1} - \sigma_i)^2} p(\sigma_{i+1}, \sigma_i). \end{aligned} \tag{18}$$

Since $p(Y, Y)$ is equal to $Y^2 l_1 + Y l_2 + l_3$ by definition, the inequality follows for all $Y \in [v_1, v_2]$. □

By this lemma we are able, up to the introduction of the new parameter Σ, to describe a set of linear constraints on the coefficients of the cubic spline which ensure that the body associated to h_d by (7) is of constant width.

To conclude the description of our optimisation approach, we recall from [1] that the surface area $|\partial K|$ of a body of constant width defined by its support function h can be evaluated by the formula:

$$|\partial K| = \int_\Omega \left(\lambda^{-2} h^2 - \frac{1}{2} |\nabla h|^2 \right) + \pi \alpha^2, \tag{19}$$

By the last equality, the surface area associated to the constant width body defined by h_d is a quadratic form of its Berstein/Bezier coefficients. This observation complete the description of our internal approximation of problem (1) as a large scale quadratic problem. We describe below the numerical optimal conditions which have been used to solve that quadratic problem.

IV.3.3 Numerical results

All the discrete constraints that have been considered up to now are the discretisation of local constraints. This matter of fact has a crucial impact on the complexity of the discrete optimisation problem since we have to deal only with sparse constraints.

Computer implementation

In order to take benefit of this sparsity we used the efficient large scale optimisation software LANCELOT of the GALAHAD library developed by N. Gould D.Orban and P. Toint (see [4] and [3]).

Let us describe more precisely the local optimality conditions that the LANCELOT module tries to reach. As explained in [3], a general nonlinear constrained optimisation problem can be reformulated in the form:

$$\min_{x\in\mathbb{R}^N} f(x)$$

where x is subject to the equality constraints

$$c_j(x) = 0 \quad 1 \le j \le m,$$

and the simple bounds

$$l_i \le x_i \le u_i, \quad 1 \le j \le N.$$

The algorithm implemented in LANCELOT is based on an Augmented Lagrangian Method. At each step, an approximate minimiser of the augmented Lagrangian function

$$\Phi(x, \lambda, S, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i c_i(x) + \frac{1}{2\nu} \sum_{i=1}^{m} s_{ii} c_i(x)^2$$

is found (the parameter ν and the factors s_{ii} are adjusted by the program). Let P be the projection operator on the bound constraints, namely:

$$P(x, l, u)_i = \begin{cases} l_i & \text{if} \quad x_i < l_i \\ u_i & \text{if} \quad x_i > u_i \\ x_i & \text{otherwise.} \end{cases}$$

The algorithm stops when the two conditions

$$\|x - P(x - \nabla_x L(x, \lambda), l, u)\|_\infty \le \varepsilon_l \tag{20}$$

and

$$\|c(x)\|_\infty \le \varepsilon_c \tag{21}$$

where ε_l and ε_c are precision factors which are prescribed by the user.

Our results

We present in the following figures the results of our approach. The first point we are interested in is to check the local optimality of Meissner's tetrahedron. We then compute analytically the h function (see [1] for the complete expression) which defines this body and project h on the grid we are working on. This set of values is the starting point of our numerical process. We present in figure IV.1 the starting h function and Meissner's body.

Actually, it has not been possible to distinguish the initial shape and the result produced by the optimisation: Meissner's body is, at least in a numerical way, a local minimiser of the surface area among constant width bodies. We give in table IV.1 numerical details of the precision reached with Meissner's tetrahedron as initial guess on different grids.

The same experiment has been carried out starting from the Reuleaux's rotated triangle. This body is known to be the body of least surface area among constant width body of revolution. By this experiment we wanted to study the optimality of that body in the larger class of sets of constant width. As it is reported in figures IV.2 and IV.3, this body is not numerically speaking a

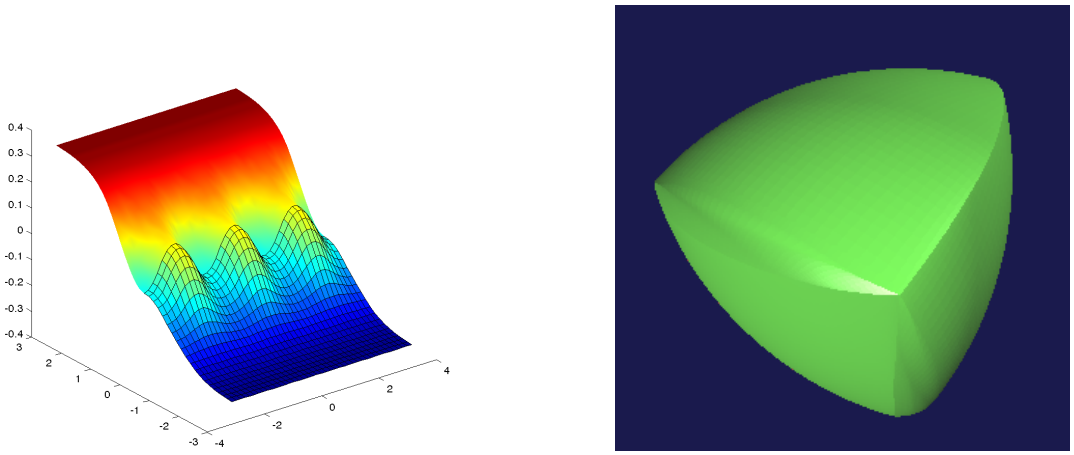

Figure IV.1: Meissner's h function and the associated body

Nb of variables (with the gap variables)	Nb of active bounds	Constraints in $\|.\|_\infty$	Projected gradient in $\|.\|_\infty$
3772	965	2.6509E-04	6.8437E-04
5967	1692	7.8426E-04	7.8741E-04
8662	2495	9.6875E-04	5.3467E-04

Table IV.1: Precision obtained with the grids 41x20, 51x25, 61x30

local optima: the critical shape that has been found seems to be build on a Reuleaux pentagon by the process described in [8]. To conclude, as reported in the introduction, notice that the shape of figure IV.2 satisfies the weak optimality condition which has been described in [1]: for any $\omega \in \mathbb{S}^2$ small enough, the part of the boundary of that body which normals are ω and the opposite part which normals are $-\omega$ are not both regular.

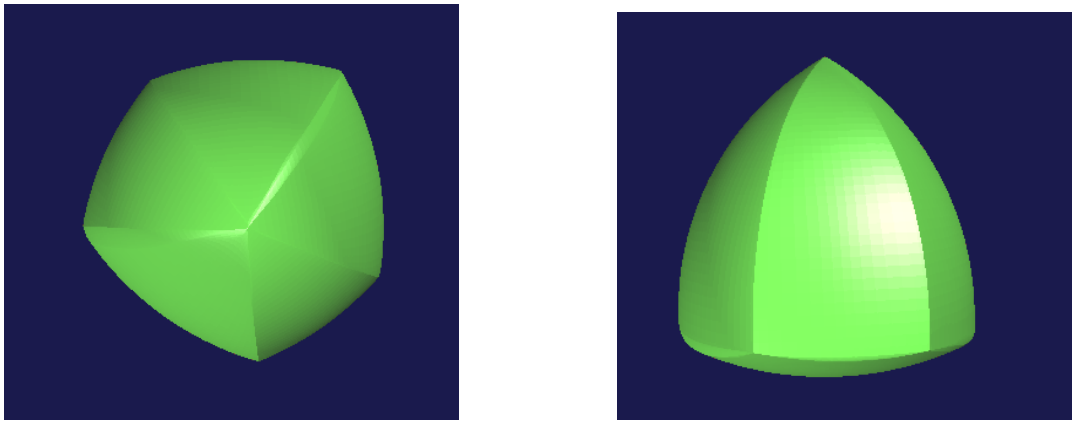

Figure IV.2: Result of the local optimisation of reuleaux rotated body

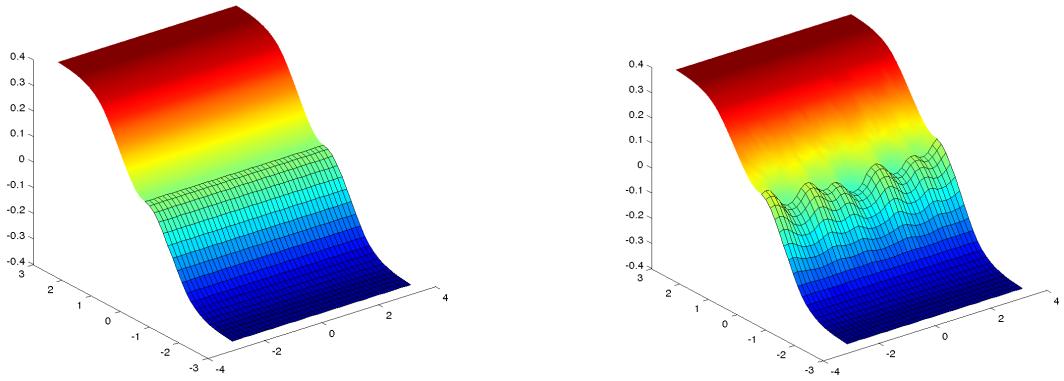

Figure IV.3: Initial and final h support functions

IV.4 Minkowski sums: an algebraic discretisation for inequality constraints

In the following we are interested in the approximation of an optimal solution of the following problem:

$$\min_{K \in \mathcal{K}} S_K, \tag{22}$$

where $\mathcal{K} = \{K \subset \mathbb{R}^3, \quad \text{convex}, \quad w_K(v) \geq 1, \quad \forall v \in \mathbb{S}^2\}$ and S_K stands for the surface area of the body K. Before introducing our approach, let us first recall some basic facts on Minkowski's sum of two sets $A, B \subset \mathbb{R}^3$. We define Minkowski's sum of sets A and B by

$$A + B = \{x + y, (x, y) \in A \times B\}.$$

An interesting feature related to the width of a convex set and Minkowski's sum is its almost linear behaviour. Let $\lambda, \mu \in \mathbb{R}_+^*$, A and B two convex sets of $\mathbb{R}^N$, then $\lambda A + \mu B$ is convex and its support function is given by

$$\varphi_{\lambda A + \mu B} = \lambda \varphi_A + \mu \varphi_B. \tag{23}$$

When A and B are subsets of $\mathbb{R}^3$ with nonempty interior, the surface area of the resulting body $S_{\lambda A + \mu B}$ is deduced by the formula

$$S_{\lambda A + \mu B} = \lambda^2 S_A + \mu^2 S_B + \lambda\mu(S_{A+B} - S_A - S_B). \tag{24}$$

We refer to [13] or [2] for the proof of the previous equality and many other results on convex bodies.

IV.4.1 Outline of the algorithm

Equations (23) and (24) are the starting points of our first method. Let $(K_i)_{i \in I}$ be a finite family of convex sets of $\mathbb{R}^3$ which contain the origin. Consider the approximation of $\mathcal{K}$ obtained by the cone

$$C_I = \left\{ \sum_{i \in I} \lambda_i K_i, \lambda_i \in \mathbb{R}^+ \right\} \tag{25}$$

where the positive vector $\lambda = (\lambda_1, \dots)$ is restricted to the subset of vectors which satisfy $\sum_{i\in I} \lambda_i K_i \in \mathcal{K}$. By relation (23), the constraint $\sum_{i\in I} \lambda_i K_i \in \mathcal{K}$ is equivalent to impose inequality constraints depending on polytopes (K_j) to the coefficients (λ_j). That is

$$\sum_{i\in I} \lambda_i \varphi_{K_i}(v) \geq 1 \ \forall v \in \mathbb{S}^2 \tag{26}$$

It is classical that the convex polytope $\sum_{i\in I} \lambda_i K_i$ may have a huge number of vertices. Thus it is not possible to impose exactly the previous constraints. Then, we approximate (26) by a naive discretisation of $\mathbb{S}^2$: let $v_1, \dots v_m$ be m randomly chosen vectors of the sphere. We consider the finite set of constraints:

$$\varphi_{\sum_i \lambda_i K_i}(v_k) + \varphi_{\sum_i \lambda_i K_i}(-v_k) \geq 1, \quad k = 1, \dots, m$$

which are equivalent thanks to (23) to

$$\sum_i (b^+_{ik} + b^-_{ik})\lambda_i \geq 1, \quad k = 1, \dots, m \tag{27}$$

where $b^\pm_{ik} = \varphi_{K_i}(\pm v_k)$. Thus, solutions of the sub-problem

$$\min_{K\in C_I} S_K, \tag{28}$$

may be approximated by the solutions of the quadratic program:

$$\min_\lambda \sum_{i,j} a_{ij}\lambda_i\lambda_j, \tag{29}$$

for vectors λ which satisfy (27). Moreover according to (24), the coefficients a_{ij} can be explicitly estimated by the relations

$$\begin{cases} a_{ij} = & \frac{1}{2}(S_{K_i+K_j} - S_{K_i} - S_{K_j}) & i \neq j, \\ a_{ij} = & S_{K_i} & i = j. \end{cases}$$

At this step one main difficulty remains. How do we choose the family (K_i) in an effective way in order to get a reasonable approximation of $\mathcal{K}$?

IV.4.2 The cone C_I

The algorithm

Since it is difficult, to estimate numerically quantities like $S_{K_i+K_j}$ when the convex bodies K_i or K_j have a great number of vertices, we would like to be able to approximate a convex body as a Minkowski's sum of "simple" polytopes. Whereas it is true in $\mathbb{R}^2$ that every convex polytope can be decomposed as a finite sum of triangles and segments, the situation is dramatically more complex in dimension 3. Actually, any generic convex polytope (that is a polytope every 2-faces of which are triangles) is indecomposable (see [14]). Thus, if we want to generate a sequence of bodies (K_i) whose associated cone (25) converges to $\mathcal{K}$, we do not have to restrict ourselves to simplices. We propose the following iterative process to handle this difficulty.

Fix l the maximum number of extremal points of an element of the sequence (K^0_i) and n the number of elements of this family. Let $\varepsilon > 0$ be a precision parameter and j_{max} the maximum number of iterations.

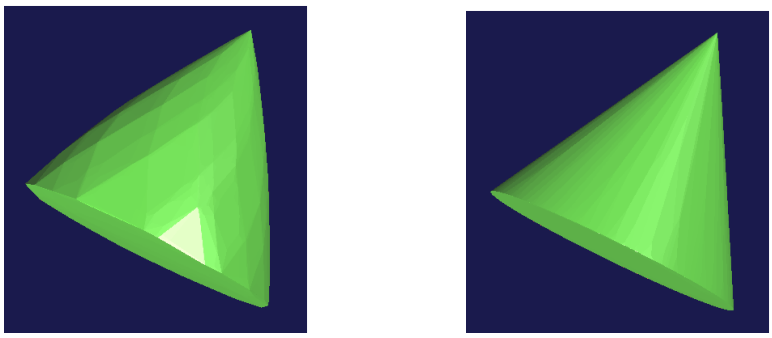

Figure IV.4: Approximation of a cone by Minkowski sums

0. Set $j = 0$. Choose randomly n convex polytopes (K_i^0) with at most l extremal points.

1. Solve the optimisation problem (29) associated to the family (K_i^j).

2. Let I^j be the subset of indices of the optimal vector λ^0 whose components are greater in absolute value than ε. Construct a new family (K_i^{j+1}) keeping the bodies $(K_i^j)_{i\in I^j}$ and choosing randomly the others.

3. Let $j \leftarrow j + 1$. If $j > j_{max}$ or $I = I^j$ stop, otherwise go to 1.

Of course the parameter j_{max} has to be adjusted in relation with the CPU time needed for solving the optimisation step 1. There is no simple way to give convergence estimates with respect to the choice of the parameters n, l and ε. We propose hereafter one simple numerical experiment to adjust those parameters.

Numerical tests

In order to choose relevant values for parameters n, l and ε, we test our discretisation by Minkowski's sum to approximate a truncated cone. The cone seems to us difficult to approximate by sum of random polytopes since its normal directions cover only a subset of $\mathbb{S}^2$ of dimension 1. To measure the quality of the approximation we introduce a cost functional based on the Euclidian distance between support functions. Consider one truncated cone K and its support function φ_K. In order to observe if our algorithm is able to generate a sequence K^j of bodies which converges to K we define the following cost function:

$$D_K(K^j) = \sum_k (\varphi_K(v_k) - \varphi_{K^j}(v_k))^2,$$

where (v_k) is a fixed list of arbitrary vectors of $\mathbf{S}^2$. Thanks to (23), the auxiliary optimisation problem that we solve at step 1. is the quadratic problem in λ:

$$\min_{\lambda\geq 0} \sum_k (\varphi_K(v_k) - \sum_i \lambda_i \varphi_{K_i^j}(v_k))^2.$$

We present in Figures IV.4 the results we obtained for K equal to a regular cone. The values that have been used to obtain this approximation, are $\#I = 100$, $\varepsilon = 10^{-6}$, $j_{max} = 10^5$, $l = 10$ and 10^4 normal vectors v_k. Notice that a large number of iterations are required in order to get a satisfactory sequence of bodies. This constraint requires an efficient and fast solver for the optimisation step.

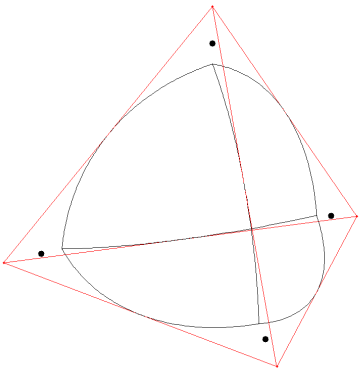

Figure IV.5: The body of E. Heil

IV.4.3 The relaxed problem and the conjecture of E. Heil

In this section we apply the previous method to a more realistic situation which was addressed by E. Heil in [5] p. 261. We look for a solution of

$$\min_{K\in\mathcal{K}} S_K, \tag{30}$$

where $\mathcal{K} = \{K \subset \mathbb{R}^3, \quad \text{convex}, \quad w_K(v) \geq 1, \quad \forall v \in \mathbb{S}^2\}$.

As it has been explained, the latter problem can be approximated by a sequence of quadratic problems. Exactly the same method applied on the problem of minimising the volume would lead to solve a sequence of cubic problems. Up to now, there is no efficient way to solve numerically dense and large cubic problems which makes our method irrelevant in this situation.

E. Heil propose the following construction of its optimal body: Consider the regular tetrahedron of edge-length 1 and replace each edge by a circular arc of radius $\sqrt{2}/2$ and center in the middle of the opposite edge. Take the four points of distance $\sqrt{2}/2$ to the facets of the tetrahedron which are on the line between a vertex and the center of the opposite facet of the tetrahedron. E. Heil claims that the convex hull of the previous 4 points and 6 arcs is a set of width greater or equal than $\sqrt{2}/2$ (see Figure IV.5). Moreover, he observed that its volume and its surface are smaller than the ones of standard convex shapes of same minimal width (such the regular tetrahedron, the circular cone, the ball, Meissner's tetrahedron and Reuleaux's tetrahedron). Does this set minimise the volume and the surface among convex bodies of fixed minimal width ?

Due to the approximation made by (29), our algorithm does not always provide us a polytope which is precisely a member of $\mathcal{K}$. The resulting body satisfies only the width constraint at a discrete level. Notice that for a polytope,

$$\Delta = \min_{v\in\mathbb{S}^2} w_K(v),$$

is equal to the finite number of conditions

$$\min_{v_k} w_K(v_k) \tag{31}$$

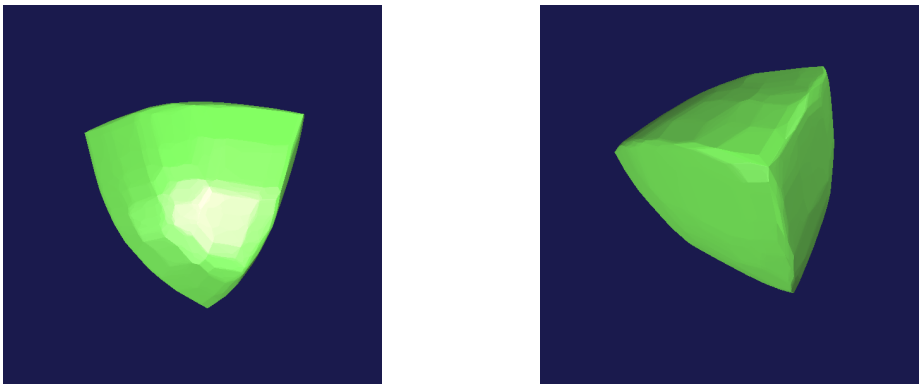

Figure IV.6: Approximation of the body of fixed minimal width with smallest surface area

where ν_k are the normal vectors of the polytope K. In order to get an element of $\mathcal{K}$, we apply the following post-processing: starting from the result of our optimisation process, we first compute its normal vectors. Then thanks to (23) and (31), the polytope defined by $\frac{1}{\Delta}P$ is in $\mathcal{K}$.

We present in figure IV.6 two different views of the resulting body. Hereafter are the values related to surfaces area and volumes of the our optimal shape and the body of E. Heil (for a minimal width equal to 1):

	Surface area	Volume
E. Heil body	2.9306	0.2983
Computed shape	2.9249	0.3862

Notice that the polytope generated by our algorithm has a significantly smaller surface area than the shape proposed by E. Heil but has a greater volume.

Bibliography

[1] T. Bayen, T. Lachand-Robert, and É. Oudet. Analytic parametrization of three-dimensional bodies of constant width. *Arch. Ration. Mech. Anal.*, 186(2):225–249, 2007.

[2] Marcel Berger. *Géométrie. Vol. 3*. CEDIC, Paris, 1977. Convexes et polytopes, polyèdres réguliers, aires et volumes. [Convexes and polytopes, regular polyhedra, areas and volumes].

[3] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *LANCELOT*, volume 17 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1992. A Fortran package for large-scale nonlinear optimization (release A).

[4] Nicholas I. M. Gould, Dominique Orban, and Philippe L. Toint. GALAHAD, a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization. *ACM Trans. Math. Software*, 29(4):353–372, 2003.

[5] P. M. Gruber and R. Schneider. Problems in geometric convexity. In *Contributions to geometry (Proc. Geom. Sympos., Siegen, 1978)*, pages 255–278. Birkhäuser, Basel, 1979.

[6] Bert Jüttler. Surface fitting using convex tensor-product splines. *J. Comput. Appl. Math.*, 84(1):23–44, 1997.

[7] Thomas Lachand-Robert and Édouard Oudet. Minimizing within convex bodies using a convex hull method. *SIAM J. Optim.*, 16(2):368–379 (electronic), 2005.

[8] Thomas Lachand-Robert and Édouard Oudet. Bodies of constant width in arbitrary dimension. *Math. Nachr.*, 280(7):740–750, 2007.

[9] E Meissner. über die anwendung von fourierreihen auf einige aufgaben der geometrie und kinematik. *Vierteljahresschr. Naturfor. Ges. Zürich.*, 54:309–329, 1909.

[10] E Meissner. über punktmengen konstanter breite. *Vierteljahresschr. Naturfor. Ges. Zürich.*, 56:42–50, 1911.

[11] E Meissner. über punktmengen konstanter breitedrei gipsmodelle von flächen konstanter breite. *Zeitschrift der Mathematik und Physik*, 60:92–94, 1912.

[12] E Meissner. über die durch regulare polyeder nicht stautzbaren körper. *Vierteljahresschr. Naturfor. Ges. Zürich.*, 63:544–551, 1918.

[13] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[14] G. C. Shephard. Decomposable convex polyhedra. *Mathematika*, 10:89–95, 1963.

PART 1.

Deuxième partie

Optimisation de forme à plusieurs phases

Local minimizers of functionals with multiple volume constraints

Édouard Oudet & M . O. Rieger

V.1 Introduction

Let Ω be a bounded open set in $\mathbb{R}^n$. The general form of a variational problem on Ω with two level set constraints is given by the minimization of

$$\begin{aligned} \text{Minimize } E(u) := \int_\Omega f(x, u(x), \nabla u(x))\, dx, \\ |\{x \in \Omega,\ u(x) = a\}| = \alpha, \\ |\{x \in \Omega,\ u(x) = b\}| = \beta, \end{aligned} \tag{1}$$

where $u \in H^1(\Omega)$ and $\alpha, \beta > 0$, $\alpha + \beta < |\Omega|$. Problems of this class have been encountered in the context of immissible fluids [8] and mixtures of micromagnetic materials [1]. The difficulty of such problems is the special structure of their constraints: A sequence of functions satisfying these constraints can have a limit which fails to satisfy the constraints.
Such minimization problems but with only one volume constraint have been studied by various authors, see e.g. [2]. Problems with two or more constraints have a very different nature than problems with only one volume constraint: In the case of one volume constraint, only additional boundary conditions or the design of the energy can induce transitions of the solution between different values. Two or more volume constraints, on the other hand, force transitions of the solution by their very nature. Such problems have been studied starting from the fundamental work by Ambrosio, Marcellini, Fonseca and Tartar [3]. Their results have been generalized by various authors, compare e.g., [11, 10, 15]. It turned out that existence can only be guaranteed for functions f satisfying quite specific conditions, and that there are easy examples of nonexistence, e.g. if $n = 1$, $f(x, u, u') = |u'|^2 + |u|$ and $|\Omega| - \alpha - \beta$ sufficiently large [10]. Whereas the one dimensional case by now is relatively well understood (compare [10, 15]), there are few sharp results on existence in the higher dimensional case [16]. There are in addition some results on local minimizers in the one-dimensional case [10], but there were so far no rigorous results in the higher dimensional case. By computing the shape derivative of the functional it is, however, possible to give a necessary condition for minimizers, as has been done in [3]:

Theorem 16 *Let $u \in W^{1,2}(\Omega, [0, 1])$ be a solution of (1). Assume that $S := \partial\{u = 0\} \cap \Omega$ is C^1, then $\frac{\partial u}{\partial n}$ is locally constant on S.*

There is also very little known about explicit examples of minimizers in two dimensions, compare [3, 15].

In this article we are introducing a numerical method for the approximation of local minimizers of (1). We apply this method to various examples and obtain a first picture of the shape of local and global minimizers for some simple domains in $\mathbb{R}^2$. Guided by the numerical results, we prove rigorously that even on the unit square solutions are not depending continuously on the parameter α and β and illustrate this with numerical results. Moreover, we show that even on convex domains in $\mathbb{R}^2$ nontrivial local minimizers can exist.

V.2 Numerical approximations

V.2.1 General approach and level-set methods

We suppose in this section the existence of a solution of (1), i.e. that there exists a function $u \in H^1(\Omega)$ minimizing the problem (1). Our goal is to find a numerical method for the computation of this solution.

We will first explain our ideas in the simplest situation where $f(x, u(x), \nabla u(x)) = |\nabla u(x)|^2$. In this situation existence of a solution for problem (1) has been already found in [3]. Our approach is based on the following fact: Let u^* be an optimal function for the problem, and denote

$$\Omega_a = \{x \in \Omega,\ u^*(x) = a\}, \quad \Omega_b = \{x \in \Omega,\ u^*(x) = b\}.$$

Ω_a and Ω_b are closed sets, since u is Hölder continuous, for a proof see [11, Theorem 3.3]. Then, it is possible to reconstruct u^* by solving the elliptic boundary value problem:

$$\begin{cases} \Delta u = 0, & \text{in} \quad \Omega\backslash(\Omega_a \cup \Omega_b), \\ u = \alpha & \text{on} \quad \partial\Omega_a, \\ u = \beta & \text{on} \quad \partial\Omega_b, \\ \frac{\partial u}{\partial n} = 0 & \text{on} \quad \partial\Omega\backslash(\Omega_a \cup \Omega_b). \end{cases} \tag{2}$$

The numerical approximation of an optimal function u^* is hence reduced to an optimization problem for the two sets Ω_a and Ω_b. Unfortunately, very few results are known concerning the optimal sets Ω_a and Ω_b. In particular, it is not possible to restrict the optimization process to connected sets since disconnected sets can be optimal. We propose below an approach based on level set methods which makes it possible to generate also disconnected sets.

Before this, we recall briefly the standard tools of level set methods in a simplified context where only one single shape is unknown (see for instance [12] for numerical details closely related to our approach). We explain later how to deal with more than one unknown shape.

Let Ω be a subset of $\mathbb{R}^2$, we consider an optimization problem where we want to find an optimal set $O \subset \Omega$ for a given functional. The main idea of the method is to parametrize O by a function Φ, the so-called *level set function*, that satisfies

$$\begin{cases} \Phi(x) < 0 & \text{if } x \in O, \\ \Phi(x) > 0 & \text{if } x \in \Omega\backslash\overline{O}, \\ \Phi(x) = 0 & \text{if } x \in \partial O. \end{cases}$$

For numerical convenience which will be explain below, the level set function Φ is always defined on a cartesian grid defined on a square containaing the set Ω.

As suggested in [13], such a function will be initialized with the signed-distance which is given by

$$\begin{cases} \Phi(x) = & -\text{dist}(x, \partial O) & \text{if } x \in O, \\ \Phi(x) = & \text{dist}(x, \partial O) & \text{if } x \in \Omega \backslash O. \end{cases}$$

We remark that the constructed distance is generally not easy to compute. In our case, for the cartesian mesh on Ω, deduced by the cartesian grid where Φ is defined, we choose an approximate signed-distance function which is constant on each triangle of the mesh. Its value in the triangle T is computed by evaluating the distance between the center of mass of T and the center of mass of the closest triangle lying on the boundary of the initial shape.

Once Φ is defined, we can let its level set at 0 (i.e. ∂O) fluctuate with time under the vector field vn (where v is a real-valued function and n is the normal vector on ∂O). In other words, if $x(t)$ describes the evolution of a point on ∂O under such a transformation, it has to satisfy

$$\Phi(t, x(t)) = 0$$

for all t. Differentiating this expression, we obtain

$$\frac{\partial \Phi}{\partial t}(t, x(t)) + v(x(t)) n(x(t)) \cdot \nabla_x \Phi(t, x(t)) = 0. \tag{3}$$

Now the normal to a level set in a non-stationary point is given by

$$n(x(t)) = \frac{\nabla_x \Phi}{|\nabla_x \Phi|}(t, x(t)).$$

Hence, using (3), we derive

$$\frac{\partial \Phi}{\partial t}(t, x(t)) + v(x(t)) \, |\nabla_x \Phi| \, (t, x(t)) = 0. \tag{4}$$

In order to compute the evolution of Φ, we thus have to solve a Hamilton-Jacobi equation. We remark that the computation we have presented only concerns the level set 0, but since in practice the vector field vn has a natural extension on Ω, we solve the equation (4) in the whole set Ω.

We want to find a good velocity field vn for the shape optimization problem under investigation. Therefore we follow an approach which has been first introduced in [7] and choose vn as the vector field obtained by boundary variations. Let $O \subset \Omega$ be a connected set with C^2-boundary and u a solution of the problem

$$\begin{cases} \Delta u = 0, & \text{in} \quad \Omega \backslash O, \\ u = \alpha & \text{on} \quad \partial O, \\ \frac{\partial u}{\partial n} = 0 & \text{on} \quad \partial \Omega \backslash O. \end{cases} \tag{5}$$

It is well known in shape optimization (see for instance [9, 5, 17]) that the shape derivative of the energy of u in the direction of a vector field V localized around ∂O is given by Hadamard's formula

$$\frac{dE}{dV} = - \int_{\partial O} \left(\frac{\partial u}{\partial n} \right)^2 Vn \, d\sigma.$$

This computation suggests that the steepest descent direction is given by the normal vector field

$$-\left(\frac{\partial u}{\partial n}\right)^2 n.$$

Moreover, since u is by definition constant along ∂O this vector field has a natural extension to the domain Ω using the relation:

$$n = \pm\frac{\nabla\Phi}{|\nabla\Phi|}.$$

In order to avoid the computation of a new mesh at each iteration, we compute an approximation of the solution of (5) via a penalization method introduced in [14].

V.2.2 A multi-level set method

As explained before, the numerical approximation of (1) can be reduced to the approximation of the two sets $\{x \in \Omega,\ u(x) = a\}$ and $\{x \in \Omega,\ u(x) = b\}$. In that case, two shapes are unknown and we propose to parametrize those sets with two different level set functions, namely Φ_a and Φ_b. At each step of the algorithm the two sets evolve under the local vector field given by the shape derivative. The only point that we have to worry about is the possibility of crossing of those level sets. Several approaches have already been investigated for dealing with this kind of difficulty. The most standard way to avoid the crossing of the level sets is to add a penalization term like

$$\int_\Omega (H(\Phi_a(x)) + H(\Phi_b(x)) - 1)^+ \, dx = 0$$

to the functional , where $H(y)$ is equal to 1 for $y < 0$ and equal to 0 otherwise and $(y)^+$ stands for the positive part of y. Although we are not able to prove that the crossing of level sets will never happen during the optimization, we did not need to implement the previous method, since in our simulations, we never observed a crossing of level sets. This fact is probably a result of the fact that such crossing (or even touching) of the level sets cannot occur in the limit, i.e. for minimizers of (1) as the following theorem states:

Theorem 17 *Let u be a minimizer of (1). Then* $\mathrm{dist}\,(\{u = a\}, \{u = b\}) > 0$.

Proof. This is an immediate consequence of a regularity result by Mosconi and Tilli [11] that ensures that u is Hölder continuous. □

Of course, this idea can be extended to arbitrary numbers of level sets.

We now compute the solution of the above Hamilton-Jacobi equation. Our description will be limited to a simple algorithm reported in [13] designed to approach the weak viscosity solution of Hamilton-Jacobi equation problem. Let us consider the first order Cauchy system:

$$\begin{cases} \frac{\partial\Phi}{\partial t}(t,x) - F(x)\,|\nabla\Phi(t,x)| = 0 & \text{in } \mathbb{R}_+ \times D, \\ \Phi(0,x) = u_0(x) \text{ in } D, \end{cases}$$

where D is a bounded rectangle of $\mathbb{R}^2$ and u_0 and F are given functions. From now on we shall use the classical notations for finite difference schemes on regular meshes of points indexed by i, j. Starting from $\Phi(0,x) = u_0(x)$, then the evolution of Φ after one time step Δt is given by

$$\Phi_{ij}^{n+1} = \Phi_{ij}^n - \Delta t(\max(F_{ij},0)\nabla^+\Phi + \min(F_{ij},0)\nabla^-\Phi),$$

where

$$\nabla^+\Phi = \left[\max(D_{ij}^{-x}\Phi,0)^2 + \min(D_{ij}^{+x}\Phi,0)^2 + \max(D_{ij}^{-y}\Phi,0)^2 + \min(D_{ij}^{+y}\Phi,0)^2\right]^{1/2}$$

and

$$\nabla^-\Phi = \left[\max(D_{ij}^{+x}\Phi,0)^2 + \min(D_{ij}^{-x}\Phi,0)^2 + \max(D_{ij}^{+y}\Phi,0)^2 + \min(D_{ij}^{-y}\Phi,0)^2\right]^{1/2},$$

with

$$D_{ij}^{+x}\Phi = \frac{\Phi_{i+1,j} - \Phi_{i,j}}{\Delta x}$$

for a space step equal to Δx. The quantities $D_{ij}^{-x}\Phi$, $D_{ij}^{+y}\Phi$ and $D_{ij}^{-y}\Phi$ are easily deduced. Finally, to define completely our problem, we add the boundary condition

$$\frac{\partial\nabla\Phi(t,x)}{\partial n} = 0 \text{ on } \partial D.$$

The volume of the level set function Φ_a at the discrete level is by definition the volume of all the elements of the mesh where Φ_a is less or equal than zero. In order to preserve this volume equal to α along the iterations, we use the Lagrange multiplier technique reported in [12]. According to the derivative computed in (4), the level set function Φ_a satisfies the Hamilton-Jacobi equation

$$\frac{\partial\Phi_a}{\partial t}(t,x) - \left(-|\nabla u|^2(t,x) + \mu\right)|\nabla\Phi_a(t,x)| = 0 \text{ in } \mathbb{R}_+ \times D \tag{6}$$

where $u(t,.)$ is the solution of the system (2) associated to $\Phi_a(t,.)$ and $\Phi_b(t,.)$. As suggested by Osher and Santosa [12], at each iteration we adapt the Lagrange multiplier μ to preserve the volume constraint. The same projection method is of course reproduced for the level set function Φ_b, in case of two volume constraints.

It is now possible to describe all the steps of our algorithm:

1. Initialization of Φ_a and Φ_b by the signed distance on a cartesian grid containing Ω.

2. Computation of the velocity field by a penalization method introduced in [14] on the fixed triangular mesh deduced from the cartesian grid. Checking of an exit criterion.

3. Propagation of the level sets solving the Hamilton-Jacobi equations (6) preserving the volume constraints.

4. Evaluation of the cost function. If the cost decreases then go to step 5. Otherwise divide the time step by 1.5 and go to step 3.

5. Redefinition of Φ_a and Φ_b.

6. Eventually, reinitialization of Φ_a and Φ_b with the signed distance. Back to step 2.

For more details on the computation of the solution of the state equation associated to Φ_a and Φ_b (in the context of one level set constraint) see [7] or [14].

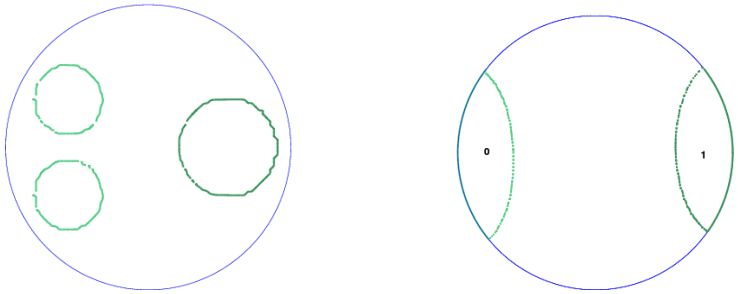

Figure V.1: Initial and optimized level sets for a problem with two constraints

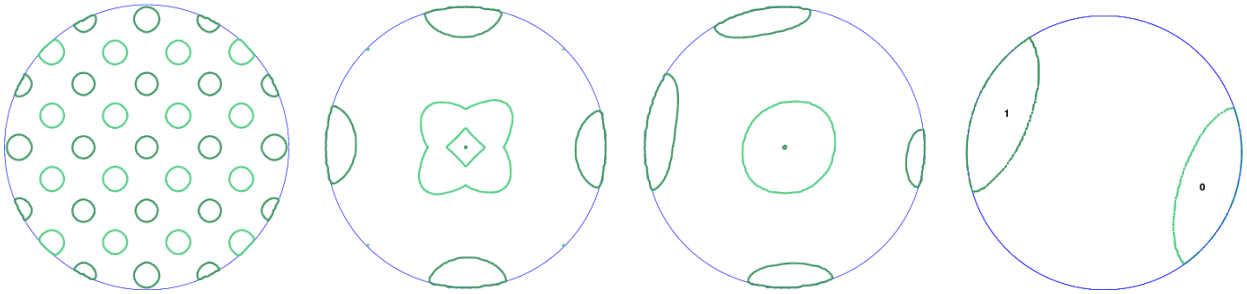

Figure V.2: Evolution of the level sets for a problem with two constraints (the same as the ones of the Figure V.1)

V.2.3 Examples

We present the result of our optimization process in the next figures. We first study the problem (1) with Ω a disc of radius 0.45, $\alpha = \beta = 0.15^2\pi$, $a = 0$ and $b = 1$. We obtain the same optimal shape with different initial guesses presented in Figures V.1 and V.2. The algorithm which has been presented in the case of two constraints can easily be adapted to a situation with more constraints. We present in Figure V.3 our results for a problem with three constraints of equal volume $0.15^2/2$.

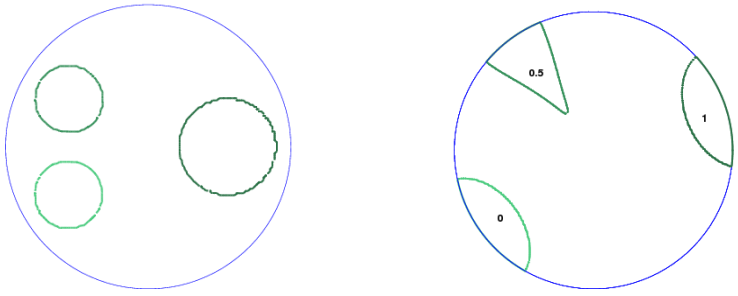

Figure V.3: Initial and optimized level sets for a problem with three constraints

V.3 Solution properties

V.3.1 Illustration of nonexistence results

It had been pointed out in [11, 10] that problems of the type (1) in general do not have solutions. However, the relaxed problem

$$\begin{aligned} \text{Minimize } E(u) := \int_\Omega f(x, u(x), \nabla u(x))\, dx, \\ |\{x \in \Omega,\ u(x) = a\}| \geq \alpha, \\ |\{x \in \Omega,\ u(x) = b\}| \geq \beta, \end{aligned} \tag{7}$$

admits a solutions whenever f satisfies some standard convexity and growth conditions [3]. Our previous numerical computations solve (7), and in the case of $f(x, u, \nabla u) = |\nabla u|^2$ it has been proved already in [3] that any solution of (7) also solves (1).

In this subsection we want to consider a situation where existence of a solution for (1) fails. To this aim we choose $f(x, u, \nabla u) = |\nabla u|^2 + |u|$ and try to compute numerically a solution of the ill-posed problem (1) for $a = 0$, $b = 1$ and $\alpha = \beta = \pi(0.15)^2$ on the unit disk Ω. As we can observe on Fig. V.4, the resulting level set of the constraint corresponding to $a = 0$ is strictly larger than the one which is prescribed. Actually, the area of that level set is approximatively equal to $0.0872 > \pi(0.15)^2$. In that sense, our numerical simulation illustrates the fact that non existence can occur for problem (1).

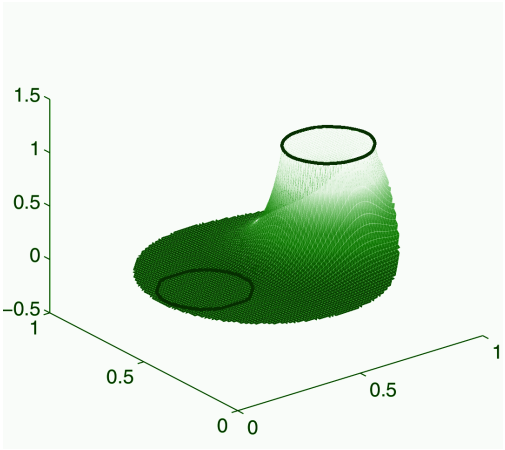

Figure V.4: Computed minimizer u of a relaxed problem (7) which does not satisfy the constraints of the exact problem (1), since its zero level set is too big. This illustrates the nonexistence of solutions for (1) in the two-dimensional case (see text for details).

V.3.2 Discontinuous parameter dependence

If $u^{\alpha,\beta}$ denotes the solution to a volume constrained problem of the type (1) then it is a natural question whether $u^{\alpha,\beta}$ depends (in an appropriate sense) continuously on α and β. It turns out that this is in general not the case, in fact we have the following result:

Theorem 18 *If we set $f(u, \nabla u) = |\nabla u|^2$ and $\Omega = (0,1)^2$ then the minimizers $u^{\alpha,\beta}$ of the problem (1) do not depend continuously on α and β, more precisely: There is an $\varepsilon > 0$ such that $\alpha \mapsto u^{\alpha,1-\alpha-\varepsilon}$ is not continuous in α with respect to the L^1-norm.*

To prove this result we use the Γ-limit of the problem (1). We briefly recall the definition of Γ-convergence and refer the reader for any details to the books of Braides and Dal Maso [4, 6]:

Definition 19 (Γ-convergence) *Let F_n be a sequence of functionals on a Banach space X. Then we say that F_n is Γ-converging in X to the functional F and denote $X-\Gamma-\lim F_n = F$ (or $F_n \overset{\Gamma}{\to} F$) if*

(i) For every $u \in X$ and for all $u_n \to u$ in X we have

$$\liminf_{n\to\infty} F_n(u_n) \geq F(u). \tag{8}$$

(ii) For every $u \in X$ there exists a sequence $u_n \subset X$ such that $u_n \to u$ and

$$\limsup_{n\to\infty} F_n(u_n) \leq F(u). \tag{9}$$

Inequality (8) is called Γ-*liminf inequality* and (9) is called Γ-*limsup* inequality. Such a Γ-limit has been derived for the case $\alpha + \beta \to 1$ and $f(u, \nabla u) = |\nabla u|^2$ in [3]. A generalization can be found in [16]. Let $\Omega \subset \mathbb{R}^N$ be an bounded open set. For fixed $\alpha, \beta \in (0, |\Omega|)$, we define the following functional

$$F_{\alpha,\beta} := \begin{cases} \gamma \displaystyle\int_\Omega |\nabla u|^2 \, dx & \text{if } u \in \mathcal{A}_{\alpha,\beta}, \\ +\infty & \text{elsewhere in } L^1(\Omega), \end{cases}$$

where $\gamma := |\Omega| - (\alpha + \beta)$ and

$$\mathcal{A}_{\alpha,\beta} := \{u \in H^1(\Omega) : \ |\{u = 0\}| = \alpha \text{ and } |\{u = 1\}| = \beta\}.$$

Then we can state the theorem from [3] as follows:

Theorem 20 *Let $\bar{\alpha} \in (0, |\Omega|)$. Then*

$$\Gamma(L^1)\text{-} \lim_{\substack{\alpha\to\bar{\alpha} \\ \beta\to|\Omega|-\bar{\alpha}}} F_{\alpha,\beta} = G_{\bar{\alpha}},$$

with $G_{\bar{\alpha}}$ given by

$$G_{\bar{\alpha}} := \begin{cases} \mathscr{H}^1(\{u = 0\})^2 & \text{if } u \in BV(\Omega, \{0,1\}) \text{ and } |\{u = 0\}| = \bar{\alpha}, \\ +\infty & \text{elsewhere in } L^1(\Omega). \end{cases} \tag{10}$$

This limit problem is much more accessible to analytical investigations. In particular we can set $A := \{u = 0\}$ and $B := \{u = 1\}$) and then the minimizers of G_α correspond to minimizers of the Dido's problem [18]: Minimize $\mathscr{H}^1(\Gamma)$ such that Γ separates Ω in open sets A and B with $|A| = \alpha$ and $|B| = |\Omega| - \alpha$. The solutions of this problem can be explicitly computed. In the following lemma we summarize the situation on the unit square:

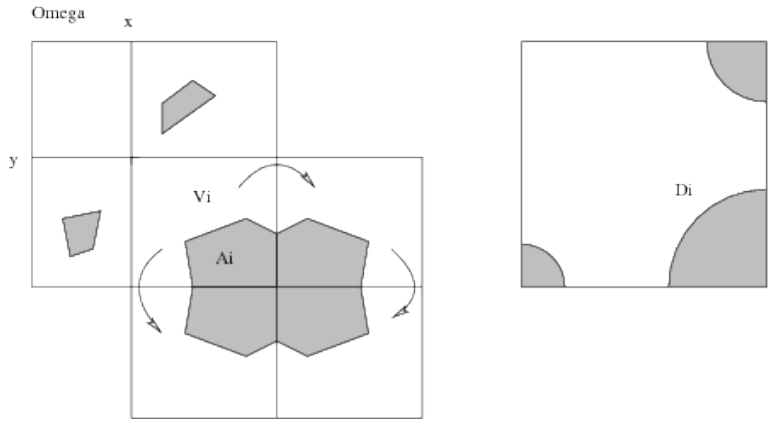

Figure V.5: The construction for the proof of Lemma 21.

Lemma 21 *Let $\Omega = (0,1)^2$, $\alpha > 0$, then there exists a set $\Gamma \subset \Omega$ minimizing $\mathscr{H}^1(\Gamma)$ among all sets with the property that there exist disjoint open sets $A, B \subset \Omega \setminus \Gamma$ with $|A| = \alpha$, $|B| = 1 - \alpha$ and $\Omega = A \cup B \cup \Gamma$.*

(i) If $\alpha < 1/\pi$ or $\alpha > 1 - 1/\pi$ then Γ is the segment of a circle with center in one of the corner points of Ω. (Type I solution, see Fig. V.6.)

(ii) If $1/\pi < \alpha < 1 - 1/\pi$ then Γ is a straight line parallel to a side of Ω. (Type II solution, see Fig. V.6.)

(iii) If $\alpha = 1/\pi$ or $\alpha = 1 - 1/\pi$ then Γ is either a circle segment or a straight line.

This Lemma seems to be folklore, but for the reader's convenience we give a proof using the isoperimetric inequality:

Proof. By symmetry we can assume that Γ is a solution of the problem for $\alpha \in (0, 1/2]$, moreover we assume first that $\ell := \mathscr{H}^1(\Gamma) < 1$. Denote the four corner points in the square Ω by Q_i and the sides by S_i. Since $\ell < 1$ the set projection π_i of Γ onto S_i satisfies $\pi_i(\Gamma) \neq S_i$. Let $x \in S_1 \setminus \pi_1(\Gamma)$ and $y \in S_2 \setminus \pi_2(\Gamma)$. Then the cross-shaped set $\{(x_1, x_2) \in \Omega \,|\, x_1 = x \text{ or } y_1 = y\}$ does not intersect with A, therefore we can decompose Ω along this cross into four disjoint connected open sets $V_1, \ldots, V_4$ such that $\bigcup_i \bar{V}_i = \bar{\Omega}$ and each $\bar{V}_i$ contains the corner point Q_i and none of the other corner points. We observe that since V_i open, $\partial V_i \cap A \subset \partial\Omega$. We can now mirror V_i and $A \cap V_i$ three times along the adjacent sides of the square Ω (see Fig. V.5) to obtain a larger set $A_i \subset \mathbb{R}^2$. Since $\partial A \cap \partial V_i$ was a subset of the mirror axis, we can now neglect the boundary and apply the isoperimetric inequality on the sets A_i, hence proving that they minimize their boundary length (under fixed volume) when they are discs. We can center these disks without loss of generality on Q_i and denote them by D_i and $D := \bigcup_i D_i$. Due to the minimality property of the boundary length, we have $\ell = \mathscr{H}^1(\Gamma) \geq \frac{1}{4} \sum_i \mathscr{H}^1(\partial D_i)$. Since $\ell < 1$, the disks D_i must be disjoint. (Otherwise the sum of two of their radii r_i would have to exceed the distance between two corner points, i.e. 1, but that would imply $1 > \ell \geq (r_1 + r_2)2\pi/4 > \pi/2$.) Since the disks are disjoint, we have $|D| = \sum_i |D_i| = |A|$. For the boundary length we have seen that $\mathscr{H}^1(\Gamma) \geq \frac{1}{4} \sum_i \mathscr{H}^1(\partial D_i)$ with

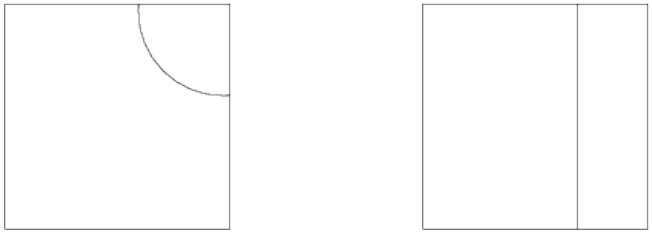

Figure V.6: Global minimizers for the parameters $\alpha = 0.55, \beta = 0.15$ (Type I) and $\alpha = 0.5, \beta = 0.2$ (Type II) on a square with side length 0.9. Although the parameters are very close, the solutions are not.

equality if and only if Γ consists of at most four arcs with centers in Q_i. It is now easy to check that the optimal configuration among these sets is given by exactly one arc with center in some Q_i. Since our initial assumption $\ell < 1$ is feasible if $\alpha < 1/\pi$, we have proved the first point of the theorem.

The last two points of the theorem follow easily: We know that in both cases there exists a Γ with $\mathscr{H}^1(\Gamma) = 1$. Suppose we could do better, then Γ would satisfy $\mathscr{H}^1(\Gamma) < 1$ and we could apply the argument above, proving that Γ must be an arc with center in some Q_i. Such an arc, however, would have a length larger than 1 (or in the case $\alpha = 1/\pi$ at least not less) which contradicts the assumption.

Proof of Theorem 18: Assume that for all $\varepsilon > 0$ the function $h_\varepsilon(\alpha) := u^{\alpha,1-\alpha-\varepsilon}$ is continuous in the L^1-norm. We know by the Γ-convergence that $u^{\alpha,1-\alpha-\varepsilon} \to u^\alpha$ in L^1 where u^α denotes the minimizer of the Γ-limit problem. Hence, for $\alpha < 1/\pi$ the functions $h_\varepsilon(\alpha)$ converge to a limit function $h(\alpha)$ of the type I as $\varepsilon \to 0$ (see Fig. V.6), for $1/\pi\alpha < 1 - 1/\pi$, however, the functions $h_\varepsilon(\alpha)$ converge to a function of the type II (see Fig. V.6). For $\alpha = 1/\pi$ we denote the two possible solutions of the limit problem by u^I and u^{II}. The L^1-distance between u^I and u^{II} is larger than 0.6 (as a small computation shows). We do not necessarily have uniform convergence of h_ε as $\varepsilon \to 0$, hence we need the following construction:

Let us fix α^1, α^2 such that $\alpha^1 < 1/\pi < \alpha^2$ and

$$\|h(\alpha^1) - u^I\|, \ \|h(\alpha^2) - u^{II}\| < 1/100 \tag{11}$$

(We can ensure this by choosing α^1 and α^2 close to $1/\pi$ since the minimizers of the limit problem are continuous outside $1/\pi$.)

Next, we choose sequences α_n^1, α_n^2 and ε_n, such that $\varepsilon_n < 1/n$, $\alpha_n^1 \to \alpha^1$, $\alpha_n^2 \to \alpha^2$ and $\|h_{\varepsilon_n}(\alpha_n^1) - h(\alpha^1)\| < 1/n$, $\|h_{\varepsilon_n}(\alpha_n^2) - h(\alpha^2)\| < 1/n$. (By the Γ-convergence we know that minimizers of the volume constraint problem converge for $\varepsilon \to 0$ to minimizers of the limit problem, hence we can find such sequences.)

Now we choose a sequence of α_n^0 that lies in between α_n^1 and α_n^2 and prove that the corresponding solutions of the volume constrained problem cannot converge to a solution of the limit problem:

Let α_n^0 satisfy $\alpha_n^1 < \alpha_n^0 < \alpha_n^2$. Using the (supposed) continuity of h we can apply the intermediate value theorem to find such an α_n^0 such that $\|h_{\varepsilon_n}(\alpha_n^0) - h_{\varepsilon_n}(\alpha_n^1)\| > 1/10$ and $\|h_{\varepsilon_n}(\alpha_n^0) - h_{\varepsilon_n}(\alpha_n^2)\| > 1/10$. Since the sequence α_n^0 is uniformly bounded, we can select a converging subsequence and, using the Γ-converge, its limit α^0 satisfies $\|h(\alpha^0) - h(\alpha^1)\| \geq 1/10$ and $\|h(\alpha^0) - h(\alpha^2)\| \geq 1/10$.

Figure V.7: Global (left) and local (right) minimizer on a nonconvex domain.

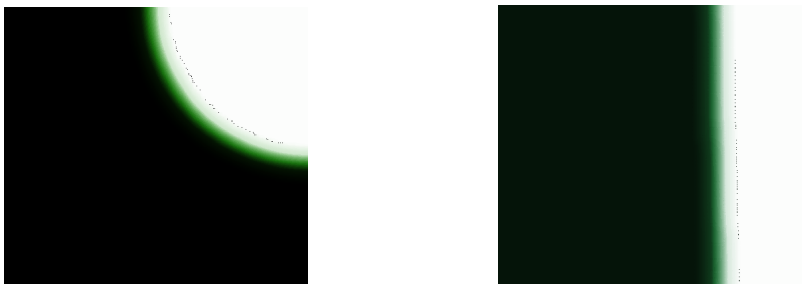

Figure V.8: Global (left) and local minimizer of the same problem as shown in the left side of Fig. V.6. This example demonstrates that there are genuinely local minimizers on a convex domain, in this case a square.

Using this together with (11) and $\|u^I - u^{II}\| > 0.6$ leads to a contradiction. Hence at least for sufficiently small $\varepsilon > 0$ the function h_ε cannot be continuous.

We illustrate this behavior with numerical computations (Fig. V.6) using the algorithm introduced in Section 2.

V.3.3 Existence of local minimizers

Our algorithm searches for minimizers which are not necessarily *global* minimizers. In one dimension it was possible to characterize local minimizer completely with analytical methods [10]. However, on convex domains of dimension $n \geq 2$ these methods do not work and it had been conjectured that in fact every minimizer is global. It is relatively simple to see examples of local minimizers in nonconvex domains (compare Fig. V.7 for a numerical computation). However, our computation hinted that also on the square there can be genuinely local minimizers, compare Fig. V.8.

In the following we present a proof of the existence of genuinely local minimizers on a square.

Theorem 22 (Existence of local minimizer) *There are convex domains $\Omega \subset \mathbb{R}^2$ such that the volume-constrained minimization problem (1) with $f(x, u, \nabla u) = |\nabla u|^2$ admits (for appropriate parameters) local minimizers (with respect to the L^∞-distance) which are not global.*

Proof. Let Ω be the unit square $(0,1)\times(0,1)$. For simplicity, $a=0$ and $b=1$. We choose $\alpha < \frac{1}{\pi}$ and $\beta = 1-\alpha-\gamma$ where $\gamma > 0$ is chosen small enough such that

$$\gamma < \frac{\alpha}{2}. \tag{12}$$

We define our candidate v for a local minimizer by a one-dimensional piecewise affine construction:

$$v(x,y) := \begin{cases} 1 & , \quad x < \beta \\ \frac{1-\alpha-x}{\gamma} & , \quad \beta \le x < 1-\alpha \\ 0 & , \quad 1-\alpha \le x \end{cases}.$$

We compute the energy of v as

$$\int_\Omega |\nabla v|^2 = \int_0^\gamma \left|\frac{d}{dx}\frac{x}{\gamma}\right|^2 = \frac{1}{\gamma}. \tag{13}$$

For $\gamma \to 0$, the function v converges in L^1 to a local minimizer of the Γ-limit functional which is not a global minimizer, compare Lemma 21. Therefore, for $\gamma > 0$ sufficiently small, v cannot be a global minimizer. It is therefore sufficient to prove that it is a local minimizer.

Let us suppose that there is another function w in the neighborhood of v with a smaller energy, more precisely suppose

$$\|w - v\|_{L^\infty} < 1/3 \tag{14}$$

and $\int_\Omega |\nabla w|^2 < \int_\Omega |\nabla v|^2 - \varepsilon$ for some $\varepsilon > 0$. Assume furthermore that w satisfies the same volume constraint as v. A priori, w does not need to be continuous. For the further construction it is, however, pivotal to work with a continuous function. Therefore we show that it is possible to construct a continuous function $\tilde{w}$ with the same properties:

We observe first, that w cannot have a "jump from zero to one", i.e. there cannot be a point $x \in \Omega$ such that there are sequences x_n and x_n', both converging to x with $w(x_n) \to 0$ and $w(x_n') \to 1$: if such a point existed, then (thanks to the continuity of v) we have $|w(x_n) - v(x_n) + v(x_n') - w(x_n')| \to 1$. On the other hand, using (14), we have $|w(x_n) - v(x_n)| < 1/3$ and $|w(x_n') - v(x_n')| < 1/3$. Together with the triangle inequality, this leads to a contradiction.

We denote $\Omega_0 := \{x \in \Omega;\ w(x) = 0\}$ and $\Omega_1 := \{x \in \Omega;\ w(x) = 1\}$. Since there is no jump from zero to one, we have $\bar{\Omega}_0 \cap \bar{\Omega}_1 = \emptyset$ and we can therefore define

$$\bar{w}(x) \quad := \quad \begin{cases} 0, & x \in \bar{\Omega}_0, \\ 1, & x \in \bar{\Omega}_1, \\ w(x), & x \in \Omega \setminus \left(\bar{\Omega}_0 \cup \bar{\Omega}_1\right) =: T. \end{cases}$$

The set T is open by construction. For each $x \in \partial T \setminus \partial\Omega$ there is *either* a sequence $x_n \to x$ such that $w(x_n) \to 0$ *or* a sequence $x_n' \to x$ such that $w(x_n') \to 1$. Denote the corresponding sets of boundary points by D_0 and D_1, then D_0 and D_1 form a disjoint union of $\partial T \setminus \partial\Omega$. Moreover, given that w has no jump from zero to one, D_0 and D_1 must be apart from each other, i.e. $\bar{D}_0 \cap \bar{D}_1 = \emptyset$. In other words, on $\partial T \setminus \partial\Omega$, $\bar{w}$ is locally constant.

The function $\bar{w}$ is by construction in $H^1(T)$, where T is open. Thus we can approximate $\bar{w}$ on T by continuous functions in the H^1-norm, where we respect the boundary conditions on $\partial T \setminus \partial\Omega$. Let w_n be such an approximating sequence, then for n large enough, $\|w_n - \bar{w}\|_{H^1(T)} < \varepsilon/2$.

We can now define $\tilde{w}$ by

$$\tilde{w}(x) \;:=\; \begin{cases} 0, & x \in \bar{\Omega}_0, \\ 1, & x \in \bar{\Omega}_1, \\ w_n(x), & x \in T. \end{cases}$$

$\tilde{w}$ is continuous by construction. Moreover, its energy is still lower than the energy of v:

$$\int_\Omega |\nabla \tilde{w}|^2 \;=\; \int_T |\nabla w_n|^2 \;<\; \int_T |\nabla \bar{w}|^2 + \frac{\varepsilon}{2} \;\le\; \int_\Omega |\nabla \bar{w}|^2 + \frac{\varepsilon}{2} \;<\; \int_\Omega |\nabla v|^2.$$

To ease notation, we will write w instead of $\tilde{w}$ in what follows.

The L^∞-constraint obviously forbids w to take a value of one where v is zero and vice versa, in other words:

$$w > 0 \text{ on } (0,\beta) \times (0,1) \text{ and } w < 1 \text{ on } (1-\alpha, 1) \times (0,1). \tag{15}$$

We define $L(y) := (0,1) \times \{y\}$ and $T := \{w \in (0,1)\}$ (the transition layer of w). Then

$$\int_0^1 |L(y) \cap \{(x,y) \in \Omega \,|\, w(x,y) \in (0,1)\}| \; dy = |T| = \gamma,$$

where the last inequality follows from the assumption that w satisfies the volume constraint.

We denote

$$G := \left\{ y \in (0,1) \,\middle|\, L(y) \cap \{w = 0\} \neq \emptyset \text{ and } L(y) \cap \{w = 1\} \neq \emptyset \right\}$$

and define on G the functions

$$B(y) := \max \left\{ |a-b| \,\middle|\, w(a,y) = 0,\; w(b,y) = 1,\; w(t,y) \in (0,1) \text{ for all } t \in (a,b) \right\}.$$

and $a(y)$, $b(y)$ as the values of a and b maximizing $|a-b|$ in the above definition of $B(y)$.

In other words: $B(y)$ is the maximal width of a transition between zero and one on the line $L(y)$ and the boundary points of this transition are given by $(a(y), y)$ and $(b(y), y)$, compare Fig. V.9 for an illustration.

If we integrate over all such maximal transitions, we get a lower bound for the total area of the transition layer:

$$\int_G b(y)\, dy \le |T|.$$

We estimate the gradient of w by its partial derivative in x-direction, as we did in (13), to get the following estimate:

$$\begin{aligned} \int_\Omega |\nabla w|^2 \;&=\; \int_T |\nabla w|^2 \;\ge\; \int_G \int_0^1 |\nabla w(x,y)|^2 \, dx\, dy \\ &\ge\; \int_G \int_0^1 \left| \frac{\partial}{\partial x} w(x,y) \right|^2 \, dx\, dy. \end{aligned}$$

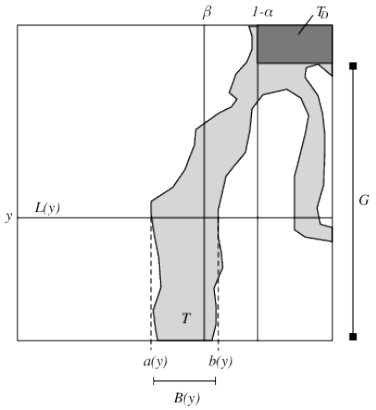

Figure V.9: Illustration of the sets T, T_D and G, the lines $L(y)$ and the maximal transitions from $a(y)$ to $b(y)$ with width $B(y) = |a(y) - b(y)|$.

Now, instead of integrating from 0 to 1, we just integrate over the largest transition layer, i.e. from $a(y)$ to $b(y)$. We recall that $|a(y) - b(y)| = B(y)$. Using Jensen's Inequality on the inner integral, we obtain therefore

$$\int_\Omega |\nabla w|^2 \;\geq\; \int_G \frac{1}{B(y)}\, dy.$$

This estimate is only useful if we find a relation between B and the set G. Otherwise, we can choose the set G small or B large to reduce the energy. Therefore we want to estimate the size of G. Let us define some area of the transition layer T that is situated outside $(0,1) \times G$ by

$$T_D := (0,1) \times ((0,1) \setminus G) \cap T,$$

compare again Fig. V.9 where this set is shaded in dark grey. Let $\delta := |T_D|$ be the size of this area.

Since for $y \in (0,1) \setminus G$ we cannot have $w(x_1, y) = 0$ and $w(x_2, y) = 1$ for two values $x_1, x_2 \in (0,1)$, and on the other hand $w(x,y) < 1$ for $x > 1 - \alpha$ and $w(x,y) > 0$ for $x < \beta$, see (15), we need to "cover" either $(0,\beta) \times ((0,1) \setminus G)$ or $(1-\alpha, 1) \times ((0,1) \setminus G)$ by the transition layer. Thus we get a lower bound for δ (taking into account that $\alpha < \beta$):

$$\delta \geq \alpha(1 - |G|).$$

Resolved for G, we obtain

$$|G| \geq 1 - \frac{\delta}{\alpha}. \tag{16}$$

Now we can continue estimating the energy of w. We first apply the Jensen Inequality with $\bar{B}$ being the average over B on G:

$$\int_\Omega |\nabla w|^2 \;\geq\; \int_G \frac{1}{B(y)}\, dy \;\geq\; |G| \frac{1}{\bar{B}}. \tag{17}$$

Let $T_G := T|_{(0,1)\times G}$ be the transition layers on $(0,1) \times G$. Since $T_G \cup T_D \subset T$ and T_G and T_D are disjoint, we have $|T_G| \leq |T| - |T_D|$. Using that $\delta = |T_D|$ and that $|T| = \gamma$ (volume constraint), we have $|T_G| \leq \gamma - \delta$.

On the other hand, $\int_G B(y)\,dy \le |T_G|$, thus $\bar{B}|G| \le \gamma - \delta$ or in other words $\bar{B} \le (\gamma - \delta)/|G|$. This provides us with the necessary relation between B and the size of G.

Together with (17) we obtain

$$\int_\Omega |\nabla w|^2 \ge |G|^2 \frac{1}{\gamma - \delta}.$$

Inserting (16), gives

$$\int_\Omega |\nabla w|^2 \quad \ge \quad \frac{(1 - \delta/\alpha)^2}{\gamma - \delta}.$$

We calculate the difference between this energy and the energy of v, as computed in (13):

$$\begin{aligned} \int_\Omega |\nabla w|^2 - \int_\Omega |\nabla v|^2 \quad &\ge \quad \frac{(1 - \delta/\alpha)^2}{\gamma - \delta} - \frac{1}{\gamma} \quad = \quad \frac{-2\frac{\delta}{\alpha}\gamma + \frac{\delta^2}{\alpha}\gamma + \delta}{\gamma(\gamma - \delta)} \\ &\ge \quad \frac{\delta}{\gamma(\gamma - \delta)}\left(1 - 2\frac{\gamma}{\alpha}\right). \end{aligned}$$

Using (12), we see that the right hand side is larger or equal than zero. This proves that w cannot have a smaller energy than v, thus v is a local minimizer.

Acknowledgement

We thank Giuseppe Buttazzo for his suggestions which helped to initiate this work.

Bibliography

[1] Emilio Acerbi, Irene Fonseca, and Giuseppe Mingione. Existence and regularity for mixtures of micromagnetic materials. *Proc. Royal Soc. London Sect. A*. To appear.

[2] N. Aguilera, H. W. Alt, and L. A. Caffarelli. An optimization problem with volume constraint. *SIAM J. Control Optim.*, 24(2):191–198, 1986.

[3] Luigi Ambrosio, Irene Fonseca, Paolo Marcellini, and Luc Tartar. On a volume-constrained variational problem. *Arch. Ration. Mech. Anal.*, 149(1):23–47, 1999.

[4] Andrea Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[5] G. Buttazzo D. Bucur. *Variational methods in shape optimization problems*, volume 65. Birkhäuser Boston, 2005.

[6] Gianni Dal Maso. *An introduction to* Γ-*convergence*. Birkhäuser Boston Inc., Boston, MA, 1993.

[7] F. Jouve G. Allaire and A. M. Toader. A level-set method for shape optimization. *C. R. Acad. Sci. Paris*, 334:1125–1130, 2002.

[8] M. E. Gurtin, D. Polignone, and J. Vinals. Two-phase binary fluids and immissible fluids described by an order parameter. *Math. Models Methods Appl. Sci.*, 6:815–831, 1996.

[9] A. Henrot and M. Pierre. *Variation et optimisation de formes, une analyse géométrique*, volume 48. Springer-Verlag Paris, 2005.

[10] Massimiliano Morini and Marc Oliver Rieger. On a volume constrained variational problem with lower order terms. *Applied Mathematics and Optimization*, 48:21–38, 2003.

[11] S. Mosconi and P. Tilli. Variational problems with several volume constraints on the level sets. *Calc. Var. Partial Differential Equations*, 14(2):233–247, 2002.

[12] S. Osher and F. Santosa. Level set methods for optimization problems involving geometry and constraints: frequencies of a two-density inhomogeneous drum. *J. Comput. Phys*, 171:272–288, 2001.

[13] S. Osher and J. A. Sethian. Front propagation with curvature-dependant speed: Algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[14] E. Oudet. Numerical minimization of eigenmodes of a membrane with respect to the domain. *ESAIM COCV*, 10:315–335, 2004.

[15] Marc Oliver Rieger. Abstract variational problems with volume constraints. *ESAIM: Control, Optimisation and Calculus of Variations*, 10(1):84–98, 2004.

[16] Marc Oliver Rieger. Higher dimensional variational problems with volume constraints – existence results and Γ-convergence. CVGMT preprint, Pisa, 2006.

[17] J. Sokolowski and J. P. Zolesio. Introduction to shape optimization: shape sensitivity analysis. *Springer Series in Computational Mathematics*, 10, 1992.

[18] Publius Vergilius Maro. *Aeneidum I.* 29–19 BC.

Optimal partitions for eigenvalues

Blaise Bourdin, Dorin Bucur & Édouard Oudet

VI.1 Introduction and motivation

This paper deals with the optimal partition problem for Dirichlet-Laplacian eigenvalues. Precisely, given a bounded open set $D \subset \mathbb{R}^2$, we are looking for a family of subsets $\{\Omega_i\}_{i=1}^n$ such that

$$\Omega_1 \cup \cdots \cup \Omega_n \subseteq D,\ \Omega_i \cap \Omega_j = \emptyset \text{ for } i \neq j$$

and which minimizes

$$\mathcal{J}_n(\Omega_1, \ldots, \Omega_n) = \sum_{i=1}^{n} \lambda_k(\Omega_i) \tag{1}$$

among all possible such partitions. Above, $\lambda_k(\Omega)$ denotes the k-th eigenvalue of the Dirichlet-laplacian on Ω, counted with multiplicity.

Existence of optimal partitions for problem (1) in the class of quasi-open sets was proved in [7]. For $k = 1$ regularity and qualitative studies of the optimal partitions were obtained by Conti, Terracini, and Verzini in [12] and Caffarelli, and Lin in [10]. Caffarelli and Lin obtained regularity results for the optimal partition and estimates for the asymptotic behavior of (1) when $n \to +\infty$. In particular, they conjectured that for the optimal partition $\{\Omega_i^*\}_{i=1}^n$

$$\sum_{i=1}^{n} \lambda_1(\Omega_i^*) \simeq \frac{n^2}{|D|} \lambda_1(H), \tag{2}$$

where H is the regular hexagon of area 1 in $\mathbb{R}^2$. Roughly speaking this estimate says that, far from ∂D, a tiling by regular hexagons of area $\frac{|D|}{n}$ is asymptotically close to the optimal partition.

A close problem, still for $k = 1$, was considered by Bonnaillie-Noël, Helffer and Vial in [4], where the cost functional is replaced by

$$\mathcal{L}_n(\Omega_1, \ldots, \Omega_n) = \max_{i=1\ldots n} \lambda_1(\Omega_i). \tag{3}$$

We notice that for fixed n, problems (1) and (3) may have different solutions (see [7] for remarks in relation with Payne conjecture). Nevertheless, Van den Berg conjectured the following asymptotic behavior :

$$\lim_{n \to +\infty} \frac{\mathcal{L}_n(\Omega_1^*, \ldots, \Omega_n^*)}{n} = \frac{\lambda_1(H)}{|D|} \tag{4}$$

It is quite easy to notice that, at least for smooth sets D, the asymptotic estimate (2) implies (4). The main feature of the case $k = 1$ is that the cost function (1) is of energy type. Namely, it can be written as:

$$\min_{u_1,\dots,u_n} \left\{ \sum_{i=1}^{n} \int_D |\nabla u_l|^2 : u_i \in H_0^1(D),\ \int_D u_i^2 = 1,\ u_i u_j = 0 \text{ for } 1 \le i < j \le n \right\}.$$

This kind of energy formulation was used by Chang [11] (see also [9]) to carry out a numerical study of optimal partitions of the disk. As expected, for m large enough, a regular hexagon tiling was observed.

The main purpose of this paper is to propose a numerical scheme for the approximation of the optimal partitions of problem (1) for any k. Our method relies on the approximation of "true domains" by positive Borel measures, the relaxation process introduced by Dal Maso and Mosco (see [13] and also Buttazzo and Timofte [8]). Based on a density argument, we replace the unknown m-upple of domains $(\Omega_1, \dots, \Omega_n)$ by an n-upple of functions $(\varphi_1, \dots, \varphi_n)$ such that

$$\varphi_l : D \longmapsto [0, 1],\ \sum_{i=1}^{n} \varphi_i(x) = 1,\ \text{a.e. } x \in D$$

For each index i, the k-th eigenvalue associated to φ_i is defined by the k-th eigenvalue of

$$\begin{cases} -\Delta u + C(1 - \varphi_i)u &= \lambda_k(\varphi_i)u \text{ in } D, \\ u \in H_0^1(D). \end{cases}$$

We notice that if φ_i equals the characteristic function 1_{Ω_i} of a smooth set Ω_i and $C \to +\infty$, then $\lambda_k(\varphi_l) \to \lambda_k(\Omega_l)$.

In this paper we propose a rigorous proof of the equivalence between problem (1) and our relaxed formulation when $C \to +\infty$ providing a complete justification of our numerical approach. Based on this method, we performed numerical simulations for $k = 1, 2, 3$ and large values of n. As expected, and up to boundary effects, in our numerical experiments, we obtain partitions that are very close to a tiling by regular hexagons in the case $k = 1$. Provided that the conjecture (2) is true, it can be easily proved that the asymptotic optimal partition for $k = 2$ is made of unions of pairs of regular hexagons (of measure $\frac{|D|}{2n}$). Again our numerical computations illustrate this fact.

Surprisingly as a consequence of our theoretical analysis, for every $k \in \mathbb{N}$ we prove the existence of an optimal partition with a mild regularity property, precisely : it is not consisting of quasi-open but open sets. Usually, the gain of regularity from quasi-open to open is a quite difficult task working only for energy functionals (see [5]).

VI.2 Analysis of the optimal partition problem

Let $d \ge 2$ and $D \subseteq \mathbb{R}^d$ be a bounded open connected set. For every open (or quasi-open) subset $A \subseteq D$ we denote by $\lambda_k(A)$ the k-th Dirichlet eigenvalue of the Laplace operator (multiplicities are counted)

$$\begin{cases} -\Delta u &= \lambda_k(A)u \text{ in } A \\ u &= 0 \text{ on } \partial A. \end{cases}$$

The previous equation has to be understood in a weak sense:

$$u \in H_0^1(A),\ \forall \varphi \in H_0^1(A) \int_A \nabla u \cdot \nabla \varphi dx = \lambda_k(A) \int_A u\varphi dx,$$

the eigenvalues being given by the Courant Fischer formula

$$\lambda_k(A) = \min_{S \in \mathcal{S}_k} \max_{u \in S} \frac{\int_A |\nabla u|^2 dx}{\int_A u^2 dx},$$

where $\mathcal{S}_k$ denotes the family of subspaces of dimension k of $H_0^1(A)$. Let

$$O_n = \{(\Omega_1, .., \Omega_n) : \Omega_i \text{ open}, \Omega_i \subseteq D, \Omega_i \cap \Omega_j = \emptyset,\ i \neq j\}.$$

Given $k, n \in \mathbb{N}$, the optimal partition problem reads

$$\inf_{(\Omega_1,..,\Omega_n) \in O_n} \sum_{i=1}^{n} \lambda_k(\Omega_i) := O(k, n). \tag{5}$$

In order to justify the numerical computations, we first introduce a relaxed version of the problem. Let

$$Q_n = \{(A_1, .., A_n) : A_i \text{ quasi-open}, A_i \subseteq D, \text{cap}(A_i \cap A_j) = 0,\ i \neq j\},$$

where $\text{cap}(U)$ stands for the capacity of U, and consider the problem

$$\inf_{(A_1,..,A_n) \in Q_n} \sum_{i=1}^{k} \lambda_k(A_i) := Q(k, n). \tag{6}$$

For every $k \geq 1$, the existence of a solution of problem (6) was proved in [7].

We begin with a first result asserting that problem (6) is indeed a relaxed version of problem (5). We rely on the γ-convergence which is a suitable topology in the family of quasi-open sets for which the eigenvalues are continuous (see [6]).

Theorem 23 *The set O_n is dense in Q_n for the γ-convergence. As a consequence, for every $k, n \in \mathbb{N}$ we have*

$$O(k, n) = Q(k, n).$$

Proof. Clearly, $O_n \subseteq Q_n$. In order to prove the density for the γ-convergence, we consider $(A_1, .., A_n) \in Q_n$. For every A_i, there exists a sequence of open sets U_i^j such that

$$A_i \subseteq U_i^j, a.e., \text{ and } \text{cap}(U_j^i \setminus A_i) \to 0 \text{ when } j \to \infty.$$

For each U_1^j there exists a smooth open subset V_1^j such that

$$\overline{V}_1^j \subseteq U_1^j,\ d_\gamma(U_1^j, V_1^j) \leq 1/j.$$

We set $\Omega_1^j = V_1^j$ and observe that $\Omega_1^j \xrightarrow{\gamma} A_1$, since

$$d_\gamma(A_1, \Omega_1^j) \leq d_\gamma(A_1, U_1^j) + d_\gamma(U_1^j, V_1^j).$$

For U_2^j there exists a smooth open subset V_2^j such that

$$\overline{V}_2^j \subseteq U_2^j,\ d_\gamma(U_2^j \setminus \overline{V}_1^j, V_2^j \setminus \overline{V}_1^j) \leq 1/j.$$

We set $\Omega_2^j = V_2^j \setminus \overline{V}_1^j$ and observe that $\Omega_2^j \xrightarrow{\gamma} A_2$. Indeed,

$$d_\gamma(A_2, \Omega_2^j) \leq d_\gamma(A_2, U_2^j \setminus \overline{V}_1^j) + d_\gamma(U_2^j \setminus \overline{V}_1^j, V_2^j \setminus \overline{V}_1^j).$$

The second term on the right hand is no greater than $1/j$, while for the first term we notice that

$$\mathrm{cap}(A_2 \setminus (U_2^j \setminus \overline{V}_1^j)) = \mathrm{cap}(A_2 \cap \overline{V}_1^j) \leq \mathrm{cap}(A_2 \cap U_1^j) \leq \mathrm{cap}(U_1^j \setminus A_1) \to 0,$$

and

$$\mathrm{cap}((U_2^j \setminus \overline{V}_1^j) \setminus A_2) \leq \mathrm{cap}(U_2^j \setminus A_2) \to 0.$$

Since in general $\mathrm{cap}(A_n \Delta A) \to 0$ implies $A_n \xrightarrow{\gamma} A$, we get that $\Omega_2^j \xrightarrow{\gamma} A_2$.

We continue the same procedure taking $\Omega_3^j = V_3^j \setminus (\overline{V}_1^j \cup \overline{V}_2^j)$, where V_3^j is chosen such that

$$d_\gamma(V_j^3 \setminus (\overline{V}_1^j \cup \overline{V}_2^j), U_j^3 \setminus (\overline{V}_1^j \cup \overline{V}_2^j)) \leq 1/j,$$

and iterating the same construction, we obtain that $(\Omega_1^j, .., \Omega_n^j) \in O_n$ and

$$(\Omega_1^j, .., \Omega_n^j) \xrightarrow{\gamma^n} (A_1, .., A_n).$$

The second assertion of the theorem is an immediate consequence of the density result. □

Let M be a measurable subset of D. There exists a quasi-open set A such that

$$H_0^1(A) = \{u \in H_0^1(D) : \quad u = 0 \text{ a.e. on } D \setminus M\}.$$

This set is precisely the union of all finely open sets U such that

$$1_U \leq 1_M \ a.e.$$

This remark provides a natural way to extend the optimal partition problem to partitions of n measurable, pairwise disjoint sets. Let $\varphi : D \to [0,1]$ be a measurable function. For any $C > 0$, by $\lambda_k(\varphi, C)$, we denote the k-th eigenvalue (counting multiplicity) of $-\Delta u + C(1-\varphi)u$, *i.e.*

$$\begin{cases} -\Delta u + C(1-\varphi)u &= \lambda_k(\varphi, C)u \text{ in } D \\ u &\in H_0^1(D) \end{cases} \tag{7}$$

Again, we have

$$\lambda_k(\varphi, C) = \min_{S \in \mathcal{S}_k} \max_{u \in S} \frac{\int_D |\nabla u|^2 + C(1-\varphi)u^2 dx}{\int_D u^2 dx},$$

$\mathcal{S}_k$ being the family of subspaces of $H_0^1(D)$ of dimension k. We introduce the set

$$M = \{(\varphi_1, .., \varphi_n) | \varphi : D \to [0,1] \text{ measurable } \sum_{i=1}^n \varphi_i = 1 \ a.e.\ D\},$$

and the problem

$$\inf_{(\varphi_1, .., \varphi_n) \in M} \sum_{i=1}^n \lambda_k(\varphi_i, C) := M(C, k, n). \tag{8}$$

Property 7 *Problem* (8) *admits at least one solution* $(\varphi_1^C, .., \varphi_n^C)$.

Proof. The existence of a solution is a consequence of the weak * $L^\infty(D)$ sequential compactness of M and of the fact that if $\varphi_h \overset{w*-L^\infty(D)}{\longrightarrow} \varphi$ then $C(1-\varphi_h)dx \xrightarrow{\gamma} C(1-\varphi)dx$. □

Theorem 24 *Let* $k = 1$. *The mapping*

$$\varphi \longrightarrow \lambda_1(\varphi, C)$$

is concave and every solution of problem (8) *is an extremal point of* M.

Proof. We give the details of the proof for $n = 2$. It is straightforward to generalize the following arguments for $n > 2$.

Let us first establish the concavity of

$$\varphi \longrightarrow \lambda_1(\varphi, C)$$

Let $\varphi_1, \varphi_2 \in L^\infty(D, [0,1])$ and $\theta \in (0,1)$. Then

$$\lambda_1(\theta\varphi_1 + (1-\theta)\varphi_2, C) = \frac{\int_D |\nabla u|^2 + C[1 - \theta\varphi_1 - (1-\theta)\varphi_2]u^2 dx}{\int_D u^2 dx}$$

where u is a non zero first eigenfunction associated to $\lambda_1(\theta\varphi_1+(1-\theta)\varphi_2, C)$. Moreover, by definition of the Rayleigh quotient we have

$$\lambda_1(\theta\varphi_1 + (1-\theta)\varphi_2, C) = \theta\frac{\int_D |\nabla u|^2 + C(1-\varphi_1)u^2 dx}{\int_D u^2 dx} + (1-\theta)\frac{\int_D |\nabla u|^2 + C(1-\varphi_2)u^2 dx}{\int_D u^2 dx},$$

so that

$$\lambda_1(\theta\varphi_1 + (1-\theta)\varphi_2, C) \geq \theta\lambda_1(\varphi_1, C) + (1-\theta)\lambda_1(\varphi_2, C), \tag{9}$$

which proves the concavity of the functional.

Let us prove now that every solution of problem (8) is an extremal point of M. First we notice that if equality occurs in (9), then $\varphi_1 - \varphi_2$ must be a constant function. Indeed, if equality occurs, the eigenfunction u associated to $\lambda_1(\theta\varphi_1 + (1-\theta)\varphi_2, C)$ is also a first eigenfunction of $\lambda_1(\varphi_1, C)$ and $\lambda_1(\varphi_2, C)$. Subtracting the two equations of type (7) satisfied by u with $\varphi = \varphi_1$ and $\varphi = \varphi_2$ we get

$$\varphi_1(x) - \varphi_2(x) = \frac{\lambda_1(\varphi_2, C) - \lambda_1(\varphi_1, C)}{C} \text{ a.e. } x \in D$$

since $u \neq 0$ a.e. on D.

Assume now that $(\varphi_1, .., \varphi_n)$ is an optimal solution for problem (8) and not an extremal point. We may assume the existence of $\varepsilon > 0$, a measurable set A such that $0 < |A| < |D|$ and

$$A \subseteq \{\varepsilon < \varphi_1 < 1 - \varepsilon\} \cap \{\varepsilon < \varphi_2 < 1 - \varepsilon\}.$$

We have from the concavity property

$$\lambda_1(\varphi_1, C) \geq \frac{1}{2}\lambda_1(\varphi_1 + \varepsilon 1_A, C) + \frac{1}{2}\lambda_1(\varphi_1 - \varepsilon 1_A, C),$$

$$\lambda_1(\varphi_2, C) \geq \frac{1}{2}\lambda_1(\varphi_2 - \varepsilon 1_A, C) + \frac{1}{2}\lambda_1(\varphi_2 + \varepsilon 1_A, C). \tag{10}$$

or

$$\lambda_1(\varphi_1, C) + \lambda_1(\varphi_2, C) \geq \min\{\lambda_1(\varphi_1 + \varepsilon 1_A, C) + \lambda_1(\varphi_2 - \varepsilon 1_A, C), \lambda_1(\varphi_1 - \varepsilon 1_A, C) + \lambda_1(\varphi_2 + \varepsilon 1_A, C)\}.$$

Finally, we have

$$\lambda_1(\varphi_1, C) + \lambda_1(\varphi_2, C) = \lambda_1(\varphi_1 + \varepsilon 1_A, C) + \lambda_1(\varphi_2 - \varepsilon 1_A, C) = \lambda_1(\varphi_1 - \varepsilon 1_A, C) + \lambda_1(\varphi_2 + \varepsilon 1_A, C).$$

Since equality holds in all previous inequalities we should have that $\varphi_1 + \varepsilon 1_A - (\varphi_1 - \varepsilon 1_A) = 2\varepsilon 1_A$ is constant in D. This last assertion is only possible only $A = D$, in contradiction with the assumption $|A| < |D|$. □

Theorem 25 *We have*

$$\lim_{C\to\infty} M(C, k, n) = O(k, n). \tag{11}$$

Moreover, if $(\varphi_1^C, .., \varphi_n^C)$ *is an optimal solution for problem* (8) *and* $\varphi_i^C \overset{w*L^\infty}{\longrightarrow} \varphi_i$ *then there exists an optimal solution* $(A_i)_{i=1,..,n}$ *for problem* (6) *such that* $A_i \subseteq \{\varphi_i = 1\}$ *a.e.*

Proof. There exists a constant K such that for every $C > 0$ and for every $i = 1, .., n$

$$\int_D C(1 - \varphi_i^C) w_i^C dx \leq K \text{ and } \|w_i^C\| \leq K$$

where w_i^C is the solution of

$$\begin{cases} -\Delta w_i^C + C(1 - \varphi_i^C) w_i^C &= 1 \text{ in } D \\ w_i^C &\in H_0^1(D) \end{cases}$$

Up to extracting a subsequence we have

$$w_i^C \overset{H_0^1(D)}{\rightharpoonup} w_i,$$

and we get

$$\int_D (1 - \varphi_i) w_i dx = 0$$

hence

$$w_i = 0 \text{ a.e. on } \{\varphi_i < 1\}.$$

We define the quasi-open sets $A_i = \{w_i > 0\}$ and notice that $(A_i)_i$ satisfy

$$\sum_{i=1}^n \lambda_1(A_i) \leq \lim_{C\to\infty} M(C, k, n). \tag{12}$$

For the converse inequality, we fix a partition $(\Omega_1, .., \Omega_n)$ consisting of open, smooth and disjoint sets. We take

$$\varphi_i = 1_{\Omega_i}$$

and observe that

$$M(C, k, n) \leq \lim_{C\to\infty} \sum_{i=1}^n \lambda_1(C, \varphi_i) = \sum_{i=1}^n \lambda_1(\Omega_i).$$

Using Theorem 23, and taking the infimum in the right hand side, we get (11).

The second assertion of the theorem is a consequence of inequality (12).

□

Theorem 26 *If $d = 2$, for every $k \geq 1$ there exists a solution of* (5) *consisting of open sets.*

Proof. Thanks to Theorem 23, we may take a minimizing sequence $(\Omega_1^h, .., \Omega_n^h)$ indexed by h consisting on polygonal disjoint sets. Assume that $\mathbb{R}^2 \setminus \Omega_1^h$ has more than $k(n-1)+1$ connected components. Since for every $i = 2, .., n$, the $k-th$ eigenvalue on Ω_i^h is given by at most k connected components, one can take the unused connected components of $\mathbb{R}^2 \setminus \Omega_1^h$ and add them to Ω_1^h in such a way that the cost functional decreases. The same procedure is repeated for every Ω_i^h, and finally we may assume that in the minimizing sequence every $\mathbb{R}^2 \setminus \Omega_i^h$ has at most $k(n-1)+1$ connected components.

Using Šverák's result (which is only valid in $\mathbb{R}^2$, see [17]) and the compactness of the Hausdorff complementary topology (see [6]), we can extract a subsequence (still denoted using the same index) such that

$$\Omega_i^h \xrightarrow{H^c} \Omega_i \quad \text{and} \quad \lambda_k(\Omega_i^h) \to \lambda_k(\Omega_i).$$

Since the Ω_i are pairwise disjoint open sets, they form a solution of problem (5). □

VI.3 Implementation and numerical results

The key to our numerical approach is the approximation Theorem 25. In order to obtain an approximation of the minimizers of (1), we fix C "large enough", and try to solve problem (8). In all the numerical experiments presented below, we assume that $\Omega = (0,1) \times (0,1)$, and use first order finite differences to represent the functions φ_l and their associated eigenvectors u_l. We decompose the domain D into a $N \times N$ grid with spacing $h = 1/(N-1)$. In order to simplify notations, we consider a renumbering operator $I : (0, N-1) \times (0, N-1) \mapsto 0, N^2 - 1$ such $I(i,j) = jN + i$. We refer to the components of a discrete field U as $U_{i,j}$ or $U_{I(i,j)}$ (which we abbreviate as U_I when there is no risk of confusion) depending on wether we want to insist on the spatial relation between the components or U or not. More precisely, to any $\varphi_l \in H_0^1(D)$, we associate a vector $\Phi_l \in \mathbb{R}^{N \times N}$ such that $[\Phi_l]_{i,j} = \varphi_l((i-1)h, (j-1)h)$, $1 \leq i, j \leq N$. By δ_x^2 and δ_y^2, we denote the classical finite difference operators, *i.e.* for any vector $U \in \mathbb{R}^{N \times N}$

$$\begin{aligned} \left[\delta_x^2 U\right]_{i,j} &= \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2}, \\ \left[\delta_y^2 U\right]_{i,j} &= \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2}. \end{aligned}$$

To each Φ_l, we associate the k-th Dirichlet eigenpair $(\lambda_{k,l}(\Phi_l), U_{k,l}(\Phi_l))$ (which we will denote by $(\lambda_{k,l}, U_{k,l})$ when there are no confusion possible) of the discrete operator $\mathrm{A}(\Phi_l)$ defined by

$$\mathrm{A}(\Phi)U := \left[-(\delta_x^2 + \delta_y^2) + C\mathrm{Id}\right] U - C\mathrm{M}(\Phi)U,$$

where $[\mathrm{M}(\Phi)]_{I,J} = \delta_{I,J} [\varphi]_I$, for any $0 \leq I \leq N^2 - 1$, and Id denotes the identity matrix of dimension $N \times N$.

Accounting for the homogeneous Dirichlet boundary conditions, we have then

$$[\mathrm{A}(\Phi_l) U_{k,l}(\Phi_l)]_I = \lambda_{k,l}(\Phi_l) [U_{k,l}(\Phi_l)]_I, \tag{13}$$

for any I corresponding to an interior node $I = I(i, j), 1 \leq i, j < N - 1$, and $U_{k,l}(\Phi_l)$ otherwise, and our discrete problem is

$$\inf \left\{ J_n(\Phi_1, \ldots, \Phi_n) : \Phi_l \in \mathbb{R}^{N \times N},\ 0 \leq [\Phi_l]_I \leq 1,\ \sum_{l=1}^{n} [\Phi_l]_I = 1,\ 0 \leq I < N^2,\ 1 \leq l \leq n \right\}, \quad (14)$$

where the discrete objective function J_n is defined by

$$J_n(\Phi_1, \ldots, \Phi_n) := \sum_{l=1}^{n} \lambda_{k,l}(\Phi_l).$$

The main difficulty in tailoring a numerical method for this problem is due the non-convexity of J_n, as stated in Theorem 24. As we are interested in the asymptotic behavior of the partitions function when n becomes large, the total number of degrees of freedom in the problem can become quite large (in the experiment presented in Figure VI.6, we have $N = 505$ and $n = 512$, leading to over 130,000,000 degrees of freedom), and to our knowledge, there are no global optimization algorithm capable of solving non-convex problems of this size. We note that the derivative of the objective function J_n with respect to the components of each of the Φ_l are easily obtained using a classical method in optimal design (see [3], for instance) assuming that all eigenvalues are simple. We first differentiate (13) with respect to the I-th component of Φ_l (I corresponding to an interior node of the discrete domain):

$$A(\Phi_l)\frac{\partial U_{k,l}(\Phi_l)}{\partial [\Phi]_I} - C\frac{\partial \mathrm{M}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}(\Phi_l) = \frac{\partial \lambda_{k,l}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}(\Phi_l) + \lambda_{k,l}(\Phi_l)\frac{\partial U_{k,l}(\Phi_l)}{\partial [\Phi]_I}.$$

Taking the dot product with $U_{k,l}(\Phi_l)$ on both side gives

$$U_{k,l}^t(\Phi_l)A(\Phi_l)\frac{\partial U_{k,l}(\Phi_l)}{\partial [\Phi]_I} - CU_{k,l}^t(\Phi_l)\frac{\partial \mathrm{M}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}(\Phi_l) = \frac{\partial \lambda_{k,l}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}^t(\Phi_l)U_{k,l}(\Phi_l) + \lambda_{k,l}(\Phi_l)U_{k,l}^t(\Phi_l)\frac{\partial U_{k,l}(\Phi_l)}{\partial [\Phi]_I}.$$

Noticing now that the operator $A(\Phi)$ is self-adjoint, and using (13) we obtain

$$-CU_{k,l}^t(\Phi_l)\frac{\partial \mathrm{M}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}(\Phi_l) = \frac{\partial \lambda_{k,l}(\Phi_l)}{\partial [\Phi]_I} U_{k,l}^t(\Phi_l)U_{k,l}(\Phi_l).$$

Last, we notice that $\left[U_{k,l}^t(\Phi_l)\frac{\partial \mathrm{M}(\Phi_l)}{\partial \Phi_I} U_{k,l}(\Phi_l)\right]_J = [U_{k,l}(\Phi_l)]_I^2\, \delta_{I,J}$, so that

$$\left[\frac{\partial \lambda_{k,l}(\Phi_l)}{\partial [\Phi]_I}\right]_J = -C\frac{[U_{k,l}(\Phi_l)]_I^2\, \delta_{I,J}}{U_{k,l}^t(\Phi_l)U_{k,l}(\Phi_l)},$$

and with the convention that the eigenvectors $U_{k,l}$ are normalized, we obtain the final expression for the sensitivity of $\lambda_{k,l}$ with respect to each component of each Φ field:

$$\left[\frac{\partial \lambda_{k,l}(\Phi_p)}{\partial [\Phi]_I}\right]_J = \begin{cases} -C\,[U_{k,l}(\Phi_l)]_I^2 & \text{if } l = p \text{ and } I = J, \\ 0 & \text{otherwise.} \end{cases}$$

VI.3.1 Minimization algorithm

From Theorem (24), we know that the functional J_n is concave, (at least when $k = 1$) and expect therefore that it admits many local minima. Due to the overall size of the problem, global minimization approaches are not practical. Instead, our numerical method is based on a projected-gradient descent with adaptive step described in Algorithm VI.3.1, where $\Pi_{\mathbb{S}^{n-1}}$ denotes a projection operator over the $n-1$ dimensional unit simplex $\mathbb{S}^{n-1}$ defined by

$$\mathbb{S}^{n-1} = \left\{ X = (X_1, \ldots, X_n) \in [0,1]^n : \sum_{l=1}^{n} X_l = 1 \right\}.$$

Note that since each $\lambda_{k,l}$ depends only on Φ_l, the parallelization of (14) is very natural. In our

Algorithm 1 General form of the projected gradient algorithm

Require: α (step), α_{min}, α_{max}, ω, ε (tolerance), p_{max}

$p = 1$
repeat
  for $l = 1$ to n **do**
    Compute the eigenpair $(\lambda_{k,l}, U_{k,l})$ of $\mathrm{A}(\Phi_l)$
    $\Phi_l \leftarrow \Phi_l - \alpha \nabla_{\Phi_l} \lambda_{k,l}$
  end for
  $\Phi_l \leftarrow \Pi_{\mathbb{S}^{n-1}} \Phi_l$, $l = 1, \ldots, n$.
  Compute $J^n := J_n(\Phi_1, \ldots, \Phi_n)$
  if $J^p \leq J^{p-1}$ **then**
    $\alpha \leftarrow \min((1+\omega)\alpha, \alpha_{max})$
  else
    $\alpha \leftarrow \max(\alpha_{min}, (1-\omega)\alpha)$
  end if
  $p \leftarrow p + 1$
until $p = p_{max}$ or $\sup_{i,j,l} \left| \alpha \Pi_{S_n} (\Phi_l)_I \right| \leq \varepsilon$

implementation, we distributed each partition function Φ_l on its own processor. We relied on PETSc [2, 1] for the main parallel infrastructure and distributed linear algebra operations, and used m uncoupled eigenvalues solvers provided by SLEPc [15]. The most computationally intensive part of this algorithm is the evaluation of the eigenpair $(\lambda_{k,l}(\Phi_l), U_{k,l}(\Phi_l))$, which does not require any inter-processor communication. In Algorithm VI.3.1, the time spent in this step is virtually independent of the number of cells m. The I/O operations can also be distributed in a trivial way. The most communication intensive part of the algorithm is the projection step, which can be achieved using a fixed number of *all-to-one* operations on the partition functions Φ_l, so the overall implementation of perfectly scalable.

Of course, we cannot guaranty that such a method will lead to the global minimizer of a non-convex energy. In particular, the concavity of J_m implies that the global minimizers of (8) lie on the boundary of the admissible simplex, which by definition is not a regular set. Roughly speaking, this means that in the course of the minimization algorithm, the Φ_l evolve rapidly toward the closest vertex of $\mathbb{S}^n$ at which point they cannot move anymore, so that the outcome of the minimization algorithm depends strongly on the initial guess. Figure VI.1 illustrates this sensitivity. We used an

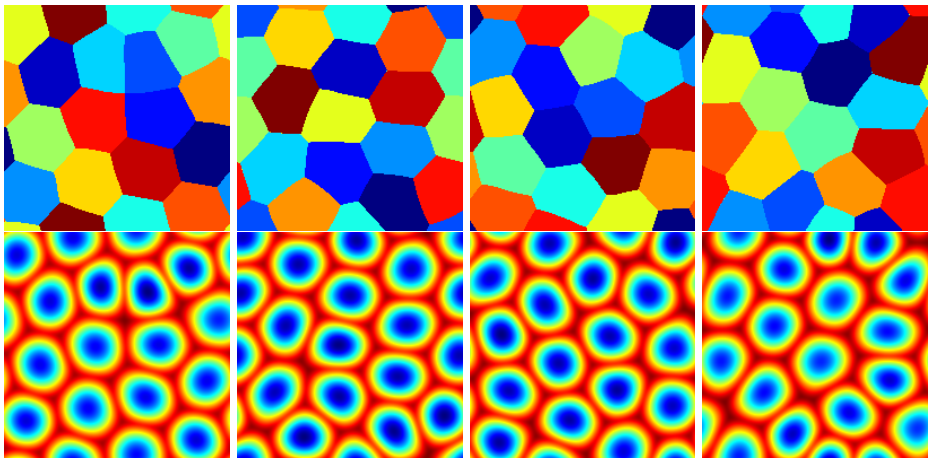

Figure VI.1: Dependence on the initial guess, using an orthogonal projection step. The initial values of the fields Φ are chosen randomly. The value of the objective function upon convergence is (left to right) 2,095.2, 2,108.5, 2,100.7, and 2,146.3

orthogonal projection operator over the unit simplex devised in [16]. In order to simulate the effect of a large number of cells on a reasonably sized domain, we used periodic boundary conditions for the Φ and U fields, and 16 cells.[1] The domain size is the unit square discretized in 200×200 nodes, and the parameter C is 10,000. We solved the same problem several times, using randomly generated initial fields. The first row represents a composite map of the functions Φ_l obtained by plotting $\sum_l l\Phi_l$, the second represents the sum of associated eigenvalues.

In order to partially alleviate this effect, we then implemented the *simple* projection operator defined by

$$[\Pi_{\mathbb{S}^{n-1}}\Phi_l]_I = \frac{|[\Phi_l]_I|}{\sum_{i=1}^n |[\Phi_i]_I|}.$$

Note that this operator is not an orthogonal projection operator and instead tend to keep the Φ in the middle of the faces of the target simplex (see the comparison of the effect of both projection in Figure VI.2). The effect of such an operator is double edged: it tends to prevent the Φ's from becoming "stuck" at the vertices of the unit simplices, but at the same time makes the actual minimizers virtually unreachable.

We then combined both operators: in step 7 of Algorithm VI.3.1, we used the simple algorithm until we reach convergence, then restart the computation using the orthogonal projection step. Figure VI.3 displays the outcome of this approach. The parameters are that of Figure VI.1, and the initial guess for the Φ_l is the same as in the leftmost experiment of the aforementioned figure. Upon convergence, we still obtain a non-regular tiling, whose energy is lesser than that obtained using only orthogonal projection. As the size of the search space is very large, convergence to a local minimizer is very likely. Our final algorithm uses a implemented a multi-level approach akin to a continuation method to address that issue. We use the simple projection algorithm and upon

[1] This choice is not innocent. It is of course impossible to construct a periodic paving of $\mathbb{R}^2$ by regular hexagons with periodicity cell the unit square. However, it is possible to do so using $4n^2$, $n \in \mathbb{N}$ slightly flattened regular hexagons. If conjecture 2 holds, it is reasonable to expect that such a paving realizes the global minimizer of J_m in this setting.

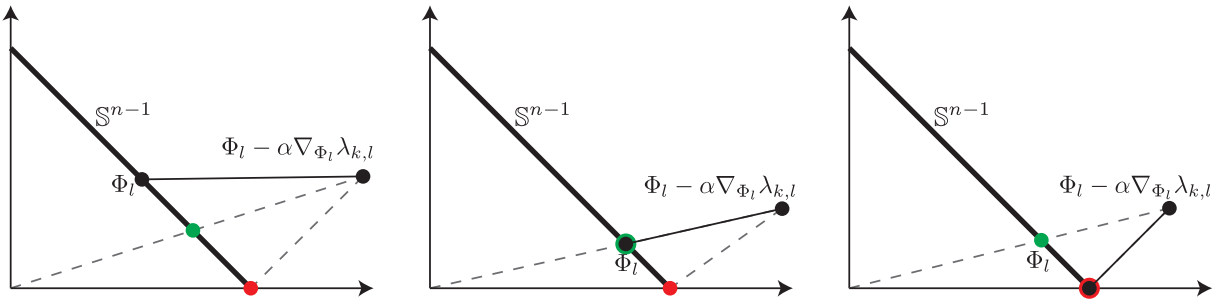

Figure VI.2: Behavior of the projection operators. The black dots represent Φ_l and $\Phi_l - \alpha\nabla_{\Phi_l}\lambda_{k,l}$ as labeled. The red ones are the orthogonal projection of $\Phi_l - \alpha\nabla_{\Phi_l}\lambda_{k,l}$, the green ones its *simple* projection. Simple projection has a lesser tendency to "send" the functions Φ_l towards the vertices of $\mathbb{S}^{n-1}$.

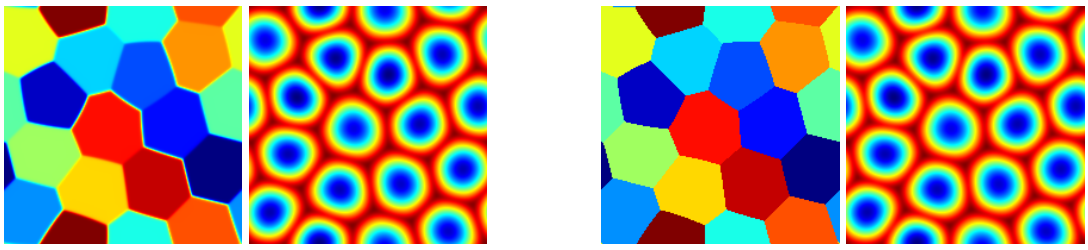

Figure VI.3: The problem from Figure VI.1(left) solved using a combination of simple and orthogonal projection. The leftmost figures represents the Φ and U fields upon convergence of the minimization algorithm using simple projection. Note how the functions Φ are not piecewise constant with values in $\{0, 1\}$. The rightmost figure corresponds to the final result obtained by using the orthogonal projection, starting from the configuration in th left. Compare the value of the objective function at 2,145.0 (left) and 2,073.8 (right) to that of the previous computations.

convergence of Algorithm VI.3.1 project the solution onto a finer grid, and iterate this process. After several grid refinement, we switch to the orthogonal projection. Figure VI.4 displays the numerical results obtained using this approach for the problem solved in Figures VI.1 and VI.3. We tested this approach using several initial conditions. In each case, we obtained a regular paving by hexagons, as expected. All the experiments presented below were obtained using the multi-level algorithm.

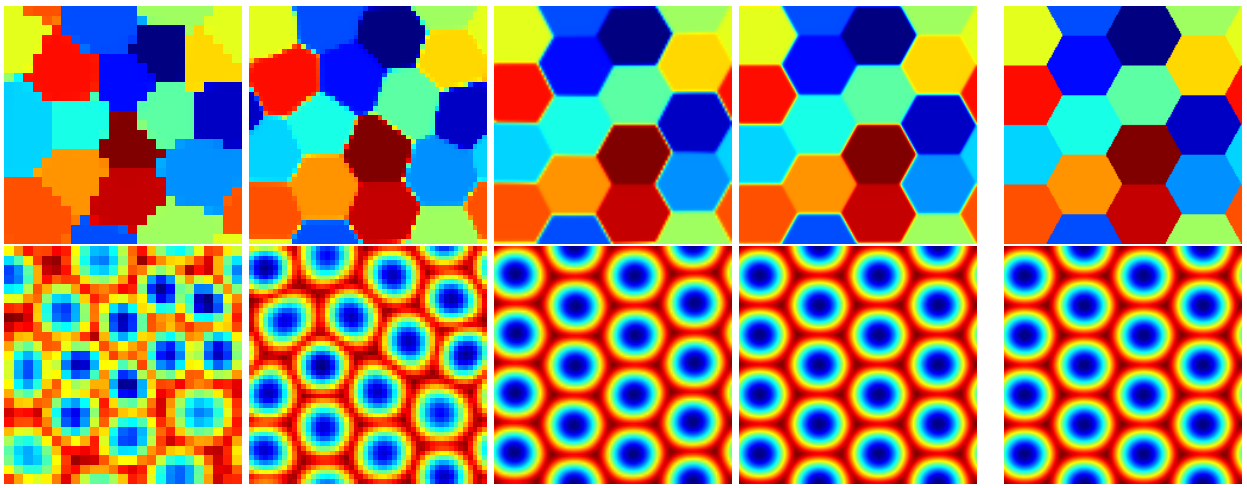

Figure VI.4: The same problem is solved again using the simple projection on increasingly refined grids (4 leftmost figures) then using the orthogonal projection on the final grid (right). The grid sizes are (from left to right) 25×25, 50×50, 100×100, and 200×200. The objective function upon convergence is (from left to right) 1,902.1, 2,033.8, 2,095.7, 2,124.6, and 2,048.8

VI.3.2 Numerical experiments

We were able to run a series of large computations on parallel supercomputers at the Texas Advanced Computing Center. In Figure VI.5, the domain is again the unit square. Periodicity boundary conditions are not used, as the number of cell ($n = 384$) is large enough that we expect that the effect of boundary conditions vanishes in the center of the domain. The computations were run on four layers of recursively refined grid of respective dimension (64×64), (127×127), (253×253), and (505×505). The parameter C is 10^5, the tolerance parameter $\varepsilon = 10^6$, the bounds on the admissible steps are $\alpha_{min} = 1$, $\alpha_{max} = 10^4$. We used only the simple projection operator, and the final objective functions on each grid are $1.602\,10^6$, $1.248\,10^6$, $1.176\,10^6$, and $1.189\,10^6$. We observe that the solution corresponds to local patches of tiling by regular hexagons, as we would expect from a "good" local minimizer.

We obtained similar results while running the same computation of 512 processors, for 512 cells. The fields Φ and U are represented using the usual convention and the final energies are $2.342\,10^6$, $2.243\,10^6$, $2.024\,10^6$, and $2.051\,10^6$. Again, the local geometry away from the edges of the domain is that of a network of regular hexagons.

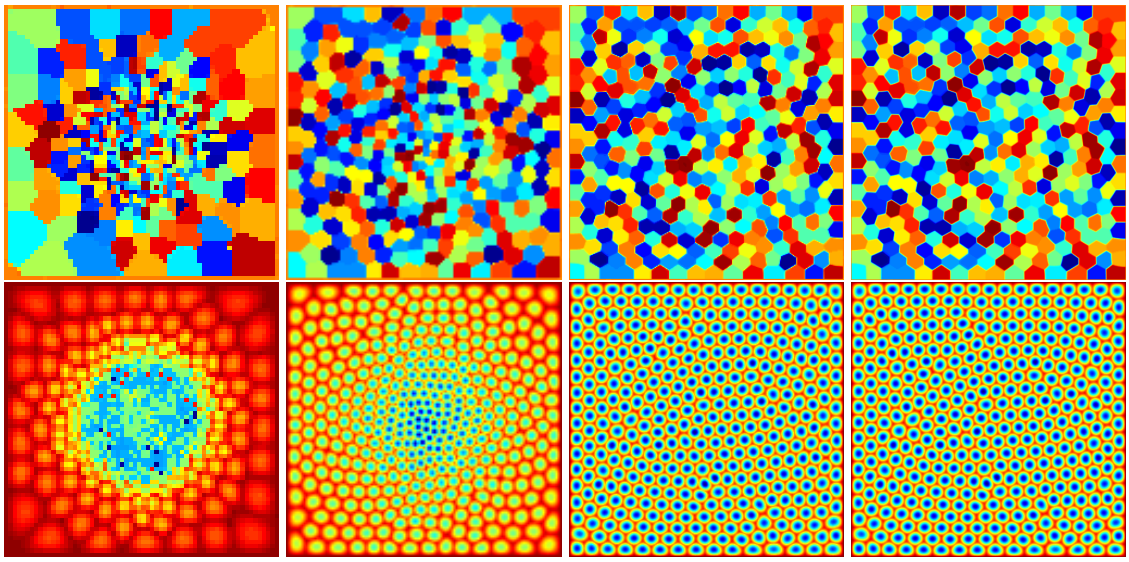

Figure VI.5: Optimization of the sum of the first eigenvalue of the Dirichlet Laplacian on 384 cells with $C = 10^5$. First row: cell shape on recursively refined grids (64×64), (127×127), (253×253), and (505×505). Second row: sum of the first eigenfunctions on the same grids.

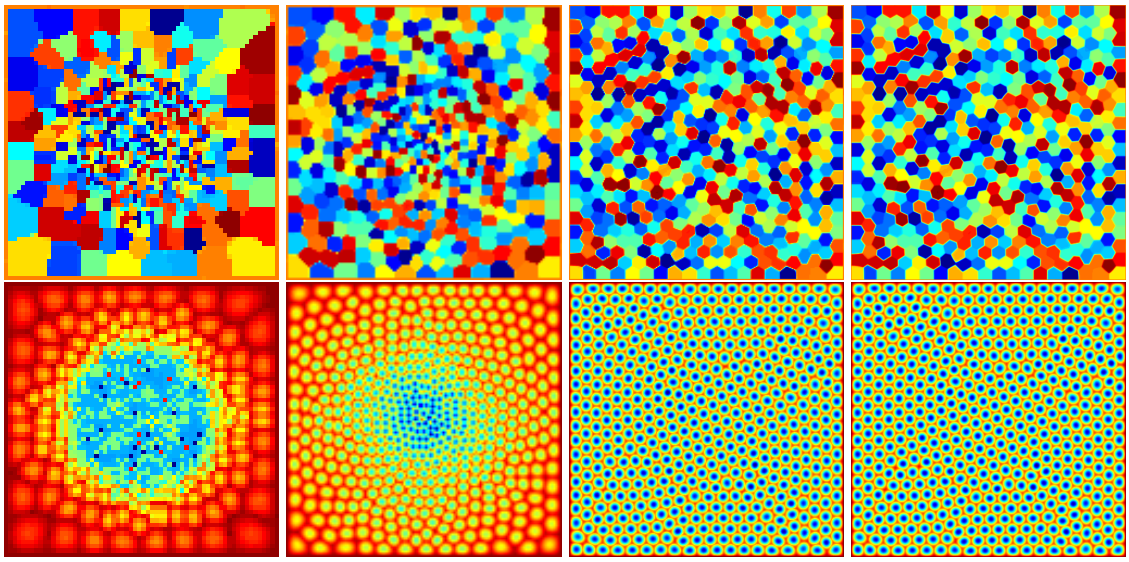

Figure VI.6: Optimization of the sum of the first eigenvalue of the Dirichlet Laplacian of 512 cells with $C = 10^5$. First row: cell shape on recursively refined grids (64×64), (127×127), (253×253), and (505×505). Second row: sum of the first eigenfunctions on the same grids.

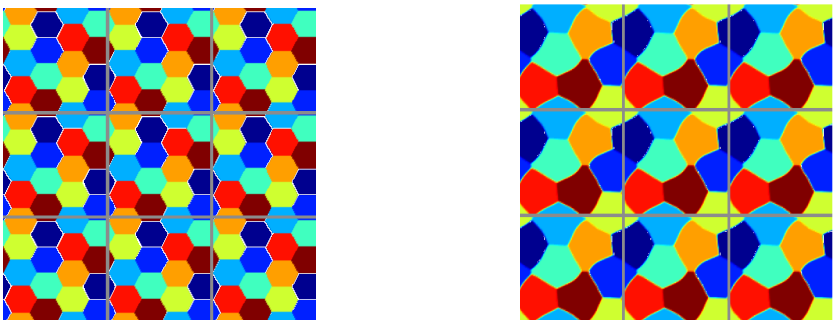

Figure VI.7: Optimal partitions of the sum of the second (left) and third (right) eigenvalues of the Dirichlet Laplacian for $n = 8$ cells. The periodicity is highlighted by repeating the unit cell 9 times on a two dimensional lattice.

VI.3.3 Extensions and conclusions

Our algorithm can easily be adapted to objective function involving higher order eigenvalues of linear combination of eigenvalues of different order. A classical numerical issue in this case comes from the potential non-differentiability of multiple eigenvalues with respect to changes of the function Φ. We did not try to address this problem, but obtained interesting results nevertheless. Figure (VI.7) represent the Φ fields obtained with $n = 8$ for $k = 2$ and $k = 3$, respectively, using periodic boundary conditions. As explained in the introduction if (2) holds, the optimal partition for $k = 2$ is obtained by a partition made of pairs of regular hexagons. Again, modulo the flattening necessary to achieve periodicity on a unit cell, this is the configuration that we observe. For $k = 3$ (Figure VI.7-right), we obtain a periodic tiling by non-regular hexagons, which can be proven to be a sub-optimal solution, as a tiling by regular hexagons would lead to a lower energy. Again, this is most certainly due to the fact that our objective function admits a great deal of local minima, which are difficult to avoid in optimization problems of this size. An additional difficulty when $k \geq 2$ is that the k-th eigenvalue of an optimal cell is expected to have multiplicity greater than 1 hence and may not be differentiable.

Noticing that the analysis and algorithm are not restricted to the two–dimensional case, we ported our program to the 3D case, but were unable to obtain any meaningful results. We believe that the convergence rate of our primitive algorithms is too slow to converge to a decent local minimizer in a reasonable time in 3D, when the dimension of the space of admissible fields Φ becomes very large, and the eigenvalue computation cannot be performed on a single processor in an acceptable time. Perhaps the current implementation needs to be improved by associating groups of processors to each function Φ (so as to improve the performance of the eigenvalue solver), and implement a more efficient minimization algorithm in order to reduce the number of necessary function evaluations in the minimization loop.

Acknowledgements

This work was initiated in an open-problems session during the Mini-Workshop “Shape Analysis for Eigenvalues” at the Mathematisches Forschungsinstitut Oberwolfach, organized by D. Bucur, G. Buttazzo, and A. Henrot. The authors would like to thank the MFO and the organizers of the mini-workshop for that.

Support for the first author was provided in part by the National Science Foundation grant DMS-0605320. The numerical experiments were performed using the National Science Foundation TeraGrid resources [14] provided by NCSA at the University of Illinois at Urbana-Champaign and TACC at the University of Texas under the Resource Allocation TG-DMS060011N.

Bibliography

[1] S. Balay, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 2.1.5, Argonne National Laboratory, 2004.

[2] S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc Web page, 2001. http://www.mcs.anl.gov/petsc.

[3] M.P. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications*. Springer, 2nd ed edition, 2003.

[4] V. Bonnaillie-Noël, B. Helffer, and G. Vial. Numerical simulations for nodal domains and spectral minimal partitions. Preprint Hal 00150455, 2007.

[5] T. Briançon, M. Hayouni, and M. Pierre. Lipschitz continuity of state functions in some optimal shaping. *Calc. Var. Partial Differential Equations*, 23(1):13–32, 2005.

[6] D. Bucur and G. Buttazzo. *Variational methods in shape optimization problems*. Progress in Nonlinear Differential Equations and their Applications, 65. Birkhäuser Verlag, Basel, 2005.

[7] D. Bucur, G. Buttazzo, and A. Henrot. Existence results for some optimal partition problems. *Adv. Math. Sci. Appl.*, 8(2):571–579, 1998.

[8] G. Buttazzo and C. Timofte. On the relaxation of some optimal partition problems. *Adv. Math. Sci. Appl.*, 12(2):509–520, 2002.

[9] L. A. Caffarelli and Fang-Hua Lin. Singularly perturbed elliptic systems and multi-valued harmonic functions with free boundaries. *J. Amer. Math. Soc.*, 21(3):847–862, 2008.

[10] L. A. Cafferelli and Fang Hua Lin. An optimal partition problem for eigenvalues. *J. Sci. Comput.*, 31(1-2):5–18, 2007.

[11] Shu-Ming Chang, Chang-Shou Lin, Tai-Chia Lin, and Wen-Wei Lin. Segregated nodal domains of two-dimensional multispecies Bose-Einstein condensates. *Phys. D*, 196(3-4):341–361, 2004.

[12] M. Conti, S. Terracini, and G. Verzini. An optimal partition problem related to nonlinear eigenvalues. *Journal of Functional Analysis*, 198(1):160–196, 2003.

[13] G. Dal Maso. Γ-convergence and μ-capacities. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 14(3):423–464 (1988), 1987.

[14] C. Catlett et al. *High Performance Computing and Grids in Action*, volume 16 of *Advances in Parallel Computing*, chapter TeraGrid: Analysis of Organization, System Architecture, and Middleware Enabling New Types of Applications. IOS Press, Amsterdam, 2008.

[15] V. Hernandez, J. E. Roman, and V. Vidal. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Transactions on Mathematical Software*, 31(3):351–362, 2005.

[16] C. Michelot. A finite algorithm for finding the projection of a point onto the canonical simplex of $\mathbf{R}^n$. *Journal of Optimization Theory and Applications*, 50(1):195–200, 1986.

[17] V. Šverák. On optimal shape design. *J. Math. Pures Appl. (9)*, 72(6):537–551, 1993.

Approximation of partitions of least perimeter by Γ-convergence : around Kelvin's conjecture

Édouard Oudet

VII.1 Introduction

We study in this article the problem of dividing a region $C \subset \mathbb{R}^N$ into pieces of equal volume such as to minimise the surface of the boundary of the partition. Physically this problem can be reformulated in: what is the most efficient soap bubble foam of C (see [14]) ?

If $C = \mathbb{R}^2$, Hales proved in 1999 that any partition of the plane made of regions of equal area has a perimeter at least equal that of the regular hexagonal honeycomb tiling (see [8] or [13]).

The problem when $C = \mathbb{R}^3$ has been first raised by Lord Kelvin in 1894. He conjectured that a tiling made of shapes which are closed from truncated octahedra may be optimal. This conjecture was motivated by the fact that this tiling satisfies Plateau's first order optimality conditions (see for instance the book of Plateau [9] translated by K. Brakke). Ten years ago, the two physicists D. Weaire and P. Phelan found a better tiling than the one of Kelvin (see [15]). This tiling is made of two kinds of cells: one with 14 sides and the other with 12. This last structure is up to now the best candidate for solving Kelvin's problem.

In this paper we propose a numerical process to approximate optimal partitions in any dimension. The key idea of our method is to relax the problem into a functional framework based on the famous result of Γ-convergence obtained by Modica and Mortolla (see [12], [11] or [1] for a different approach).

In the first section we give a rigorous mathematical framework to the question of dividing a bounded set C into pieces of equal volume with the smallest boundary measure. In a second section we extend this framework to the case $C = \mathbb{R}^3$. In both situations, we prove by a direct approach the well-posedness of our problems. In a third part, we describe how the result of Modica and Mortolla on phase transitions leads to a numerical algorithm to approximate optimal partitions. To conclude we illustrate the efficiency of our numerical process on different geometrical situations. In our experiments, we were able to recover both Kelvin's and Weaire and Phelan's tilings starting with uniform random distribution of densities.

VII.2 Dividing a bounded subset of $\mathbb{R}^N$

Let $n \in \mathbb{N}$ and C a compact regular subset of $\mathbb{R}^N$. We are first going to give a rigorous mathematical framework to the question of dividing C into n peaces of equal volume such that the boundary of the partition has the smallest measure. For this purpose, let us consider the following natural partitioning problem:

$$\inf_{(\Omega_i)_{i=1}^n \in \mathcal{O}_n} \mathcal{J}_n(\Omega_1, \dots, \Omega_n) \tag{1}$$

with

$$\mathcal{J}_n(\Omega_1, \dots, \Omega_n) = \sum_{i=1}^n \mathcal{H}^{N-1}(\partial\Omega_i) \tag{2}$$

where $\mathcal{H}^{N-1}$ stands for the $(N-1)$-dimensional Hausdorff measure and $\mathcal{O}_n$ defined by

$$\mathcal{O}_n = \{(\Omega_i) \text{ measurables } | \cup_{i=1}^n \Omega_i = C,\ \Omega_i \cap \Omega_j = \emptyset \text{ if } i \neq j \text{ and } |\Omega_i| = \frac{|C|}{n} \text{ for } i = 1 \dots n\} \tag{3}$$

where $|\Omega_i|$ is the Lebesgue measure of the set Ω_i. Notice that the two first equalities in (3) have to be understood up to a set of measure zero. We claim that the problem (1) is well posed:

Theorem 27 *It exists at least one family $(\Omega_i^*)_{i=1}^n \in \mathcal{O}_n$ such that:*

$$\mathcal{J}_n(\Omega_1^*, \dots, \Omega_n^*) = \inf_{(\Omega_i)_{i=1}^n \in \mathcal{O}_n} \mathcal{J}_n(\Omega_1, \dots, \Omega_n)$$

.

Proof. We notice first that it is equivalent to show that the problem of minimising

$$\hat{\mathcal{J}}_n(\Omega_1, \dots, \Omega_n) = \sum_{i=1}^n \mathcal{H}^{N-1}(\partial\Omega_i \backslash \partial C) \tag{4}$$

among sets of $\mathcal{O}_n$ has a solution since $\hat{\mathcal{J}}_n - \mathcal{J}_n$ is equal to the constant $\mathcal{H}^{N-1}(\partial C)$. Now, we apply the standard direct method of the calculus of variations: Consider a minimising sequence $((\Omega_i^k)_{i=1}^n)_k$ of partitions. That is

$$\lim_{k \to +\infty} \hat{\mathcal{J}}_n(\Omega_1^k, \dots, \Omega_n^k) = \inf_{(\Omega_i)_{i=1}^n \in \mathcal{O}_n} \hat{\mathcal{J}}_n(\Omega_1, \dots, \Omega_n).$$

It is clear from the previous limit that for k large enough, every set Ω_i^k has a finite perimeter with respect to the $N-1$ Hausdorff measure. This implies classically that every such set Ω_i^k is a set of Cacciopoli's type. More precisely, the characteristic function $\chi_{\Omega_i^k}$ is in the space $BV(C)$, the normed space of functions of bounded variations in C (for a precise definition of $BV(C)$ and its main properties, see [6] and [2]). Additionally, we have

$$\|\chi_{\Omega_i^k}\|_{BV(C)} = \mathcal{H}^{N-1}(\partial\Omega_i^k \backslash \partial C).$$

By a standard compactness argument (see for instance [6] page 176), there exists a subsequence of $(\Omega_i^k)_{i=1}^n$ (still denoted using the same index) that converges in $L^1(C)^n$ to a n-tuple $(\Omega_i^*)_{i=1}^n$. By the $L^1(C)^n$ convergence, every limit set Ω_i^* is still of volume $|C|/n$. Let us prove that $(\Omega_i^*)_{i=1}^n$ is optimal

for our problem. The convergence in $L^1(C)$ implies the convergence almost everywhere in C of each $\chi_{\Omega_i^k}$. As a consequence the following constraints are still satisfied at the limit:

$$\cup_{i=1}^n \Omega_i^* = C,\ \Omega_i^* \cap \Omega_j^* = \emptyset \text{ if } i \neq j. \tag{5}$$

Moreover, the norm of $BV(C)$ is lower semi-continuous, that is

$$\forall i = 1 \dots n,\ \mathcal{H}^{N-1}(\partial\Omega_i^* \backslash \partial C) \leq \liminf_k \mathcal{H}^{N-1}(\partial\Omega_i^k \backslash \partial C). \tag{6}$$

Equations (5) and (6) prove the theorem. □

From the previous proof, we deduce that problem (1) is equivalent to the functional optimisation problem:

$$\inf_{(u_i)_{i=1}^n \in \mathcal{X}_n} J_n(u_1, \dots, u_n) \tag{7}$$

where

$$J_n(u_1, \dots, u_n) = \sum_{i=1}^n \int_C |Du_i| \tag{8}$$

is the sum of all the BV norms of each function u_i and

$$\mathcal{X}_n = \{(u_i) \,|\, \forall i = 1 \dots n,\ u_i \in BV(C, \{0,1\}),\ \int_C u_i = \frac{|C|}{n},\ \sum_{i=1}^n u_i(x) = 1 \text{ a.e. in } C\}. \tag{9}$$

We will establish in section 4 a relaxed functional formulation also based on BV spaces which will be the key point of our numerical approach.

VII.3 Dividing a torus: a sub-problem of Kelvin's conjecture

In this section we would like to extend the previous optimisation problem restricted to bounded domains to partitions of all $\mathbb{R}^N$. We first recall an existence result obtained by F. Morgan in [7] which gives a rigorous mathematical formulation of Kelvin's problem in $\mathbb{R}^N$:

Theorem 28 *Consider the partitions of $\mathbb{R}^N$ into countable measurable sets (Ω_i) of unit volume. For all such partitions, we define:*

$$F((\Omega_i)) = \limsup_{r \to +\infty} \frac{\mathcal{H}^{N-1}(B(0,r) \cap (\cup_i \partial\Omega_i))}{|B(0,r)|} \tag{10}$$

where $|B(0,r)|$ is the volume of the ball of radius r centered at the origin. Then, there exists a partition which minimises F among all admissible partitions.

As noticed by F. Morgan, such a partition is not unique: a compact perturbation around the origin does not change the previous superior limit. We describe below how we are going to parametrise partitions of $\mathbb{R}^N$. In order to approximate numerically a solution of Kelvin's problem we will focus on a sub-problem involving only a finite number of sets having some property of periodicity. Consider the unit cube $C = [0,1]^N$ and $(\Omega_i)_{i=1}^n$ a finite partition of C in n measurable sets which satisfy:

$$\forall i = 1 \dots n, \quad \forall x \in \partial C, \quad \chi_{\Omega_i}(x) = \chi_{\Omega_i}(\hat{x}) \tag{11}$$

where $\hat{x}$ is roughly speaking x modulus 1. More formally, $\hat{x}$ is by definition the unique element of $[0,1[^N$ which is in the class of x in $(\mathbb{R}/\mathbb{Z})^N$. To every family $(\Omega_i)_{i=1}^n$ having the property (11) we associate the set:

$$E = \mathbb{R}^N \backslash \left(\bigcup_{l \in \mathbb{Z}^N} \tau_l \left(\bigcup_{i=1}^n \partial \Omega_i \right) \right) \tag{12}$$

where τ_l is the translation of vector l. If we assume that every connected components of E is of volume $\frac{|C|}{n}$, we obtain up to an homothecy an admissible partition for Kelvin's problem. Moreover the cost F introduced by Morgan of this homothetic partition (O_i) can be easily computed and we have:

$$F((O_i)) = \frac{\mathcal{J}_n^{per}(\Omega_1, \dots, \Omega_n)}{n^{1/3}}$$

where

$$\mathcal{J}_n^{per}(\Omega_1, \dots, \Omega_n) = \mathcal{H}^{N-1}(\partial E \cap C). \tag{13}$$

Let us point out some important facts. First, every partition of $\mathbb{R}^N$ can not be described in the previous way. Nevertheless, it is clear that letting n tend to infinity, it is possible to approximate (in the sense of Morgan's cost functional) every partition by the previous construction. Second, it is not true that every family $(\Omega_i)_{i=1}^n$ of sets of volume $\frac{|C|}{n}$ which satisfies (11) produces always by (12) a set which connected components are all of volume $\frac{|C|}{n}$. A family of parallel strips may satisfy (11) and produces a set E with unbounded connected components. It is intuitively clear that this kind of partition would not be optimal for $\mathcal{J}_n^{per}$, at least for n large. We will not consider this difficulty in the following and we will observe in section 6 that those cases do not appear numerically. Finally, notice that in the definition (13), the pieces of ∂E which are included in ∂C are counted. This detail makes an important difference with the one presented in the previous section where the standard norm of the space BV was enough to compute the perimeter associated to each set $(\Omega_i)_{i=1}^n$. This technical aspect will have a major importance regarding the relaxed formulations that we will introduce in the next section.

As in the previous section, we give a rigorous mathematical formulation in a functional context of the previous construction. Let $\hat{C} = [-1, 2]^N$, and consider the space

$$\mathcal{X}_n^{per} = \{(u_i) \,|\, \forall i = 1 \dots n,\ u_i \in BV^{per}(\hat{C}, \{0,1\}) \ \int_C u_i = \frac{|C|}{n},\ \sum_{i=1}^n u_i(x) = 1 \ \text{ a.e. in } C\} \tag{14}$$

where

$$BV^{per}(\hat{C}) = \{u \in BV(\hat{C}) \,|\, u(x) = u(\hat{x}), \ \text{ a.e. x in } \hat{C}\} \tag{15}$$

and $\hat{x}$ is defined as before. In order to optimise an energy similar to (13) we define

$$J_n^{per}(u_1, \dots, u_n) = \sum_{i=1}^n \int_C |Du_i|. \tag{16}$$

Since C is a closed set, notice that the jumps of u_i which are on the boundary of C are counted in the cost (16). Based on the same arguments as the proof of theorem 27 we have the existence result:

Theorem 29 *There exists at least one family* $(u_i^*)_{i=1}^n \in \mathcal{X}_n^{per}$ *such as:*

$$\mathcal{J}_n^{per}(u_1^*, \ldots, u_n^*) = \inf_{(u_i)_{i=1}^n \in \mathcal{X}_n^{per}} \mathcal{J}_n(u_1, \ldots, u_n).$$

VII.4 Relaxation of the perimeter and Γ-convergence

The main difficulty in solving numerically problems (7) or (16) is related to the approximation of irregular functions which are characteristic functions. In order to tackle this point we introduce a relaxation of those problems based on the famous Γ−convergence result of Modica and Mortola. The main feature of this relaxation is to make it possible to approximate optimal "true partitions" in n pieces by an n-tuple of regular functions optimal for some relaxed functionals. We first recall Modica and Mortola's theorem which will be used to establish our relaxed formulations.

Theorem 30 *(L. Modica and S. Mortola see [11] and [12]) Let* $0 < V < |C|$ *and* W *a continuous positive function which vanishes only at 0 and 1 and set* $\sigma = 2\int_0^1 \sqrt{(W(u))}\, du$. *For all* $\varepsilon > 0$, *consider*

$$F^\varepsilon(u) := \begin{cases} \varepsilon \displaystyle\int_C |\nabla u|^2 + \frac{1}{\varepsilon}\int_C W(u) \;\; \textit{if} \;\; u \in W^{1,2}(C) \cap X, \\ +\infty \;\; \textit{otherwise} \end{cases} \tag{17}$$

and

$$F(u) := \begin{cases} \sigma \mathcal{H}^{N-1}(S\,u) \;\; \textit{if} \;\; u \in BV(C, \{0,1\}) \cap X, \\ +\infty \;\; \textit{otherwise} \end{cases} \tag{18}$$

where X *is the set of functions* $u \in L^1(C)$ *which satisfy* $\int_C u = V$ *and* $S\,u$ *is the set of essential singularities of* u *(see [6] or [2]). Then the functionals* F^ε Γ*-converge to* F *in* X *and every sequence of minimisers* (u_ε) *is precompact in* X *(endowed with the* L^1 *norm).*

We establish a simple relaxation of problem (7) which is easily obtained from previous theorem and [3]. Let us point out that Baldo in [3] already proposed a vectorial formulation of Modica and Mortola's result very close from our setting. The main difference between his approach and our formulation is that we only consider scalar potentials w under the additional linear constraint $\sum_{i=1}^n u_i(x) = 1$ almost everywhere. In that way we avoid to deal with polynomials of high degree which could create important difficulties from the numerical point of view.

Theorem 31 *(Relaxation of problem (7)) Consider* C *a bounded open set of* $\mathbb{R}^n$ *and* W *a continuous positive function which vanishes only at 0 and 1 and set* $\sigma = 2\int_0^1 \sqrt{(W(u))}\, du$. *For* $n \in \mathbb{N}^*$, *let* X *be the space of functions* $u = (u_i) \in L^1(C)^n$ *which satisfy* $\int_C u_i = \frac{|C|}{n}$, $\forall i = 1 \ldots n$ *and* $\sum_{i=1}^n u_i(x) = 1$ *almost everywhere* x *in* C. *For all* $\varepsilon > 0$, *consider*

$$F^\varepsilon(u) := \begin{cases} \varepsilon \displaystyle\sum_{i=1}^n \int_C |\nabla u_i|^2 + \frac{1}{\varepsilon}\sum_{i=1}^n \int_C W(u_i) \;\; \textit{if} \;\; u \in (W^{1,2}(C))^n \cap X, \\ +\infty \;\; \textit{otherwise} \end{cases} \tag{19}$$

and

$$F(u) := \begin{cases} \sigma \sum_{i=1}^{n} \mathcal{H}^{N-1}(S\,u_i) \;\; \textit{if} \;\; u \in BV(C,\{0,1\})^n \cap X, \\ +\infty \;\; \textit{otherwise} \end{cases} \tag{20}$$

where $S\,u_i$ is the set of essential singularities of u_i. Then the functionals F^ε Γ-converge to F in X and every sequence of minimisers u^ε is precompact in X (endowed with the L^1 norm).

Proof. We follow the classical proof of Modica and Mortola. First we establish the compactness part of the theorem: suppose that (u^ε) is a sequence of minimisers of the functionals F^ε. For each $i = 1 \dots n$, we apply the compactness result of theorem 30 to the sequence u_i^ε. Classically, the precompacity of each components of the sequence u^ε gives the precompacity of the sequence (u^ε) by a diagonal argument.

As in the standard proof we decompose the Γ-convergence results into two steps: let (u^ε) converging in X to u. We have to show first that

$$\liminf F^\varepsilon(u^\varepsilon) \geq F(u).$$

Again we apply theorem 30 to each sequence u_i^ε for $i = 1 \dots n$. Since the lim inf of a finite sum is greater than the sum of the lim inf of each sequence, we have

$$\begin{aligned} \liminf F^\varepsilon(u^\varepsilon) &= \liminf \sum_{i=1}^{n} \left(\varepsilon \int_C |\nabla u_i|^2 + \frac{1}{\varepsilon} \sum_{i=1}^{n} \int_C W(u_i) \right) \\ &\geq \sum_{i=1}^{n} \liminf \varepsilon \int_C |\nabla u_i|^2 + \frac{1}{\varepsilon} \sum_{i=1}^{n} \int_C W(u_i) \\ &\geq F(u). \end{aligned} \tag{21}$$

Finally, let us prove that every value obtained by the Γ-limit can be approximated by a sequence of values obtained by F^ε. Let $u \in BV(C,\{0,1\})^n \cap X$, we look for a sequence $(u^\varepsilon) \subset (W^{1,2}(C))^n \cap X$ such as

$$\limsup F^\varepsilon(u^\varepsilon) \leq F(u).$$

This none trivial regularisation of a partition can be constructed with the same ideas as Baldo's in [3]. The main point is to restrict the study to polygonal partitions of finite perimeter which satisfy the same volume constraints. More precisely, for all $u \in BV(C,\{0,1\})^n$ and for all $i = 1 \dots n$ we define $S_i = u_i^{-1}(1/2)$. The family S_i is sometimes called a Caccioppoli partition that is a partition of C into sets (S_i) of finite perimeters. From [3] lemma 3.1, we deduce that there exists a sequence of polygonal partitions (S_i^ε) such as $\forall i = 1 \dots n$,

- $|S_i^\varepsilon| = \frac{|C|}{n}$,
- $\mathcal{H}^{N-1}(\partial S_i^\varepsilon \cap \partial C) = 0$,
- $\mathcal{H}^{N-1}(\partial S_i^\varepsilon \cap \partial C) \to \mathcal{H}^{N-1}(\partial S_i \cap \partial C)$ when $\varepsilon \to 0$.

Now, for a given polygonal partitions we can use a standard regularisation process (see [12] or [3]) to construct a sequence (u^ε) which satisfies the volume constraints, the equality $\sum_{i=1}^n u_i^\varepsilon(x) = 1$ almost everywhere x in C and also the inequality

$$\limsup F^\varepsilon(u^\varepsilon) \leq F(u). \tag{22}$$

The inequalities (21) and (22) prove the Γ-convergence. □

We now give a relaxation result for the periodic case:

Theorem 32 *(Relaxation of problem (16)) Consider $C = [0,1]^n$, $\hat{C} = [-1,2]^n$ and W a continuous positive function which vanishes only at 0 and 1 and set $\sigma = 2\int_0^1 \sqrt{(W(u))}\,du$. For $n \in \mathbb{N}^*$, let X be the space of functions $u = (u_i) \in L^1(C)^n$ which satisfy $\int_C u_i = \frac{|C|}{n}$, $\forall i = 1 \dots n$ and $\sum_{i=1}^n u_i(x) = 1$ for almost everywhere x in C. For all $\varepsilon > 0$, consider*

$$F^\varepsilon(u) := \begin{cases} \varepsilon \sum_{i=1}^n \int_C |\nabla u_i|^2 + \frac{1}{\varepsilon} \sum_{i=1}^n \int_C W(u_i) \;\; \text{if} \;\; u \in (W^{1,2}(C))^n \cap X,\; \hat{u} \in (W^{1,2}(\hat{C}))^n \\ +\infty \;\; \text{otherwise} \end{cases} \tag{23}$$

and

$$F(u) := \begin{cases} \sigma \sum_{i=1}^n \int_C |Du_i| \;\; \text{if} \;\; u \in BV(C,\{0,1\})^n \cap X,\; \hat{u} \in BV(\hat{C},\{0,1\})^n \\ +\infty \;\; \text{otherwise} \end{cases} \tag{24}$$

where $S\,u_i$ is the set of essential singularities of u_i and $\hat{u}$ is the 1-periodic extension of u to $\hat{C}$. Then the functionals F^ε Γ-converge to F in X and every sequence of minimisers (u^ε) is precompact in X (endowed with the L^1 norm).

Proof. Let (u^ε) be a sequence of minimisers for functionals F^ε. As in the previous theorem, we use the compactness part of theorem 30 applied to the sequence of 1-periodic extensions $(\hat{u}^\varepsilon)$ to obtain the precompactity in X. Now we consider (u^ε) converging in X to u. We want to prove that:

$$\liminf F^\varepsilon(u^\varepsilon) \geq F(u).$$

Notice that this fact is not an immediate consequence of theorem 30. The main difference comes from the fact that the jumps of u on ∂C are counted in the cost functional F. The idea is to move a little bit the set C in order to avoid this "bad" situation and then apply the standard Modica-Mortola's theorem. We first establish that up to a small translation of vector a, the measure $D\hat{u}$ has a support intersected with $a + \partial C$ which is negligible with respect to the $\mathcal{H}^{N-1}$ measure. Since u is a characteristic function of a set of finite perimeter, the structure theorem on the reduced boundary (which is exactly the jump set of u) claims that the measure $D\hat{u}$ has a support which is contained (up to a set of 0 $\mathcal{H}^{N-1}$ measure) in a union of countable C^1 compact hypersurfaces. Let $\delta > 0$, F_a be a face of the cube C of normal vector a and E one of those smooth hypersurfaces. Since F_a and E are both manifolds of dimension $N-1$ we can apply a classical consequence of Thom's transversality theorem which asserts that for almost all δ the two manifolds $F_a + \delta n_a$ and E are transverse (see [5] for instance). As a consequence $(F_a + \delta n_a) \cap E$ is an empty set or a smooth manifold of dimension exactly $N-2$. Then $(F_a + \delta n_a) \cap E$ is negligible with respect to the measure $\mathcal{H}^{N-1}$ for almost all $\delta > 0$. We can apply the previous arguments to each hypersurface which covers the support of $D\hat{u}$ and to all the faces of C. In that way we prove that there exists a vector a such as

$$\begin{cases} (C + a) \subset \hat{C} \\ \int_{\partial(C+a)} |Du| = 0. \end{cases} \tag{25}$$

Now setting $C_a = C + a$, we have

$$\begin{aligned}
\liminf F^\varepsilon(u^\varepsilon) &= \liminf \varepsilon \sum_{i=1}^n \int_C |\nabla u_i^\varepsilon|^2 + \frac{1}{\varepsilon} \sum_{i=1}^n \int_C W(u_i^\varepsilon) \\
&= \liminf \varepsilon \sum_{i=1}^n \int_{C_a} |\nabla u_i^\varepsilon|^2 + \frac{1}{\varepsilon} \sum_{i=1}^n \int_{C_a} W(u_i^\varepsilon) \\
&\geq \sum_{i=1}^n \int_{C_a} |Du_i| \\
&= \sum_{i=1}^n \int_{\bar{C}_a} |Du_i| \\
&= \sum_{i=1}^n \int_{\bar{C}} |Du_i|
\end{aligned}$$

where the second and the last equalities are a consequence of the periodicity of the functions (u_ε) and u. The inequality is obtained using the lim sup part of the theorem 30 applied to the open set C_a and the third equality comes from (25).

The lim sup part of the proof can be established exactly with the same ideas as in the non-periodic case. The only difference is that the elements of the sequence must be in $W^{1,2}(\hat{C})^n$, which can be achieved with very small modifications of the energy F_ε associated to the element. □

VII.5 The minimisation algorithm

The two previous theorems have two major advantages to approximate optimal partitions. First it makes it possible to work with regular functions under linear constraints. Additionally, it gives us the opportunity to replace a strongly not convex problem by a smooth sequence of optimisation problems depending of ε which are close from being convex for $\varepsilon >> 1$. We base our optimisation strategy on this observation. We start to solve the relaxed problems (19) or (23) with ε large. Since in this case those problems are almost convex, we can expect to find by standard descent method a good approximation u_ε of the solution. Then we increase the value of ε step by step and solve the new optimisation problems starting the optimisation process with the previous numerical solution. Observe that our strategy does not give any warranty to identify in the end of the process a global optima of the original problem since branching in a wrong direction may occur when ε tends to 0. Nevertheless, we observe in our experiments that this approach is surprisingly efficient for our problems.

Based on the above ideas we can now describe our optimisation algorithm. In order to simplify the notations we restrict our description to the dimension $N = 2$ and $C = [0, 1]^2$. It is straightforward to adapt our method to the case $N = 3$. We decompose the domain C into a M^2 grid with spacing $h = 1/(M-1)$. Consider a renumbering operator $K : (0, M-1)\times(0, M-1) \mapsto (0, M^2-1)$ such $K(k, l) = lM + k$. Our unknowns are the components of the discrete fields $(U_i^\varepsilon)_{k,l}$ as $(U_i^\varepsilon)_{K(k,l)}$ (which we abbreviate as $(U_i^\varepsilon)_K$ when there is no risk of confusion) depending on whether we want to insist on the spatial relation between the components. We approximate the gradient of functions u_i^ε by standard first order finite difference operators δ_x and δ_y, defined for any discrete vector field

U by:

$$[\delta_x U]_{k,l} = \frac{U_{k+1,l} - U_{k,l}}{h}, \tag{26}$$

$$\left[\delta_y U\right]_{k,l} = \frac{U_{k,l+1} - U_{k,l}}{h}. \tag{27}$$

If the index (k, l) corresponds to a boundary point, the previous gradient is computed considering the boundary conditions of the problem. In the case of a bounded domain we simply use Dirichlet conditions whereas in the torus case we use the periodicity of the grid. The discretisation of cost functionals (19) and (23) are directly deduced from the expression (26). Let us call F_d^ε that discrete cost functional.

To complete the description of our discretisation we describe now the linear constraints imposed on the discrete values $(U_i^\varepsilon)_{k,l}$. On one hand we have the volume constraints imposed on the functions u_i^ε

$$\sum_{k,l} (U_i^\varepsilon)_{K(k,l)} = \frac{M^2}{n}, \ \forall i = 1 \dots n, \tag{28}$$

and the pointwise non-overlapping constraints

$$\sum_{i} (U_i^\varepsilon)_{K(k,l)} = 1, \ \forall k, l = 0 \dots M-1. \tag{29}$$

Let us denote by Π the linear projection operator on the constraints (28) and (29). More precisely, regarding the unknown as an array of size $M^2 \times n$, the constraints on that array $(a_{i,j})$ may be written:

$$\begin{cases} \sum_j a_{i,j} = c_i \ \forall i = 1 \dots n \\ \sum_i a_{i,j} = d_j \ \forall j = 0 \dots M^2 - 1 \end{cases} \tag{30}$$

where $c_i = 1$ for all $i = 1 \dots n$ and $d_j = \frac{M^2}{n}$ for all $j = 1 \dots M^2$. Let us note that the previous constraints must satisfy the compatibility condition

$$\sum_i c_i = \sum_j d_j \tag{31}$$

which is true in our case since $\sum_i c_i = M^2$ and $\sum_j d_j = n\frac{M^2}{n} = M^2$. One consequence of the previous compatibility condition is that the set of all $n + M^2$ constraints of (30) is not of maximal rank. It is not difficult to see that keeping the $n + (M^2 - 1)$ first constraints gives a free system of constraints.

We describe in the first Algorithm a few step to compute in an efficient way the projected array $(b_{i,j}) := \Pi((a_{i,j}))$ when $n << M^2$ for any fixed vectors (c_i), (d_j) which satisfy (31). Notice that the more time consuming step in the previous algorithm is the resolution of the linear system $C|_{(n-1)\times(n-1)}(\lambda_j)|_{n-1} = (d_j)|_{n-1}$ which is only of size $(n-1)^2$. In all the experiments that we carried out, n was always less than $1e2$ which leads to a fast projection algorithm.

To finish our description, we give the successive steps of our optimisation in the second Algorithm (we refer to [10] for technical details on the conjugated gradient algorithm and the choice of the line search methods).

Algorithm 2 Projection on the linear constraints

1. $(e_i) := (2\sum_j a_{i,j} - 2c_i)$
2. $(f_j) := (2\sum_i a_{i,j} - 2d_j)$
3. Define the matrix $C = (c_{k,l})$ of size $n \times n$ by
$$\begin{cases} c_{k,l} = -\frac{M^2}{n} \text{ if } k \neq l \\ c_{k,k} = M^2 - \frac{M^2}{n} \end{cases}$$
4. $(d_j) := (f_j) - \frac{2}{n}\sum_i e_i$
5. Compute the unique vector (λ_j) of size $n\times 1$ with $\lambda_n = 0$ such as $C|_{(n-1)\times(n-1)}(\lambda_j)|_{n-1} = (d_j)|_{n-1}$ where the notation $C|_{(n-1)\times(n-1)}$ stands for the matrix of size $(n-1)\times(n-1)$ obtained from C by extracting the $n-1$ first rows and $n-1$ first columns. The definitions of $(\lambda_j)|_{n-1}$ and $(d_j)|_{n-1}$ are similar.
6. $S := \sum_j \lambda_j$
7. $(\eta_i) := \frac{(e_i)-S}{n}$
8. $A_{\text{orth}} := (\eta_i) * 1_{1\times n} + 1_{M^2\times 1} * \text{Transpose}((\lambda_j))$ where $1_{k\times l}$ is the matrix of size $k \times l$ which coefficients are all equal to 1 and $*$ is the standard matrix multiplication.
9. $B := A - A_{\text{orth}}$

Algorithm 3 Numerical optimisation by Γ-convergence

Require: $\varepsilon_{\text{initial}}$, $\varepsilon_{\text{final}}$, $(U_i^{\varepsilon_{\text{initial}}})$, ω, $\delta > 1$ (tolerance)

1: $\varepsilon := \varepsilon_{\text{initial}}$, $(U_i^{\varepsilon}) := (U_i^{\varepsilon_{\text{initial}}})$
2: **repeat**
3: Compute (V_i^{ε}) the solution of $\min F_d^{\varepsilon}((V_i))$ among arrays (V_i) which satisfy constraints (28) and (29) (up to a tolerance δ). This step is carried out by a standard projected conjugated gradient algorithm (based on the previous projection algorithm) starting from (U_i^{ε}).
4: $(U_i^{\varepsilon/\omega}) := (V_i^{\varepsilon})$, $\varepsilon{:=}\varepsilon/\omega$
5: **until** $\varepsilon > \varepsilon_{\text{final}}$

n	Morgan's cost, see (10)	n	Bounded convex polyhedra C	Morgan's cost
8	2.644175	6	Truncated octahedron	2.852505
16	2.653171	10	Truncated octahedron	2.924930
20	2.655404	6	Rhombic dodecahedron	2.934629
21	2.657727	8	Truncated octahedron	2.942078
22	2.666318	8	Rhombic dodecahedron	2.945360
12	2.671376	10	Rhombic dodecahedron	2.956432
17	2.675445	4	Rhombic dodecahedron	2.984274
19	2.680236	2	Rhombic dodecahedron	2.987346
18	2.681586	2	Truncated octahedron	3.004914
13	2.683315	3	Truncated octahedron	3.009927
15	2.689541	4	Truncated octahedron	3.014228
10	2.692954	4	Hexagonal prism	3.021674
9	2.693281	6	Hexagonal prism	3.051920
14	2.694757	8	Triangular prism	3.061425
11	2.695891	2	Hexagonal prism	3.078461

Table VII.1: Optimal values for the periodic case (2 first columns) and different polyhedral cuttings (three last columns).

Finally, if the domain C is not a square or a cube, we simply consider a squared or cubic domain which contains C and impose the additional Dirichlet constraints:

$$(U_i)_K = 0, \ \forall i = 1 \dots n$$

if K corresponds to a grid point which is outside of C. The previous algorithms are easily adapted to this more general situation.

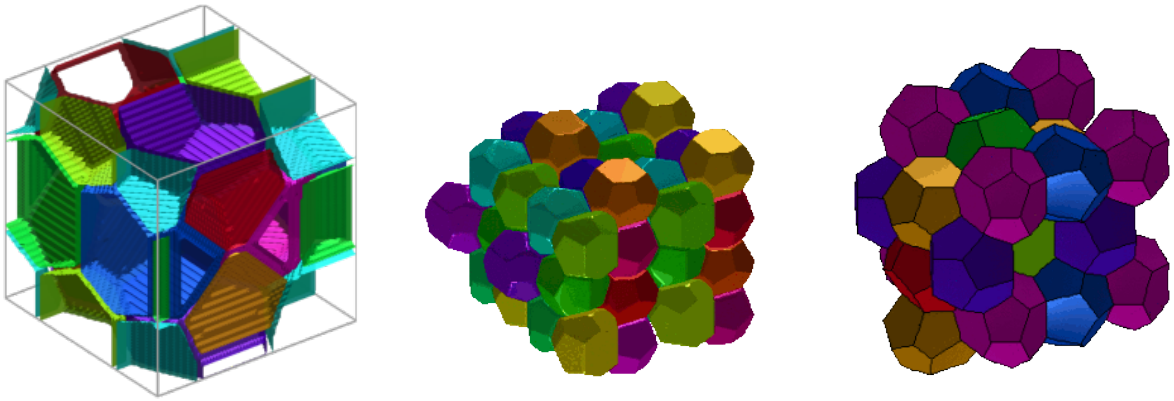

Figure VII.1: Switching from a density representation to a boundary description

VII.6 Numerical results

We were able to run a series of large computations on 2D and 3D problems. We first address problem (1) when C is a disk (see figure VII.2) and a triangle (figure VII.3). All the 2D computations have been done on a grid of dimension (253×253). We set $\varepsilon_{\text{initial}} = 1$, $\varepsilon_{\text{final}} = 1e-3$, the tolerance parameter $\delta = 1e-6$ and $\omega = 1.1$. We always start our optimisation process with an array $(U_i^{\varepsilon_{\text{initial}}})$ made of uniform random values in $[0, 1]$. As expected, our numerical solutions are made of local patches satisfying the 120 degrees angular conditions. Moreover some symmetries of the set C are preserved for small values of n.

We performed 3D computation for problem (13) with n from 8 to 21 (see figure VII.4) on grids of dimension $(128 \times 128 \times 128)$. As a post treatment, we used the very efficient local optimisation software "Evolver" (see [4]) developed by Ken Brakke to obtain a finer description of optimal tilings. Let us point out that most of the geometrical structure was already contained in the parametrisation of the tiling given by the density functions (U_i) at the end of our algorithm. In figure VII.1 we represent in the first picture the level sets $\{U_i = \frac{1}{2}\}$ for $i = 1 \dots n$. In the second picture we draw the periodic reconstruction of the densities without any surface optimisation. Notice that a small gap remains between the level sets. In the last picture, we display the result of the optimisation performed by "Evolver".

With $n = 16$ we observe that we obtain Kelvin's tiling only made of truncated octahedra. With $n = 8$, starting again from a complete random array, we recover the famous tiling obtained by D. Weaire and P. Phelan which is made of exactly two kinds of cells. We give below the values corresponding to the cost functional ...for $n = 8$ to 21. Unfortunately we were not able to find a better tiling than the one discovered by D. Weaire and P. Phelan.

Finally, we tried to beat Weaire and Phelan's tiling by considering optimal cutting of sets C which already tile the space. Namely, we approximated optimal cuttings of a truncated octahedron, a triangular prism, a rhombic dodecahedron and one hexagonal prism (see figure VII.5). We then computed the cost (13) associated to the tilling deduced from the previous optimal cutting. The array below sum up the optimal values in the periodic and non-periodic cases of the functional.

We sum up our results in table VII.1. The first column gives different values of Morgan's cost functional obtained by the periodic tilings and the second one gives the values obtained by the optimal cutting of sets which already tile the space. We observe that none of such tiling gave a better cost than the ones obtained by periodic boundary conditions.

Acknowledgements

Alessandro Giacomini is warmly thanked for his help in the proof of the relaxation result of theorem 32.

Professor K. Brakke contributed to the current version of the paper in many fruitful comments on the use of his software "Evolver".

Bibliography

[1] G. Alberti. Variational models for phase transitions, an approach via Γ-convergence. In *Calculus of variations and partial differential equations (Pisa, 1996)*, pages 95–114. Springer,

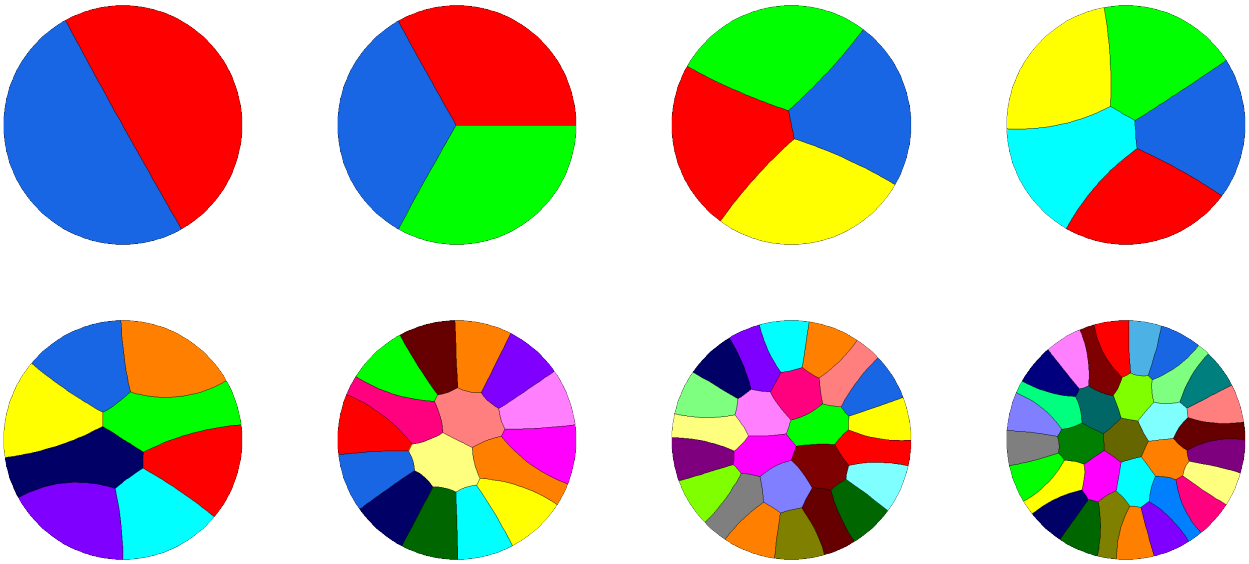

Figure VII.2: Tiling of the disk with 2, 3, 4, 5, 8, 16, 24, 32 cells

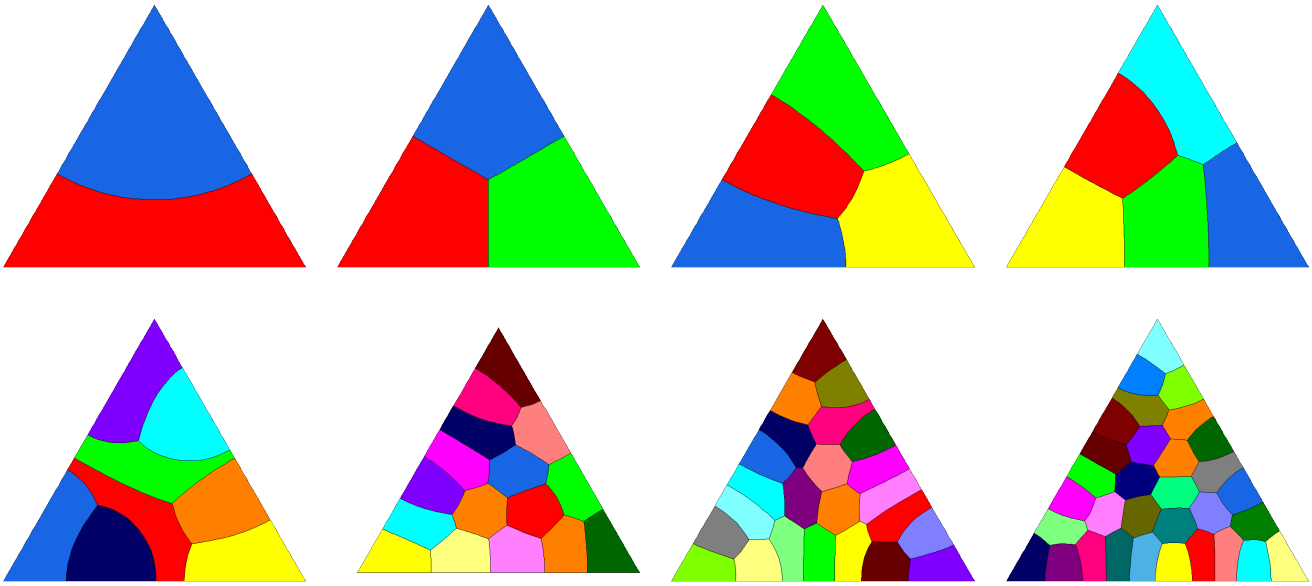

Figure VII.3: Tiling of the triangle with 2, 3, 4, 5, 8, 16, 24, 32 cells

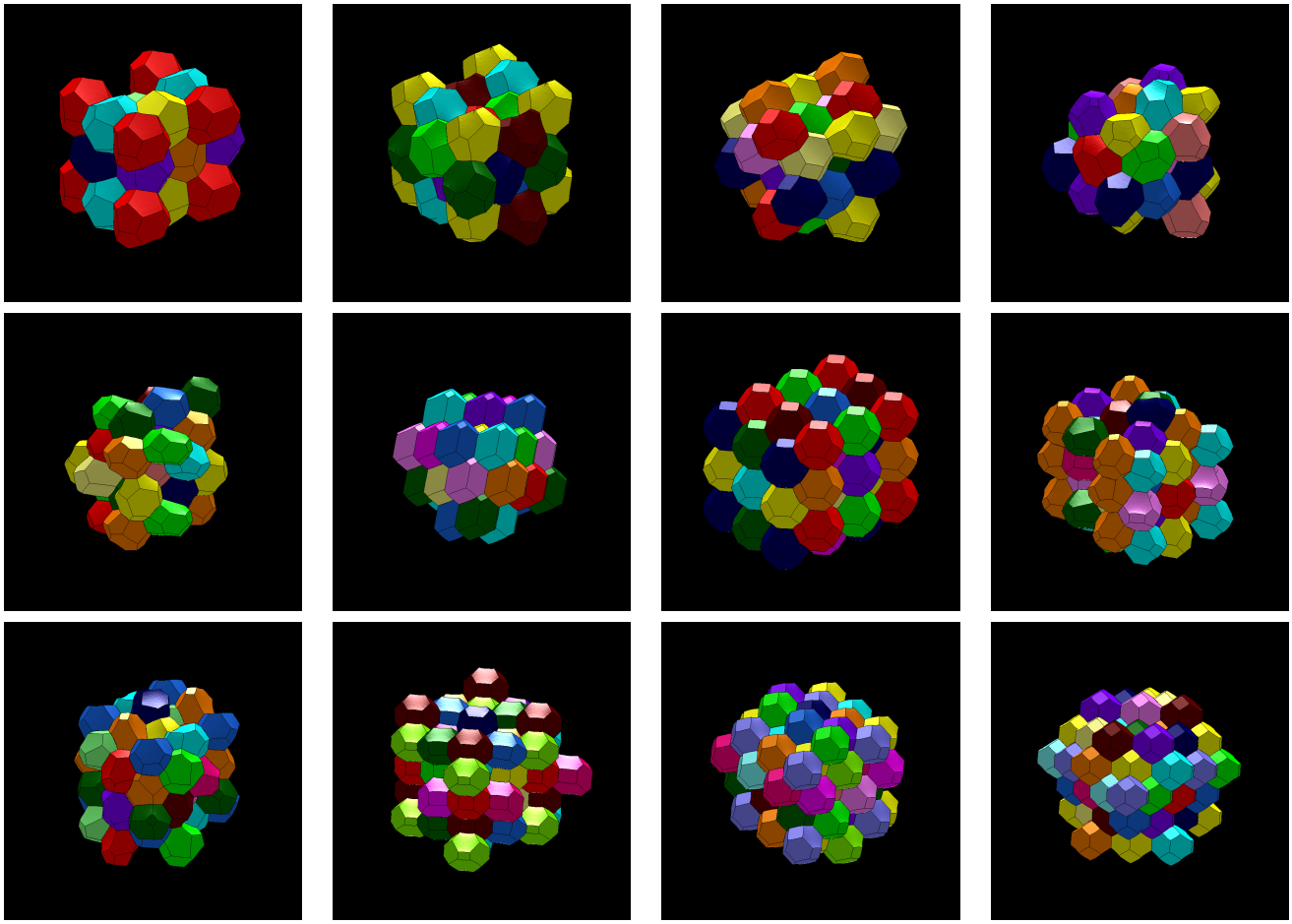

Figure VII.4: Periodic tilings of the space by 8, 10, 12, 13, 14, 15, 16, 17, 18, 19 20, 21 cells

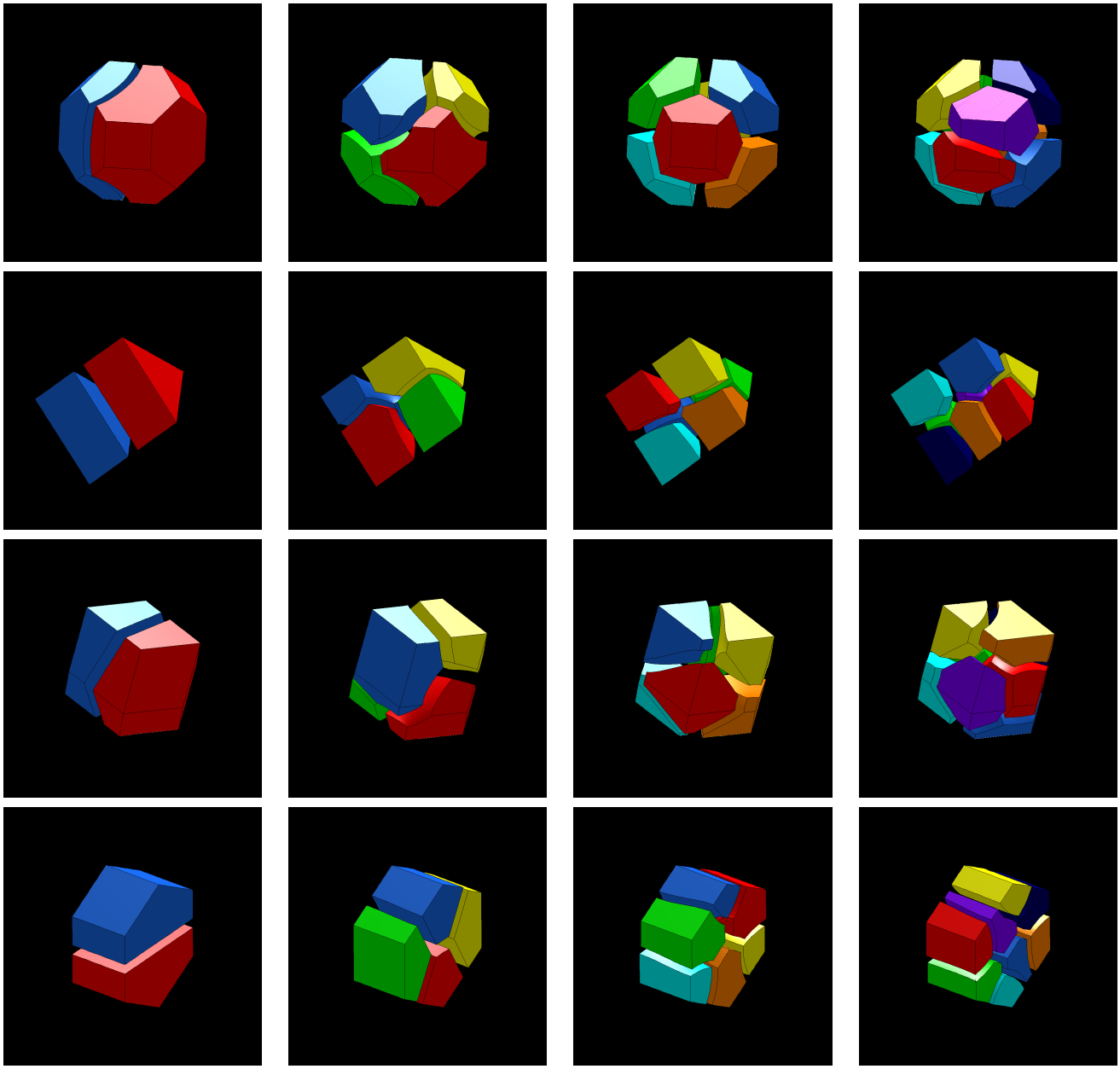

Figure VII.5: Non-periodic tilings

Berlin, 2000.

[2] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[3] Sisto Baldo. Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 7(2):67–90, 1990.

[4] Kenneth A. Brakke. The surface evolver. *Experiment. Math.*, 1(2):141–165, 1992.

[5] Michel Demazure. *Bifurcations and catastrophes*. Universitext. Springer-Verlag, Berlin, 2000. Geometry of solutions to nonlinear problems, Translated from the 1989 French original by David Chillingworth.

[6] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[7] Morgan .F. Existence of least-perimeter partitions. *Philos. Mag. Lett.*, 2008.

[8] T. C. Hales. The honeycomb conjecture. *Discrete Comput. Geom.*, 25(1):1–22, 2001.

[9] Plateau Joseph. *Statique Expérimentale et Théorique des Liquides soumis aux Seules Forces Moléculaires*. 1873. Translated by K. Brakke, `http://www.susqu.edu/brakke/aux/downloads/Plateau-Fr.pdf`.

[10] C. T. Kelley. *Iterative methods for optimization*, volume 18 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[11] Luciano Modica. The gradient theory of phase transitions and the minimal interface criterion. *Arch. Rational Mech. Anal.*, 98(2):123–142, 1987.

[12] Luciano Modica and Stefano Mortola. Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.

[13] Frank Morgan. *Geometric measure theory*. Elsevier/Academic Press, Amsterdam, fourth edition, 2009. A beginner's guide.

[14] Kelvin William Thomson. On the division of space with minimum partitional area. *Philos. Mag. Lett.*, 24(151):503, 1887. `http://zapatopi.net/kelvin/papers/on_the_division_of_space.html`.

[15] D. Weaire and R. Phelan. A counter-example to Kelvin's conjecture on minimal surfaces. *Forma*, 11(3):209–213, 1996. Reprint of Philos. Mag. Lett. **6**9 (1994), no. 2, 107–110.

Troisième partie

Transport optimal et irrigation optimale

An optimization problem for mass transportation with congested dynamics

G. Buttazzo , C. Jimenez & E. Oudet

VIII.1 Introduction

Mass transportation theory received much attention in the mathematical community in the last years. Starting from the initial setting by Monge where, given two mass densities ρ_0 and ρ_1, a *transport map* $T : \mathbb{R}^d \to \mathbb{R}^d$ was searched among the admissible maps transporting ρ_0 onto ρ_1 in order to minimize the total transportation cost

$$\int_{\mathbb{R}^d} |x - T(x)|\, d\rho_0(x) \ ,$$

several other equivalent formulations have been provided (see for instance [24], [18], [5]). In particular, the formulation given in [10] is the one which motivated our study: the goal in [10] was to introduce a "dynamic" formulation of the mass transportation problem providing a map $\rho : [0,1] \to \mathcal{P}(\overline{\Omega})$ which describes the motion of ρ_0 onto ρ_1 as a function of a parameter $t \in [0,1]$, where Ω is the space constraint that all the densities $\rho(t,\cdot)$ have to fulfill.

The set of applications of mass transportation theory is also very rich: many urban planning models have been studied, searching e.g. for the best design of public transportation networks (see [9], [11]), for the optimal pricing policies of their use (see [12]), for the best distribution of residential and working areas in a city (see [13]). We also mention the strict link between mass transportation theory and shape optimization in elasticity, as was shown in [7], [5].

The general framework we consider is the one of functionals defined on the space of measures acting on a time-space domain $\overline{Q} \subset \mathbb{R}^{1+d}$; the minimization problem we are interested in is then written in the form

$$\min \left\{ \Psi(\sigma) \ : \ -\mathrm{div}\sigma = f \text{ in } \overline{Q},\ \sigma \cdot \nu = 0 \text{ on } \partial Q \right\} \tag{1}$$

where Ψ is an integral functional on the $\mathbb{R}^{1+d}$-valued measures defined on $\overline{Q}$. Writing $\sigma = (\rho, E)$ the classical Monge case is then related to the cost function

$$\Psi(\sigma) = \int_{\overline{Q}} d|E| \ ,$$

while the case considered by Brenier in [10] is represented by the cost function

$$\Psi(\sigma) = \int_{\overline{Q}} \left| \frac{dE}{d\rho} \right|^2 d\rho \ .$$

As shown in [10], [2], [22] all these cases are related to the Wasserstein distances $W_p(\rho_0, \rho_1)$, where each particle x in the source ρ_0 moves to its final point $T(x)$ in the target ρ_1 following a line segment, or a geodesic line in case the space constraint Ω is not convex. However, in many problems where a high number of particles (or a probability density) is involved, other effects are present which may deviate the trajectories from straight lines: in particular we are interested in the *congestion* effects that occur when the density $\rho(t, x)$ is high, slowing the ideal mass transportation and increasing the cost.

Modelling the congestion effects has been considered by several authors (see for instance [15], [23]); here we simply consider the Brenier formulation (1) assuming that the functional Ψ has a term which has a superlinear growth with respect to ρ.

In Sections 2 and 3 we discuss the general formulation (1) and its dual problem, with the primal-dual optimality conditions. In Section 4 we provide a numerical scheme to treat this kind of problems: the scheme is based on the one by Benamou and Brenier [2], adapted to include the congestion terms. In the cases we present the domain Ω is always nonconvex, having some obstacles at its interior, and the mass moves from ρ_0 onto ρ_1 according to:

- the Wasserstein distance W_2, so minimizing the cost $\int_{\overline{Q}} \left|\frac{dE}{d\rho}\right|^2 d\rho$;
- the Wasserstein distance W_2 with the addition of the congestion term $\int_{\overline{Q}} \rho^2 \, dt \, dx$;
- the Wasserstein distance W_2 with the addition of the constraint $\{\rho \leq M\}$ which for instance occurs when a crowd of individuals moves and two different individuals cannot stay too close.

VIII.2 The general setting

In this section we consider an open bounded subset Q of $\mathbb{R}^{d+1}$ ($d \geq 1$). We assume Q has a Lipschitz boundary and denote by $\nu(x)$ the outward pointing normal vector to x in the boundary ∂Q of Q, defined almost everywhere. Let $\mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$ be the space of vectorial Borel measures supported on $\overline{Q}$.
We also consider a functional Ψ on $\mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$ and we assume Ψ is lower semicontinuous for the weak* convergence of measures.
Let $f \in \mathcal{M}(\overline{Q})$ be a Borel measure of zero total mass that is $\int_{\overline{Q}} df = 0$. We deal with the following optimization problem:

$$\inf_{\sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})} \Psi(\sigma) \tag{2}$$

with the constraint:

$$\begin{cases} -\operatorname{div} \sigma &= f \quad \text{in } \overline{Q} \\ \sigma \cdot \nu &= 0 \quad \text{on } \partial Q. \end{cases} \tag{3}$$

The condition (3) is intended in the weak sense i.e. for every $\varphi \in C^1(\overline{Q})$:

$$\int_{\overline{Q}} D\varphi \cdot d\sigma(x) = \int_{\overline{Q}} \varphi(y) \, df(y). \tag{4}$$

The following general existence result holds:

Theorem 33 *Let $\Psi : \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1}) \to [0, +\infty]$ be lower semicontinuous for the weak* convergence and such that:*

$$\Psi(\sigma) \geq C|\sigma|(\overline{Q}) - \frac{1}{C} \quad \forall \sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1}) \tag{5}$$

for a suitable constant $C > 0$, where $|\sigma|$ denotes the total variation of σ. We assume that $\Psi(\sigma_0) < +\infty$ for at least one measure σ_0 satisfying (3). Then the problem

$$\min \left\{ \Psi(\sigma) : \ -\mathrm{div}\, \sigma = f \text{ in } \overline{Q},\ \sigma \cdot \nu = 0 \text{ on } \partial Q \right\} \tag{6}$$

admits a solution. Moreover if Ψ is strictly convex, this solution is unique.

Proof. Let $(\sigma_n)_{n\in\mathbb{N}}$ be a minimizing sequence for problem (6). By assumption (5), this sequence is bounded and by consequence it admits a subsequence $(\sigma_{n_k})_{k\in\mathbb{N}}$ which converges weakly* to a measure $\sigma \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^{d+1})$. By writing the constraint (4) for any σ_{n_k} and passing to the limit as $k \to +\infty$, we get the admissibility of σ. Then, by the lower semicontinuity of Ψ, we get

$$\inf(6) = \lim_{k\to+\infty} \Psi(\sigma_{n_k}) \geq \Psi(\sigma)$$

which shows that σ is a solution of (6). □

In case Ψ is convex, problem (6) also admits a dual formulation. Indeed, if $A : C(\overline{Q}) \to C(\overline{Q}, \mathbb{R}^{d+1})$ denotes the operator given by:

$$A(\varphi) = D\varphi \text{ for all } \varphi \text{ in its domain } C^1(\overline{Q}),$$

we have the convex analysis formula for the dual formulation of (6) (see [6]):

$$\begin{aligned} (\Psi^* \circ A)^*(f) &= \min_{\sigma} \left\{ \Psi(\sigma) : \ -\mathrm{div}\, \sigma = f \text{ in } \overline{Q},\ \sigma \cdot \nu = 0 \text{ on } \partial Q \right\} \\ &= \sup \left\{ \int_{\overline{Q}} \varphi(x)\, df(x) - \Psi^*(D\varphi) : \quad \varphi \in C^1(\overline{Q}) \right\}. \end{aligned} \tag{7}$$

This formula holds if Ψ^* is continuous at least at a point of the image of A.
For any set C, we denote by χ_C the function which is 0 inside C and $+\infty$ outside. The primal-dual optimality condition then reads as

$$\min \Psi(\sigma) + \chi_{\left\{ \begin{smallmatrix} -\mathrm{div}\sigma = f \text{ in } \overline{Q} \\ \sigma \cdot \nu = 0 \text{ on } \partial Q \end{smallmatrix} \right\}} = \max \int \varphi\, df(x) - \Psi^*(D\varphi)$$

which, if a solution φ_{opt} of (7) exists, yields

$$\int D\varphi_{\mathrm{opt}} \cdot d\sigma_{\mathrm{opt}} = \Psi(\sigma_{\mathrm{opt}}) + \Psi^*(D\varphi_{\mathrm{opt}}) \tag{8}$$

where σ_{opt} is any solution of (6). The point is that, in general, the maximizers φ_{opt} in (7) are not in $C^1(\overline{Q})$. As we will see in the next section, for a large class of cost functions Ψ, (7) can be relaxed so that the primal-dual optimality condition will be explicitly identified.

VIII.3 The Transportation model

In order to introduce a model for the description of the dynamics of a crowd in a given domain, it is convenient to particularize the framework above as follows:

$Q =$ $]0,1[\times\Omega$ where Ω is a bounded Lipschitz open subset of $\mathbb{R}^d$ with outward normal vector denoted by ν_Ω. The set Ω represents the domain the crowd is constrained to stay inside, including possible obstacles that cannot be crossed. The current variable in Q will be denoted by (t,x) $(t\in]0,1[,\ x\in\Omega)$.

$\sigma =$ (ρ,E) where $\rho(t,x)$ represents the mass density at position x and time t and E is the flux at (t,x). In the usual mass transportation cases we have $E\ll\rho$ so that $E=\rho v$ being $v(t,x)$ the velocity field at (t,x). We assume the constraint $\rho\geq 0$ so that the set of admissible variables is:

$$\mathcal{D} := \{(\rho,E):\ \rho\in\mathcal{M}_b(\overline{Q},\mathbb{R}^+), E\in\mathcal{M}_b(\overline{Q},\mathbb{R}^d)\}.$$

$f =$ $\delta_1(t)\otimes\rho_1(x)-\delta_0(t)\otimes\rho_0(x)$ where $\rho_0(x),\rho_1(x)$ represent the crowd densities at $t=0$ and $t=1$ respectively, both prescribed as probabilities on $\overline{\Omega}$. Then equation (3) reads as:

$$\begin{cases} -\partial_t\rho-\mathrm{div}_x E=0 \text{ in } \overline{Q}\\ \rho(0,x)=\rho_0(x),\quad \rho(1,x)=\rho_1(x),\\ E\cdot\nu_\Omega=0 \text{ on }]0,1[\times\partial\Omega \end{cases} \tag{9}$$

as it is easy to see using the weak formulation (4). Note that (9) is the continuity equation of our mass transportation model.

Our problem is then

$$\min\{\Psi(\rho,E):\ (\rho,E) \text{ verifies (9)}\}$$

and we denote by $\mathcal{W}_\Psi(\rho_0,\rho_1)$ its minimal value.
We may deduce from (9) that for a.e. $t\in]0,1[$, $\rho(t,\cdot)$ is a probability on $\overline{\Omega}$. Indeed, disintegrating the measure ρ on $\overline{Q}$ we obtain

$$\rho(t,x)=m(t)\otimes\rho^t(x)$$

where m is the marginal of ρ with respect to t and $\rho^t(\cdot)$ is a probability for $m-$a.e. $t\in[0,1]$. Taking in (9) a test function $\alpha(t)\in C^1_c(Q)$ depending only on t we have

$$0=\int_{\overline{Q}}\alpha'(t)\,d\rho(t,x)=\int_0^1\alpha'(t)\,dm(t)$$

which gives $m=cdt$ for a suitable constant c. Using the conservation of the mass gives that $c=1$.

We now discuss the choice of Ψ. We may take for Ψ any local lower semicontinuous function on $\mathcal{M}_b(\overline{Q},\mathbb{R}^{d+1})$. By the results that can be found in [3] and [4], these functions can be represented in the following form:

$$\Psi(\sigma)=\int_{\overline{Q}}\psi\left(\frac{d\sigma}{dm}\right)dm+\int_{\overline{Q}\setminus A_\sigma}\psi^\infty\left(\frac{d\sigma^s}{d|\sigma^s|}\right)d|\sigma^s|+\int_{A_\sigma}g(\sigma(x))\,d\sharp(x)$$

where

- m is a positive non-atomic Borel measure on Q;
- $d\sigma/dm$ is the Radon-Nikodym derivative of σ with respect to m;
- $\psi : \mathbb{R}^{d+1} \to [0, +\infty]$ is convex, lower semicontinuous and proper;
- ψ^∞ is the recession function $\psi^\infty(z) := \lim_{t\to+\infty} \frac{\psi(z_0+tz)}{t}$ (the limit is independent of the choice of z_0 in the domain of ψ);
- A_σ is the set of atoms of σ i.e. $A_\sigma := \{x : \ \sigma(x) := \sigma(\{x\}) \neq 0\}$;
- $g : \mathbb{R}^{d+1} \to [0, +\infty]$ is a lower semicontinuous subadditive function such that $g(0) = 0$ and $g_0(z) := \sup_{t>0} \frac{g(tz)}{t} = \psi^\infty(z)$;
- $\sharp$ is the counting measure.

In the sequel we assume the convexity of Ψ i.e. g is asked to be positively 1-homogeneous.

An interesting choice is the one of Benamou and Brenier (see [2], [10]):

$$\psi(r, e) = \begin{cases} \frac{|e|^2}{r} & \text{if } (r, e) \in]0, +\infty[\times\mathbb{R}^d, \\ 0 & \text{if } (r, e) = (0, 0), \\ +\infty & \text{otherwise.} \end{cases}$$

This is a positively 1-homogeneous function so $\psi^\infty = \psi = g$ and Ψ does not depend on the choice of the measure m so that

$$\Psi(\rho, E) = \begin{cases} \displaystyle\int_{[0,1]\times\Omega} \psi(d\rho/dm, dE/dm)\, dm(t, x) & \text{if } \rho \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Note that since $\psi(0, e)$ is infinite for any $e \neq 0$, it holds:

$$\Psi(\rho, E) < +\infty \Rightarrow E \ll \rho \tag{10}$$

so for any (ρ, E) in the domain of Ψ, we may write:

$$E(t, x) = v(t, x)\rho(t, x), \qquad \text{with } \rho(t, x) \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+) \text{ and } v(t, x) \in \mathrm{L}^1_\rho(\overline{Q}, \mathbb{R}^d).$$

The measure $\rho(t, x)$ can be viewed as the quantity of mass in time and space whereas $v(t, x)$ is the velocity of the mass transiting at x at time t. Moreover Ψ can be written in the simpler form:

$$\Psi(\rho, E) = \begin{cases} \displaystyle\int_{[0,1]\times\overline{\Omega}} \frac{|E|^2}{\rho} := \int_{[0,1]\times\overline{\Omega}} |v|^2 \, d\rho(t, x) & \text{if } \rho \geq 0 \text{ and } E = v\rho, \\ +\infty & \text{otherwise.} \end{cases}$$

As shown in [10], in this case we have:

$$\mathcal{W}_\Psi(\rho_0, \rho_1) = (W_2(\rho_0, \rho_1))^2$$

where W_2 is the classical 2-Wasserstein distance (see for instance [25]). Indeed, in the formula above, the Wasserstein distance is intended as:

$$(W_2(\rho_0,\rho_1))^2 = \min\left\{\int_{\overline{\Omega}\times\overline{\Omega}} |x_1-x_2|^2\, d\gamma(x_1,x_2) :\ \gamma \text{ has marginals } \rho_0,\ \rho_1\right\}$$

when Ω is convex, while the Euclidean distance has to be replaced by the geodesic distance when Ω is not convex.

It has been proved in [22] that the same result can be reached with any p-Wasserstein distance ($p > 1$) by choosing the function:

$$\psi_p(r,e) = \begin{cases} \frac{|e|^p}{r^{p-1}} & \text{if } (r,e) \in]0,+\infty[\times\mathbb{R}^d, \\ 0 & \text{if } (r,e) = (0,0), \\ +\infty & \text{otherwise.} \end{cases} \tag{11}$$

In the case $p = 1$ we simply take $\psi(r,e) = |e|$.
As in the previous case (10) is satisfied, whenever $p \geq 1$, together with

$$\mathcal{W}_\Psi(\rho_0,\rho_1) = \Big(W_p(\rho_0,\rho_1)\Big)^p$$

where W_p is the p-th Wasserstein distance:

$$\Big(W_p(\rho_0,\rho_1)\Big)^p = \min\left\{\int_{\overline{\Omega}\times\overline{\Omega}} |x_1-x_2|^p\, d\gamma(x_1,x_2) :\ \gamma \text{ has marginals } \rho_0,\ \rho_1\right\}.$$

An important remark is that, in this setting, a solution of problem (1) can be built using the idea that *masses should move along straight lines* when Ω is convex and *along geodesic curves* when Ω is not convex. More precisely, if we denote by $\gamma \in \mathcal{M}_b(\overline{\Omega}^2,\mathbb{R}^+)$ an optimal transport plan for W_p and by ξ_{x_1,x_2} a geodesic curve parametrized by $t \in [0,1]$ joining x_1 to x_2 for γ-almost every (x_1,x_2), then, an optimal $\sigma = (\rho, E)$ is given by:

$$\begin{array}{rcl} \displaystyle\int \varphi\, d\rho &=& \displaystyle\int_{\overline{\Omega}^2}\int_0^1 \varphi(t,\xi_{x_1,x_2}(t))\, dt\, d\gamma(x_1,x_2) \quad \forall \varphi \in C(\overline{Q}) \\ \displaystyle\int \varphi\cdot d\sigma &=& \displaystyle\int_{\overline{\Omega}^2}\int_0^1 \varphi(t,\xi_{x_1,x_2}(t))\cdot(1,\dot{\xi}_{x_1,x_2}(t))\, dt\, d\gamma(x_1,x_2)\ \forall \varphi \in C(\overline{Q})^{d+1}. \end{array} \tag{12}$$

Indeed, for this choice of σ, the decomposition $E = v\rho$ holds and we have:

$$\int_{\overline{Q}} \frac{|E|^p}{\rho^{p-1}} = \int_{\overline{Q}} |v|^p\, d\rho = \int_{\overline{\Omega}^2}\int_0^1 |\dot{\xi}_{x,y}(t)|^p\, dt\, d\gamma(x,y) = \Big(W_p(\rho_0,\rho_1)\Big)^p.$$

Even if this is not the purpose of the paper, we notice that in general the condition $f_0 \ll dx$ does not imply in the case of p-Wasserstein distance (11) that the optimal σ is unique as the following example shows.

Example VIII.3.1 *Take Ω be the complement of a disc K, $f_0 = dx\llcorner S$ and $f_1 = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B$ as in figure 1; where S is a disc of area 1 and A, B are two points at the same geodesic distance from*

P. It is clear that all geodesics joining a point of S to either A or B must pass through P. **For this reason, any admissible transport plan between f_0 and f_1 is optimal.**

We denote by Γ a line whose points are at the same distance from P, which separates S in two parts S^+ and S^- with the same area. **Two admissible (and optimal) transport plans are given by** *γ_1 which sends S^- to A and S^+ to B, and γ_2 which does the opposite. Formula (12) provides σ_1 and σ_2 associated to γ_1 and γ_2. Since every particle of S travels with constant speed and since they are at different distances form P, it is easy to see that the corresponding ρ_1 and ρ_2 cannot coincide. For instance there exists a time $\bar{t}$ such that the corresponding ρ_1 loads the geodesic from P to B but not the geodesic from P to A, while at the same time $\bar{t}$, the density ρ_2 does the opposite.* **The non-uniqueness of the optimal pair (v, ρ) may seem an easy consequence of the non-uniqueness of the optimal transport plan. However, we warn the reader from hurried conclusion. Think of the very similar and instructive case of transport densities in classical Monge transportation ($p = 1$) starting from an absolutely continuous measure. In this setting, there are many optimal transport maps but the transport density - that can be built from any transport map- is unique (see for instance [19]).**

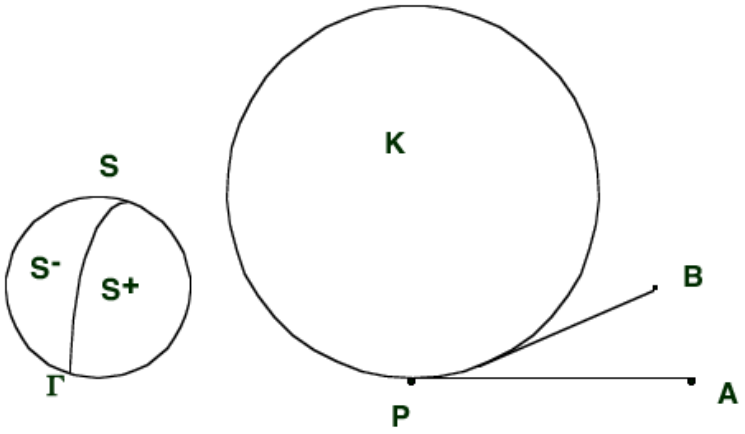

Figure VIII.1: An example of non-uniqueness.

When Ω is convex, $p > 1$ and $f_0 \ll dx$, there is only one optimal transport plan $\bar{\gamma}$ and the unique $\bar{\sigma}$ associated to $\bar{\gamma}$ by use of (12) is the only solution of problem (1). Let us give a quick scheme of a proof of this uniqueness. Take $\sigma = (\rho, v\rho)$ another solution. Using a result by Ambrosio, Gigli and Savaré (see [1], Theorem 8.2.1.), we can write σ as a superposition of generalized curves. More precisely, it exists some probability measure Γ on the set of absolutely continuous curves $G := W^{1,1}([0,1], \mathbb{R}^N)$ such that:

$$\begin{aligned} \int \varphi \, d\rho &= \int_G \int_0^1 \varphi(t, \alpha(t)) \, dt \, d\Gamma(\alpha) \quad \forall \varphi \in C(\overline{Q}) \\ \int \varphi \cdot d\sigma &= \int_G \int_0^1 \varphi(t, \alpha(t)) \cdot (1, \dot{\alpha}(t)) \, dt \, d\Gamma(\alpha) \; \forall \varphi \in C(\overline{Q})^{N+1} \\ v(t, \alpha(t)) &= \dot{\alpha}(t) \text{ a.e. } t \in [0,1]. \end{aligned}$$

The measure Γ is associated to a transport plan π by the following formula:

$$\int_{\overline{\Omega}^2} \varphi(x,y) d\pi(x,y) := \int_G \varphi(\alpha(0), \alpha(1)) d\Gamma(\alpha).$$

Now, it can easily be seen that the optimality of σ implies the optimality of π and that, Γ-almost everywhere, $\alpha([0,1])$ is the straight line $[\alpha(0), \alpha(1)]$. By uniqueness of the optimal transport plan we have $\pi = \bar{\gamma}$ which yields that σ and $\overline{\sigma}$ coincides.

This does not remain true for $p = 1$. In this case the uniqueness of the optimal transport plan is not insured. Moreover, not all the solutions of problem (1) are of the type (12). Actually the following measure (teletransport) happens to be optimal too:

$$\rho(t,x) = (1-t)\rho_0 + t\rho_1,$$

$$\int \Phi \cdot dE = \int_0^1 \int_{\overline{\Omega}^2} \int_{[x_0,x_1]} \Phi(t,x) \cdot \frac{x_1 - x_0}{|x_1 - x_0|}\, d\mathcal{L}^1(x) d\gamma(x_0,x_1) dt,\ \forall \Phi \in C(\overline{\Omega})^{d+1}.$$

However, the choice of Benamou and Brenier does not take into account congestion effects which are crucial in problems of crowd dynamics. Indeed there is a wide choice (see also [17]) for the cost function Ψ, the congestion effect being due to the superlinear terms. For instance the following are prototypical examples:

- $\psi(r,e) = \frac{|e|^p}{pr^{p-1}} + kr^2$ $(k > 0)$ which gives the cost

$$\Psi(\rho, E) = \int_0^1 \int_\Omega \left[\frac{|E|^p}{p\rho^{p-1}} + k\rho^2 \right] dt\, dx$$

  intending that $\Psi(\rho, E) = +\infty$ if ρ is not absolutely continuous with respect to $dt \otimes dx$ or ρ is not positive. In this case the high concentrations of ρ are penalized providing a lower congestion during the mass transportation from ρ_0 to ρ_1. Note that, in this case, as ψ is strictly convex in r, the optimal ρ is unique without any other assumption. If, in addition, we have $p > 1$, then E is of the form $E = v\rho$ and, the functional being also strictly convex in v, we have the uniqueness of the optimal measure (E, ρ).

- $\psi(r,e) = \frac{|e|^p}{pr^{p-1}} + \chi_{\{0 \le r \le M\}}(r)$ which gives the cost:

$$\Psi(\rho, E) = \begin{cases} \displaystyle\int_0^1 \int_\Omega \frac{|E|^p}{p\rho^{p-1}}\, dt\, dx & \text{if } 0 \le \rho \le M, \\ +\infty & \text{otherwise,} \end{cases}$$

  In this case the density ρ is constrained to remain below M, which is for instance the case when the model takes into account that two different individuals of the crowd cannot get too close.

We now study the dual problem (7) for general functionals $\Psi(\sigma)$ of the form above.

For the computation of Ψ^* we use a result by Bouchitté and Valadier (Theorem 1 of [8]) on the interchange between sup and integral; we get:

$$\Psi^*(\varphi) = \begin{cases} \displaystyle\int_{\overline{Q}} \psi^*(\varphi)\, dm & \text{if } (\psi^\infty)^*(\varphi(t,x)) + g^*(\varphi(t,x)) = 0 \quad \forall (t,x) \in \overline{Q}, \\ +\infty & \text{otherwise,} \end{cases}$$

for all $\varphi \in C(\overline{Q}, \mathbb{R}^{d+1})$, so that (7) writes as:

$$\sup_{\varphi \in C^1(\overline{Q})} \left\{ \int \varphi \, df - \int_{\overline{Q}} \psi^*(D\varphi) \, dm : \ (\psi^\infty)^*(D\varphi) + g^*(D\varphi) = 0 \right\}. \tag{13}$$

Note that, as g and Ψ^∞ are positively 1-homogeneous, the constraint of (13) can be reformulated saying that $D\varphi(t,x)$ belongs to a convex set K:

$$K := \{u \in \mathbb{R}^{d+1} : \ u \cdot z \leq \min(g(z), \psi^\infty(z)) \ \forall z \in \mathbb{R}^{d+1} \text{ such that } |z| = 1\}.$$

As we have already said, Problem (13) has to be relaxed in order to make the primal-dual optimality condition meaningful.

To that aim, we need to choose an appropriate space for the dual variable φ and give a sense to the gradient $D\varphi$ appearing in (13) and in (8) which will write as:

$$\begin{aligned} \int D\varphi_{\text{opt}} \cdot d\sigma_{\text{opt}} \quad = \quad & \int_{\overline{Q}} \psi\left(\frac{d\sigma_{\text{opt}}}{dm}\right) dm + \int_{\overline{Q}\setminus A_\sigma} \psi^\infty\left(\frac{\sigma^s_{\text{opt}}}{|\sigma^s_{\text{opt}}|}\right) d|\sigma^s_{\text{opt}}| \\ & + \int_{A_\sigma} g(\sigma_{\text{opt}}(x)) \, d\sharp(x) + \int_{\overline{Q}} \psi^*(D\varphi_{\text{opt}}) dm \\ \text{with the constraint } & (\psi^\infty)^*(D\varphi_{\text{opt}}) + g^*(D\varphi_{\text{opt}}) = 0. \end{aligned} \tag{14}$$

The space X of the dual variables φ and its topology must be chosen according to the properties of ψ. Then the idea will be to approach φ by a sequence of regular functions $(\varphi_n)_n$ tending to φ. The problem is that the vectorial function η obtained as the limit – in a weak sense – of the sequence $(D\varphi_n)_n$ is not unique in the sense that it depends on the choice of the sequence $(\varphi_n)_n$. Uniqueness can be recovered by making locally the projection of $D\psi_n(t,x)$ on an appropriate tangent space to a measure μ at (t,x) (see [6] and [7]) which has to be chosen in a proper way. In the following, we give some references for some particular cases.

- In [6] a relaxation result is given in case ψ satisfies the following assumption for some $p \in]1, +\infty[$:

  $$c_1|(r,e)|^p - \frac{1}{c_1} \leq \psi(r,e) \leq c_2(|(r,e)|^p + 1) \quad \forall (r,e) \in \mathbb{R}^{d+1}$$

  for suitable $c_1, \ c_2 > 0$. Therefore for a fixed measure $m \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+)$ the functionals Ψ and Ψ^* are:

  $$\Psi(\sigma) = \begin{cases} \displaystyle\int_{\overline{Q}} \psi\left(\frac{d\sigma}{dm}\right) dm & \text{if } \sigma \ll m \\ +\infty & \text{otherwise,} \end{cases}$$

  $$\Psi^*(\varphi) = \int_{\overline{Q}} \psi^*(\varphi) \ dm,$$

  where in the definition of Ψ, we have taken $g \equiv +\infty$. The dual variable φ then belongs to the Sobolev space $W^{1,p'}_m(\overline{Q})$ with $1/p + 1/p' = 1$ made with respect to the measure m (see [6]). Following [6] and [7], the gradient $D_m\varphi(t,x)$ has to be intended as an element of the tangent space $T^{p'}_m(t,x)$ for m-almost every (t,x). Then, as shown in [6], the relaxed dual problem can be expressed as:

  $$\sup_{\varphi \in W^{1,p'}_m(Q)} \left\{ \int \varphi df - \int_{\overline{Q}} \psi^*_m(D_m\varphi) \, dm \right\}$$

where

$$\psi_m^*(r,e) = \inf\{\psi^*(r,e+\eta) : \ \eta \in (T_m^{p'}(r,e))^\perp\}.$$

Finally the primal-dual optimality condition reads as:

$$\begin{cases} \displaystyle\int D_m\varphi_{\rm opt}\cdot d\sigma_{\rm opt} = \int_{\overline{Q}} \psi\left(\frac{d\sigma_{\rm opt}}{dm}\right) dm + \int_{\overline{Q}} \psi_m^*(D_m\varphi_{\rm opt})dm \\ \sigma_{\rm opt} \ll m. \end{cases}$$

- In case $\psi(r,e) = \frac{|e|^p}{pr^{p-1}}$ with $p \geq 1$ (see [22]), the functional Ψ^* becomes:

$$\Psi^*(\varphi) = \begin{cases} 0 & \text{if } \varphi_1 + \frac{|(\varphi_2,\dots\varphi_{d+1})|^{p'}}{p'} \leq 0 \text{ a.e.} \\ +\infty & \text{otherwise} \end{cases}$$

where p' is such that $1/p+1/p' = 1$. For $p = 1$, $\frac{|(\varphi_2,\dots\varphi_{d+1})|^{p'}}{p'}$ has to be intended as $\chi_{\{|(\varphi_2,\dots\varphi_{d+1})|\leq 1\}}$. The dual variable then is Lipschitz continuous and the relaxed dual problem becomes:

$$\sup_{\varphi \text{ Lipschitz}} \left\{ \int \varphi \, df : \ \partial_t\varphi(t,x) + \frac{|\nabla_x\varphi(t,x)|^{p'}}{p'} \leq 0 \quad \text{a.e. } (t,x) \right\}.$$

In case $p > 1$, the primal-dual optimality condition can be written as:

$$\begin{cases} \displaystyle\int D_{\rho_{\rm opt}}\varphi_{\rm opt}\cdot(1,v_{\rm opt}(t,x))d\rho_{\rm opt} = \int_{\overline{Q}} \frac{|v_{\rm opt}(t,x)|^p}{p}\, d\rho_{\rm opt}(t,x), \\ \partial_t\varphi_{\rm opt}(t,x) + \frac{|\nabla_x\varphi_{\rm opt}(t,x)|^{p'}}{p'} \leq 0 \quad \text{a.e. } (t,x), \end{cases} \tag{15}$$

where the gradient $D_{\rho_{\rm opt}}\varphi(t,x) = (\partial_{(\rho_{\rm opt},t)}\varphi(t,x), \nabla_{(\rho_{\rm opt},x)}\varphi(t,x))$ is an element of the tangent space $T^\infty_{\rho_{\rm opt}}(t,x)$ for $\rho_{\rm opt}$-almost every (t,x). As it can be seen in [22], we have

$$D_{\rho_{\rm opt}}\varphi_{\rm opt}(t,x) - D\varphi_{\rm opt}(t,x) \in T^\perp_{\rho_{\rm opt}}(t,x) \quad \rho_{\rm opt} - \text{a.e.}$$

and thanks to (9):

$$(1, v_{\rm opt}(t,x)) \in T_{\rho_{\rm opt}}(t,x) \quad \rho_{\rm opt} - \text{a.e.}$$

so that the inequality (15) gives:

$$\begin{aligned} D_{\rho_{\rm opt}}\varphi_{\rm opt}\cdot(1,v_{\rm opt}) &= D\varphi_{\rm opt}\cdot(1,v_{\rm opt}) \\ &\leq -\frac{|\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}|^{p'}}{p'} + v_{\rm opt}(t,x)\cdot\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt} \\ &\leq \sup_{\omega\in\mathbb{R}^d}\left\{ v_{\rm opt}(t,x)\cdot\omega - \frac{|\omega|^{p'}}{p'}\right\} = \frac{|v_{\rm opt}|^p}{p}. \end{aligned}$$

Then, the equality in (15) gives that all the previous inequalities happen to be equalities that is to say

$$\begin{aligned} &\partial_{(\rho_{\rm opt},t)}\varphi_{\rm opt} = -\frac{|\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}|^{p'}}{p'}, \\ &\nabla_{(\rho_{\rm opt},x)}\varphi_{\rm opt}(x,t) \in \operatorname{argmax}\left\{ \omega \mapsto v_{\rm opt}(t,x)\cdot\omega - \frac{|\omega|^{p'}}{p'}\right\}. \end{aligned} \tag{16}$$

By making an easy computation, we get:

$$\begin{aligned} \nabla_{(\rho_{\text{opt}},x)}\varphi_{\text{opt}} &= |v_{\text{opt}}|^{p-2}v_{\text{opt}}, \\ \partial_{(\rho_{\text{opt}},t)}\varphi_{\text{opt}} &= \frac{-|v_{\text{opt}}|^{p}}{p'}. \end{aligned} \tag{17}$$

If $p = 1$, we make the computation in the similar way by writing $(\rho_{\text{opt}}, E_{\text{opt}})$ as

$$(\rho_{\text{opt}}(t,x), E_{\text{opt}}(t,x)) = (h_{\text{opt}}(t,x), v_{\text{opt}}(t,x))d\mu_{\text{opt}}$$

where $\mu_{\text{opt}} \in \mathcal{M}_b(\overline{Q}, \mathbb{R}^+)$ and $(h_{\text{opt}}, v_{\text{opt}}) \in \mathrm{L}^1_{\mu_{\text{opt}}}(\overline{Q}) \times \mathrm{L}^1_{\mu_{\text{opt}}}(\overline{Q}, \mathbb{R}^d)$. Then the primal-dual optimality condition writes as:

$$\begin{cases} \displaystyle\int D_{\mu_{\text{opt}}}\varphi_{\text{opt}} \cdot (h_{\text{opt}}(t,x), v_{\text{opt}}(t,x))d\mu_{\text{opt}} = \int_{\overline{Q}} |v_{\text{opt}}(t,x)|\, d\mu_{\text{opt}}(t,x), \\ \partial_t\varphi_{\text{opt}}(t,x) \le 0 \text{ and } |\nabla_x\varphi_{\text{opt}}(t,x)| \le 1 \quad \text{a.e. } (t,x), \end{cases} \tag{18}$$

which leads to:

$$\begin{aligned} \nabla_{(\rho_{\text{opt}},x)}\varphi_{\text{opt}} &= \frac{v_{\text{opt}}}{|v_{\text{opt}}|}, \\ \partial_{(\rho_{\text{opt}},t)}\varphi_{\text{opt}} &= 0. \end{aligned} \tag{19}$$

VIII.4 Numerical computation

We describe in the present section an algorithm to approximate problem (1). This method is directly adapted from the augmented Lagrangian method presented in [2]. For the reader convenience, we recall below in our formalism the main steps of this algorithm.

First, solving problem (1) is equivalent to solve the saddle point problem:

$$\min_{\sigma} \max_{\varphi \in C(Q)} L(\sigma, \varphi) \tag{20}$$

where $L(\sigma, \varphi)$ is the Lagrangian defined by:

$$L(\sigma, \varphi) = \Psi(\sigma) - \int D\varphi \cdot d\sigma + \int \varphi\, df.$$

Following [2], for all $r > 0$, we introduce the augmented Lagrangian

$$L_r(\sigma, \sigma^*, \varphi) := \Psi^*(\sigma^*) + \int (D\varphi - \sigma^*) \cdot d\sigma - \int \varphi\, df + \frac{r}{2}\int |D\varphi - \sigma^*|^2 dy.$$

Using the identity $\Psi^*(\sigma^*) + \Psi(\sigma) = \int \sigma^* \cdot d\sigma$ it can easily be established that the saddle point problem (20) is equivalent to the new problem:

$$\max_{\sigma} \min_{\sigma^*, \varphi} L_r(\sigma, \sigma^*, \varphi). \tag{21}$$

As reported in [2], the simple algorithm ALG2 (see [20]), which is a classical relaxation of Uzawa's method, can be used to approximate problem (21). Let us recall with our notation this iterative process:

- let $(\sigma_n, \sigma^*_{n-1}, \varphi_{n-1})$ be given;
- Step A: find φ_n such that:

$$L_r(\sigma_n, \sigma^*_{n-1}, \varphi_n) \leq L_r(\sigma_n, \sigma^*_{n-1}, \varphi), \quad \forall \varphi \in C^1(\overline{Q});$$

- Step B: find σ^*_n such that:

$$L_r(\sigma_n, \sigma^*_n, \varphi_n) \leq L_r(\sigma_n, \sigma^*, \varphi_n), \quad \forall \sigma^* \in C(\overline{Q}, \mathbb{R}^{d+1});$$

- Step C: set $\sigma_{n+1} = \sigma_n + r(D\varphi_n - \sigma^*_n)$;
- go back to Step A.

Note that the variables (v, q) in [2] are renamed (σ, σ^*) in the previous description of the algorithm. Let us now underline the two main differences of our approach.
First, Step A consists in solving the Euler-Lagrange equation:

$$\int D\varphi \cdot d\sigma_n - \int \varphi \, df + r \int D\varphi(-\sigma^*_{n-1} + D\varphi_n) \, dy = 0, \qquad \forall \varphi \in C^1(\overline{Q}).$$

This variational formulation is nothing else than the weak form of the partial differential equation:

$$\begin{cases} -r\Delta\varphi_n = \operatorname{div}(\sigma_n - r\sigma^*_{n-1}) + f & \text{in } \overline{Q} \\ r\frac{\partial \varphi_n}{\partial n} = (\sigma_n - r\sigma^*) \cdot \nu & \text{on } \partial Q. \end{cases}$$

The resolution of the previous PDE has been achieved with the very efficient software freeFEM3D (see [21] and [16]) provided by S. Del Pino and O. Pironneau. As in [2], for computational stability, we perturbed the previous Laplace equation in:

$$-r\Delta\varphi_n + r\varepsilon\varphi_n = \operatorname{div}(\sigma_n - r\sigma^*_{n-1}) + f$$

with $\varepsilon = 10^{-4}$.
Second, since in our general framework, Ψ^* is not always a characteristic function, step B consists in minimizing the following quantity with respect to σ^*:

$$\Psi^*(\sigma^*) + \int (D\varphi_n - \sigma^*) \cdot d\sigma_n + \frac{r}{2} \int |D\varphi_n - \sigma^*|^2 dy.$$

In all the test cases presented below, it has been possible to solve this problem analytically. Indeed, this pointwise optimization problem reduces to the numerical computation of the roots of a polynomial with real coefficients.

Example VIII.4.1 We consider here a transportation domain $\Omega = [-1, 1]^2$ in which there are spatial obstacles that the mass cannot cross. This is for instance the case of a subway gate that a mass of individuals has to cross to reach a final destination. In this first example, the transportation is described simply by the Wasserstein distance W_2 which turns out, setting $\sigma = (\rho, E)$, to consider the convex function

$$\Psi(\sigma) = \int_Q \frac{|E|^2}{2\rho}$$

in the sense precised in Section VIII.3. The Fenchel transform Ψ^* can be easily computed and we have:

$$\Psi^*(\Phi) = \begin{cases} 0 & \text{if } \Phi_1 + \frac{|(\Phi_2,\Phi_3)|^2}{2} \leq 0 \text{ a.e.} \\ +\infty & \text{otherwise.} \end{cases}$$

Notice that, since Ψ is homogeneous of degree 1, the function Ψ^* is the indicator of a convex set. Here below, we plot the mass density ρ_t at various instants of time. The initial configuration ρ_0 is taken as a Gaussian distribution centered at the point $(-0.65, 0)$ and the final measure ρ_1 is taken as $\rho_1(x_1, x_2) = \rho_0(x_1 - 1.3, x_2)$. Notice that without the obstacle gate, the mass density $\rho(t, \cdot)$ would simply be the translation $\rho(t, x_1, x_2) = \rho_0(x_1 - 1.3t, x_2)$. In general, in presence of obstacles, the mass density ρ will follow the geodesic paths and by consequence the supports of all $\rho(t, \cdot)$ have to be contained in the geodesic envelope of ρ_0 and ρ_1; this is why most of the mass passes through the central gate. Our computation done on a regular grid of $70 \times 70 \times 70$ (from which cells corresponding to the obstacles have been removed) and presented in Figure 2 is in agreement with that observation. Convergence with respect to the criterium proposed in [2] has been achieved in 150 iterations.

Example VIII.4.2 We consider the same geometrical configuration as in the previous example. In this case, we add a diffusion term in order to penalize mass congestion which is described in our case by high values of ρ. The function Ψ we consider is:

$$\int_0^1 \int_\Omega \frac{|E|^2}{2\rho} + c\rho^2 \, dt \, dx$$

with $c = 0.1$. The Fenchel transform is given by:

$$\Psi^*(\Phi) = \frac{1}{2c} \int_Q \left(\left(\Phi_1 + \frac{|(\Phi_2, \Phi_3)|^2}{2} \right)^+ \right)^2 (y) \, dy.$$

Notice that, due to the addition of the diffusion term, the dual function Ψ^* is now finite everywhere. This fact could explain the improvement in the convergence of the iteration scheme: in that example, convergence is reached in only 50 iterations.
As expected (see Figure 3), the mass crosses the obstacle by using several gates.

Example VIII.4.3 In our last example we consider again the same geometrical configuration and a new term which takes congestion into account. More precisely, we consider the cost functional

$$\int_0^1 \int_\Omega \frac{|E|^2}{2\rho} + \chi_{\rho \leq 1}.$$

The Fenchel transform is given by:

$$\Psi^*(\Phi) = \int_Q \left(\Phi_1 + \frac{|(\Phi_2, \Phi_3)|^2}{2} \right)^+ (y) \, dy.$$

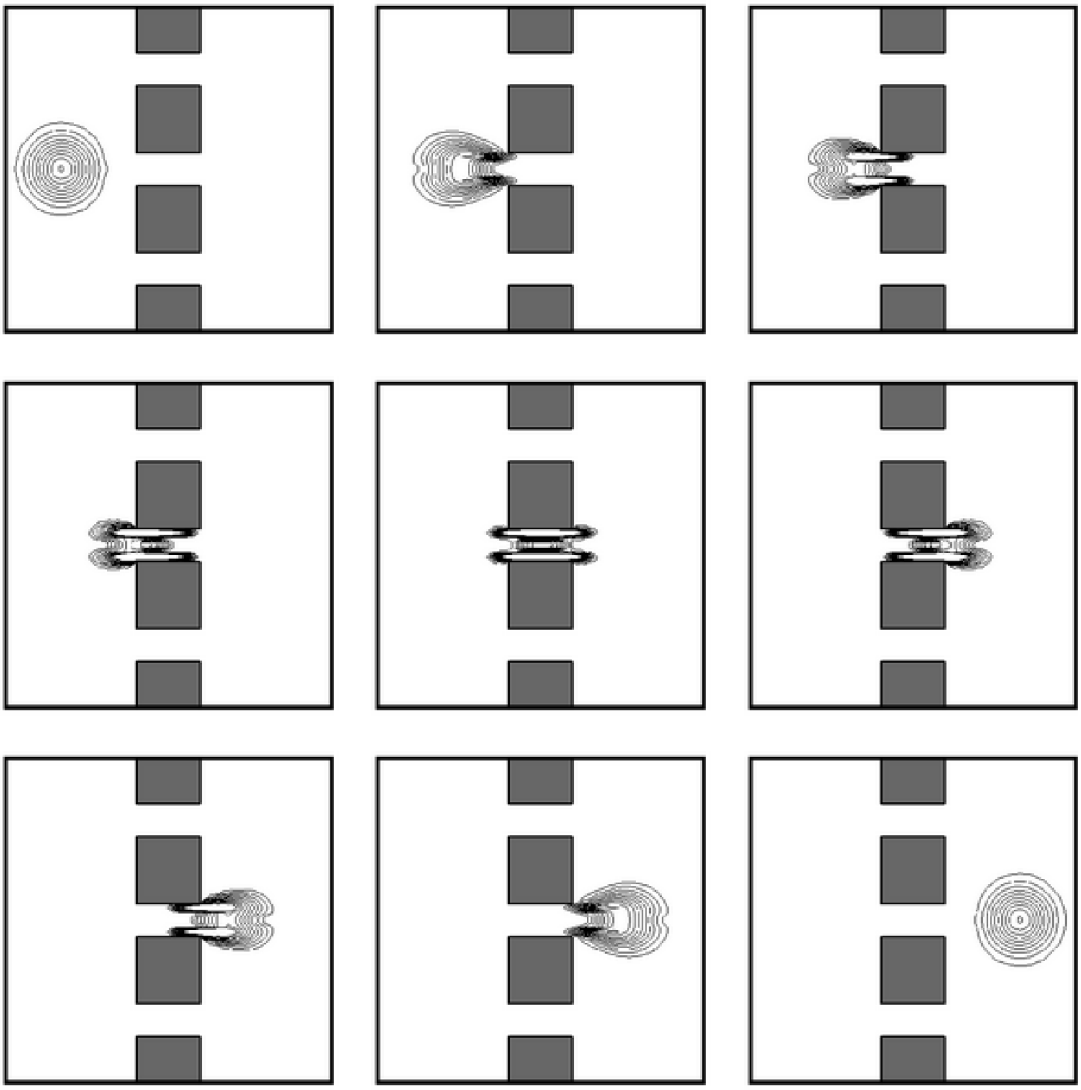

Figure VIII.2: Plot of $\rho(t,\cdot)$ for 9 values of t.

At a first glance (see Figure 4 where level lines are plotted), the result seems to be very similar to our first situation where the congestion effect was not considered. Again, most of the mass passes through the central gate, but contrary to the first case the density in the front gate is spread all over the channel and not only near the boundaries of the obstacles.

Acknowledgements: This work has been done during a visit of Edouard Oudet to Pisa in the framework of a GNAMPA program for scientific visits. Chloé Jimenez gratefully acknowledges the Scuola Normale Superiore di Pisa for the post-doc fellowship provided.

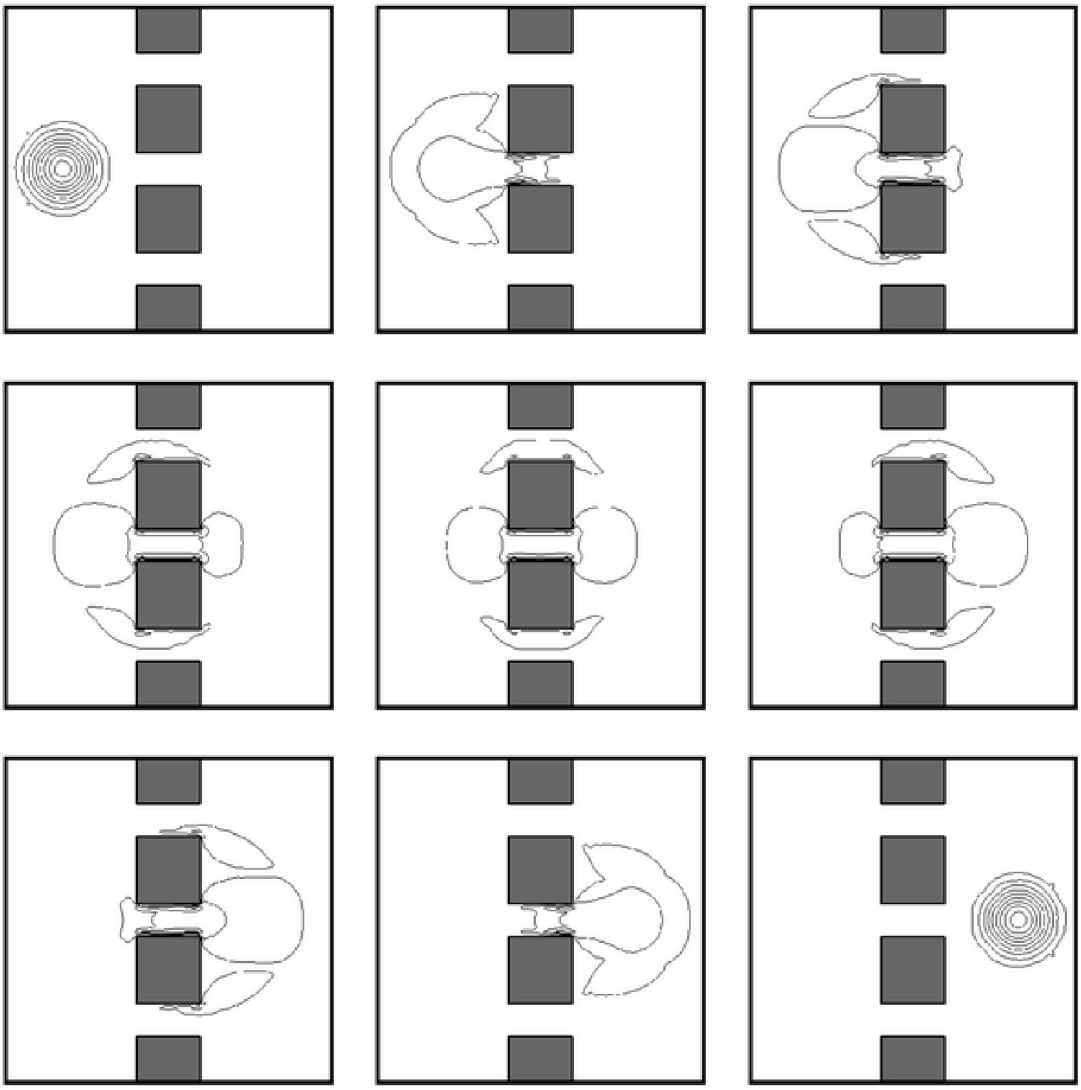

Figure VIII.3: Plot of $\rho(t,\cdot)$ for 9 values of t.

Bibliography

[1] L. Ambrosio, N. Gigli, G. Savaré: *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005. viii+333 pp.

[2] J.-D. Benamou, Y. Brenier: *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem.* Numer. Math., **84** (3) (2000), 375–393.

[3] G. Bouchitté, G. Buttazzo: *New lower semicontinuity results for nonconvex functionals defined on measures.* Nonlinear Anal., **15** (7) (1990), 679–692.

[4] G. Bouchitté, G. Buttazzo: *Integral representation of nonconvex functionals defined on measures.* Ann. Inst. H. Poincaré Anal. Non Linéaire, **9** (1) (1992), 101–117.

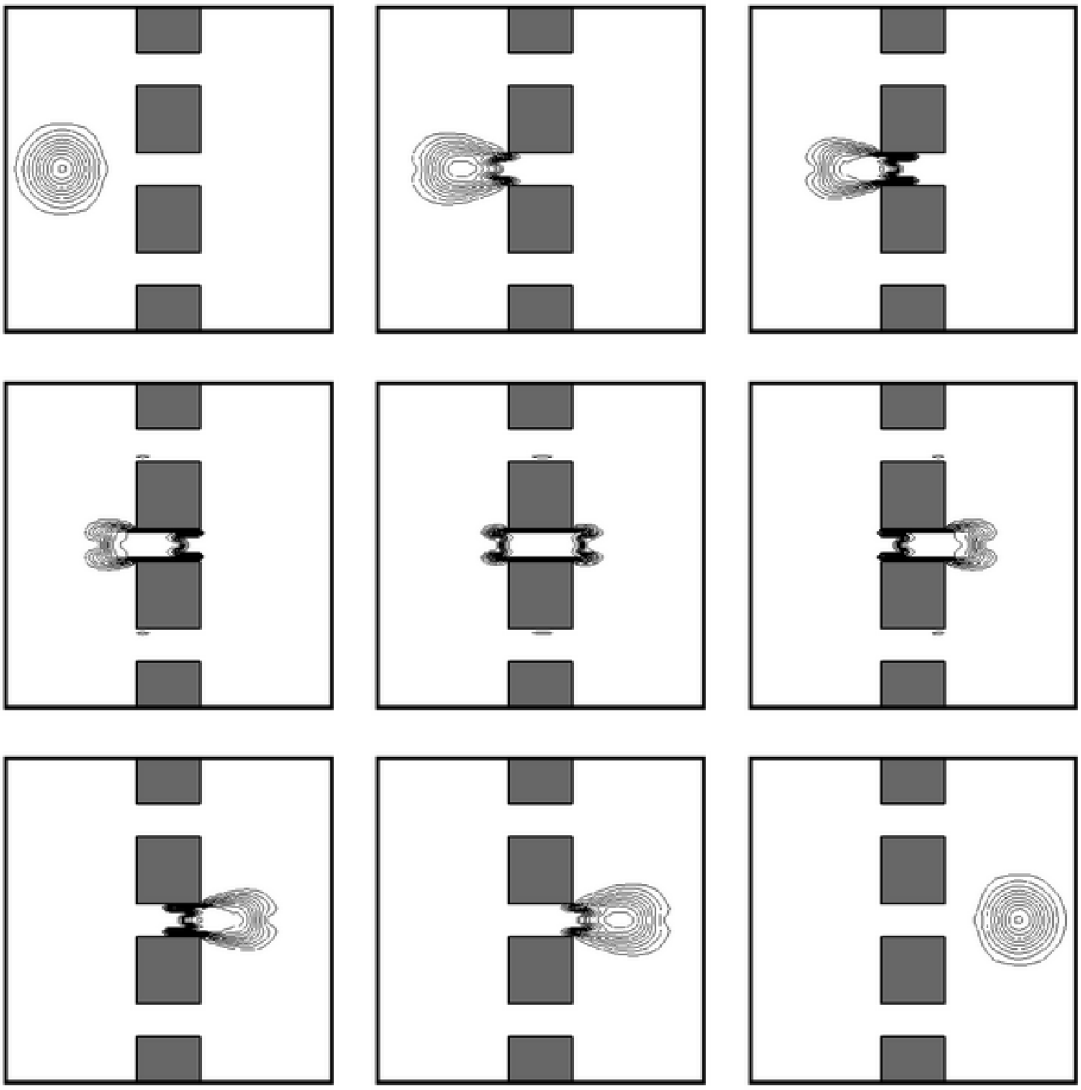

Figure VIII.4: Plot of $\rho(t,\cdot)$ for 9 values of t.

[5] G. Bouchitté, G. Buttazzo: *Characterization of optimal shapes and masses through Monge-Kantorovich equation.* J. Eur. Math. Soc., **3** (2) (2001). 139–168.

[6] G. Bouchitté, G. Buttazzo, P. Seppecher: *Energies with respect to a measure and applications to low dimensional structures.* Calc. Var., **5** (1997), 37–54.

[7] G. Bouchitté, G. Buttazzo, P. Seppecher: *Shape optimization solutions via Monge-Kantorovich equation.* C. R. Acad. Sci. Paris Sér. I Math., **324** (10) (1997), 1185–1191.

[8] G. Bouchitté, M. Valadier: *Integral representation of convex functionals on a space of measures.* J. Funct. Anal. 80 (1988), no. 2, 398–420.

[9] A. Brancolini, G. Buttazzo: *Optimal networks for mass transportation problems.* ESAIM Control Optim. Calc. Var., **11** (1) (2005), 88–101.

[10] Y. Brenier: *Extended Monge-Kantorovich theory.* In "Optimal Transportation and Applications" (Martina Franca 2001), Lecture Notes in Math. **1813**, Springer-Verlag, Berlin (2003), 91–121.

[11] G. Buttazzo, A. Pratelli, S. Solimini, E. Stepanov: *Optimal urban networks via mass transportation.* Lecture Notes in Math. **1961**, Springer-Verlag, (in print).

[12] G. Buttazzo, A. Pratelli, E. Stepanov: *Optimal pricing policies for public transportation networks.* SIAM J. Optim., **16** (3) (2006), 826–853.

[13] G. Buttazzo, F. Santambrogio: *A model for the optimal planning of an urban area.* SIAM J. Math. Anal., **37** (2) (2005), 514–530.

[14] G. Carlier, P. Cardaliaguet, C. Jimenez: *Optimal transport with convex constraints.* Work in progress.

[15] G. Carlier, C. Jimenez, F. Santambrogio: *Optimal transportation with traffic congestion and Wardrop equilibria.* SIAM J. Control Optim. 47 (2008), no. 3, 1330–1350.

[16] S. Del Pino, O. Pironneau: *A Fictitious domain based general PDE solver.* In "Numerical Methods for Scientific Computing", conf. METSO-ECCOMAS, E. Heikkola ed., CIMNE, Barcelona (2003).

[17] J. Dolbeault, B. Nazaret, G. Savaré: *A new class of transport distances between measures.* Calc. Var. PDE, **34** (2) (2009), 193-231.

[18] L. C. Evans, W.Gangbo: *Differential equation methods for the Monge-Kantorovich mass transfer problem.* Mem. Amer. Math. Soc., **658** (1999).

[19] M. Feldman, R. J. McCann: *Uniqueness and transport density in Monge's mass transportation problem.* Calc. Var. Partial Differential Equations 15 (2002), no. 1, 81–113.

[20] M. Fortin, R. Glowinski: *Augmented Lagrangian Methods. Applications to the Numerical Solution of Boundary Value Problems.* Stud. Math. Appl. **15**, North-Holland, Amsterdam (1983).

[21] freeFEM3D: available at `http://www.freefem.org/ff3d`.

[22] C. Jimenez: *Dynamic formulation of optimal transport problems.* Dynamic formulation of optimal transport problems. J. Convex Anal. 15 (2008), no. 3, 593–622.

[23] B. Maury, J. Venel: *Un modèle de mouvements de foule.* Paris-Sud Working Group on Modelling and Scientific Computing 2006–2007, 143–152, ESAIM Proc., 18, EDP Sci., Les Ulis, 2007.

[24] L. V. Kantorovich: *On the transfer of masses.* Dokl. Akad. Nauk. SSSR, **37** (1942), 227–229.

[25] C. Villani: *Topics in Optimal Transportation*. Grad. Stud. Math. **58**, Amer. Math. Soc., Providence (2003).

PART 3.

Branched transport

Édouard Oudet & F. Santambrogio

IX.1 Introduction

This last brief chapter is an introduction to some preliminary results related to optimal branched transport : let us introduce this kind of problem in a discrete setting. Consider a compact convex domain $\Omega \subset \mathbb{R}^N$ and two measures which are sum of dirac masses :

$$s = \sum_{i=1}^{m} a_i \, \delta_{x_i} \text{ and } g = \sum_{j=1}^{n} b_j \, \delta_{y_j}$$

where (a_i) and (b_j) are positive numbers. We ask to s and g to have the same total mass, that is $\sum a_i = \sum b_j$. Following [3] we define a transport path (G, w) from s to g as both a weighted directed graph G which vertices contains the points (x_i) and (y_j) and a weight function

$$w : E(G) \to \mathbb{R}_+$$

where $E(G)$ is the set of directed edges of G. Moreover we ask w to satisfy Kirchhoff's law that is for all vertex v of G we have:

$$\sum_{e \in E(G),\, e^- = v} w(e) = \sum_{e \in E(G),\, e^+ = v} w(e) + \begin{cases} a_i & \text{if } v = x_i \text{ for some } i \\ -b_j & \text{if } v = y_j \text{ for some } j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where e^- and e^+ denote the starting and ending points of each directed edge $e \in E(G)$. To every transport path (G, w) we associate a cost of transportation defined by

$$M_\alpha(G) = \sum_{e \in E(G)} w(e)^\alpha \, \text{length}(e) \tag{2}$$

for some fixed parameter $\alpha \in [0, 1]$.

Such optimal irrigation networks received a large attention last years (see for instance the articles [1, 3, 2], of Xia, Bernot, Maddalena, Morel, Santambrogio and Solimini). Many qualitative results and generalization of the problem have been discussed. Nevertheless, it is surprising that very few has been done on the numerical approximation of optimal networks. As explained in [4] the exact identification of global optimal networks, in this combinatorial context is NP hard (with respect to the number of dirac measures). In order to tackle this difficulty we introduce in the next section a continuous framework based on a Γ-convergence result obtained recently by F. Santambrogio which leads to a numerical procedure similar to the one introduced in chapter VII.

IX.2 F. Santambrogio's Γ-convergence result

Let (G, w) be a transport path associated to the two measures s and g. In order to introduce a relaxed formulation to the previous problem we associate to every (G, w) the vectorial measure

$$u_G = \sum_{e \in E(G)} w(e) \frac{e^+ - e^-}{||e^+ - e^-||} \mathcal{H}^1_{|e}.$$

By definition, the measure u is in the class of vectorial measures of the type $u(M, \theta, \xi) = \theta\xi.\mathcal{H}^1_M$, where θ and ξ are respectively a positive function and a unit vector field defined on M a set of finite $\mathcal{H}^1$-Hausdorff measure. Moreover Kirchhoff's law in this continuous setting is equivalent to the divergence constraint

$$\nabla.u = g - s.$$

In a analogous way to (2) we define for every vectorial measure u the cost functional:

$$M^\alpha(u) = \begin{cases} \int_M \theta^\alpha \, d\mathcal{H}^1 \text{ if } u \text{ is of the type } u(M, \theta, \xi) \\ +\infty \text{ otherwise} \end{cases} \tag{3}$$

Thus, our relaxed optimisation problem is to minimise M^α under the previous (weak) divergence constraint. The Γ-convergence result we are going to present is based on functionals very similar to the ones of Modica and Mortola (see chapter VII) but of the form

$$E_\varepsilon(u) = \varepsilon^{\gamma_1} \int_\Omega |u|^\beta + \varepsilon^{\gamma_2} \int_\Omega |\nabla u|^2$$

defined on $H^1(\Omega; \mathbb{R}^N)$ with $\gamma_1 < 0 < \gamma_2$. The main difference with Modica and Mortola's functionnal is the fact that the double-well potential has been replaced by a concave power β which forces the modulus of the vector field u to tends to 0 or $+\infty$. The idea is now to choose the parameters γ_1, γ_2 and β to obtain a functional equivalent (asymptotically when ε tends to 0) to the cost (3). Considering that the support of u is concentrated on a segment, an heuristic argument leads to the following choice of the parameters (see [2] for the details) :

$$\beta = \frac{2 - 2N + 2\alpha N}{3 - N + \alpha(N-1)}, \quad , \frac{\gamma_1}{\gamma_2} = \frac{(N-1)(\alpha - 1)}{3 - N + \alpha(N-1)} \tag{4}$$

Let us call M^α_ε the functional associated to a set of parameters satisfying conditions (4). F.

Theorem 34 *Suppose $N = 2$ and $\alpha \in]1/2, 1[$. Then M^α_ε Γ-converges to cM^α with respect to the convergence of measures for some suitable constant c when ε tends to 0.*

We describe below how this result can be used to propose an efficient algorithm regarding the numerical approximation of optimal irrigation networks.

IX.3 An efficient numerical approximation and some preliminary results

As noticed in chapter VII, previous Γ-convergence result makes it possible to replace an hard discrete problem by a sequence of optimisation problems under linear constraints. In addition, we observed that for $\varepsilon >> 1$ the functional M_ε^α is close from being convex. This observation was the starting point of our optimisation strategy described in details in chapter VII. The main difference between this situation and the previous one is related to the divergence constraint. Due to the very simple structure of the constraints we had to deal with, it was straightforward to compute a projection on the linear constraints in the context of chapter VII. Here, the divergence constraint requires a more careful numerical treatment. By the classical Helmholtz's decomposition, computing the projection is equivalent to solve a problem of Poisson's type. In order to compute the solution of Poisson's problem efficiently, we implemented a fast Fourier approach which has an almost linear complexity with respect to the number of points of the grid.

We present below the first results obtained with our simple approach. The following figures are the results of four different experiments with two different values of the parameter α. On the first rows of the figures, we represent two views of the graph of the given density $g - s$. The second rows represent two views of the graph of the norm of the optimal vector field for each value of α. As expected, "Kirchhoff's law is approximatively satisfied" by the support of the vector field which converges to a one dimensional set. Moreover we observe that two different values of α may lead to very different optimal structures.

IX.4 Some perspectives

To conclude this chapter, we list below some theoretical and numerical questions we are going to investigate in the future:

- Is it possible to generalize theorem 34 to $\alpha \in [0, 1[$? For $\alpha = 0$, our approach would lead to an original algorithm to solve Steiner's problem.
- Can we extend theorem 34 to the dimension $N = 3$?
- Is there a way to localise the numerical optimisation process in order to increase its efficiency in dimension 2 and 3 ? Multigrid methods should be consider in this context to keep a fast projection operator.
- What could be a relevant mathematical framework to describe the growth of one dimensional structures ? Is it possible to adapt our cost functional to some more realistic situations like the growth of vascular networks in the context of angiogenesis ?

Bibliography

[1] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. *Optimal transportation networks*, volume 1955 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Models and theory.

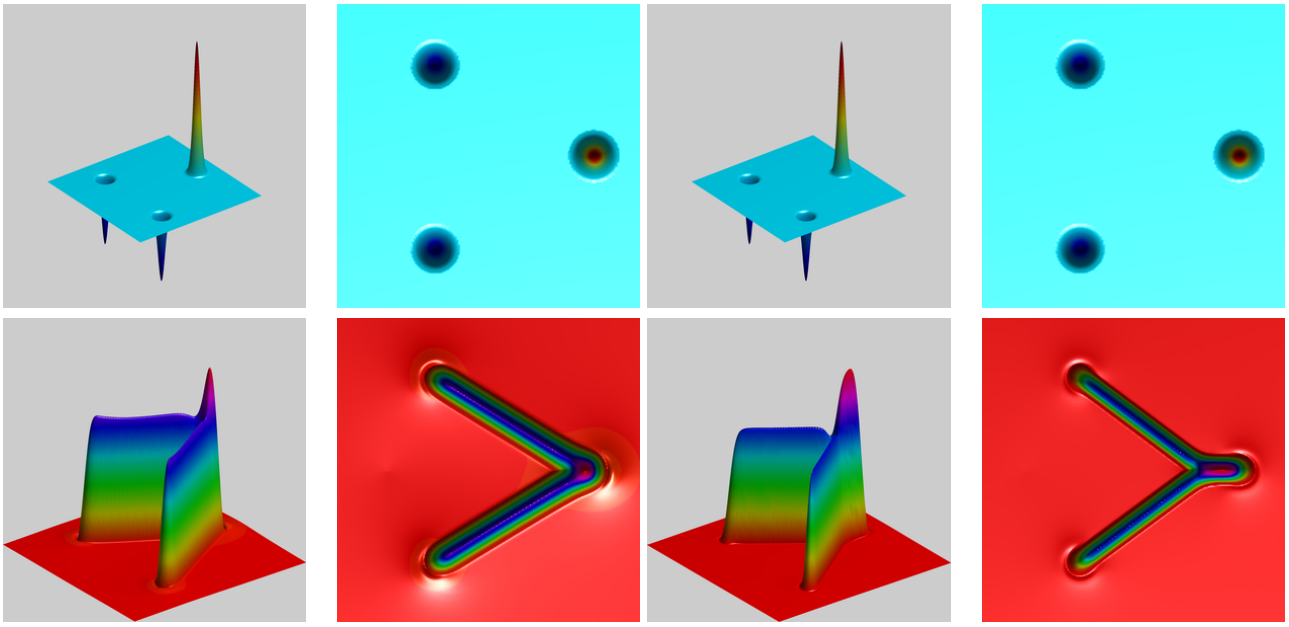

Figure IX.1: Optimal irrigation with $\alpha = 1/2 + 1/10$ and $\alpha = 1/2 + 1/4$

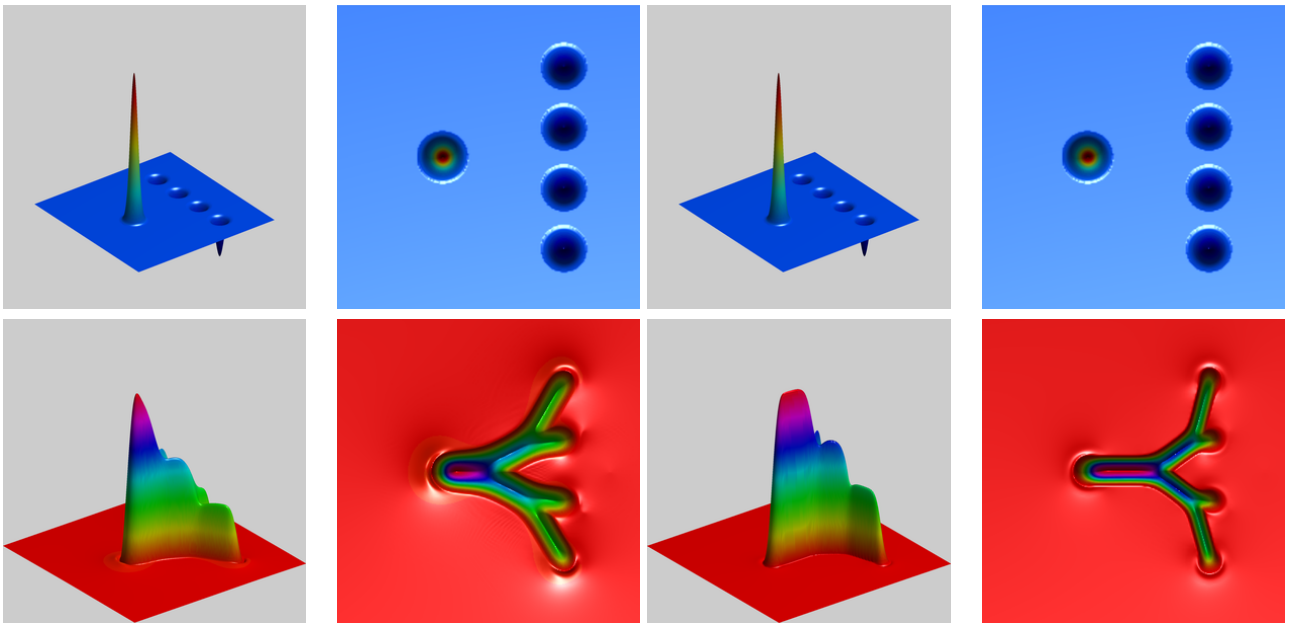

Figure IX.2: Optimal irrigation with $\alpha = 1/2 + 1/10$ and $\alpha = 1/2 + 1/4$

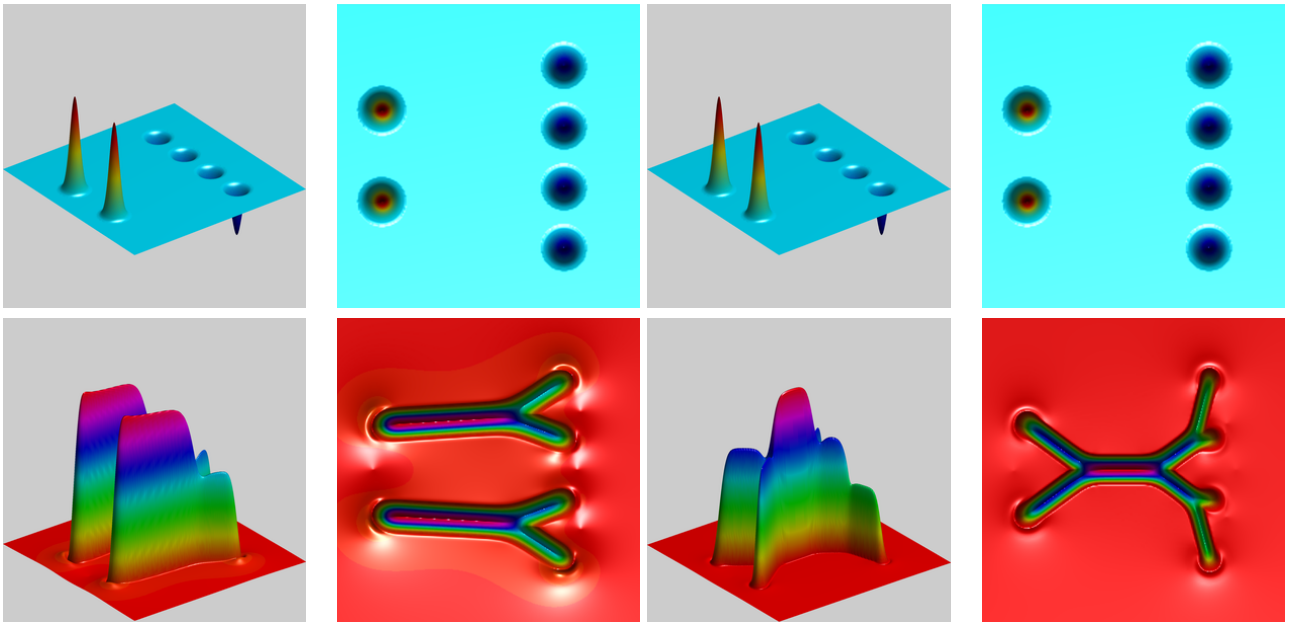

Figure IX.3: Optimal irrigation with $\alpha = 1/2 + 1/10$ and $\alpha = 1/2 + 1/4$

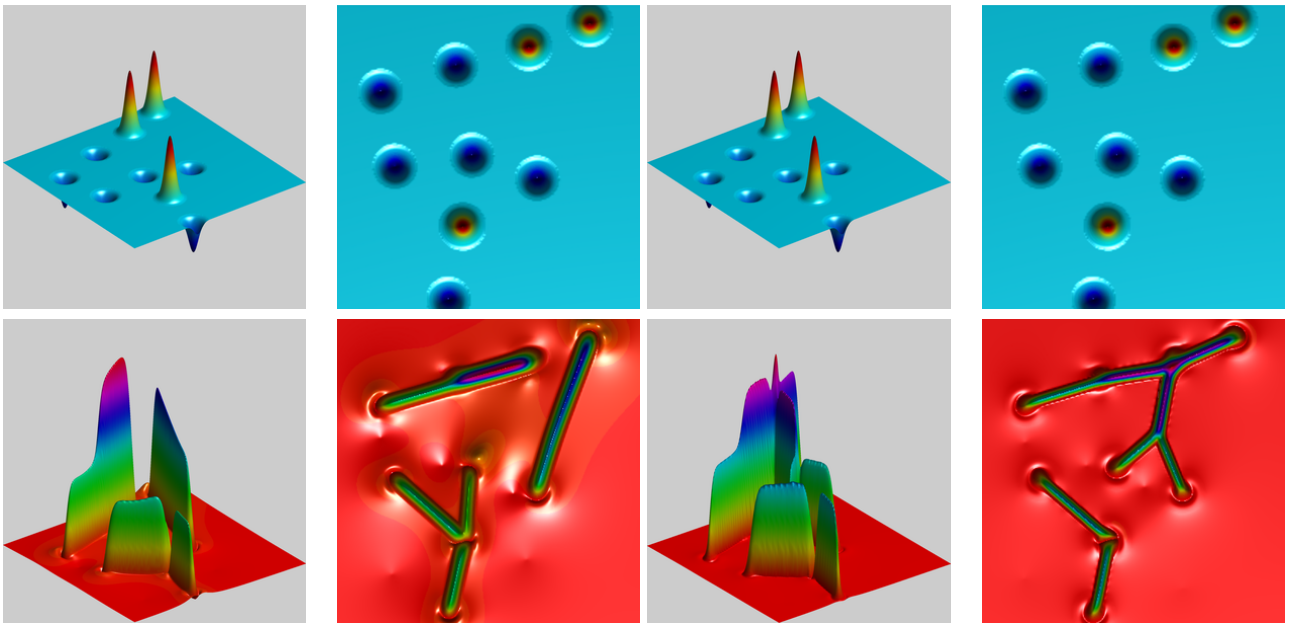

Figure IX.4: Optimal irrigation with $\alpha = 1/2 + 1/10$ and $\alpha = 1/2 + 1/4$

[2] F. Santambrogio. A modica-mortola approximation for branched transport. 2009.

[3] Qinglan Xia. The formation of a tree leaf. *ESAIM Control Optim. Calc. Var.*, 13(2):359–377 (electronic), 2007.

[4] Guoliang Xue, Theodore P. Lillys, and David E. Dougherty. Computing the minimum cost pipe network interconnecting one sink and many sources. *SIAM J. Optim.*, 10(1):22–42 (electronic), 1999.