Modèle hydrostatique pour les écoulements à surface libre tridimensionnels et schémas numériques

Astrid Decoene

THESE DE DOCTORAT DE L'UNIVERSITE PARIS 6

PIERRE ET MARIE CURIE

- Ecole Doctorale de Sciences Mathématiques -
de Paris Centre

Spécialité

Mathématiques Appliquées

Présentée par

Astrid DECOENE

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITE PARIS 6

Sujet de la thèse :

Modèle hydrostatique pour les écoulements à surface libre tridimensionnels et schémas numériques

soutenue le 24 mai 2006 devant le jury composé de :

M. Jean-Frédéric Gerbeau, Directeur de thèse
M. Benoît Perthame, Directeur de thèse
M. Fausto Saleri, Rapporteur
M. Charles-Henri Bruneau, Rapporteur
M. Frédéric Hecht, Président du jury
M. Jean-Michel Hervouet, Examinateur
Mme. Marie-Odile Bristeau, Examinateur.

REMERCIEMENTS

Je souhaite remercier toutes les personnes qui ont contribué à l'engagement et à l'aboutissement de ce long parcours qu'a été ma thèse.

Merci donc à Benoît Perthame et Jean-Michel Hervouet de m'avoir proposé ce sujet.

Ma profonde gratitude va à Jean-Frédéric Gerbeau qui m'a patiemment guidé pendant ces trois années (et quelques...). Merci pour son exigence, sa disponibilité et son soutien.

Je remercie sincèrement Emile Razafindrakoto, en particulier pour la confiance qu'il a toujours placée en moi.

Je ne voudrais surtout pas oublier Marie-Odile Bristeau, qui m'a si chaleureusement acceuillie et dont les conseils et l'expérience m'ont été précieux.

Je tiens également à exprimer toute ma gratitude à mes rapporteurs, Charles-Henri Bruneau et Fausto Saleri, pour avoir accepté de lire et d'évaluer ma thèse, ainsi qu'à Frédéric Hecht pour avoir présidé ce jury.

Vincent Martin m'a permis de collaborer sous le soleil marseillais puis dans les sombres bureaux rocquencourtois à son travail autour des écoulements de sang dans les artères : cette petite parenthèse dans ma thèse s'est avérée extrêment enrichissante et satisfaisante, et je l'en remercie.

Je tiens finalement à saluer et remercier mes collègues de l'INRIA et du LNHE : Emmanuel, Victor (por donde andas ?), Claire, Paola, Manu, Thierry (merci encore), Reza, Florence, Didier, Mohammed, Cécile, Miguel, Amel, Nicolas, Iria, Nuno, Paula, Jacques, Loïc, Hatem... pour ne citer qu'eux.

Merci bien sûr à mon père et ma mère pour leur soutien inconditionnel, ainsi qu'à mes soeurettes adorées.

Merci enfin et surtout à Laurent.

Modèle hydrostatique pour les écoulements à surface libre tridimensionnels et schémas numériques.

Résumé : Cette thèse a pour objectif l'approfondissement de l'étude des équations régissant les écoulements à surface libre en dimension trois. Nous proposons d'une part une nouvelle formulation variationnelle du problème hydrostatique aboutissant à un problème semi-discretisé en temps bien posé. Nous en faisons l'analyse mathématique et nous montrons quelques résultats numériques obtenus après programmation de l'approximation de ce problème dans le logiciel Telemac-3D développé au Laboratoire National d'Hydraulique et Environnement (LNHE) d'edf. D'autre part, nous étudions la réinterprétation dans le cadre ALE de la méthode de discrétisation verticale de domaines tridimensionnels appelée transformation sigma, et nous en proposons une généralisation permettant d'améliorer la représentation des stratifications dans un écoulement. Finalement, nous présentons un schéma ALE-MURD conservatif pour la résolution des équations de convection linéaires posées sur un domaine mobile. Une condition particulière doit être vérifiée afin que le schéma soit conservatif lorsque le domain bouge effectivement. Nous montrons comment assurer cette contrainte dans le cas particulier où le domaine est tridimensionnel et ne bouge que selon la verticale. Ce résultat est illustré dans le cadre des écoulements à surface libre en dimension trois.

Mots-clés : équations de Navier-Stokes à surface libre, écoulements tridimensionnels, Telemac-3D, modèle hydrostatique, éléments finis, ALE, transformation sigma, problèmes de convection linéaire, schémas distribués, lois de conservation géométrique.

Hydrostatic model for the three-dimensional free surface flow problem and numerical schemes.

Abstract : This PhD thesis aims to deepen the analysis of the equations governing the three-dimensional free surface flows. On one hand we present a new weak formulation of the hydrostatic problem leading to a well-posed time-discrete problem. This problem is analysed mathematically and its resolution is implemented into the Telemac-3D system, developed at the Laboratoire National d'Hydraulique et Environnement (LNHE), edf. Some numerical results are shown. On the other hand, we study the ALE interpretation of the sigma transformation method for the vertical discretization of three-dimensional domains. Especially we propose a generalization allowing to improve the representation of stratifications in a flow. Finally, we

introduce an ALE-MURD scheme for the linear advection problem posed on moving domains. A particular constraint must be satisfied for the scheme to be conservative when the domain moves. We show how to ensure this constraint in the particular case where the domain is three-dimensional and only moves in the vertical direction. This result is illustrated numerically in the framework of the three-dimensional free surface flow problem.

Key words : free-surface Navier-Stokes equations, three-dimensional flows, Telemac-3D, hydrostatic model, finite elements, ALE, sigma transformation, linear advection problem, residual distributive schemes, geometric conservation laws.

Table des matières

Chapitre 1

Introduction

1.1 Motivation et objectif de la thèse

Cette thèse a été proposée dans le cadre d'une collaboration entre l'Institut National de Recherche en Informatique et Automatique (INRIA) et le Laboratoire National d'Hydraulique et Environnement (LNHE). Ce dernier fait partie de la division Recherche et Développement d'EDF. Il a notamment développé depuis 1987 un système hydro-informatique dédié aux écoulements incompressibles à surface libre. Initialement basé sur un modèle bidimensionnel (2D) découlant du système de Saint-Venant, ce code, baptisé Telemac et basé sur les travaux de Hervouet *et al.* [50, 66], a plus récemment intégré un module résolvant le problème en dimension 3 [63].

Suivant cette évolution, cette thèse a pour objectif l'approfondissement de l'étude des équations régissant les écoulements à surface libre en dimension 3, parallèlement à l'amélioration de leur résolution par le code Telemac-3D.

Nous nous sommes intéressés aux équations de Navier-Stokes décrivant le mouvement d'un fluide Newtonien incompressible dans un domaine tridimensionnel (3D) à surface libre. Celles-ci s'appliquent à un nombre incalculable d'écoulements. Citons par exemple les eaux des lacs, des rivières, des océans, ainsi que les écoulements dans l'atmosphère. Leur étude est donc indispensable dans des domaines d'application tels que l'environnement ou la climatologie. De fait, la maîtrise de ces équations est extrêment utile à la protection de l'environnement, grâce par exemple aux études de qualité de l'eau, à la création et protection des ouvrages qui interagissent avec lui (impact et stabilité des aménagements) et à la protection de l'homme face aux catastrophes naturelles. Elle joue en outre un rôle primordial en météo ou dans la prévision des changements climatiques.

Ces équations, pour être bien connues, n'en sont pas moins très complexes, en raison notamment de leur caractère instationnaire, de la mobilité

et de l'inconnaissance a priori du domaine sur lequel elles sont posées, ainsi que des phénomènes de turbulence. Elles ont fait et font toujours l'objet de recherches soutenues en analyse mathématique théorique et numérique. Hormis pour quelques cas particuliers, l'existence de solutions régulières aux équations de Navier-Stokes 3D n'a d'ailleurs pour l'instant jamais été démontrée. Soulignons ici l'importance de l'analyse mathématique détaillée des équations régissant un problème physique, à la fois pour sa compréhension, sa résolution, mais aussi pour le développement de nouveaux modèles plus adaptés aux besoins de l'ingénierie.

La maîtrise de la mécanique des fluides à surface libre dans l'ingénierie était initialement assurée par les maquettes et modèles physiques. Mais avec l'avènement de la modélisation numérique, ces outils trop coûteux et dépassés ne sont plus utilisés aujourd'hui que pour le calibrage et la validation des modèles, ou pour des applications très particulières comme les évacuateurs de crues. En effet, ces 30 dernières années ont connu le grand essor de la simulation numérique, grâce d'une part au progrès technologique (la progression de la capacité de calcul et de mémoire des ordinateurs), et d'autre part au développement considérable des méthodes numériques en mécanique des fluides. Sans oublier le rôle important joué dans cet avènement par d'autres outils essentiels comme les logiciels de visualisation et les sytèmes d'acquisition de données et d'information géographique qui ont considérablement simplifié et cadencé la position des problèmes et l'analyse des résultats. C'est l'ensemble de ces raisons qui ont conduit les modèles numériques à s'imposer aujourd'hui comme un outil indispensable à la modélisation d'écoulements à l'échelle géophysique.

Cependant, les outils accessibles aujourd'hui sont encore loin d'être entièrement satisfaisants, et de nombreux phénomènes ne peuvent encore être simulés, tout particulièrement certains écoulements tridimensionnels à surface libre. L'étude des différents modèles ainsi que la recherche de nouveaux algorithmes doivent donc être poursuivies.

Le code Telemac-3D résout les équations de Navier-Stokes *Reynolds-averaged* en dimension 3 pour les écoulements incompressibles, turbulents, à surface libre et avec approximation de Boussinesq, par la méthode de discrétisation spatiale des éléments finis. Il s'applique aux écoulements géophysiques à surface libre dans des géométries compliquées. Couplé à d'autres modèles dans le système global Telemac, il s'étend au-delà de l'hydrodynamique, permettant le traitement de problèmes liés à l'environnement comme l'étude de la qualité de l'eau. Il a été validé par de nombreuses applications pratiques, la plupart en hydraulique fluviale ou maritime, comme par exemple le transport de sédiments dans des estuaires [83, 84]. Il est notamment connu pour sa bonne gestion des bancs découvrants et des zones sèches, assurant la conser-

vation exacte de la masse d'eau.

Le système présente néanmoins encore de nombreuses faiblesses et son développement doit être approfondi. Le temps de calcul ainsi que la précision des résultats pourraient être améliorés, mais on pourrait surtout envisager la simulation de nouveaux phénomènes, dans de nouvelles géométries.

Aux premières heures de cette thèse, notre attention a été attirée sur la supériorité du temps de calcul de Telemac-3D par rapport à d'autres systèmes connus résolvant les mêmes équations. Cependant, l'étude plus approfondie de ces systèmes et de leur technique de discrétisation nous ont montré que certaines simplifications étaient faites, qui rendent les algorithmes plus efficaces. Ainsi, un modèle basé sur la discrétisation en escalier du fond et de la surface du domaine, et utilisant un maillage fixe, comme par exemple les modèles développés par Casulli *et al.* [24, 27] ou par Saleri, Miglio *et al.* [42, 97], gagne naturellement en simplicité et rapidité de calcul. Mais il gagne dans ces domaines ce qu'il perd en précision (cf. [54, 93]). Signalons toutefois que la force de ces modèles ne se résume pas à ces seuls éléments : les algorithmes utilisés sont particulièrement performants même indépendamment du maillage.

Un défaut du code Telemac-3D signalé par certains utilisateurs est la difficulté à représenter des stratifications, notamment les stratifications horizontales dans des domaines à fond fortement variable. Ce problème est en fait lié à la *transformation sigma classique*, méthode choisie dans ce code pour définir le maillage ainsi que son déplacement. Nous avons proposé une extension de cette méthode permettant d'adapter davantage le maillage aux besoins particuliers de chaque application, et notamment à la simulation de stratifications horizontales.

Cette extension a mis en évidence le défaut de conservativité d'un schéma de convection de Telemac-3D lorsque la *transformation sigma classique* n'est pas utilisée. Cela nous a engagés dans un travail autour des schémas numériques distribués pour la convection linéaire d'un scalaire, et plus particulièrement de leur application aux domaines mobiles. Les résultats obtenus peuvent s'appliquer au module de convection du code Telemac-3D, ce qui nous a permis de corriger son défaut de conservativité.

Nous avons donc orienté notre travail dans le sens des attentes formulées initialement. Cependant, nos recherches nous ont également mené dans d'autres directions. Nous avons notamment proposé une nouvelle formulation variationnelle du problème étudié. Celle-ci permet d'obtenir, après discrétisation en temps et division en pas fractionnaires, un sous-problème bien posé couplant la vitesse horizontale du fluide et la surface libre de l'écoulement. En effet, ce sous-problème est linéaire, symétrique et surtout coercif. Nous en avons fait l'analyse mathématique et avons mis en oeuvre sa résolution dans le code Telemac-3D.

Indiquons que certains de nos travaux ont d'ores et déjà été intégrés dans le code Telemac-3D. D'autres sont restés au stade d'étude de faisabilité et seront peut-être intégrés au système une fois optimisés.

1.2 Les écoulements à surface libre : modélisation et résolution

1.2.1 Les équations fondamentales de l'hydrodynamique

Les équations de Navier-Stokes sont en fait la forme adaptée aux fluides de la relation fondamentale de la dynamique. Elles ont été établies par Sir Gabriel Stokes [133] en 1845, se basant en partie sur les travaux de l'ingénieur Henri Navier [104] datés de 1823, près de deux siècles après la publication des travaux de Newton en 1687. Elles décrivent tout écoulement de fluide dans un domaine tridimensionnel à travers sa vitesse $\boldsymbol{U}(x,y,z,t)$, sa densité $\rho(x,y,z,t)$ et sa pression $p(x,y,z,t)$, ainsi que le domaine $\Omega(t)$ dans lequel il se meut. Elles sont posées à tout instant t sur $\Omega(t)$ et s'écrivent :

$$\begin{aligned} \frac{d\,(\,\rho\,\boldsymbol{U}\,)}{dt} \;-\; \mathrm{div}\,(\,\boldsymbol{\sigma_T}\,) \;&=\; \rho\,\boldsymbol{f} \;+\; \rho\,\boldsymbol{g}, \\ \frac{\partial \rho}{\partial t} \;+\; \mathrm{div}\,(\,\rho\,\boldsymbol{U}\,) \;&=\; 0, \end{aligned} \tag{1.1}$$

où $\boldsymbol{\sigma_T}$ désigne le tenseur des contraintes, $\boldsymbol{f}(x,y,z,t)$ la somme des forces extérieures pouvant s'appliquer sur le fluide, et $\boldsymbol{g} = (0,0,-g)$ la gravité.

Dans le cas des écoulements à surface libre, le domaine $\Omega(t)$ varie effectivement en fonction du temps, et cette variation fait partie des inconnues du problème, ce qui comme nous le verrons complique considérablement l'étude et la résolution de ce dernier.

1.2.2 Principaux modèles

Afin de simplifier les équations de l'hydrodynamique à surface libre, diverses approximations ont été proposées. Nous ne présentons ici que les trois d'entre elles qui sont le plus souvent utilisées.

Les variations de densité sont relativement faibles dans les milieux aqueux naturels. L'approximation de Boussinesq (cf. [130]) consiste alors à négliger leur effet sur la masse du fluide, mais à tenir compte de leur influence sur la force de pesanteur appliquée au fluide. Cette approximation permet de supposer la densité constante dans l'équation de continuité, et donc de considérer un écoulement incompressible.

Pour faciliter le traitement de la turbulence, les équations de Navier-Stokes peuvent être moyennées en temps : ce sont les équations dont l'appellation anglaise est *Reynolds-averaged Navier Stokes* (RANS) [118]. L'effet des fluctuations turbulentes sur l'écoulement moyenné est alors pris en compte par des modèles de turbulence (cf. par exemple [140]).

Enfin, l'approximation hydrostatique consiste à négliger les accélérations de la vitesse verticale, ainsi que la diffusion et le terme source dans l'équation de quantité de mouvement sur w, si bien que cette équation se réduit à :

$$\frac{\partial p}{\partial z} = -\rho g\,. \tag{1.2}$$

La pression est alors considérée hydrostatique, c'est-à-dire qu'elle ne varie sur la verticale qu'en fonction du poids de la colonne d'eau. Typiquement, cette hypothèse est appliquée dans les milieux peu profonds, pour des écoulements dont la longueur d'onde est grande par rapport à la profondeur.

Ces approximations ont permis la dérivation de modèles simplifiés, plus accessibles à la résolution, mais dont l'analyse mathématique n'est pas, notons-le, d'autant plus aisée.

Les modèles bidimensionnels existants sont le plus souvent basés sur *le système de Saint-Venant*, dont l'appellation anglaise est *shallow water equations*, de par son application aux écoulements en eaux peu profondes. Ce système, introduit de manière indépendante en 1871 par Jean-Claude Barré de Saint-Venant, peut être vu comme une intégration sur la verticale des équations de Navier-Stokes tridimensionnelles avec hypothèse hydrostatique et approximation de Boussinesq. Cependant, l'existence de termes non linéaires oblige à quelques hypothèses et approximations supplémentaires. Il s'agit dans tous les cas d'un système 2D portant sur la hauteur d'eau et la vitesse horizontale moyennée sur la verticale. L'intérêt majeur du système de Saint-Venant est de permettre d'aborder des problèmes de physique 3D et instationnaires par un système posé sur un domaine 2D et invariant en temps. Sa validité expérimentale et sa robustesse reconnues, ainsi que la grande quantité de méthodes numériques efficaces développées en font, aujourd'hui encore, le modèle le plus utilisé en mécanique des fluides à surface libre. Cependant, son domaine de validité reste limité aux écoulements en milieux peu profonds à faible variation de la bathymétrie, à faible accélération de la vitesse verticale, à viscosité et coefficient de frottement également faibles, et dont le gradient de surface libre reste borné (cf. par exemple l'étude asymptotique du système de Saint-Venant effectuée par Perthame et Gerbeau dans [110]). Notons également que le système de Saint-Venant implique une perte d'information de par l'utilisation d'une vitesse moyennée. La liste des références sur le sujet étant infiniment longue, nous n'indiquons ici que les travaux récents de Audusse [7], où de nombreuses références pourront être

trouvées. Notons que le système Telemac comprend un module 2D qui résout le système de Saint-Venant. Il en est de même pour le code industriel TRIM et sa version pour maillages non structurés UNTRIM, développés par Casulli *et al.* [24, 27], et largement utilisés pour des études d'ingénierie notamment dans des régions côtières avec effet de marée.

L'un des premiers modèles numériques tridimensionnels pour les écoulements à surface libre a été développé à partir de 1970 [33].

Il existe un type de modèle 3D bien établi, basé sur ce qu'on appelle souvent les *équations de Saint-Venant tridimensionnelles*. Celles-ci dérivent des équations de Navier-Stokes avec hypothèse hydrostatique et approximation de Boussinesq. L'équation de continuité est intégrée sur la verticale, permettant d'obtenir une équation supplémentaire sur la variable surface libre. L'équation de continuité tridimensionnelle est toutefois conservée dans le système : elle permet de coupler les composantes horizontale et verticale de la vitesse. Remarquons que cette manipulation sous-entend que la surface libre est une fonction univoque, qu'elle est donc définie de façon unique sur un domaine bidimensionnel. Ceci représente une approximation supplémentaire, puisque l'on rencontre dans la nature des écoulements ne vérifiant pas cette hypothèse, comme par exemple les vagues déferlantes.

Dans leur version pour fluide Newtonien à densité constante, les *équations de Saint-Venant tridimensionnelles* portent sur la vitesse horizontale $\boldsymbol{u}(x,y,z,t)$, la surface libre $\eta(x,y,t)$ et la vitesse verticale $w(x,y,z,t)$ du fluide. Elles sont posées à tout instant t sur un domaine $\Omega(t)$ décrit par la variable surface libre $\eta(x,y,t)$ et par le fond $b(x,y)$, supposé constant :

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{U}\cdot\boldsymbol{\nabla}\boldsymbol{u} - \operatorname{div}(\nu\boldsymbol{\nabla}\boldsymbol{u}) + g\,\boldsymbol{\nabla}_2\eta = \boldsymbol{f}^{hor}, \tag{1.3}$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\int_b^\eta \boldsymbol{u}\,dz + \frac{\partial}{\partial y}\int_b^\eta \boldsymbol{u}\,dz = 0, \tag{1.4}$$

$$\operatorname{div}_2\,\boldsymbol{u} + \frac{\partial w}{\partial z} = 0, \tag{1.5}$$

où $\boldsymbol{f}^{hor}$ désigne le terme source horizontal et $\boldsymbol{\nabla}_2$, div_2 sont les opérateurs de gradient et divergence horizontaux.

Les modèles basés sur ces équations (voir par exemple Heaps [62], Casulli *et al.* [25], Miglio, Saleri *et al.* [42], ainsi que Hervouet [66]), que nous appellerons dans toute la suite modèles hydrostatiques 3D, sont plus riches que les modèles 2D type Saint-Venant, puisqu'ils donnent un profil vertical de la vitesse horizontale et permettent de calculer la vitesse verticale. De plus, ils permettent de tenir compte de façon naturelle des conditions limites au fond et à la surface, et de simuler certains effets 3D comme les stratifications

et les recirculations verticales. Cependant, de par l'approximation hydrostatique, leur domaine de validité est également limité aux écoulements en eau peu profonde et à faible accélération de la vitesse verticale, dans lesquels les gradients de pression et de fond restent faibles.

L'utilisation de modèles hydrostatiques tridimensionnels dans l'ingénierie est longtemps restée limitée à cause de leur besoin élevé en temps de calcul. Ils sont pourtant désormais très utilisés, et appliqués dans des situations allant au-delà même des limites de leur domaine de validité. La plupart des logiciels utilisant ce modèle ont été développés par extension d'un modèle bidimensionnel de type Saint-Venant, afin de décrire des effets tridimensionnels. Nous pouvons citer les modèles océanographiques POM [19] et ROMS [123], ainsi que les codes industriels Delft3D-Flow [38] et ADCIRC [81], dédiés aux océans, aux milieux côtiers et aux estuaires. Notons que les codes industriels tridimensionnels (UN)TRIM-3D [24, 27] et Telemac-3D [63] étaient initialement basés sur le modèle hydrostatique et en proposent toujours l'utilisation.

De nombreux auteurs ont également travaillé sur la résolution complète des équations de Navier-Stokes à surface libre, c'est-à-dire sans supposer la pression hydrostatique.

En 1995, Casulli et Stelling présentent dans [26] le premier algorithme complet pour la résolution de ces équations. Ils proposent pour cela une décomposition de la pression p en partie hydrostatique p_h et correction dynamique p_d : $p = p_h + p_d$. Cette décomposition permet de faire apparaître explicitement la variable surface libre. Les modèles basés sur cette idée sont souvent appelés quasihydrostatiques, et peuvent être considérés comme une extension des modèles hydrostatiques 3D présentés précédemment : ils comprennent une étape hydrostatique, dans laquelle sont résolues les *équations de Saint-Venant tridimensionnelles* afin de calculer la surface libre, ainsi qu'une étape nonhydrostatique, dans laquelle les vitesses sont corrigées par la pression dynamique afin d'assurer l'équation de continuité. Les équations de Navier-Stokes incompressibles s'écrivent alors comme suit à tout instant t sur un domaine $\Omega(t)$ décrit par la variable surface libre $\eta(x, y, t)$ et par le fond $b(x, y)$, supposé constant :

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{U} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \operatorname{div}(\nu \boldsymbol{\nabla} \boldsymbol{u}) + g \boldsymbol{\nabla}_2 \eta + \boldsymbol{\nabla} p_d = \boldsymbol{f}^{hor}, \tag{1.6}$$

$$\frac{\partial w}{\partial t} + \boldsymbol{U} \cdot \boldsymbol{\nabla} w - \operatorname{div}(\nu \boldsymbol{\nabla} w) + \boldsymbol{\nabla} p_d = f_z, \tag{1.7}$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} \int_b^\eta \boldsymbol{u} \, dz + \frac{\partial}{\partial y} \int_b^\eta \boldsymbol{u} \, dz = 0, \tag{1.8}$$

$$\operatorname{div}_2 \boldsymbol{u} + \frac{\partial w}{\partial z} = 0, \tag{1.9}$$

où $\boldsymbol{u}$ et w sont les composantes horizontale et verticale de la vitesse et f_z le terme source vertical. Mahadevan *et al.* [82] ont développé un modèle océanographique nonhydrostatique basé sur cette approche. Ils ont de plus démontré que le problème nonhydrostatique est bien posé, et qu'il permet de traiter les conditions limites physiques. Plusieurs modèles quasihydrostatiques 3D ont été developpés : citons par exemple les travaux de Miglio, Saleri *et al.* [97] ainqi que Kocyigit *et al.* [77], et le modèle 2D vertical de Zhou et Stansby [144]. Le code Telemac-3D propose depuis 1998 la résolution des équations nonhydrostatiques avec décomposition de la pression (voir Jankowski [75]), de même que le système (UN)TRIM-3D [28].

D'autres approches ont été développées pour traiter le problème des écoulements tridimensionnels à surface libre, notamment dans des cas particuliers. Le mouvement du domaine est alors établi indépendamment de la résolution des équations de Navier-Stokes 3D, grâce à des méthodes très variées. Nous en détaillerons certaines plus loin dans cette introduction. Citons en particulier Yost [142], Maury [89], Hodges [69], ainsi que les travaux plus récents de Namin *et al.* [103] et Yuan *et al.* [141].

Les modèles nonhydrostatiques permettent de traiter les écoulements auxquels l'approximation hydrostatique ne s'applique pas. Ceci à cause d'un fort gradient de surface libre ou de fond, ou de petites longueurs d'onde. Ils prennent en compte des phénomènes 3D plus variés, comme les mouvements orbitaux, les recirculations verticales intensives, les écoulements autour d'obstacles, dans lesquels la pression est fortement nonhydrostatique.

De nouveaux modèles sont néanmoins recherchés. Un modèle intermédiaire entre Saint-Venant et Navier-Stokes 3D a par exemple été développé sous le nom de *modèle Saint-Venant multi-couches*. Le domaine tridimensionnel est décomposé en plusieurs couches, dont l'interface est calculée grâce à la résolution du système de Saint-Venant sur chacune de ces couches. Cette résolution permet également d'obtenir une vitesse horizontale moyennée par couche. Plusieures approches existent, notamment celle de Audusse [8]. Des intercomparaisons entre les résultats fournis par ce modèle et ceux fournis par le modèle hydrostatique 3D de Telemac3D ont été réalisées par Bristeau *et al.* dans [9].

1.2.3 Aperçu des principales méthodes numériques

Trois méthodes d'approximation basées sur la construction de maillages sont le plus souvent utilisées en mécanique des fluides. La plus ancienne est celle des différences finies. Elle s'avère très simple et efficace, permettant la construction de schémas peu coûteux. Cependant, elle n'est utilisable qu'avec des maillages structurés et ne s'applique donc qu'aux problèmes posés sur des géométries simples. Elle pose en particulier des problèmes quant à l'imposition des conditions limites et à la conservation de la masse. Cette

méthode est notamment utilisée en océanographie, où il n'est pas nécessaire de discrétiser finement les bords latéraux du domaine, par exemple dans les modèles POM [19] et ROMS [123]. Mais elle est également utilisée dans les logiciels TRIM3D [24] et Delft-3D-Flow [38].

La méthode des éléments finis [113] et celle des volumes finis [138] ont émergé peu de temps après, et leur intérêt a immédiatement été reconnu puisqu'elles sont applicables aux maillages non structurés. Elles offrent une grande souplesse pour discrétiser des géométries complexes et imposer les conditions limites. D'une part, la méthode des éléments finis offre un cadre théorique rigoureux et une mise en oeuvre relativement simple. Les modèles hydro- et nonhydrostatiques proposés par Miglio, Saleri *et al.* dans [42, 97] pour les écoulements 3D à surface libre sont par exemple basés sur la méthode des éléments finis, de même que les travaux de Yost [142] et Maury [89]. Le code Telemac est entièrement basé sur cette méthode. D'autre part, les volumes finis permettent de résoudre simplement les équations sous forme de bilan, et d'assurer ainsi la conservation au niveau local des quantités calculées. Les éléments finis, eux, ne permettent le plus souvent de l'assurer qu'au niveau global. L'imposition des conditions limites est également parfois facilitée par l'approche volumes finis. Le modèle différence finies TRIM a par exemple été étendu à la méthode des volumes finis (cf. le modèle UNTRIM [27]) afin de permettre le traitement de maillages non structurés. Mahadevan a également opté pour cette méthode dans [82].

Aujourd'hui, la tendance est à la combinaison de ces méthodes de discrétisation, afin de tirer parti des avantages de chacune d'entre elles. Les schémas distributifs [1, 34, 108] par exemple, qui connaissent un fort engouement en mécanique des fluides, sont basés sur la méthode des volumes finis mais utilisent la structure de données des éléments finis.

Il existe cependant une alternative, plus récente, à ces trois méthodes d'approximation : la méthode *Smooth Particle Hydrodynamics* (SPH), initialement développée en astrophysique puis appliquée aux écoulements à surface libre [99]. Il s'agit d'une méthode entièrement lagrangienne et sans maillage, utilisant des particules qui se déplacent sous l'influence des forces de pression, de viscosité et de pesanteur. Elle permet de simuler des cas complexes, en particulier les écoulements mixtes, le déferlement, et tous les cas où la surface libre n'est pas univoque. De plus, elle traite avec grande précision les fortes déformations de surface libre. Encore peu utilisée en hydrodynamique, elle a néanmoins été récemment appliquée à un écoulement turbulent tridimensionnel dans un canal ainsi qu'à un cas de rupture de barrage [139].

L'une des grandes difficultées des écoulements qui nous occupent vient du mouvement de la surface libre. Plusieurs approches existent pour le traitement des équations dans un domaine mobile. Il y a d'une part l'approche lagrangienne, dans laquelle le maillage est délacé à la vitesse du fluide. Elle

permet de représenter de manière précise la surface libre et d'y imposer facilement les conditions limites, mais sa mise en oeuvre est compliquée et son coût élevé. De plus, pour de larges distorsions dans l'écoulement les éléments du maillage peuvent devenir singuliers.

L'approche eulérienne traite au contraire le problème dans un maillage fixe. Cette approche est plus simple puisqu'elle ne nécessite pas de remaillage et que le problème des éléments déformés ne se pose pas. Cependant, il est difficile de suivre la surface libre de façon précise avec ce type de maillages et d'appliquer les conditions limites naturelles. Cette approche est par exemple appliquée aux écoulements à surface libre dans [102] et dans les travaux de Miglio, Saleri *et al.* [42, 97], ainsi que dans les systèmes TRIM-3D et UNTRIM-3D [24, 27].

Une nouvelle approche a été introduite en 1971 : la formulation *Arbitrary Lagrangian Eulerian* (ALE) [67]. Elle consiste à déplacer le maillage à une vitesse arbitraire, différente du fluide, permettant ainsi un réarrangement continu des mailles. De plus, ce formalisme rend la mise à jour du maillage totalement implicite et permet de s'affranchir de l'interpolation des variables à chaque pas de temps sur le nouveau maillage. La formulation ALE permet également de discrétiser plus facilement les dérivées partielles en temps. Elle a été appliquée aux écoulements à surface libre par de nombreux auteurs, comme par exemple Huerta et Liu [70], Maury [89], Soulaïmani *et al.* [129] et Cairncross *et al.* [23]. La *transformation sigma* est une forme simple de méthode ALE [66, 144, 77], qui consiste à déplacer le maillage uniquement selon la verticale, en suivant la forme et les mouvement du fond et de la surface libre du domaine. L'utilisation de cette méthode est très répandue, notamment en océanographie et en climatologie (cf. les modèles POM [19] et ROMS [123]), son rapport avec le formalisme ALE reste néanmoins encore peu connu. Elle assure la représentation précise et continue de la surface libre et du fond, simplifie la prescription des conditions limites et permet d'incorporer facilement des couches limites. Cependant, les formes classiques de la transformation sigma, dont l'une est utilisée dans le code Telemac-3D, limitent le choix de la discrétisation verticale du domaine et peuvent entraîner des erreurs graves. Notons que la première partie de cette thèse est consacrée à la description détaillée du formalisme ALE et de son application en éléments finis, à l'extension d'un type particulier de schémas convectifs aux domaines variables grâce au formalisme ALE, ainsi qu'au développement d'une forme plus générale de la *transformation sigma* pour les codes à surface libre.

Mais la difficulté des écoulements à surface libre ne se limite pas au simple mouvement du domaine. Celui-ci est en fait lui-même une inconnue du problème puisqu'il est défini à chaque instant par la position de la surface libre. Plusieurs techniques ont été développées pour surmonter cette difficulté supplémentaire. Les modèles 3D hydrostatiques et quasihydrostatiques utilisent

ce qu'on appelle le suivi de surface (*surface tracking*). Les équations résolues font intervenir de façon explicite la variable surface libre, soit à travers une équation 2D obtenue par intégration sur la verticale de l'équation de continuité, soit à travers la condition limite cinématique à la surface libre. Cette technique permet d'approcher de façon précise la surface libre, mais elle sous-entend que celle-ci est une fonction univoque. Le suivi de surface est également appliqué dans les modèles proposés par Yost [142],Hodges [69] et Yuan [141]. Notons que dans [89], Maury utilise la méthode des caractéristiques pour résoudre la condition cinématique à la surface libre, mais celle-ci n'est définie de façon univoque qu'au voisinage de chaque point.

Une alternative à cette méthode consiste à suivre plutôt le volume de fluide, d'où son appellation anglaise *fluid volume tracking*. Il existe aujourd'hui deux variantes de cette méthode. D'une part, la méthode MAC (Marker and Cell [61]) consiste à suivre le domaine grâce à des particules sans masse se déplaçant à la vitesse du fluide dans un maillage fixe. La position de la surface libre est définie par la position des marqueurs dans le domaine. Le principal inconvénient de cette méthode est qu'elle requiert une importante place mémoire car un grand nombre de particules est nécessaire pour les simulations. D'autre part, la méthode VOF (Volume of Fluid [68]) consiste à ajouter une inconnue ϕ au système, valant un ou zéro selon la présence ou l'absence de fluide dans chaque élément du maillage, ainsi qu'une équation caractérisant le transport de cette fonction par la vitesse du fluide $\boldsymbol{U}$,

$$\frac{\partial \phi}{\partial t} \; + \; \boldsymbol{U} \cdot \boldsymbol{\nabla} \phi \; = \; 0 \, . \tag{1.10}$$

La résolution de cette équation exige la mise au point de schémas numériques parfaitement conservatifs, monotones et à diffusion numérique très limitée. Cette méthode est largement répandue (cf. ar exemple [101, 86]), car elle est très robuste et permet de représenter une surface libre non-univoque et des fortes déformations. Mais elle fournit une représentation de la surface libre moins précise que les méthodes de suivi de surface, et elle pose des difficultés quant à l'imposition des conditions limites. Notons qu'il existe des méthodes mixtes.

En ce qui concerne la résolution proprement dite des équations de Navier-Stokes, plusieurs choix se présentent. Le schéma en temps établit tout d'abord le traitement implicite ou explicite de chacun des termes, permettant par exemple de linéariser les équations. Casulli *et al.* ont développé un algorithme semi-implicite en différences finies (étendu plus tard aux volumes finis) pour la résolution des équations hydrostatiques [25] et nonhydrostatiques [28]. Ils montrent que cet algorithme est inconditionnellement stable. Cette approche semi-implicite a très souvent été appliquée [66, 42, 97, 144, 77], mais certains auteurs ont également traité le problème de façon complètement implicite [103, 141]. Le système peut ensuite être résolu de façon directe, grâce

à des solveurs non-linéaires et itératifs [70, 142, 103], ou en utilisant plutôt une approche découplée. Cette approche consiste à résoudre le problème en plusieurs petits sous-problèmes linéaires, en appliquant les méthodes numériques les plus adaptées aux caractéristiques mathématiques de chacun d'entre eux. Il existe diverses méthodes permettant le découplage des équations, comme la méthode des pas fractionnaires (utilisée par exemple dans [26, 42, 97, 66]), ou les méthodes de projection (appliquées notamment dans [69, 75]).

Nous nous permettons d'en rester à cette description très globale des méthodes numériques existantes pour la résolution des équations de Navier-Stokes, leur description exhaustive étant aujourd'hui presque impossible.

1.3 Travail effectué

Cette thèse a tout naturellement débuté par une phase de recherche bibliographique autour du problème qui nous intéresse, les écoulement à surface libre en dimension trois. Nous avons notamment étudié de façon détaillée les différents modèles traités par le code Telemac-3D, ainsi que les méthodes numériques utilisées (cf. Hervouet [66]). Nous avons également étudié de façon approfondie les modèles 3D développés par Casulli *et al.* en différences finies [25, 26] et en volumes finis [27, 28], ainsi que ceux développés par Miglio, Saleri *et al.* en éléments finis [42, 97].

Les travaux de ces auteurs ont tout particulièrement suscité notre intérêt car leur approche est semblable à celle adoptée par Hervouet pour les modèles 3D de Telemac. Ils adoptent notamment la méthode de *suivi de surface* et proposent, comme Telemac-3D, la résolution des modèles 3D hydrostatique et quasihydrostatique. Ils utilisent enfin une méthode semi-implicite pour linéariser et stabiliser les équations, avant de les découpler afin de simplifier leur résolution. Cependant, deux éléments différencient ces deux modèles du code Telemac-3D : ils utilisent d'une part une approche eulérienne pour traiter le mouvement du domaine, et d'autre part la surface libre ainsi que le fond du domaine sont discrétisés en escalier. Dans le code Telemac-3D, au contraire, le formalisme ALE est adopté, plus précisément la *transformation sigma*, qui permet de suivre de manière très précise la topographie et le déplacement du domaine. Ceci est d'une importance extrême car l'approche eulérienne simplifie considérablement la mise en oeuvre d'un tel modèle, et la discrétisation en escalier permet la construction d'algorithmes très performants puisque les éléments du maillage sont réguliers. Cependant, cette approche requiert une résolution très fine du maillage afin de représenter de façon précise la surface libre et le fond. De plus, elle pose de sérieux problèmes quant à l'imposition des conditions limites naturelles. Notons que de nombreuses comparaisons ont été effectuées entre des mo-

dèles utilisant une discrétisation en escalier de la verticale du domaine et d'autres modèles utilisant la tranformation sigma [22, 54, 88, 93]. Celles-ci ont montré que les résultats obtenus avec la *transformation sigma* sont plus précis. Nous pensons pour cela que l'approche utilisée dans Telemac-3D est préférable, bien qu'elle alourdisse considérablement les calculs. Par contre, l'utilisation de l'élément *Raviart-Thomas* [117] pour discrétiser la vitesse du fluide, comme cela est fait dans [42, 97], serait intéressante car cet élément facilite l'imposition des conditions limites aux surfaces imperméables ainsi que la conservation de la masse au niveau local.

Nous nous sommes alors intéressés de plus près à la formulation variationnelle du problème hydrostatique 3D à surface libre proposé par Miglio, Saleri *et al.* dans [98]. Notons que les termes de diffusion horizontale sont négligés dans cette formulation. Après discrétisation en temps et explicitation des termes de convection, ils obtiennent un sous-problème couplant la vitesse horizontale du fluide et sa surface libre. Ce problème étant non symétrique, l'une des ses équations est modifiée. Les auteurs font ensuite une analyse mathématique du problème rendu symétrique et montrent l'existence et l'unicité d'une solution au problème (continu en espace) dans le cadre des problèmes mixtes.

1.3.1 Formulation variationnelle du modèle hydrostatique et analyse du problème discrétisé en temps.

Nous avons proposé une nouvelle formulation variationnelle du problème hydrostatique. Celle-ci inclut les termes de diffusion horizontale et permet notamment, après discrétisation en temps et explicitation des termes de convection, d'aboutir à un *problème symétrique et bien posé* couplant la vitesse horizontale et la surface libre, *sans modification des équations physiques*. Ce problème, discret en temps, sera nommé dans toute la thèse le *problème* $\boldsymbol{u} - \eta$, par référence à la vitesse horizontale notée $\boldsymbol{u}$ et à la surface libre notée η. Nous avons fait l'analyse mathématique du problème $\boldsymbol{u} - \eta$ et cela nous a permis de montrer qu'il admet une solution unique. Nous avons ensuite proposé deux approximations de ce problème : la première correspond à un traitement essentiel de la condition limite d'imperméabilité aux parois latérales du domaine, l'autre correspond au traitement naturel de cette condition limite. Ces deux approximations du problème admettent une solution unique lorsque le coefficient de diffusion horizontale est non négligeable. Dans le cas contraire, il se pourrait que la vérification d'une condition de type *inf-sup* soit nécessaire à la stabilité des solutions de ces problèmes discrets. Nous avons pour ce cas proposé un couple d'espaces d'éléments finis approchant la vitesse horizontale du fluide et la surface libre, et qui garantit la vérification de cette condition. Notons que l'intérêt du problème $\boldsymbol{u} - \eta$ ne se limite pas au modèle hydrostatique. Il peut également dériver de

la discrétisation en temps du modèle quasihydrostatique pour le traitement complet des équations de navier-Stokes 3D, puisque celui-ci comprend une étape hydrostatique.

1.3.2 Nouvel algorithme hydrostatique dans le code Telemac-3D.

L'algorithme qui résout les équations du modèle hydrostatique dans le code Telemac-3D, également utilisé, notons-le, pour le modèle nonhydrostatique, comprend la résolution d'un problème similaire au problème $\boldsymbol{u} - \eta$. Ce problème couple également la vitesse horizontale du fluide et la surface libre, mais il ne présente pas les mêmes propriétés mathématiques puisqu'il n'inclut pas les termes de diffusion (qui sont traités dans une étape préalable), et que les équations sont moyennées sur la verticale. Notons que ces manipulations détruisent la propriété de symétrie et de coercivité du problème $\boldsymbol{u} - \eta$. Nous avons donc modifié l'algorithme hydrostatique dans le code Telemac-3D afin de résoudre exactement le problème $\boldsymbol{u} - \eta$ que nous avons analysé, et nous avons mis en oeuvre sa résolution. Ce nouvel algorithme hydrostatique a été validé en comparant ses résultats avec la solution analytique des équations linéaires dans un cas particulier. Il a ensuite été appliqué à de nombreux cas test, et les résultats obtenus on été comparés à ceux fournis par la version standard de Telemac-3D, c'est-à-dire en utilisant l'algorithme hydrostatique standard. Dans la grande majorité des tests réalisés, les résultats fournis par l'utilisation des deux algorithmes sont quasi-identiques. Nous espérions pourtant que les résultats seraient, sinon plus précis (notre algorithme comprend un pas fractionnaire en moins, il devrait donc être plus précis en temps), du moins plus stables : en effet, le problème couplant la vitesse horizontale et la surface libre dans l'algorithme standard n'est pas coercif, alors que le problème $\boldsymbol{u} - \eta$ discret résolu dans le nouvel algorithme, lui, l'est. Nous avons obtenu des résultats plus stables dans deux cas particuliers, que nous présentons dans ce manuscrit, mais ceux-ci ne nous semblent pas assez probants.

Ce nouvel algorithme a cependant également été motivé par la réduction du nombre de pas fractionnaires dans la résolution des équations hydrostatiques, et parcequ'il permet un traitement entièrement tridimensionnel du problème à surface libre, ouvrant ainsi la voie à l'utilisation dans le code Telemac-3D de maillages non structurés sur la verticale.

1.3.3 ALE, transformation sigma et généralisation.

Comme nous l'avons indiqué au début de cette introduction, l'un des objectifs de cette thèse a été l'amélioration de la représentation de stratifications par le code Telemac-3D. Des erreurs relativement graves ont en effet été observées par certains utilisateurs lors de simulations faisant intervenir

des stratifications de densité.

Ce problème est intimement lié à l'utilisation de la *tranformation sigma classique* [111] dans le code Telemac-3D. Il intervient plus généralement lorsque la densité du fluide a une forte influence sur l'écoulement et que le domaine de calcul présente de forts gradients de bathymétrie. D'importantes erreurs peuvent alors apparaître dans le calcul du gradient horizontal de pression, dues à la déformation des éléments du maillage inhérente à la transformation sigma. Ce problème, que l'on appelle l'*erreur de gradient horizontal de pression* (HPG), a été étudié par de nombreux auteurs, notamment en océanographie et en modélisation atmosphérique [51, 83, 91, 92, 125, 132].

Etant donné que le lien entre la transformation sigma et la formulation ALE reste flou dans certaines communautés, nous nous sommes tout d'abord efforcés de montrer que la première n'est qu'un cas particulier de la deuxième. Nous avons pour cela introduit l'approche ALE-Sigma (ALES), qui est une interprétation ALE de la transformation sigma. Nous avons ensuite établi que les variantes classiques de la transformation sigma, puisqu'elles limitent considérablement le choix de la définition verticale du maillage et de son déplacement, empêchent de l'adapter aux besoins spécifiques de chaque simulation. Plus particulièrement, elles ne permettent pas de s'affranchir de l'erreur de HPG, ni de toutes les erreurs dues au déformement trop accru du maillage dans certains cas d'écoulements. C'est pourquoi nous avons proposé, grâce à l'approche ALES, une forme bien plus générale de transformation sigma. Celle-ci permet d'adapter davantage la discrétisation du domaine sur la verticale. Cette transformation sigma généralisée a été mise en oeuvre dans le code Telemac-3D, ce qui nous a permis d'illustrer numériquement certains des avantages de son utilisation, comme par exemple l'amélioration de la représentation de stratifications horizontales.

1.3.4 Schémas conservatifs pour les équations de convection à domaine mobile.

Un défaut de conservativité dans deux schémas de convection du code Tele- mac-3D a été mis en évidence par l'utilisation de la transformation sigma généralisée. Il s'agit de deux schémas MURD (*Multidimensional Residual Distributive scheme*) bien connus, étendus aux maillages 3D prismatiques et mis en oeuvre dans Telemac-3D par Jean-Marc Janin [74] pour la convection linéaire de traceurs dans un champ de vitesses à divergence nulle.

Les schémas MURD, bien connus pour leur propriétés de conservation, monotonie et stabilité, ont néanmoins été développés pour des problèmes posés sur un domaine fixe. La préservation de ces propriétés sur un domaine mobile n'est donc pas garantie. Plusieurs auteurs ont d'ailleurs montré que le mouvement du domaine peut avoir des conséquences négatives sur certains schémas numériques (voir par exemple [43, 57, 58]). Dans certains cas, la vérification d'une condition particulière assure la préservation de certaines

propriétés comme la conservation, la stabilité ou même la précision des schémas. Cette condition est mieux connue sous le nom de *loi de Conservation Géométrique* (GCL).

De fait, Janin a montré dans [74] qu'une contrainte particulière devait être vérifiée afin d'assurer la bonne conservation des quantités convectées dans les schémas MURD mis en oeuvre dans Telemac-3D. Cependant, son raisonnement est limité à l'utilisation de la transformation sigma classique, et les schémas cessent d'être conservatifs dès qu'une autre méthode ALE est utilisée.

Nous avons généralisé les travaux de Janin grâce à la dérivation dans le cadre ALE d'une forme générique de schémas MURD pour la résolution des équations de convection linéaires posées sur un domaine mobile. Nous avons ensuite formulé la condition supplémentaire que ces schémas doivent satisfaire pour rester conservatifs quand le domaine bouge effectivement. Celle-ci est d'ailleurs fortement liée à la notion de GCL. Nous avons enfin constaté que lorsque le problème est posé sur un domaine tridimensionnel qui ne se déplace que dans la direction verticale, cette condition peut être assurée très simplement. Il suffit pour cela de considérer une vitesse de maillage constante par pas de temps. Des résultats numériques ont été obtenus qui illustrent ce résultat.

Celui-ci peut notamment être appliqué aux équations de convection linéaires dans le cadre des écoulements tridimensionnels à surface libre en dimension 3. Il nous a notamment permis de corriger le défaut de conservativité dans Telemac-3D.

1.3.5 Estimation de paramètres dans un modèle 1D pour les écoulements sanguins dans les artères.

Nous présentons finalement en annexe de ce manuscrit un travail effectué à l'école d'été du CEMRACS 2004, dont le thème était *Mathématiques et application en biologie et en médecine*. Il a été réalisé en collaboration avec Vincent Martin, François Clément et Jean-Frédéric Gerbeau. Il a été publié dans *ESAIM Proceedings* [87] en septembre 2005.

Le but de ce travail est d'identifier certains des paramètres d'un modèle 1D pour les écoulements sanguins dans les artères [116, 45]. Notons qu'il s'agit d'un modèle hyperbolique très proche du système de Saint-Venant décrit au début de cette introduction. Les paramètres du modèle permettent d'approcher plus ou moins bien des configurations géométriques réalistes ou des données expérimentales. Il est donc important d'en avoir une bonne estimation.

L'approche adoptée pour l'estimation des paramètres est celle des moindres carrés non-linéaires (cf. [13, 14, 15]), basée sur l'optimisation d'une certaine fonction coût. Notons que la résolution d'un tel problème de mini- misation requiert le calcul efficace et précis du gradient de la fonction coût par rapport

aux paramètres. Nous avons choisi d'approcher ce gradient par la méthode de l'état adjoint.

Le problème direct utilisé est une discrétisation par le schema de Taylor-Galerkin du modèle hyperbolique 1D. Notons que, suivant les conseils donnés dans [13, 14, 15], nous avons choisi d'inverser le problème discret et non pas de discrétiser le problème inverse. Nous avons donc dérivé (analytiquement) le problème adjoint, permettant de calculer le gradient de la fonction coût, à partir du problème 1D discret. Remarquons que le problème adjoint est un problème hyperbolique 1D très similaire au problème direct.

De premiers résultats probants ont été obtenus pour l'estimation d'un paramètre lié aux propriétés mécaniques de la paroi artérielle. Les données utilisées pour cette estimation sont synthétiques. Elles ont été obtenues à partir d'un modèle beaucoup plus raffiné : un modèle 3D d'interaction fluide-structure [48, 53]. Les résultats obtenus semblent intéressants notamment du fait que le paramètre estimé est assez différent de sa valeur *a priori*.

1.4 Plan de la thèse

PARTIE I. Cette partie est dédiée à la formulation ALE et aux propriétés de conservation des schémas appliqués aux problèmes à domaine mobile.

Au chapitre 1 nous décrivons de façon détaillée la formulation ALE et nous donnons quelques outils pour son application dans le cadre des Eléments Finis. Nous introduisons également la notion de Loi de Conservation Géométrique (GCL).

Au chapitre 2 est dérivée une forme générique de schémas ALE-MURD pour les équations de convection linéaires posées sur un domaine mobile. La contrainte de conservation de ces schémas est formulée, et il est montré que cette contrainte est facilement vérifiable dans le cas de problèmes posés sur des domaines tridimensionnels se déplaçant uniquement selon la direction verticale. Une illustration numérique de ce résultat est donnée.

Le chapitre 3 est dédié à la description détaillée de la transformation sigma et à sa réinterprétation comme forme particulière de méthode ALE. Les avantages et inconvénients de cette transformation et de ses différentes variantes sont exposés. Une forme très générale de ce type de transformation est enfin proposée et mise en oeuvre dans le code Telemac-3D. Des tests numériques sont finalement présentés, qui montrent l'amélioration des résultats dans deux cas de simulation grâce à la transformée sigma généralisée.

PARTIE II. Cette partie est consacrée au modèle hydrostatique pour les écoulements tridimensionnels à surface libre.

Dans le chapitre 4 nous présentons le problème des écoulements tridimensionnels à surface libre, ainsi que les modèles 3D hydrostatiques et nonhydrostatiques qui seront considérés par la suite. Une formulation variationnelle du problème hydrostatique est proposée. Celle-ci est ensuite discrétisée en temps et linéarisée ; les termes de convections sont traités de façon explicite. Ceci permet d'obtenir trois sous-problèmes bien distincts, dont le problème $\boldsymbol{u} - \eta$.

Au chapitre 5 nous analysons le problème semi-discret en temps $\boldsymbol{u} - \eta$, ainsi que deux approximations de ce problème. Nous proposons également un couple d'éléments finis stables pour ce problème.

Le chapitre 6 est consacré à l'application au code Telemac-3D des travaux présentés aux chapitres 2, 3 et 5. Nous détaillons dans une première sous-partie le nouvel algorithme hydrostatique mis en oeuvre dans le code et nous présentons quelques résultats. Puis nous décrivons comment les schémas de convection ALE-MURD présentés au chapitre 2 peuvent être appliqués dans le cadre du modèle hydrostatique 3D présenté au chapitre 4, afin que la contrainte de conservation imposée sur les domaines mobiles soit vérifiée. Nous présentons finalement des résultats obtenus après adaptation, grâce à l'approche ALE-MURD, des schémas de convection MURD préexistants dans le code afin qu'ils soient conservatifs pour toute transformation ALE.

Dans l'annexe A nous présentons le travail autour de l'estimation de paramètres dans un modèle 1D pour les écoulements sanguins dans les artères.

1.5 Conclusion et perspectives

Ce travail nous a menés à travers plusieurs aspects des écoulements tridimensionnels à surface libre.

D'une part, nous avons proposé une nouvelle formulation variationnelle des équations régissant ces écoulements. Celle-ci permet d'aboutir à un sous-problème bien posé couplant la vitesse horizontale et la surface libre. Mais les résultats numériques obtenus après mise en oeuvre de la résolution de ce sous-problème dans le logiciel Telemac-3D ne sont pas encore assez probants. Nous comptons donc poursuivre notre travail sur ce système. Il serait notamment intéressant de tester la discrétisation de la vitesse horizontale et de la surface libre par le couple d'éléments finis stables proposés. Nous envisageons également d'évaluer les conséquences de l'explicitation du terme de diffusion verticale sur la stabilité des solutions du problème, car ce terme n'est pas nécessaire à la coercivité du problème.

Nous avons d'autre part proposé une forme très générale de transformation ALE-sigma pour les domaines 3D se déplaçant uniquement selon

la verticale. Celle-ci permet, davantage que la *transformation sigma classique*, d'adapter le maillage aux besoins de chaque simulation, et notamment d'améliorer la représentation de stratifications. Cette transformation est dorénavant utilisée dans le code Telemac-3D. Il serait maintenant intéressant d'étudier de façon plus approfondie les possibilités de discrétisation verticale du domaine offertes par cette nouvelle forme de tranformation ALE, ainsi que leurs avantages et inconvénients dans différents cas d'applications.

La travail effectué autour des schémas ALE-MURD pour la convection linéaire de scalaires, et de leurs propriétés de conservation lorsqu'ils sont posés sur des domaines mobiles est, nous semble-t-il, intéressant en soi. Il nous a également permis de mettre à jour les schémas de convection de Telemac-3D, afin qu'ils soient conservatifs pour tout cas de transformation ALE du maillage, tant que celle-ci est uniquement verticale.

Mais notre travail ouvre également la voie à l'extension des modèles basés sur des maillages se déplaçant uniquement selon la verticale, à l'utilisation de maillages tridimensionnels entièrement non structurés.

D'une part, la résolution que nous avons proposée du sous-problème couplant la vitesse horizontale et la surface libre est compatible avec un maillage non structuré sur la verticale, puisqu'elle n'implique pas l'intégration sur la verticale des variables ni des termes sources. D'autre part, les schémas ALE-MURD proposés dans cette thèse sont applicables aux problèmes de convection linéaire de scalaires sur tout domaine mobile, pour toute forme de déplacement du maillage. L'expression de la contrainte de conservativité de ces schémas étant donnée dans le cas général, nous pourrions envisager de la vérifier également dans le cas de maillages 3D ne se déplaçant pas uniquement selon la verticale, autrement dit dans le cas de maillages entièrement non structurés.

Cette perspective nous semble extrêment intéressante car elle permettrait la prise en compte de tout obstacle tridimensionnel dans l'écoulement à surface libre étudié.

Finalement, nous aimerions également poursuivre le travail entamé autour de l'estimation de paramètres dans un modèle 1D pour les écoulements sanguins dans les artères. Nous pensons cependant abandonner la métho- de du problème adjoint pour le calcul du gradient de la fonction coût à minimiser, car cette méthode nous a semblé trop lourde. Nous envisageons d'utiliser plutôt une méthode globale, plus facile à mettre en oeuvre, combinée à une méthode locale pour accélérer la convergence de la minimisation.

Première partie

Arbitrary Lagrangian Eulerian Framework and Conservation Properties.

Chapitre 2

Arbitrary Lagrangian Eulerian formulation and Geometric Conservation Laws

In this chapter we aim to describe in details the Arbitrary Lagrangian Eulerian (ALE) formulation and its application in the Finite Element Method (FEM) framework. In addition, we wish to introduce the notion of Geometric Conservation Law (GCL), related to the preservation of the stability and conservation properties of a scheme when posed on a moving domain.

2.1 ALE formulation and finite elements

2.1.1 Principle

The ALE formulation is adopted in order to simplify the difficult treatment of problems with moving boundaries. First introduced by Hirt *et al.* [67] in 1974, it consists in defining a reference configuration – which is generally chosen to be fixed – and a mapping that gives the correspondence with the real domain at each time. The mapping can be chosen arbitrarily, but it has to be conforming to the evolution of the domain boundaries. It defines an instantaneous domain velocity, whose discrete counterpart will be the velocity of the mesh. Consequently, the approximation of a problem raised on a moving domain only requires the construction of a mesh on the reference configuration : the discretized ALE mapping provides, at any time, the coordinates of each grid node in the real domain. That allows to switch very simply between the reference frame and the real frame. Moreover, the partial time derivatives of a problem can be turned into the so-called *ALE time derivatives*, that are derivatives with respect to the time at a point constant in the reference configuration. This is very useful because partial time derivatives cannot be discretized in a moving domain using the standard finite

differences approach.

To illustrate this matter, let Ω_t be a moving domain defined on a time interval I. For a function $\psi(x,t)$ defined at any time $t \in I$ on the domain Ω_t and with values in $\mathbb{R}$, the partial time derivative at $\boldsymbol{x} \in \Omega_t$ is denoted

$$\frac{\partial \psi}{\partial t} = \frac{\partial \psi}{\partial t}(\boldsymbol{x}, t).$$

Let us divide the interval I into N time steps of equal duration Δt. We would then like to make a first order discretization of the partial time derivative at $t = t^n = n \times \Delta t$, using the finite difference approach. This would consist in making the following approximation :

$$\frac{\partial \psi}{\partial t}(\boldsymbol{x}, t^n) \approx \frac{\psi(\boldsymbol{x}, t^{n+1}) - \psi(\boldsymbol{x}, t^n)}{\Delta t} \qquad \forall\, \boldsymbol{x} \in \Omega_{t_n}. \tag{2.1}$$

The problem is that a point belonging to the domain at a given time might not belong to it anymore one time step later.

Therefore, if $\boldsymbol{x} \in \Omega_{t_n}$, $\psi(\boldsymbol{x}, t^n)$ is defined but $\psi(\boldsymbol{x}, t^{n+1})$ may not. We will see that this problem doesn't exist in the case of an ALE time derivative, because it takes into account the movement of the domain.

2.1.2 Notations

We now introduce some notations that will be used in the entire work. Let us consider a moving domain configuration $\Omega_t \subset \mathbb{R}^d$, defined on a time interval I. In order to write the ALE formulation of a problem raised on this domain, a reference configuration $\hat{\Omega} \subset \mathbb{R}^d$ is defined, as well as a mapping $\hat{\mathcal{A}}_t$ which at each time $t \in I$ associates to a point $\hat{\boldsymbol{x}}$ in $\hat{\Omega}$ a point $\boldsymbol{x}$ in Ω_t. It means, for each $t \in I$, that

$$\hat{\mathcal{A}}_t : \hat{\Omega} \longrightarrow \Omega_t, \qquad \boldsymbol{x}(\hat{\boldsymbol{x}}, t) = \hat{\mathcal{A}}_t(\hat{\boldsymbol{x}}). \tag{2.2}$$

We shall call $\hat{\boldsymbol{x}} \in \hat{\Omega}$ the *ALE coordinate* and $\boldsymbol{x} \in \Omega_t$ the *spatial* (or eulerian) *coordinate*. Throughout, $\hat{\mathcal{A}}_t$ will be assumed to be an homeomorphism, that is $\hat{\mathcal{A}}_t \in \mathcal{C}^0(\bar{\hat{\Omega}})$ is invertible with continuous inverse $\mathcal{A}_t^{-1} \in \mathcal{C}^0(\bar{\Omega}_t)$. In addition, the application

$$t \longrightarrow \boldsymbol{x}(\hat{\boldsymbol{x}}, t), \quad \hat{\boldsymbol{x}} \in \hat{\Omega}, \tag{2.3}$$

will be assumed to be differentiable almost everywhere in I.
Throughout we will denote the set

$$\{(\boldsymbol{x}, t) \mid \boldsymbol{x} \in \Omega_t,\ t \in I\}$$

by $\Omega_t \times I$, which is a slight abuse of notation since this set is not a cylinder. Let then $\psi : \Omega_t \times I \longrightarrow \mathbb{R}$ be a function defined on the Eulerian frame and $\hat{\psi} : \hat{\Omega} \times I \longrightarrow \mathbb{R}$ the corresponding function on the ALE frame, defined as

$$\hat{\psi}(\hat{\boldsymbol{x}}, t) = \psi(\boldsymbol{x}, t) \quad \text{where} \quad \boldsymbol{x} = \hat{\mathcal{A}}_t(\hat{\boldsymbol{x}}). \tag{2.4}$$

By a classical abuse of notation, the time derivative on the ALE frame, written in the spatial coordinate, will be denoted by the symbol $\frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}}$. It is defined as follows :

$$\frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}} : \Omega_t \times I \longrightarrow \mathbb{R}, \qquad \frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}}(\boldsymbol{x},t) = \frac{\partial \hat{\psi}}{\partial t}(\hat{\boldsymbol{x}},t) \quad \text{with } \hat{\boldsymbol{x}} = \mathcal{A}_t^{-1}(\boldsymbol{x}). \tag{2.5}$$

Throughout, when not explicitly mentioned, all the differential operators ($\frac{\partial}{\partial t}$, $\boldsymbol{\nabla}$, div , ...) will be taken with respect to the Eulerian frame.
The instantaneous domain velocity will be defined by

$$\boldsymbol{c}(\boldsymbol{x},t) = \hat{\boldsymbol{c}}(\hat{\boldsymbol{x}},t) = \frac{\partial \hat{\mathcal{A}}_t}{\partial t} \quad \text{with } \hat{\boldsymbol{x}} = \mathcal{A}_t^{-1}(\boldsymbol{x}), \tag{2.6}$$

which can also be written, with a convenient abuse of notation,

$$\boldsymbol{c}(\boldsymbol{x},t) = \frac{\partial \boldsymbol{x}}{\partial t}\Big|_{\hat{\boldsymbol{x}}}. \tag{2.7}$$

The Jacobian matrix of $\hat{\mathcal{A}}_t$ will be denoted by $\hat{\mathbf{J}}_t$

$$\hat{\mathbf{J}}_t \;=\; \left[\frac{\partial \hat{\mathcal{A}}_t}{\partial \hat{\boldsymbol{x}}_j}\right]$$

and by $\hat{J}_t$ its determinant.

2.1.3 The ALE time derivative

Let us now explain why, for the ALE time derivative, the discretization problem mentioned in section 2.1.1 does not arise. Let us divide the interval I into N time steps Δt and define, for every integer $n \leq N$, $t^n = n \times \Delta t$. A first order approximation of $\frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}}$ at $t = t^n$ and $\boldsymbol{x} \in \Omega_{t_n}$ is then

$$\frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}}(\boldsymbol{x},t^n) \;=\; \frac{\partial \hat{\psi}}{\partial t}(\hat{\boldsymbol{x}},t^n) \;\approx\; \frac{\hat{\psi}(\hat{\boldsymbol{x}},t^{n+1}) - \hat{\psi}(\hat{\boldsymbol{x}},t^n)}{\Delta t}, \tag{2.8}$$

$$\text{with} \quad \hat{\boldsymbol{x}} = \mathcal{A}_n^{-1}(\boldsymbol{x}).$$

In the Eulerian frame, this approximation is equivalent to :

$$\frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}}(\boldsymbol{x},t^n) \;\approx\; \frac{\psi(\,\hat{\mathcal{A}}_{t^{n+1}} \circ \mathcal{A}_n^{-1}(\boldsymbol{x},t^{n+1})\,,\,t^{n+1}\,) - \psi(\,\boldsymbol{x}\,,\,t^n\,)}{\Delta t}. \tag{2.9}$$

By construction, if $\boldsymbol{x} \in \Omega_{t_n}$ then $\hat{\mathcal{A}}_{t^{n+1}} \circ \mathcal{A}_n^{-1}(\boldsymbol{x},t^{n+1})$ belongs to $\Omega_{t_{n+1}}$. All the terms are therefore well-defined.

2.1.4 Some standard formulae

Let $\psi : \Omega_t \times I \longrightarrow \mathbb{R}$ be regular enough, the chain rule applied to the time derivative gives :

$$\left.\frac{\partial \psi}{\partial t}\right|_{\hat{\boldsymbol{x}}}(\boldsymbol{x},t) \;=\; \frac{\partial \hat{\psi}}{\partial t}(\hat{\boldsymbol{x}},t) \;=\; \frac{\partial \psi}{\partial t}(\boldsymbol{x},t) \;+\; \boldsymbol{c}\cdot\boldsymbol{\nabla}\psi(\boldsymbol{x},t). \tag{2.10}$$

The *Euler expansion formula* (see [5]) relates the time evolution of $\hat{J}_t$ to the divergence of the domain velocity field, according to the following differential equation written in the ALE frame :

$$\frac{\partial \hat{J}_t}{\partial t}\left(\mathcal{A}_t^{-1}(\boldsymbol{x}),\, t\right) \;=\; \hat{J}_t\left(\mathcal{A}_t^{-1}(\boldsymbol{x}), t\right)\; \operatorname{div}\boldsymbol{c} \tag{2.11}$$

and valid for any $t \in I$ and $\hat{\boldsymbol{x}} \in \hat{\Omega}$. This formula can be interpreted as an evolution law for the Jacobian determinant. We will see later on that it is strongly related to the so-called Geometric Conservation Laws that guarantee a specific property to schemes set on moving domains.

Using this formula the following relation is obtained :

$$\frac{d}{dt}\int_{\Omega_t} \psi(\boldsymbol{x},t)\; d\boldsymbol{x} = \int_{\Omega_t}\left(\left.\frac{\partial \psi}{\partial t}\right|_{\hat{\boldsymbol{x}}} + \psi(\boldsymbol{x},t)\operatorname{div}\boldsymbol{c}\right)\; d\boldsymbol{x}, \tag{2.12}$$

which is a variant of the *Reynolds transport formula*.

2.1.5 Functional spaces in the ALE frame

We introduce here a general type of functional spaces, compatible with the ALE mapping. They allow to define functions that depend on time and are defined on a moving domain.

Let us consider a space of functions $\hat{\mathcal{X}}$, defined on the reference domain, and made of functions $\hat{\psi} : \hat{\Omega} \longrightarrow \mathbb{R}$ that are smooth enough. The ALE mapping then identifies the corresponding function space on the current domain :

$$\mathcal{X}(\Omega_t) \;=\; \left\{\; \psi : \Omega_t \times I \to \mathbb{R},\; \psi(\boldsymbol{x},t) \;=\; \hat{\psi}(\mathcal{A}_t^{-1}(\boldsymbol{x})),\; \hat{\psi} \in \hat{\mathcal{X}} \;\right\}. \tag{2.13}$$

Note that functions of $\mathcal{X}(\Omega_t)$ depend on time *only* through the ALE mapping. Indeed,

$$\forall\, \psi \in \mathcal{X}(\Omega_t) \qquad \left.\frac{\partial \psi}{\partial t}\right|_{\hat{\boldsymbol{x}}} \;=\; 0. \tag{2.14}$$

This space will be used to define the test functions of a problem raised on a

moving domain. The time dependent solutions will rather be defined in the following functional space :

$$\mathcal{X}(\Omega_t,t) \;=\; \Big\{\; \psi : \Omega_t \times I \to \mathbb{R},\; \psi(\boldsymbol{x},t) \;=\; \hat{\psi}(\mathcal{A}_t^{-1}(\boldsymbol{x}),t),\; \hat{\psi}(\cdot,t) \in \hat{\mathcal{X}} \;\Big\}. \tag{2.15}$$

In this work, the main functional spaces we will make use of are $L^2(\Omega)$ and

$$H^1(\Omega) \;=\; \Big\{\; \psi : \Omega \subset \mathbb{R}^d \to \mathbb{R} \mid \quad \psi \in L^2(\Omega) \quad \text{and} \quad \boldsymbol{\nabla}\psi \in \big(L^2(\Omega)\big)^d \;\Big\}$$

and The associated norms are, respectively,

$$\begin{aligned} \|\psi\|_{\,0,\Omega} \;&=\; \Big(\int_\Omega \psi \; d\boldsymbol{x} \Big)^{1/2}, \\ \|\psi\|_{\,1,\Omega} \;&=\; \big(\;\|\psi\|^2_{0,\omega} \;+\; \|\boldsymbol{\nabla}\psi\|^2_{0,\Omega} \;\big)^{1/2} . \end{aligned}$$

We also introduce the semi-norm on $H^1(\Omega)$:

$$|\psi|_{\,1,\Omega} \;=\; \|\boldsymbol{\nabla}\psi\|_{\,0,\Omega}\,.$$

Now, consider a problem whose solution defined on the current space must belong to $H^1(\Omega_t)$ at each time $t \in I$. A functional space $\hat{\mathcal{V}}$ defined on the reference configuration $\hat{\Omega}$ must be determined, such that the time-dependent functions in the corresponding space $\mathcal{V}(\Omega_t,t)$ belong to $H^1(\Omega_t)$ at each time t.

The following proposition has been proved in [43] :

Proposition 2.1 *Assume that the mapping $\hat{\mathcal{A}}_t$, which is invertible on $\hat{\Omega}$, satisfies the following conditions :*

$$(i) \qquad \Omega_t \;=\; \hat{\mathcal{A}}_t(\hat{\Omega}) \quad \textit{is bounded and Lipschitz continuous.} \tag{2.16}$$

$$(ii) \qquad \hat{\mathcal{A}}_t \in \Big(W^{1,\infty}(\hat{\Omega})\Big)^d \qquad \textit{and} \qquad \mathcal{A}_t^{-1} \in \big(W^{1,\infty}(\Omega_t)\big)^d. \tag{2.17}$$

Then, for any function $\psi : \Omega_t \longrightarrow \mathbb{R}$,

$$\psi \in H^1(\Omega_t) \qquad \textit{if and only if} \qquad \hat{\psi} \;=\; \psi \circ \hat{\mathcal{A}}_t \;\in H^1(\hat{\Omega}). \tag{2.18}$$

Throughout, we will assume the ALE mappings considered to satisfy conditions (2.16) and (2.17). Consequently, the following properties are satisfied

by any function space $\hat{\mathcal{V}}$ and its corresponding spaces $\mathcal{V}(\Omega_t)$ and $\mathcal{V}(\Omega_t, t)$ on the current domain :

$$\begin{aligned} \text{If} \quad \hat{\mathcal{V}} \subset L^2(\hat{\Omega}) \quad \text{then} \quad \forall\, \psi \in \mathcal{V}(\Omega_t) \quad \text{and} \quad \forall\, u \in \mathcal{V}(\Omega_t, t), \\ \psi(t) \in L^2(\Omega_t) \quad \text{and} \quad u(t) \in L^2(\Omega_t) \quad \forall t \in I. \end{aligned} \tag{2.19}$$

$$\begin{aligned} \text{If} \quad \hat{\mathcal{V}} \subset H^1(\hat{\Omega}) \quad \text{then} \quad \forall\, \psi \in \mathcal{V}(\Omega_t) \quad \text{and} \quad \forall\, u \in \mathcal{V}(\Omega_t, t), \\ \psi(t) \in H^1(\Omega_t) \quad \text{and} \quad u(t) \in H^1(\Omega_t) \quad \forall t \in I. \end{aligned} \tag{2.20}$$

$$\begin{aligned} \text{If} \quad \hat{\mathcal{V}} \subset H^1_v(\hat{\Omega}) \quad \text{then} \quad \forall\, \psi \in \mathcal{V}(\Omega_t) \quad \text{and} \quad \forall\, u \in \mathcal{V}(\Omega_t, t), \\ \psi(t) \in H^1_v(\Omega_t) \quad \text{and} \quad u(t) \in H^1_v(\Omega_t) \quad \forall t \in I. \end{aligned} \tag{2.21}$$

2.1.6 Construction of the ALE mapping and finite element discretization

As already mentioned, the ALE mapping can be chosen arbitrarily inside the domain, but at the boundaries it has to be conforming to the motion of the domain. Thus, the mapping must be such that the instantaneous domain velocity $\boldsymbol{c}$ satisfies, at each time $t \in I$,

$$\boldsymbol{c} \cdot \boldsymbol{n} \;=\; \boldsymbol{\beta} \cdot \boldsymbol{n} \tag{2.22}$$

on impermeable boundaries – *i.e.*, on impermeable solid boundaries and free surfaces –, where $\boldsymbol{\beta}$ denotes the velocity of the fluid and $\boldsymbol{n}$ the outward normal to the current domain Ω_t. In addition, on fictitious liquid boundaries with no normal movement – *i.e.*, open boundaries which do not move in the normal direction –, it must be such that

$$\boldsymbol{c} \cdot \boldsymbol{n} \;=\; 0\,. \tag{2.23}$$

Throughout, we will assume that all open boundaries have no normal movement. We will denote by $\Gamma_{imp,t}$ the impermeable boundaries of a moving domain Ω_t, and by $\Gamma_{liq,t}$ its fictitious liquid boundaries.

Moreover, some regularity properties might be required for the mapping, as for instance the satisfaction of conditions (2.16) and (2.17).

In the literature, several techniques have been proposed to construct the ALE mapping. For instance, one can solve a parabolic problem with zero source term and relations (2.22) and (2.23) as boundary condition. We won't elaborate on the construction of the continuous mapping and we will rather switch directly to its discretization.

In order to be able to write a finite element approximation of a problem in the ALE frame, the mapping has to be discretized. We point out that, throughout, we will only consider Lagrangian finite elements.

A triangulation $\hat{\mathcal{T}}_h$ of the reference configuration $\hat{\Omega}$ is first introduced. The union of its elements K_0 forms the discrete domain, which will be assumed to be identical to $\hat{\Omega}$. We will indicate by $N^{\mathcal{T}}$ the set of nodes composing the finite element mesh and by $\hat{\boldsymbol{x}}_i$, for $i = 1, \ldots, N^{\mathcal{T}}$, the coordinate of the i-th node in $\hat{\mathcal{T}}_h$.

Let us now introduce the following general finite element space defined on the reference configuration :

$$\begin{aligned} \mathcal{F}_{n,\,k}(\hat{\mathcal{T}}_h) \;=\; & \left\{ \hat{\psi}_h : \hat{\Omega} \to \mathbb{R} \mid \quad \hat{\psi}_h \in \mathcal{C}^0(\bar{\hat{\Omega}}), \quad \hat{\psi}_h\Big|_{K_0} \circ \mathcal{M}_k^{K_0} \in P_n(K_R), \right. \\ & \left. \forall K_0 \in \hat{\mathcal{T}}_h \right\} \end{aligned} \tag{2.24}$$

where $P_n(K_R)$ is the space of polynomials of degree n defined on the reference element K_R, and $\mathcal{M}_k^{K_0} \in P_k(K_R)$ is a polynomial homeomorphic mapping of degree k from K_R to K_0. In general $k \leq n$, and in particular k is either equal to 1 – in the case of an affine mapping – or to n – for an isoparametric mapping. It can be shown that

$$\mathcal{F}_{n,\,k}(\hat{\mathcal{T}}_h) \subset H^1(\hat{\Omega})\,. \tag{2.25}$$

The ALE mapping is discretized by means of functions in the isoparametric space $\mathcal{F}_{k,\,k}(\hat{\mathcal{T}}_h)$, for some specific value of k. This space is chosen for regularity considerations that will be discussed later. The discrete ALE mapping will be denoted by $\hat{\mathcal{A}}_{h,t}$.
For each $t \in I$, let then $\mathcal{T}_{h,t}$ be the image of $\hat{\mathcal{T}}_h$ by $\hat{\mathcal{A}}_{h,t}$, that is

$$\mathcal{T}_{h,t} = \left\{ K_t = \hat{\mathcal{A}}_{h,t}(K_0), \quad K_0 \in \hat{\mathcal{T}}_h \right\}.$$

Letting $\Omega_{h,t}$ be the union of the elements of $\mathcal{T}_{h,t}$, we have that

$$\hat{\mathcal{A}}_{h,t} : \hat{\Omega} \longrightarrow \Omega_{h,t},$$

that $\Omega_{h,t}$ represents the discretization of the current domain Ω_t, and that $\mathcal{T}_{h,t}$ represents a triangulation of Ω_t. At each time $t \in I$, the position $\boldsymbol{x}_i$ of the i-th node in $\mathcal{T}_{h,t}$ is then defined by

$$\boldsymbol{x}_i(t) = \hat{\mathcal{A}}_{h,t}(\hat{\boldsymbol{x}}_i), \quad \text{for } i = 1, \ldots, N^{\mathcal{T}}.$$

Moreover, the discrete domain velocity $\boldsymbol{c}_h$ is defined by

$$\boldsymbol{c}_h\left(\boldsymbol{x}, t\right) = \hat{\boldsymbol{c}}_h\left(\hat{\boldsymbol{x}}, t\right) = \frac{\partial \hat{\mathcal{A}}_{h,t}}{\partial t}(\boldsymbol{x}, t),$$

where $\hat{\boldsymbol{x}} = \mathcal{A}_{h,t}^{-1}\left(\boldsymbol{x}\right)$.

Remark 2.1.1 According to this definition of the domain triangulation, the topology of the mesh remains the same during the time evolution and only the coordinates of the mesh nodes change. Therefore, no interpolation is required between each timestep. Note that this is an important advantage of the ALE formulation for problems raised on a moving domain.

2.1.7 Finite element spaces in the ALE frame

In this section we present the finite element spaces that will be used to discretize the ALE formulation of the problems studied in this work. Let us recall that we will only use Lagrangian type finite elements and a discrete ALE mapping as described in section 2.1.6.

If a space of functions $\hat{\mathcal{X}}$ defined on the reference configuration $\hat{\Omega}$ is considered, it will be approximated by $\hat{\mathcal{X}}_h = \mathcal{F}_{n,k}(\hat{\mathcal{T}}_h)$ – defined by (2.24), for some arbitrary value of n and for a value $k \leq n$ used for all discrete spaces.

Remark 2.1.2 We point out that the value of the degree k will be the same for all the finite element spaces used to approximate a same problem. Particularly, the ALE mapping will be discretized by means of functions of the finite element function space $\mathcal{F}_{k,k}(\hat{\mathcal{T}}_h)$ – see section 2.1.6 – for exactly this value of k.

The corresponding discrete space $\mathcal{X}_h(\Omega_{h,t})$ on the current approximated domain is then :

$$\mathcal{X}_h(\Omega_{h,t}) \;=\; \left\{ \psi_h : \Omega_{h,t} \times I \to \mathbb{R}, \quad \psi_h(\boldsymbol{x},t) \;=\; \hat{\psi}_h(\mathcal{A}_{h,t}^{-1}(\boldsymbol{x})), \quad \hat{\psi}_h \in \hat{\mathcal{X}}_h \right\}. \tag{2.26}$$

We will also consider the discrete function space

$$\mathcal{X}_h(\Omega_{h,t},t) \;=\; \left\{ \psi_h : \Omega_{h,t} \times I \to \mathbb{R}, \quad \psi_h(\boldsymbol{x},t) \;=\; \hat{\psi}_h(\mathcal{A}_{h,t}^{-1}(\boldsymbol{x}),t), \quad \hat{\psi}_h \in \hat{\mathcal{X}}_h \right\}. \tag{2.27}$$

Let N_h be the dimension of $\hat{\mathcal{X}}_h$ and let $\{\hat{\psi}_i\}_{i=1,\ldots,N_h}$ be a set of basis functions for this space. Any function $\hat{v}_h$ in $\hat{\mathcal{X}}_h$ can then be decomposed as follows :

$$\hat{v}_h(\hat{\boldsymbol{x}}) \;=\; \sum_{i=1}^{N_h} v_i \; \hat{\psi}_i(\hat{\boldsymbol{x}}), \qquad \hat{\boldsymbol{x}} \in \hat{\Omega}, \tag{2.28}$$

where v_i is the coefficient of $\hat{v}_h$ associated to the i-th degree of freedom.

Clearly, the dimension of $\mathcal{X}_h(\Omega_{h,t})$ is also N_h and the set of functions $\{\ \psi_i(t)\ \}_{i=1,\dots,N_h}$ defined by

$$\psi_i(t) \ = \ \hat{\psi}_i \circ \mathcal{A}_{h,t}^{-1} \tag{2.29}$$

constitutes a set of basis functions of $\mathcal{X}_h(\Omega_{h,t})$. Any function v_h in $\mathcal{X}_h(\Omega_{h,t})$ can then be written as

$$v_h(\boldsymbol{x},t) \ = \ \sum_{i=1}^{N_h} v_i \ \psi_i(\boldsymbol{x},t), \quad \text{for } \boldsymbol{x} \in \Omega_{h,t},\ t \in I. \tag{2.30}$$

Moreover, for any function v_h in $\mathcal{X}_h(\Omega_{h,t},t)$, we have that

$$v_h(\boldsymbol{x},t) \ = \ \sum_{i=1}^{N_h} v_i(t) \ \psi_i(\boldsymbol{x},t), \quad \text{for } \boldsymbol{x} \in \Omega_{h,t},\ t \in I, \tag{2.31}$$

with time dependent coefficients $v_i(t)$. Thanks to relation (2.14) we may then write :

$$\frac{\partial v_i}{\partial t}\Big|_{\hat{\boldsymbol{x}}}(\boldsymbol{x},t) \ = \ \sum_{i=1}^{N_h} \frac{dv_i}{dt}(t) \ \psi_i(\boldsymbol{x},t), \qquad \boldsymbol{x} \in \Omega_{h,t}, \quad t \in I. \tag{2.32}$$

Remark 2.1.3 We point out that these coefficients are the same regardless of the domain on which the function is defined. Indeed, the corresponding function of $v_h(x,t)$ on the ALE frame – $\hat{v}_h(t) = v_h(t) \circ \hat{\mathcal{A}}_{h,t}$ – can be decomposed on the basis functions of its definition space $\hat{\mathcal{X}}_h$ as follows :

$$\hat{v}_h(\hat{\boldsymbol{x}},t) \ = \ \sum_{i=1}^{N_h} v_i(t) \ \hat{\psi}_i(\hat{\boldsymbol{x}}), \qquad \hat{\boldsymbol{x}} \in \hat{\Omega}. \tag{2.33}$$

Let us now introduce the following general finite element space defined on the triangulation $\mathcal{T}_{h,t}$ of the current domain :

$$\begin{aligned} \mathcal{F}_{n,k}(\mathcal{T}_{h,t}) \ = \ & \Big\{ \ \psi_h : \Omega_t \times I \to \mathbb{R} \mid \quad \psi \in \mathcal{C}^0(\bar{\Omega}_t), \quad \psi\Big|_{K_t} \circ \mathcal{M}_k^{K_t} \in P_n(K_R), \\ & \forall K_t \in \mathcal{T}_{h,t} \ \Big\}, \end{aligned} \tag{2.34}$$

where $\mathcal{M}_k^{K_t}$ can be defined on each element $K_t \in \mathcal{T}_{h,t}$ as follows :

$$\mathcal{M}_k^{K_t} : K_R \to K_t, \quad \boldsymbol{x}(\eta) \ = \ \mathcal{M}_k^{K_t}(\eta) \ = \ \sum_{i \in N^K} \boldsymbol{x}_i(t) \ \tilde{\phi}_i(\eta), \tag{2.35}$$

$$\eta \in K_R \qquad \text{and} \qquad \tilde{\phi}_i \in P_k(K_R).$$

The choice made in section 2.1.6 of the finite element function space discretizing the ALE mapping can now be justified. Indeed, it guarantees that

$$\mathcal{X}_h(\Omega_{h,t}) = \mathcal{F}_{n,k}(\mathcal{T}_{h,t}), \tag{2.36}$$

as easily proven using the following proposition stated in [43].

Proposition 2.2 *If, at any $t \in I$, the discrete ALE mapping satisfies*

$$\hat{\mathcal{A}}_{h,t}|_{K_0} \circ \mathcal{M}_k^{K_0} \;=\; \mathcal{M}_k^{K_t}, \qquad \forall K_0 \in \hat{\mathcal{T}}_h, \quad K_t = \hat{\mathcal{A}}_{h,t}(K_0),$$

and

$$\hat{\mathcal{X}}_h \;=\; \mathcal{F}_{n,k}(\hat{\mathcal{T}}_h),$$

then we have that $\mathcal{X}_h(\Omega_{h,t})$, defined in (2.26), satisfies :

$$\mathcal{X}_h(\Omega_{h,t}) \;=\; \mathcal{F}_{n,k}(\mathcal{T}_{h,t}).$$

Finally, for the time and space discretization of a problem in the ALE frame, we will use the following general finite element space defined on the triangulation $\mathcal{T}_h^n$ of the domain at any time $t = t^n$:

$$\begin{aligned} \mathcal{F}_{n',k}(\mathcal{T}_h^n) \;=\; \Big\{ & \psi_h : \Omega_h^n \to \mathbb{R} \mid \quad \psi \in \mathcal{C}^0(\bar{\Omega_h^n}),\ \psi\Big|_{K_n} \circ \mathcal{M}_k^{K_n} \in P_{n'}(K_R) \\ & \forall K_n \in \mathcal{T}_h^n \ \Big\}. \end{aligned} \tag{2.37}$$

2.2 The linear advection problem on moving domains

2.2.1 Continuous problem

Let $I = [0, T]$ be an open interval and, for each $t \in I$, let Ω_t be a time dependent domain in $\mathbb{R}^d$ with a smooth enough boundary. We consider the linear advection equation of a scalar u :

$$\begin{aligned} \frac{\partial u}{\partial t} + \text{div}\,(\boldsymbol{\beta} u) \;&=\; f \qquad \text{in} \quad \Omega_t \times I, \\ u \;&=\; u_0 \qquad \text{in} \quad \Omega_{t_0} \quad \text{at} \quad t = 0 \end{aligned} \tag{2.38}$$

where $\boldsymbol{\beta}$ is a general advection velocity, assumed to be sufficiently regular and f is a source term. The system is closed by a suitable boundary condition : u is prescribed at the inflow boundaries $\Gamma_{in,t}$ of the domain, i.e.

$$u(\boldsymbol{x},t) \;=\; u_{in}(\boldsymbol{x},t) \qquad \forall\, \boldsymbol{x} \in \Gamma_{in,t},\; t \in I, \tag{2.39}$$

where $\Gamma_{in,t}$ is defined by

$$\Gamma_{in,t} \;=\; \{\; \boldsymbol{x} \in \partial\Omega_t \;/\; \boldsymbol{\beta}\cdot\boldsymbol{n} < 0 \;\}\,. \tag{2.40}$$

Using relation (2.10) , we obtain the ALE formulation of this problem :

$$\begin{aligned} \left.\frac{\partial u}{\partial t}\right|_{\hat{\boldsymbol{x}}} + \text{div }(\boldsymbol{\beta} u) - \boldsymbol{c}\cdot\boldsymbol{\nabla} u &= f \qquad \text{in} \quad \Omega_t \times I, \\ u &= u_0 \qquad \text{in} \quad \Omega_{t_0} \quad \text{at} \quad t = 0, \end{aligned} \tag{2.41}$$

where $\boldsymbol{c}(\boldsymbol{x},t)$ is the domain velocity defined by the ALE mapping $\hat{\mathcal{A}}_t$.

Let us indicate that in the case of a divergence free advection velocity – $\text{div}\,\boldsymbol{\beta} = 0$ – the advection equation can be reformulated in the following non-conservative way :

$$\begin{aligned} \left.\frac{\partial u}{\partial t}\right|_{\hat{\boldsymbol{x}}} + (\boldsymbol{\beta} - \boldsymbol{c})\cdot\boldsymbol{\nabla} u &= f \qquad \text{in} \quad \Omega_t \times I, \\ u &= u_0 \qquad \text{in} \quad \Omega_{t_0} \quad \text{at} \quad t = 0. \end{aligned} \tag{2.42}$$

2.2.2 Weak ALE formulations

Let us consider two spaces $\hat{\mathcal{Y}}$ and $\hat{\mathcal{X}}$ of functions defined on the reference domain and initially assumed to be regular enough. At each time $t \in I$, the corresponding spaces on the current domain are respectively $\mathcal{Y}(\Omega_t)$ and $\mathcal{X}(\Omega_t)$, defined as in (2.13). Considering a function $\hat{\psi} \in \hat{\mathcal{Y}}$, we multiply the first equation in (2.38) by the time dependent function $\psi(\boldsymbol{x},t) = \hat{\psi}(\mathcal{A}_t^{-1}(\boldsymbol{x}))$ defined in $\mathcal{Y}(\Omega_t)$ and we obtain the following weak formulation :

$$\int_{\Omega_t} \psi \left(\frac{\partial u}{\partial t} + \text{div }(\boldsymbol{\beta} u) \right)\, d\boldsymbol{x} = \int_{\Omega_t} \psi\; f\; d\boldsymbol{x} \qquad \forall\, \psi \in \mathcal{Y}(\Omega_t), \quad t \in I\,. \tag{2.43}$$

The time-dependent solution $u(t)$ is sought in

$$\mathcal{X}_0(\Omega_t,t) \;=\; \{\; \psi \in \mathcal{X}(\Omega_t,t) \quad | \quad \psi(t) = u_{in}(t) \quad \text{in} \quad \Gamma_{in,t}\,,\; t \in I\} \;. \tag{2.44}$$

where $\mathcal{X}(\Omega_t,t)$ is defined by (2.15) and $\Gamma_{in,t}$ denotes the inflow boundary of the domain.

The velocity $\boldsymbol{\beta}$ is assumed to satisfy $|| \operatorname{div} \boldsymbol{\beta} ||_{L^\infty(\Omega_t)} < \infty$. Consequently, we require that

$$\forall t \in I, \qquad u(t) \in H^1(\Omega_t) \quad \text{and} \quad \psi(t) \in L^2(\Omega_t).$$

Therefore we choose

$$\hat{\mathcal{X}} \;=\; H^1(\hat{\Omega}) \quad \text{and} \quad \hat{\mathcal{Y}} \;=\; L^2(\hat{\Omega}).$$

But since condition (2.39) is prescribed, u will be sought in

A non-conservative formulation Transforming the eulerian partial time derivative in (2.43), we obtain the following non-conservative weak ALE formulation of the problem :

$$\text{find } u \in \mathcal{X}_0(\Omega_t, t) \quad \text{such that} \quad \forall\, \psi \in \mathcal{Y}(\Omega_t),\ \forall t \in I,$$

$$\begin{aligned} \int_{\Omega_t} \psi \left.\frac{\partial u}{\partial t}\right|_{\hat{\boldsymbol{x}}} d\boldsymbol{x} \;+\; \int_{\Omega_t} \psi \,(\operatorname{div}\,(\boldsymbol{\beta} u) \;-\; \boldsymbol{c}\cdot\boldsymbol{\nabla} u)\; d\boldsymbol{x} \;&=\; \int_{\Omega_t} \psi\; f\; d\boldsymbol{x}, \\ u(\boldsymbol{x},0) \;&=\; u_0(\boldsymbol{x}) \quad \text{in } \Omega_{t_0}. \end{aligned} \tag{2.45}$$

A conservative formulation A conservative formulation can also be obtained, starting again from (2.43) but using relations (2.12) and (2.14) satisfied by the test functions. The result is the following :

$$\text{find } u \in \mathcal{X}_0(\Omega_t, t) \quad \text{such that} \quad \forall\, \psi \in \mathcal{Y}(\Omega_t),\ \forall t \in I,$$

$$\begin{aligned} \frac{d}{dt}\int_{\Omega_t} \psi\, u\; d\boldsymbol{x} + \int_{\Omega_t} \psi\, \operatorname{div}\,((\boldsymbol{\beta}-\boldsymbol{c})\, u)\; d\boldsymbol{x} \;&=\; \int_{\Omega_t} \psi\; f\; d\boldsymbol{x}, \\ u(\boldsymbol{x},0) \;&=\; u_0(\boldsymbol{x}) \quad \text{in } \Omega_{t_0}. \end{aligned} \tag{2.46}$$

For the derivation of this formulation, see for instance [43].

This formulation is called *conservative* because the ALE term is itself in a conservative form. For any $V \subset \Omega_t$ with sufficiently smooth boundaries and such that $\bar{V} \subset \Omega_t$, it is then possible to derive from (2.46) , taking $\psi\Big|_V = 1_V$, that

$$\frac{d}{dt}\int_V u\; d\boldsymbol{x} + \int_{\partial V} u\,(\boldsymbol{\beta}-\boldsymbol{c})\cdot\boldsymbol{n}\; d\Gamma = \int_V f\; d\boldsymbol{x} \qquad \forall\, t \in I. \tag{2.47}$$

This relation expresses the fact that, in absence of source terms, the variation of u over V is only due to contribution coming from the boundary of V. The ALE term consists indeed of an additional flux of u through the boundary as a consequence of the movement.

Formulations (2.45) and (2.46) are formaly equivalent at the continuous level. But we will see that their discrete counterparts are different, and the discrete system arising from (2.45) may not preserve the conservation property just mentioned.

2.2.3 Semi-discrete ALE formulations

In this section, the space discretization of both weak ALE formulations are derived, using the finite element spaces introduced in sections 2.1.6 and 2.1.7.

First, a triangulation $\hat{\mathcal{T}}_h$ of the reference configuration $\hat{\Omega}$ is defined. Three degrees k, n and n' are then chosen in order to determine the finite element function spaces

$$\hat{\mathcal{X}}_h \;=\; \mathcal{F}_{n,\,k}(\hat{\mathcal{T}}_h) \quad \text{and} \quad \hat{\mathcal{Y}}_h \;=\; \mathcal{F}_{n',\,k}(\hat{\mathcal{T}}_h)\,,$$

defined by (2.24) and approximating $\hat{\mathcal{X}} = H^1(\hat{\Omega})$ and $\hat{\mathcal{Y}} = L^2(\hat{\Omega})$. The ALE mapping is discretized by functions in $\mathcal{F}_{k,\,k}(\hat{\mathcal{T}}_h)$ as described in section 2.1.6, determining the triangulation $\mathcal{T}_{h,t}$ of the current domain Ω_t and thus its approximation $\Omega_{h,t}$.

The discrete ALE mapping then identifies the discrete spaces corresponding to $\hat{\mathcal{X}}_h$ and $\hat{\mathcal{Y}}_h$ on the current domain, which according to (2.36) are

$$\mathcal{X}_h(\Omega_{h,t}) \;=\; \mathcal{F}_{n,\,k}(\mathcal{T}_{h,t}) \quad \text{and} \quad \mathcal{Y}_h(\Omega_{h,t}) \;=\; \mathcal{F}_{n,\,k}(\mathcal{T}_{h,t}).$$

These discrete functional spaces will approximate respectively $\mathcal{X}(\Omega_t)$ and $\mathcal{Y}(\Omega_t)$. The solution u will be approximated in

$$\mathcal{X}_{h,0}(\Omega_{h,t},t) \;=\; \{\; \psi_h \in \mathcal{X}_h(\Omega_{h,t},t) \quad | \quad \psi_h(t) = u_{h,in}(t) \quad \text{in} \quad \Gamma_{h,in}\,,\; t \in I\}\;, \tag{2.48}$$

where $\mathcal{X}(\Omega_t,t)$ is defined by (2.27). The two following semi-discrete weak ALE formulations are obtained :

Semi-discrete non-conservative formulation :

$$\text{find } u_h \in \mathcal{X}_{h,0}(\Omega_{h,t},t) \quad \text{such that} \quad \forall\, \psi_h \in \mathcal{Y}_h(\Omega_{h,t}) \quad \text{and} \quad \forall\, t \in I,$$

$$\left\{\begin{array}{l} \displaystyle\int_{\Omega_{h,t}} \psi_h \frac{\partial u_h}{\partial t}\Big|_{\hat{\boldsymbol{x}}}\, d\boldsymbol{x} \;+\; \int_{\Omega_{h,t}} \psi_h \left(\,\mathrm{div}\,(\boldsymbol{\beta} u_h) \;-\; \boldsymbol{c}_h \cdot \boldsymbol{\nabla} u_h\right)\, d\boldsymbol{x} \\ \qquad\qquad = \displaystyle\int_{\Omega_{h,t}} f_h \psi_h \; d\boldsymbol{x}\,, \qquad (2.49) \\ \\ u_h(\boldsymbol{x},0) \;=\; u_{h,0}(\boldsymbol{x}) \quad \text{in } \Omega_{h,0}, \end{array}\right.$$

f_h being a suitable discretization of the source term.

semi-discrete conservative formulation :

$$\text{find } u_h \in \mathcal{X}_{h,0}(\Omega_{h,t},t) \quad \text{such that} \quad \forall\, \psi_h \in \mathcal{Y}_h(\Omega_{h,t}) \quad \text{and} \quad \forall\, t \in I,$$

$$\left\{\begin{array}{l} \displaystyle\frac{d}{dt}\int_{\Omega_{h,t}} u_h \psi_h \; d\boldsymbol{x} \;+\; \int_{\Omega_{h,t}} \psi_h\, \mathrm{div}\,\left((\boldsymbol{\beta} - \boldsymbol{c}_h)\, u_h\right)\, d\boldsymbol{x} \\ \qquad\qquad = \displaystyle\int_{\Omega_{h,t}} f_h \psi_h \; d\boldsymbol{x}\,, \qquad (2.50) \\ \\ u_h(\boldsymbol{x},0) \;=\; u_{h,0}(\boldsymbol{x}) \quad \text{in } \Omega_{h,0}. \end{array}\right.$$

2.3 Conservation and stability properties of the space-time discretization

As already announced, a numerical scheme deriving from the non-conservative formulation of the advection problem doesn't automatically preserve the conservation property described in the previous section, whereas the conservative formulation does. But there is another property of the continuous problem that should be preserved at the discrete level.

Let us consider the particular case where the advection velocity $\boldsymbol{\beta}$ is assumed to be divergence free, i.e.

$$\mathrm{div}\,\boldsymbol{\beta} \;=\; 0 \qquad \text{in } \Omega_t, \quad t \in I. \qquad (2.51)$$

In this case, the continuous problem can be written as in (2.42) and admits constant solutions whenever the source term f is zero. We observe that any numerical scheme deriving from the semi-discrete non-conservative formulation (2.49), preserves this property independently of the time discretization. On the contrary, it is not automatically preserved by a scheme deriving from the semi-discrete conservative formulation (2.50). Note however that this only holds for advection problems raised on moving domains. Indeed, if $\boldsymbol{c}_h = 0$, any numerical scheme admits constant solutions.

Throughout, we will refer to the ability of a numerical scheme to admit constant solutions whenever the continuous problem does as the *constancy preservation property*. The condition for a numerical scheme set on a moving domain to verify this property is called the Geometric Conservation Law (GCL). We will next describe this notion more precisely.

In summary, on moving domains, the conservative formulation ensures a correct conservation of the advected quantity at the discrete level, whereas the non-conservative formulation ensures the constancy preservation property.

2.3.1 The Geometric Conservation Law

The notion of GCL has initially been widely investigated in the framework of the finite difference and finite volume methods applied to fluid dynamic problems, see in particular [143, 80, 58]. The aim of these laws is to guarantee that a discrete scheme respects the conservation of geometric quantities as the continuous problem does. Indeed, the domain movement can affect in a negative way the solution of a discrete scheme arising from an ALE formulation, causing a loss of the conservation properties and possible resulting instabilities. As already mentioned, the GCL is generally defined as the capability of an ALE numerical scheme to represent a constant solution. In [58] Guillard and Farhat prove that satisfying the GCL is a sufficient condition for a particular numerical scheme to be consistent on moving grids. More generally, these authors claim that a higher accuracy is obtained with schemes satisfying the GCL compared to schemes that violate it.

This notion has been extended to the finite element method by Formaggia and Nobile in [43]. They establish a clear link between the fulfilment of these laws and the degree of exactness of the time advancing scheme. Furthermore, they prove that the GCL provides a sufficient condition for the unconditional stability of the backward Euler scheme, when applied to the linear advection diffusion problem with divergence free velocity.

The work has been carried on, leading to further characterizations of the GCL. In [52], Gerbeau *et al.* studied an ALE formulation of a magnetohydrodynamic (MHD) flow problem with a moving interface. The discrete mass conservation and the global energy inequality of the numerical scheme proposed are ensured under two conditions, which are strongly related to the GCL.

However, in [57], it is proven that the GCL is neither a necessary nor a sufficient condition for an ALE numerical scheme to preserve on moving grids its order of time-accuracy established on fixed grids. Finally, Formaggia and Nobile analysed more widely the linear advection diffusion problem in

[47], and they reached the same conclusions. The time accuracy of the numerical schemes can be preserved irrespectively of the satisfaction of the GCL, provided that the domain movement is suitably interpolated. Moreover, they have shown that, except in the case of the Euler backward scheme, the GCL does not guarantee for unconditional numerical stability.

2.3.2 The GCL in this work

In the next chapter we will present a numerical scheme for the linear advection problem with divergence-free velocity, derived from the non-conservative ALE formulation (2.45). We will show that this scheme requires the verification of a particular constraint to ensure a good conservation of the advected quantity when the domain moves. The interesting feature is that the expression of this constraint is strongly related to the expression of the GCL for the backward Euler scheme applied to the linear advection diffusion problem with divergence free velocity derived in [43]. We can therefore conclude that, for some numerical schemes, a particular discrete relation must be satisfied to ensure both the conservation and the constancy preservation properties.

With this work, we aim in particular to contribute to the understanding of the notion of conserving the geometric quantities on moving domains.

Chapitre 3

A conservative multidimensional upwind residual distributive scheme for the linear advection problem on moving domains

3.1 Introduction

In this chapter we introduce a numerical scheme for the resolution of the linear advection problem set on a moving domain. The requirements for this scheme were a high accuracy and stability of the solution, as well as the monotonicity- and conservation of the advected quantity.

The idea was to use the Multidimensional Upwind Residual Distributive (MURD) approach, well-known for its ability to satisfy the properties just mentioned. However, the schemes based on this approach have been developed for problems defined on fixed domains, and their extension to moving domains is not yet well-established. Now, in section 2.3, we have explained how the domain movement can affect numerical schemes in a negative way. Indeed, it has been proven by several authors – see for instance [43, 57, 58] – that some schemes suffer from a loss of their accuracy, stability or conservation properties when applied to a problem set on a moving domain. In certain cases, the satisfaction of a particular condition – the so-called Geometric Conservation Law (GCL) – ensures the preservation of these properties.

Therefore, particular care must be taken when applying a MURD scheme to a problem set on a domain with moving boundaries, since some of the properties it satisfies on a fixed domain may be lost. In [74], Janin proposed the application of two widely-used MURD schemes to the advection of an active

tracer driven by a three-dimensional free-surface flow. Both schemes were derived for the particular case where the flow and the tracer are computed by the Telemac-3D system (see [66]), developed at the LNHE, EDF R&D. The author showed that a particular constraint must be satisfied to ensure the accurate conservation of the tracer quantity.

Our aim is to generalize this result and reformulate it in the ALE frame. In particular, we are interested in relating the additional conditions a MURD scheme must satisfy when applied to a problem set on a moving domain to the concept of the GCL. For this purpose we derive a generic form of MURD schemes for the linear advection problem with moving boundaries formulated in the ALE frame. Note that these schemes approximate the non-conservative form of the advection equation. We will refer to this class of schemes as ALE-MURD schemes for the linear advection problem in non-conservative form.

After introducing more widely the upwind residual distribution approach, the bases of the ALE-MURD schemes we propose are derived from a non-conservative ALE formulation of the continuous advection problem with divergence-free advection velocity. Then, it is shown that the properties ensured by the corresponding residual distributive schemes set on fixed domains are not altered by the domain movement, except for the conservation of the advected quantity. Thus, we formulate the additional condition the ALE-MURD schemes proposed must satisfy to be conservative when the domain moves. The strong relation between this conservation constraint and the GCL is discussed. Finally, a conservative advection ALE-MURD scheme is proposed for the particular framework of three-dimensional domains moving in the vertical direction only.

3.2 The Multidimensional Upwind Residual Distributive (MURD) schemes

It is now well established that for hyperbolic systems of conservation laws, upwind methods outperform other space discretizations in terms of accuracy, control of oscillations near discontinuities and rate of convergence to steady state. Traditionally associated to the finite volume method, this concept has also found its way into the finite element framework. In particular, the Streamline Upwind Petrov-Galerkin (SUPG) method [71] was developed, based on the modification of the weighting functions in order to stabilize the standard Galerkin method. This approach considerably reduces the numerical diffusion with respect to the classical upwind methods, but it is still not optimal because it doesn't preserve the monotonic property of an advected quantity. This means that the method does not prevent the creation of new extrema on the solution nor the amplification of existing extrema. A typical consequence of the lack of monotonicity of a numerical scheme is the possible appearing of "undershoots or overshoots" relative to the amplitude

of the initial distribution when attempting to represent a discontinuity.

Upwind Residual Distributive schemes – also called Fluctuation Splitting schemes – have emerged as an attractive alternative to the classical finite volume and finite element upwind methods.

The Residual Distribution approach was pioneered by Ni in [105]. He proposed a first scheme which was further developed by Morton *et al.* [100] and has strong similarities with the Finite Element SUPG scheme on bilinear quadrilaterals. In 1982, Roe [119] introduced a Residual Distribution framework in an upwind context under the name of Fluctuation Splitting method. A first scheme [114], developed for quadrilateral cells, was followed by the development of multidimensional upwind schemes for unstructured meshes by a long list of authors. We can note Roe, Deconinck, Paillère, Sidilkover and Struijs (see for instance [34, 106, 108, 134]), as well as Abgrall [2, 3]. These schemes present a number of attractive features – such as lower cross-diffusion due to multidimensional upwinding –, but their main advantage is that they can be made *monotonic* by design. First developed for scalar advection equations – see [107] for a review of the existing scalar advection schemes in 2 and 3 dimensions –, Fluctuation Splitting schemes have then been extended to systems of conservation laws (see for instance [1] and [35]).

The residual distributive schemes are based on the finite volume method but they use a finite element data structure. Indeed, the unknowns are defined at the vertices of linear mesh elements and are approximated by continuous finite-dimension functions. The method consists in computing, for each element of the mesh, the residual (or fluctuation) within each time step – *i.e.* the source term for the local increment of the unknown. It is then optimally distributed over the vertices of the element with weighting coefficients summing up to unity for consistency. These nodal local residuals are finally assembled in order to retrieve a global increment at each node.

There are several ways to perform the optimal distribution of the residual inside an element. A significant advantage of the approach lies in the fact that the different properties required for the numerical scheme can easily be expressed. The distribution function is then chosen such that the conditions imposed by these properties are satisfied.

MURD schemes are *multidimensionally upwind* Fluctuation Splitting schemes, in the sense that the residual is distributed to downstream nodes only. High-order, monotonic and shock-capturing advection schemes have been designed using the MURD approach, with particular concern for the minimization of cross-diffusion.

The properties required for the class of MURD schemes we derive below for the linear advection problem set on a moving domain are *conservation* and *monotonicity*, as well as the *minimization* of *oscillations* and *cross-diffusion*

in the numerical solution. Note that the generic form of these schemes will be derived from a non-conservative formulation of the advection equation, because the monotonicity is then more easily ensured.

3.3 A generic form of ALE-MURD schemes for the linear advection problem in non-conservative form.

3.3.1 Problem settings

The linear advection problem (2.38) is proposed to be solved in the particular case where the advection velocity is divergence-free, , *i.e.*

$$\operatorname{div} \boldsymbol{\beta} \;=\; 0 \qquad \text{in } \Omega_t, \quad t \in I. \tag{3.1}$$

We recall that this problem is closed by a Dirichlet boundary condition (2.39) at the inflow boundaries of the domain.

Integrating the non-conservative ALE formulation (2.42) of the advection equation over the current domain Ω_t leads to the following weak problem

$$\text{find } u \in \mathcal{X}_0(\Omega_t, t) \quad \text{such that}$$

$$\begin{aligned} \int_{\Omega_t} \left(\frac{\partial u}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + (\boldsymbol{\beta} - \boldsymbol{c}) \cdot \boldsymbol{\nabla} u \right) d\boldsymbol{x} \;&=\; 0, \qquad t \in I, \\ u(0) \;&=\; u_0 \qquad \text{in } \Omega_{t_0}, \end{aligned} \tag{3.2}$$

where $\mathcal{X}_0(\Omega_t, t)$ is tdefined by (2.44).

The problem is first semi-discretized in space as described in sections 2.1.6 and 2.1.7. We indicate that the Lagrangian finite element function spaces $\mathcal{F}_{k,\,k}(\hat{\mathcal{T}}_h)$ and $\mathcal{F}_{n',\,k}(\hat{\mathcal{T}}_h)$ – defined by (2.24) – chosen to approximate respectively the ALE mapping and the functional space $\hat{\mathcal{X}}$, are such that *the degrees of freedom are located at the vertices of the mesh elements.*

The time interval I is then divided in N_t time steps of equal length Δt. We denote by Ω_h^n the approximation of the real domain at each time $t^n = n\,\Delta t$, for $n = 0, \ldots, N_t$, and we introduce the following application, for any $n \in [0, N_t]$:

$$\mathcal{A}_{n,n+1} : \Omega_h^n \longrightarrow \Omega_h^{n+1}, \qquad \mathcal{A}_{n,n+1}(\boldsymbol{x}) = \hat{\mathcal{A}}_{h,n+1} \circ \mathcal{A}_{h,n}^{-1}(\boldsymbol{x}).$$

At each time t^n, the solution $u(t^n)$ will be approximated by u_h^n in the following time-discrete functional space :

$$\mathcal{X}_h^n \;=\; \{\; \psi_h : \Omega^n \longrightarrow \mathbb{R}, \quad \psi_h \;=\; \hat{\psi}_h \circ \mathcal{A}_{h,n}^{-1},\; \hat{\psi}_h \in \hat{\mathcal{X}}_h \;\}\,, \tag{3.3}$$

which, following Proposition 2.2, coincides with $\mathcal{F}_{n',k}(\mathcal{T}_h^n)$ defined in (2.37).

Given the initial condition u_0 we consider the following explicit Euler time-advancing scheme :

$$\text{for each} \quad n = 0, \ldots, N_t - 1 \quad \text{find} \quad u_h^{n+1} \in \mathcal{X}_{h,0}^{n+1} \quad \text{such that}$$

$$\begin{aligned} &\int_{\Omega_h^{n+1}} \left(u_h^{n+1} - u_h^n \circ \mathcal{A}_{n+1,n}\right)\, d\boldsymbol{x} \\ &\qquad + \quad \Delta t \int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\left(u_h^n \circ \mathcal{A}_{c,n}\right)\, d\boldsymbol{x} \;=\; 0, \end{aligned} \tag{3.4}$$

where Ω_h^c and $(\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c$ denote respectively the approximations of the real domain and the ALE advection velocity at some arbitrary time t^c within t^n and t^{n+1}.

Note that the advection velocity $\boldsymbol{\beta}$ is discretized at each time t^n in some particular finite element space which is a subset of $L^2(\Omega_h^n)$ and whose degrees of freedom are located at the vertices of the mesh elements.

3.3.2 Notations

Throughout the derivation of the scheme we will use an important amount of notations. For the sake of clarity, let us previously define some of them.

- $\delta\mathcal{M}_u^{n+1}$ denotes the global increment of the advected quantity u_h within times t^n and t^{n+1},
- $\delta\mathcal{M}_{u,K_0}^{n+1}$ denotes the local increment of u_h on the element K_0,
- $\Delta t\, \Phi_{K_0}^{n+1}$ denotes the local residual on the element K_0 within $[t^n, t^{n+1}]$,
- $\hat{\delta u_h}^{n+1}$ is the increment $\hat{u}_h^{n+1} - \hat{u}_h^n$ of the advected variable u_h formulated in the reference configuration,
- δu_i^{n+1} denotes the value of the increment $\hat{\delta u_h}^{n+1}$ at the i-th node,
- $\delta u_{i,K_0}^{n+1}$ is the contribution of the element K_0 containing the i-th node to the nodal increment δu_i^{n+1}.

3.3.3 Derivation of the scheme

In the particular case studied here, the global increment $\delta\mathcal{M}_u^{n+1}$ of the advected variable u_h is due to the different advective fluxes considered within $[t^n, t^{n+1}]$ as expressed by (3.4). In order to apply the *residual distribution*

principle, we must first reformulate the discrete problem locally. In this way, the local residual Φ_{K_0} – "source" of the local increment $\delta\mathcal{M}_{u,K_0}^{n+1}$ – can be expressed on each element K_0 of the mesh. Each local residual will then be distributed to the nodes of its corresponding element. Finally, the increment δu_i^{n+1} will be retrieved at each node i, determining the value of u_h at $t = t^{n+1}$.

For this purpose, let us reformulate the global increment on the reference configuration as follows :

$$\begin{aligned}
\delta\mathcal{M}_u^{n+1} &= \int_{\Omega_h^{n+1}} \left(u_h^{n+1} - u_h^n \circ \mathcal{A}_{n+1,n}\right)\, d\boldsymbol{x} \\
&= \int_{\hat{\Omega}} \hat{J}_h^{n+1} \left(\hat{u}_h^{n+1} - \hat{u}_h^n\right)\, d\hat{\boldsymbol{x}} \\
&= \int_{\hat{\Omega}} \hat{J}_h^{n+1}\, \hat{\delta u}_h^{n+1}\, d\hat{\boldsymbol{x}}\,,
\end{aligned}$$

Clearly, $\hat{u}_h^n \in \hat{\mathcal{X}}_h$ at any time t^n – see the description of the discrete function spaces in section 2.1.7. The dimension of $\hat{\mathcal{X}}_h$ is denoted by N_h and is equal to the amount of nodes in $\hat{\mathcal{T}}_h$. For $i = 1, \ldots, N_h$, $\hat{\psi}_i$ indicates the basis function associated to node i in the mesh. At each time t^n, the corresponding basis function of $\mathcal{X}_h^n$ is then $\psi_i^n = \hat{\psi}_i \circ \mathcal{A}_{h,n}^{-1}$.

Following (2.33), $\hat{\delta u}_h^{n+1}$ is decomposed on the basis functions of $\hat{\mathcal{X}}_h$ and we obtain :

$$\begin{aligned}
\delta\mathcal{M}_u^{n+1} &= \sum_{i=1}^{N_h} \delta u_i^{n+1} \int_{\hat{\Omega}} \hat{J}_h^{n+1}\, \hat{\psi}_i\, d\hat{\boldsymbol{x}} \\
&= \sum_{i=1}^{N_h} \delta u_i^{n+1}\, V_i^{n+1},
\end{aligned}$$

where

$$V_i^{n+1} = \int_{\hat{\Omega}} \hat{J}_h^{n+1}\, \hat{\psi}_i\, d\hat{\boldsymbol{x}} = \int_{\Omega_h^{n+1}} \psi_i^{n+1}\, d\boldsymbol{x}\,.$$

Each nodal increment δu_i^{n+1}, for $i = 1, \ldots, N_h$, is then assumed to result from the sum of the contributions of each element in the mesh $\hat{\mathcal{T}}_h$ containing the i-th node :

$$\delta u_i^{n+1} = \sum_{\substack{K_0 \in \hat{\mathcal{T}}_h \\ i \in K_0}} \delta u_{i,K_0}^{n+1}, \tag{3.5}$$

where, with an obvious abuse of notation, $i \in K_0$ means that node i belongs

to the element K_0.

In this way, the global increment can be formulated as the sum of local increments :

$$\delta\mathcal{M}_u^{n+1} \;=\; \sum_{K_0\in\hat{\mathcal{T}}_h} \Big(\sum_{i\in K_0} \delta u_{i,K_0}^{n+1}\; V_i^{n+1} \Big) \;=\; \sum_{K_0\in\hat{\mathcal{T}}_h} \delta\mathcal{M}_{u,K_0}^{n+1}, \tag{3.6}$$

with $\delta\mathcal{M}_{u,K_0}^{n+1} \;=\; \sum_{i\in K_0} \delta u_{i,K_0}^{n+1}\; V_i^{n+1}$.

Let us now reformulate the second term in equation (3.4), that is the global flux term. It can be expressed as a linear combination of local fluxes by decomposing u_h^n as in (2.31) :

$$\begin{aligned}
&\int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\,(u_h^n \circ \mathcal{A}_{c,n})\;\, d\boldsymbol{x} \\
&\qquad = \sum_{K_0\in\hat{\mathcal{T}}_h} \Big(\sum_{i\in K_0} u_i^n \int_{K_c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\,(\psi_i^n \circ \mathcal{A}_{c,n})\;\, d\boldsymbol{x}\; \Big)\,, \\
&\qquad = \sum_{K_0\in\hat{\mathcal{T}}_h} \Big(\sum_{i\in K_0} u_i^n \int_{K_c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\psi_i^c\; d\boldsymbol{x}\; \Big)\,,
\end{aligned}$$

where $K_c = \hat{\mathcal{A}}_{h,c}(K_0)$ and u_i^n is the value of u_h^n at the i-th node. The local fluxes considered within time step $[t^n, t^{n+1}]$, for $n = 0 \ldots N_t$, will be denoted, for each element $K_0 \in \hat{\mathcal{T}}_h$, by

$$\Phi_{K_0}^{n+1} \;=\; \sum_{i\in K_0} u_i^n \int_{K_c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\psi_i^c\; d\boldsymbol{x}\,. \tag{3.7}$$

Equation (3.4) can thus be expressed locally as follows :

$$\forall K_0 \in \hat{\mathcal{T}}_h, \qquad \delta\mathcal{M}_{u,K_0}^{n+1} \;=\; -\,\Delta t\; \Phi_{K_0}^{n+1}. \tag{3.8}$$

Thus, on each element $K_0 \in \hat{\mathcal{T}}_h$, the local residual within time step $[t^n, t^{n+1}]$ is $\Delta t\; \Phi_{K_0}^{n+1}$, since it is the source of the local increment.

The residual distribution principle can now be applied, leading to a first definition characterizing a more general class of ALE schemes, which we will refer to as *ALE Residual Distributive (ALE-RD) schemes for the linear advection problem in non-conservative form.*

Definition 3.3.1 The ALE-Residual Distributive schemes for the linear advection problem in non-conservative form consist in distributing each local residual $\Phi_{K_0}^{n+1}$, defined by (3.7), to the nodes of its corresponding element, contributing to the increment of the advected variable u_h as follows :

$$\forall K_0 \in \hat{\mathcal{T}}_h, \quad \forall i \in K_0, \qquad \delta u_{i,K_0}^{n+1} \;=\; -\,\frac{\Delta t}{V_i^{n+1}}\;\beta_{i,K_0}^{n+1}\;\Phi_{K_0}^{n+1}, \tag{3.9}$$

where the distribution coefficients β_{i,K_0}^{n+1} are chosen such that

$$\forall K_0 \in \hat{\mathcal{T}}_h, \quad \sum_{i \in K_0} \beta_{i,K_0}^{n+1} \;=\; 1 \tag{3.10}$$

in order to guarantee consistency.
The increment δu_i^{n+1} is retrieved by assembling contributions from all the elements using (3.5). Each nodal value of u_h is thus updated as

$$u_i^{n+1} \;=\; u_i^n \;-\; \frac{\Delta t}{V_i^{n+1}} \sum_{\substack{K_0 \in \hat{\mathcal{T}}_h \\ i \in K_0}} \beta_{i,K_0}^{n+1}\;\Phi_{K_0}^{n+1} \qquad \forall\, i = 1, \ldots, N_h. \tag{3.11}$$

Remark 3.3.1 We indicate that this definition is still valid if the scheme is raised on a fixed domain, considering that, for each node i and at each time t^n, V_i^n is equal to a constant value V_i.

Definition 3.3.2 A residual distributive scheme is multidimensionally upwind if the local residuals are distributed to downstream nodes only ; that is if the distribution coefficients are chosen such that, for any mesh element K_0, if i is an upstream node of K_0 then $\beta_{i,K_0} = 0$.

We call ALE-MURD scheme a multidimensionally upwind ALE-RD scheme.

Remark 3.3.2 Because of the explicit character of these schemes, the prescription of the boundary condition (2.39) can be performed independently from the resolution of the equation : at each time step, once the solution u_h^{n+1} has been computed, the boundary values are imposed strongly.

At this point we have derived a generic form of ALE-MURD schemes for the linear advection problem in non-conservative form. Different schemes can now be designed on this basis by choosing a particular set of distribution coefficients – making sure that the consistency condition (3.10) as well as the multidimensional upwinding condition are satisfied. The optimal distribution can be chosen such that particular properties are satisfied by the scheme.

3.4 Characterization of the optimal distribution

Referring to [107], four essential design criteria can be imposed on a residual distributive scheme, ensuring particular properties. We describe below these criteria, the conditions they impose on the choice of the distribution coefficients and the properties they guarantee to residual distributive schemes set on fixed domains. We also establish whether or not these properties are preserved on a moving domain.

3.4.1 Consistency

Definition 3.4.1 The consistency condition requires that the distribution coefficients associated to a same element in the mesh add up to 1, *i.e.* that

$$\sum_{i \in K_0} \beta_{i,K_0} \;=\; 1 \quad \forall K_0 \in \hat{\mathcal{T}}_h.$$

This condition is called *consistency condition* because it ensures that the solution computed by the distributive scheme is actually a solution of the original discrete equation. For instance in the case of the ALE-RD schemes defined previously, this condition ensures that relation (3.8) holds true, and therefore that the discrete equation (3.4) is solved.

Remark 3.4.1 Note that any Residual Distributive scheme can be rewritten as :

$$u_i^{n+1} \;=\; \sum_j m_j \; u_j^n \,,$$

where the consistency condition implies that $\sum_j \; m_j = 1$. Two important classes of schemes can then be identified : *linear schemes*, for which the m_j are independent of u, and *non-linear schemes* for which the m_j depend on u.

Definition 3.4.2 Let us consider an explicit numerical scheme for the linear advection problem with divergence free velocity, raised on a fixed domain Ω_h. The scheme is conservative if, at each time step,

$$\int_{\Omega_h} (u_h^{n+1} \;-\; u_h^n) \; d\boldsymbol{x} \;=\; -\; \Delta t \int_{\partial\Omega_h} u_h^n \; \boldsymbol{\beta}_h^c \cdot \boldsymbol{n}_h \; d\Gamma, \qquad (3.12)$$

where $\boldsymbol{\beta}_h^c$ is the advection velocity considered within this time step.

Proposition 3.1 *A Residual Distributive scheme for the linear advection problem in non-conservative form, with divergence free advection velocity and raised on a fixed domain Ω_h, is conservative in the sense of (3.12) if the consistency condition is satisfied.*

Proof. The Residual Distributive scheme considered is defined as described in Definition 3.3.1, considering that the domain Ω_h is fixed. Following (3.6) we have that

$$\int_{\Omega_h} (u_h^{n+1} - u_h^n)\, d\boldsymbol{x} \;=\; \delta\mathcal{M}_u^{n+1} \;=\; \sum_{K_0 \in \mathcal{T}_h} \delta\mathcal{M}_{u,K_0}^{n+1}\,.$$

Since the consistency condition is satisfied at each time step, relation (3.8) holds true and therefore :

$$\begin{aligned}
\sum_{K_0 \in \mathcal{T}_h} \delta\mathcal{M}_{u,K_0}^{n+1} &= -\,\Delta t \sum_{K_0 \in \mathcal{T}_h} \Phi_{K_0}^{n+1} \\
&= -\,\Delta t \sum_{K_0 \in \mathcal{T}_h} \sum_{i \in K_0} u_i^n \int_{K_0} \boldsymbol{\beta}_h^c \cdot \boldsymbol{\nabla}\psi_i^c \, d\boldsymbol{x} \\
&= -\,\Delta t \int_{\Omega_h} \boldsymbol{\beta}_h^c \cdot \boldsymbol{\nabla}\psi_i^c \, d\boldsymbol{x}\,.
\end{aligned}$$

Integrating by parts the flux term we obtain :

$$\int_{\Omega_h} u_h^{n+1} - u_h^n \, d\boldsymbol{x} \;=\; -\,\Delta t \int_{\partial\Omega_h} u_h^n \, \boldsymbol{\beta}_h^c \cdot \boldsymbol{n}_h \, d\Gamma \;+\; \Delta t \int_{\Omega_h} u_h^n \;\mathrm{div}\, \boldsymbol{\beta}_h^c \, d\boldsymbol{x}\,,$$

Finally, since the advection velocity is divergence-free, we can state that the scheme satisfies the conservation property (3.12). $\diamondsuit$

3.4.2 Multidimensional upwinding

Multidimensional upwinding prevents the numerical solution from unphysical oscillations and minimizes the creation of cross-diffusion. Indeed, since no fraction of a local increment is sent to an upstream node, the information is sent in the physical propagation direction.

3.4.3 Positivity

Another criterion for the design of non-oscillatory schemes is the so-called *positivity* property.

Definition 3.4.3 A scheme is said to be positive if the value of the solution at the new time can be written as a convex sum of values at the previous time, *i.e.*

$$\forall i \in [1, N_h],$$

$$u_i^{n+1} \;=\; \sum_{j=1}^{N_h} m_j \; u_j^n\,, \quad \text{with} \quad m_j \geq 0 \quad \forall j \quad \text{and} \quad \sum_j m_j = 1\,.$$

The following proposition holds :

Proposition 3.2 *The positivity of a distributive scheme ensures a maximum principle which prohibits the creation of new extrema.*

This is easy to verify since

$$\min_j u_j^n \;\leq\; u_i^{n+1} \;=\; \sum_{j=1}^{N_h} m_j \; u_j^n \;\leq\; \max_j u_j^n.$$

Another criterion – more restrictive but easier to implement – is the *local positivity.*

Definition 3.4.4 A scheme is said to be locally positive if the contribution of each element $K_0 \in \hat{\mathcal{T}}_h$ to the nodal increment δu_i^{n+1} of the solution for the new time can be written as a convex sum of values at the previous time, *i.e.*

$$\forall i \in [1, N_h], \quad \forall K_0 \text{ such that } i \in K_0,$$

$$\delta u_{i,K_0}^{n+1} \;=\; \sum_{j \in K_0} m_{j,K_0} \; u_j^n \,, \quad \text{with} \quad m_{j,K_0} \;\geq\; 0 \quad \forall j \neq i \;.$$

It is obvious that m_i can be made positive for a sufficiently small time step length Δt. This condition therefore comes to require that the contribution of each element to the nodal increment of the unknown be positive.

Proposition 3.3 *Local positivity implies global positivity under a global CFL condition.*

Proposition 3.4 *Local positivity ensures a discrete maximum principle which guarantees for the monotonicity of the scheme.*

For the proof of the last proposition and more details about local positivity, see for instance [107].

3.4.4 Linearity preservation

Definition 3.4.5 A scheme is said to be linearity-preserving ($\mathcal{L}P$) if it preserves the exact steady state solution when this is a piecewise linear polynomial function on the considered triangulation of the domain.

Proposition 3.5 *A necessary and sufficient condition for a scheme to be $\mathcal{L}P$ is that, for any element $K_0 \in \hat{\mathcal{T}}_h$,*

$$\textit{if} \quad \Phi_{K_0} \to 0 \quad \textit{then} \quad \beta_{i,K_0} \; \Phi_{K_0} \to 0 \quad \forall i \in K_0.$$

This condition is equivalent to the condition requiring that all the distribution coefficients β_{i,K_0} be bounded.

Indeed, when approaching the steady state, the local residual tends to zero. This tendance should be preserved when distributing the local residual on the element nodes. For linear schemes, the following property has been shown in [106].

Proposition 3.6 *On a regular mesh, $\mathcal{LP}$ implies the absence of cross-flow diffusion for the steady-state solution of the homogenous advection equation, and hence second-order accuracy.*

Another property has been proven by Struijs in [135]. It extends the Godunov theorem to the multidimensional case.

Proposition 3.7 *Only non-linear schemes can combine positivity and linear preservation.*

Clearly, the properties ensured by the multidimensional upwinding, the positivity – local as well as global – and the linearity preservation criteria are preserved on a moving domain. Indeed, neither the conditions they impose nor the relations they imply are altered by the domain's movement. On the contrary, the conservation property is not automatically preserved, since the expression of the required discrete conservation changes when the domain moves. We will formulate later on the additional constraint the numerical scheme must satisfy to be conservative.

3.5 Some particular coefficient distributions

Well-known MURD schemes have been developed by designing particular distribution functions combining multidimensional upwinding with other criteria among those described. We will only mention three of them. The reader may refer to [107] for a more extended list and a precise description of the existing MURD schemes.

In order to better describe the schemes we present next, we formulate the local residual on each element K_0 as a sum of values of the solution at the previous time step, that is

$$\Phi_{K_0}^{n+1} = \sum_{i \in K_0} k_{i,K_0} \, u_i^n,$$

where k_{i,K_0} are real coefficients determined by the distribution coefficients of the scheme. This is always possible for residual distributive schemes because they are explicit.

For the sake of simplicity, the distribution coefficients characterizing these schemes will only be given for the two-dimensional case. For this purpose,

we assume that the domain is discretized by triangles.

The linear positive scheme with lower cross-diffusion is the N scheme [120], given by

$$\beta_{i,K_0} \;=\; \frac{\Phi_{i,K_0}}{\Phi_{K_0}} \tag{3.13}$$

on each element K_0, with

$$\Phi_{i,K_0} \;=\; -\,\frac{max(0,k_{i,K_0})}{\displaystyle\sum_{l\in K_0} max(0,k_{l,K_0})}\;\sum_{j\in K_0} min(0,k_{j,K_0})\;(u_i^n-u_j^n). \tag{3.14}$$

It is easy to verify that this scheme is positive. But since the distribution coefficients are not bounded, the linearity preservation is not satisfied. As a result, the scheme can diffuse even when a steady-state linear solution has been found.

The MURD version of the SUPG scheme, identical to its finite element version when the advection velocity is constant, is a linear and linearity-preserving MURD scheme, but it is not positive.

Finally, we indicate the non-linear PSI (Positive Streamwise Invariant) scheme, developed by Struijs [134]. It can be written as a limited version of the N scheme, that is as

$$\beta_{i,K_0} \;=\; \Psi\left(\frac{\Phi_{i,K_0}}{\Phi_{K_0}}\right) \tag{3.15}$$

on each element K_0, with

$$\Psi(r) \;=\; max\left(0,min(r,1)\right). \tag{3.16}$$

This scheme is upwind, positive under a CFL condition, and linearity-preserving because the limiter bounds its coefficients.

3.6 Conservation property of the schemes

We have indicated previously that consistency is only a necessary condition for a distributive scheme to be conservative when set on a moving domain. Indeed, further conditions must be satisfied to make sure that the advected quantity is really conserved.

Let us first define the conservation property required. At the continuous level, the conservation of a variable u, advected by a velocity $\beta(t)$ in a domain Ω_t moving with the velocity $\boldsymbol{c}(t)$, can be expressed as follows :

$$\frac{d}{dt}\int_{\Omega_t} u\;d\boldsymbol{x} \;=\; -\int_{\partial\Omega_t} u\,(\boldsymbol{\beta}-\boldsymbol{c})\cdot\boldsymbol{n}_h\;d\Gamma, \tag{3.17}$$

where $\boldsymbol{n}_h$ is the outward normal to the current domain. By integrating in time this relation over $[t^n, t^{n+1}]$ we obtain :

$$\int_{\Omega_{t_{n+1}}} u(t^{n+1})\ d\boldsymbol{x} - \int_{\Omega_{t_n}} u(t^n)\ d\boldsymbol{x} \ = \ -\int_{t^n}^{t^{n+1}} \int_{\partial\Omega_t} u\,(\boldsymbol{\beta}-\boldsymbol{c})\cdot\boldsymbol{n}_h\ d\Gamma\ dt.$$

According to the time and space discretization (3.4) of the linear advection problem, the required conservation property is :

$$\int_{\Omega_h^{n+1}} u_h^{n+1}\ d\boldsymbol{x} - \int_{\Omega_h^n} u_h^n\ d\boldsymbol{x} \ = \ -\ \Delta t \int_{\partial\Omega_h^c} (u_h^n \circ \mathcal{A}_{c,n})\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma. \tag{3.18}$$

Moreover, we recall that the ALE mapping $\hat{\mathcal{A}}_t$ must be such that the domain velocity $\boldsymbol{c}$ satisfies

$$\boldsymbol{c}\cdot\boldsymbol{n} \ = \ \boldsymbol{\beta}\cdot\boldsymbol{n} \quad \text{on} \quad \Gamma_{imp,t} \qquad \text{and} \qquad \boldsymbol{c}\cdot\boldsymbol{n} \ = \ 0 \quad \text{on} \quad \Gamma_{liq,t},$$

where $\Gamma_{imp,t}$ denote the impermeable boundaries of the domain – *i.e.*, impermeable solid boundaries and free surfaces –, and $\Gamma_{liq,t}$ denotes its fictitious liquid boundaries, assumed to be still in the normal direction – see section 2.1.6. This condition should be preserved at the discrete level. The discrete ALE mapping $\hat{\mathcal{A}}_{h,t}$ must therefore be such that, for each $n = 1, \ldots, N_t$,

$$\boldsymbol{c}_h^n\cdot\boldsymbol{n}_h \ = \ \boldsymbol{\beta}_h^n\cdot\boldsymbol{n}_h \quad \text{on} \quad \Gamma_{imp}^n \qquad \text{and} \qquad \boldsymbol{c}_h^n\cdot\boldsymbol{n}_h \ = \ 0 \quad \text{on} \quad \Gamma_{liq}^n.$$

A stronger discrete conservation property than (3.18) can therefore be required at the discrete level, whose definition is the following.

Definition 3.6.1 According to the time and space discretization (3.4) of the linear advection problem, the *discrete conservation property* required for the numerical scheme at each time step is :

$$\int_{\Omega_h^{n+1}} u_h^{n+1}\ d\boldsymbol{x} - \int_{\Omega_h^n} u_h^n\ d\boldsymbol{x} \ = \ -\ \Delta t \int_{\Gamma_{liq}^c} (u_h^n \circ \mathcal{A}_{c,n})\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma. \tag{3.19}$$

Let us now formulate the *conservation constraint* of the ALE-RD schemes for the linear advection problem in non-conservative form, that is the condition they must satisfy to ensure the required discrete conservation property (3.19).

Assuming that the consistency condition (3.10) is satisfied, the solution provided by these schemes satisfy the discrete equation (3.4). We can therefore write the following equivalences :

$$
\begin{aligned}
(3.19) \quad \Leftrightarrow & \int_{\hat{\Omega}} \left(\hat{J}_h^{n+1}\ \hat{u}_h^{n+1} \ - \hat{J}_h^{n}\ \hat{u}_h^{n} \right)\ d\hat{\boldsymbol{x}} \\
& \qquad + \quad \Delta t \int_{\Gamma_{liq}^c} (u_h^n \circ \mathcal{A}_{c,n})\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma \ = \ 0 \\
\Leftrightarrow & \int_{\hat{\Omega}} \hat{J}_h^{n+1}\ (\hat{u}_h^{n+1} - \hat{u}_h^{n}) \ + \ (\hat{J}_h^{n+1} - \hat{J}_h^{n})\ \hat{u}_h^{n}\ d\hat{\boldsymbol{x}} \\
& \qquad + \quad \Delta t \int_{\Gamma_{liq}^c} (u_h^n \circ \mathcal{A}_{c,n})\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma \ = \ 0 \\
& (\text{ using (3.4) }) \\
\Leftrightarrow & \int_{\Omega_h^{n+1}} u_h^n\ d\boldsymbol{x} \ - \int_{\Omega_h^n} u_h^n\ d\boldsymbol{x} \\
& \qquad - \quad \Delta t \int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla} (u_h^n \circ \mathcal{A}_{c,n})\ d\boldsymbol{x} \\
& \qquad\qquad + \quad \Delta t \int_{\Gamma_{h,liq}^c} (u_h^n \circ \mathcal{A}_{c,n})\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma \ = \ 0
\end{aligned}
$$

By decomposing u_h^n on the degrees of freedom we obtain :

$$
\begin{aligned}
(3.19) \quad \Leftrightarrow \quad \sum_{i=1}^{N_h} \ u_i^n \ \Bigg\{ & \int_{\Omega_h^{n+1}} \psi_i^{n+1}\ d\boldsymbol{x} \ - \int_{\Omega_h^n} \psi_i^n\ d\boldsymbol{x} \\
- \quad \Delta t \int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla} \psi_i^c\ d\boldsymbol{x} \quad & + \quad \Delta t \int_{\Gamma_{h,liq}^c} \psi_i^c\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma \Bigg\} \ = \ 0
\end{aligned}
$$

This relation must be true at any time t^n and for any distribution of u_h. Therefore, we can state the following proposition.

Proposition 3.8 *The numerical scheme (3.4) is conservative in the sense of (3.19) if and only if*

$$
\forall\ n = 1, \ldots, N_t, \quad \forall\ i = 1, \ldots, N_h,
$$

$$
\begin{aligned}
\int_{\Omega_h^{n+1}} \psi_i^{n+1}\ d\boldsymbol{x} \ - \int_{\Omega_h^n} \psi_i^n\ d\boldsymbol{x} \ - \ \Delta t \int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla} \psi_i^c\ d\boldsymbol{x} \\
+ \ \Delta t \int_{\Gamma_{h,liq}^c} \psi_i^c\ (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h\ d\Gamma \ = \ 0, \qquad (3.20)
\end{aligned}
$$

where t^c is the time at which the advective fluxes are considered within each time step.

This expression of the conservation constraint will be usefull for the particular case of the three-dimensional hydrostatic free surface problem that will be studied later on. However, for the general case, we will rather make use of a simplier form of this constraint. For this purpose, we first make the assumption that the Green Formula can be applied to the global flux term in (3.20), *i.e.* that at any time step t^c and for each node i,

$$\int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla}\psi_i^c \; d\boldsymbol{x} \;=\; \int_{\Gamma_{h,liq}^c} \psi_i^c \; (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{n}_h \; d\Gamma - \int_{\Omega_h^c} \psi_i^c \; \mathrm{div}\,(\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \; d\boldsymbol{x}\,. \tag{3.21}$$

Note that relation (3.21) holds true provided the space integrals and the gradient in the local flux terms $\Phi_{K_0}^{n+1}$ defined by (3.7) are computed exactly for the particular finite element functions used to approximate the solution, the ALE advection velocity and the domain.

Using (3.21), relation (3.20) turns to

$$\int_{\Omega_h^{n+1}} \psi_i^{n+1} \; d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n \; d\boldsymbol{x} \;+\; \Delta t \int_{\Omega_h^c} \psi_i^c \; \mathrm{div}\;(\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \; d\boldsymbol{x} \;=\; 0.$$

This expression of the conservation constraint involves the domain velocity but also the velocity of the fluid. That means it prevents from a loss of the advected quantity which can arise from the domain movement but also from an eventual mass-loss of the fluid itself. Obviously, if the discrete velocity of the fluid $\boldsymbol{\beta}_h$ is not locally divergence-free, the advected quantity cannot be well conserved, even on a fixed domain.

Let us now assume that the advection velocity has been discretized – or computed – such that it is divergence-free at least in the sense of

$$\int_{\Omega_h^c} \psi_i^c \; \mathrm{div}\,\boldsymbol{\beta}_h^c \; d\boldsymbol{x} \;=\; 0 \qquad \forall\; i = 1, \ldots, N_h, \quad \forall\; t^c \in I. \tag{3.22}$$

The following proposition can be written.

Proposition 3.9 *If the local fluxes (3.7) are computed exactly and if the discrete advection velocity is divergence-free in the sense of (3.22), then the numerical scheme (3.4) is conservative in the sense of (3.19) if the following relation holds :*

$$\forall\; n = 1, \ldots, N_t, \quad \forall\; i = 1, \ldots, N_h,$$
$$\int_{\Omega_h^{n+1}} \psi_i^{n+1} \; d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n \; d\boldsymbol{x} \;-\; \Delta t \int_{\Omega_h^c} \psi_i^c \; div\,\boldsymbol{c}_h^c \; d\boldsymbol{x} \;=\; 0, \tag{3.23}$$

where t^c is the time at which the advective fluxes are considered within each time step.

This conservation constraint only involves terms related to the mesh movement. Clearly, (3.23) is automatically satisfied in the case where the domain doesn't move. In fact, this traduces a loss of the advected quantity arising from the domain movement when a particular condition is not satisfied. Therefore, it is strongly related to the Geometric Conservation Laws (GCL) introduced in section 2.3.

3.7 Conservation constraint and GCL

The conservation constraint (3.23) can be viewed as a particular discretization – strongly related to the one used to approximate the problem – of the weak form associated to the Euler expansion formula (2.11).

This can easily be shown by deriving a weak form for the Euler expansion formula :

$$\int_{\hat{\Omega}} \frac{\partial \hat{J}_t}{\partial t}\ \hat{\psi}\ d\hat{\boldsymbol{x}}\ =\ \int_{\hat{\Omega}} \hat{J}_t\ \hat{\psi}\ \left(\operatorname{div} \boldsymbol{c} \circ \hat{\mathcal{A}}_t\right)\ d\hat{\boldsymbol{x}}, \qquad \forall \hat{\psi} \in \hat{\mathcal{X}}.$$

Since $\hat{\psi}$ does not depend on time, the left term can be integrated over $[t^n, t^{n+1}]$, and the weak form finally obtained is

$$\int_{\Omega^{n+1}} \psi(t^{n+1})\ d\boldsymbol{x}\ -\int_{\Omega^n} \psi(t^n)\ d\boldsymbol{x}\ =\ \int_{t^n}^{t^{n+1}} \int_{\Omega_t} \psi(t)\ \operatorname{div} \boldsymbol{c}\ d\boldsymbol{x}\ dt \qquad \forall \hat{\psi} \in \hat{\mathcal{X}} \tag{3.24}$$

In section 2.1.4, we mentioned that the Euler expansion formula can be interpreted as an evolution law for the Jacobian determinant. Now we have shown that, when using the scheme presented in this chapter to solve a linear advection equation set on a moving domain, *a particular discrete counterpart of this evolution law must be verified for the advected quantity to be well-conserved.*

It may be interesting to refer to the analysis by Formaggia and Nobile, in [43], of a particular FEM numerical scheme for the resolution of the linear advection diffusion problem with divergence-free velocity raised on a moving domain. They precisely show that the expression of the GCL – see section 2.3 – for this scheme can be interpreted as a particular discretization of the Euler expansion formula. Note that the scheme they consider is derived from a conservative ALE formulation of the continuous problem, which implies that the unknow is automatically conserved. But we recall that, in the case of a problem raised on a moving domain, the conservative formulation leads

to schemes which do not necessarily preserve constant solutions whenever they are solutions of the continuous problem. The GCL is then defined as the condition for the scheme to satisfy this property, which we refer to as the constancy preservation.

We therefore conclude that, for some particular numerical schemes raised on moving domains, *the evolution law for the Jacobian determinant must be preserved at the discrete level* to allow the coexistence of both the conservation and the constancy preservation properties.

In the next section we describe the application of the ALE-MURD schemes proposed in this chapter to a particular framework : the linear advection problems set on three-dimensional domains moving in the vertical direction only. Especially, we will show how the conservation constraint of these schemes can be satisfied.

3.8 Conservation in the framework of the linear advection problem set on a three-dimensional domain moving in the vertical direction only

In this section we consider a problem in which *the domain is three-dimensional and only moves in the vertical direction.*

Note that, a part from very particular cases, the domain movement as well as the advection velocity are not known a priori ; for instance in a free surface problem or a fluid-structure interaction problem. In these cases, the linear advection problem is solved in the framework of a more global problem providing the position of the domain boundaries and the advection velocity.

The whole problem is formulated in the ALE approach, using a mapping of the form :

$$\hat{\mathcal{A}}_t \ : \ \hat{\Omega} \longrightarrow \Omega_t, \qquad \boldsymbol{x}(\hat{\boldsymbol{x}},t) = \hat{\mathcal{A}}_t(\hat{\boldsymbol{x}})$$
$$\text{with} \quad x = \hat{x}, \quad y = \hat{y} \quad \text{and} \quad z = Z\left(x,y,\hat{z},t\right), \tag{3.25}$$

where $Z\left(x,y,\hat{z},t\right)$ is an arbitrary continuous and monotonic ($\partial Z/\partial \hat{z} \geq 0$) function such that the velocity of the domain $\boldsymbol{c}$ satisfies

$$\boldsymbol{c}\cdot\boldsymbol{n} \ = \ \boldsymbol{\beta}\cdot\boldsymbol{n} \quad \text{on} \quad \Gamma_{imp,t} \qquad \text{and} \qquad \boldsymbol{c}\cdot\boldsymbol{n} \ = \ 0 \quad \text{on} \quad \Gamma_{liq,t},$$

where $\Gamma_{imp,t}$ denotes the impervious boundaries of the domain and $\Gamma_{liq,t}$ denotes its fictitious liquid boundaries – which we recall only move in the vertical direction. Figure 3.1 shows an example of such a mapping. Note that $\boldsymbol{c} = (0,0,c)^T$.

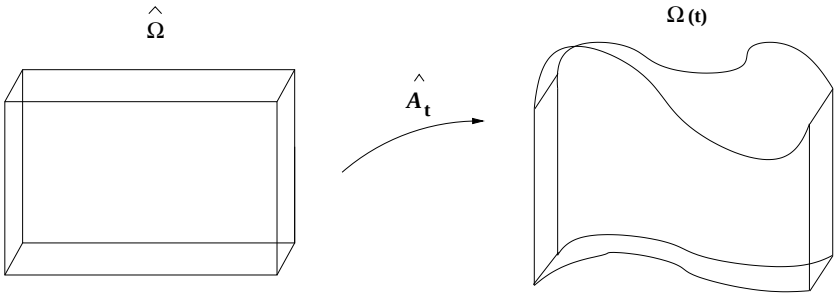


FIG. 3.1 – Example of a mapping describing a 3D domain moving in the vertical direction only.

3.8.1 Space and time discretization

The problem is semi-discretized in space using the finite element spaces described in sections 2.1.6 and 2.1.7. Then, the time interval I is divided in N_t time steps of equal length Δt and a particular time discretization is chosen. We point out that

the mesh velocity $\boldsymbol{c}_h$ is considered constant within each time step.

The motion of the mesh can thus be described by :

$$\boldsymbol{x} \;=\; \hat{\mathcal{A}}_{h,n+1}(\hat{\boldsymbol{x}}) \;=\; \hat{\mathcal{A}}_{h,n}(\hat{\boldsymbol{x}}) \;+\; \Delta t\; \hat{\boldsymbol{c}}_h^{n,n+1} \qquad \text{for} \quad \hat{\boldsymbol{x}} \in \hat{\Omega}, \tag{3.26}$$

for each $n = 1, \ldots, N_t - 1$.

The problem is solved as follows. If the domain and the advective velocity are known *a priori*, they are discretized. If not, a particular sub-problem is solved at each time step, updating the values of the advection velocity $\boldsymbol{\beta}_h$ and the domain Ω_h. Hence the domain velocity $\boldsymbol{c}_h$ during the considered time step can be computed and the ALE advection velocity $(\boldsymbol{\beta}_h - \boldsymbol{c}_h)$ is deduced. Finally, the linear advection problem can be solved on the new domain using the advection ALE-MURD schemes we have introduced in this chapter – see Definition 3.3.1.

We indicate next some essential characteristics of the global numerical scheme. On one hand, the velocity of the fluid $\boldsymbol{\beta}_h$ is computed – or discretized – such that, at each time $t^n \in I$,

$$\int_{\Omega_h^n} \psi_h \;\mathrm{div}\, \boldsymbol{\beta}_h^n \; d\boldsymbol{x} \;=\; 0 \tag{3.27}$$

for any function $\psi_h : \Omega_h^n \longrightarrow \mathbb{R}$ such that $\psi_h = \hat{\psi}_h \circ \hat{\mathcal{A}}_{h,n}$, where $\hat{\psi}_h$ is a function belonging to a particular discrete space $\hat{\mathcal{X}}_h$ defined on the reference

configuration. Note that $\hat{\mathcal{X}}_h$ is such that its degrees of freedom are located at the vertices of the mesh elements.

On the other hand, at each time t^n, the advected variable $u(t^n)$ is approximated by u_h^n in $\mathcal{X}_{h,0}^n$, where $\mathcal{X}_h^n$ is the discrete space corresponding to $\hat{\mathcal{X}}_h$ on the current domain Ω_h^n and defined by (3.3). The linear advection problem (2.42) is then discretized by (3.4), that is, given the initial condition u_0,

$$\text{for each} \quad n = 0, \ldots, N_t - 1 \quad \text{find} \quad u_h^{n+1} \in \mathcal{X}_{h,0}^{n+1} \quad \text{such that}$$

$$\begin{aligned} &\int_{\Omega_h^{n+1}} \left(u_h^{n+1} - u_h^n \circ \mathcal{A}_{n+1,n} \right) \, d\boldsymbol{x} \\ &\qquad + \quad \Delta t \int_{\Omega_h^c} (\boldsymbol{\beta}_h - \boldsymbol{c}_h)^c \cdot \boldsymbol{\nabla} \left(u_h^n \circ \mathcal{A}_{c,n} \right) \, d\boldsymbol{x} \quad = \quad 0, \end{aligned} \tag{3.28}$$

where t^c is a particular time within t^n and t^{n+1}. This discrete problem can thus be solved using an ALE-MURD sheme for the linear advection problem in non-conservative form.
Finally, we assume that the local flux terms (3.7) are computed exactly for the advection ALE-MURD scheme, so that relation (3.21) holds true.

Let us now prove that the advection ALE-MURD schemes introduced in this chapter are automatically conservative in the particular framework just described.

3.8.2 Discrete conservation of the advected quantity

Let us now explain how the advection ALE-MURD scheme can be made conservative in the particular framework just described. Let us recall that we wish the advected quantity to be conserved during the simulation in the sense of (3.19).

In the following, N_h will denote the dimension of $\hat{\mathcal{X}}_h$ and $\{\hat{\psi}_i\}_{i=1,\ldots,N_h}$ a set of basis functions. We recall that, at each time t^n, a set of basis functions $\{\psi_i^n\}_{i=1,\ldots,N_h}$ of $\mathcal{X}_h^n$ can be defined by $\psi_i^n = \hat{\psi}_i \circ \mathcal{A}_{h,n}^{-1}$ for each node i.

Let us state a first lemma.

Lemma 3.1 *For any $t^c \in [t^n, t^{n+1}]$, let α be the constant such that*

$$0 \leq \alpha \leq 1 \qquad \textit{and} \qquad t^c = \alpha\, t^{n+1} + (1-\alpha)\, t^n.$$

If the approximated domain Ω_h^c and the advection velocity $\boldsymbol{\beta}_h^c$ at $t = t^c$ are assumed to be respectively the linear combinations

$$\Omega_h^c = \alpha\, \Omega_h^n + (1-\alpha)\, \Omega_h^n \quad \textit{and} \quad \boldsymbol{\beta}_h^c = \alpha\, \boldsymbol{\beta}_h^{n+1} + (1-\alpha)\, \boldsymbol{\beta}_h^n, \tag{3.29}$$

then, for each node i, we have that

$$\int_{\Omega_h^c} \psi_i^c \ div\, \boldsymbol{\beta}_h^c \ d\boldsymbol{x} \ = \ 0. \tag{3.30}$$

Proof.

$$\begin{aligned}
&\int_{\Omega_h^c} \psi_i^c \ \mathrm{div}\, \boldsymbol{\beta}_h^c \ d\boldsymbol{x} \\
= \ & \alpha \int_{\Omega_h^{n+1}} (\psi_i^c \circ \mathcal{A}_{n+1,c}) \ \mathrm{div}\, \boldsymbol{\beta}_h^{n+1} \ d\boldsymbol{x} \ + \ (1-\alpha) \int_{\Omega_h^n} (\psi_i^c \circ \mathcal{A}_{n,c}) \ \mathrm{div}\, \boldsymbol{\beta}_h^n \ d\boldsymbol{x} \\
= \ & \alpha \int_{\Omega_h^{n+1}} \psi_i^{n+1} \ \mathrm{div}\, \boldsymbol{\beta}_h^{n+1} \ d\boldsymbol{x} \ + \ (1-\alpha) \int_{\Omega_h^n} \psi_i^n \ \mathrm{div}\, \boldsymbol{\beta}_h^n \ d\boldsymbol{x} \\
= \ & 0.
\end{aligned}$$

$\diamondsuit$

Following Proposition 3.9, if relation (3.30) is verified and if the local flux terms (3.7) are computed exactly, the conservation constraint of the advection scheme (3.28) is :

$$\forall\, n = 1, \ldots, N_t, \quad \forall\, i = 1, \ldots, N_h,$$

$$\int_{\Omega_h^{n+1}} \psi_i^{n+1} \ d\boldsymbol{x} \ - \ \int_{\Omega_h^n} \psi_i^n \ d\boldsymbol{x} \ - \ \Delta t \int_{\Omega_h^c} \psi_i^c \ \mathrm{div}\, \boldsymbol{c}_h^c \ d\boldsymbol{x} \ = \ 0. \tag{3.31}$$

This constraint is automatically satisfied by the advection scheme in the particular framework considered here. Let us state the following lemma :

Lemma 3.2 *If the approximated domain velocity within time step $[t^n, t^{n+1}]$ is constant and of the form $(0, 0, c_h^{n,n+1})^T$ then*

$$\int_{\Omega_h^{n+1}} \psi_i^{n+1} \ d\boldsymbol{x} \ - \ \int_{\Omega_h^{n+1}} \psi_i^n \ d\boldsymbol{x} \ = \ \Delta t \int_{\hat{\Omega}} \hat{\psi}_i \ \ div\, \hat{\boldsymbol{c}}_h^{n,n+1} \ d\hat{\boldsymbol{x}}\,. \tag{3.32}$$

Proof. Since the domain only moves in the vertical direction, the Jacobian determinant of the discrete mapping $\hat{\mathcal{A}}_{h,t}$ is

$$\hat{J}_{h,t} \ = \ \left| \frac{\partial z}{\partial \hat{z}} \right| \qquad \forall t \in I. \tag{3.33}$$

Therefore

$$\begin{aligned}
\int_{\Omega_h^{n+1}} \psi_i^{n+1} \ d\boldsymbol{x} \ - \ \int_{\Omega_h^{n+1}} \psi_i^n \ d\boldsymbol{x} \ &= \ \int_{\hat{\Omega}} (\hat{J}_h^{n+1} - \hat{J}_h^n) \ \hat{\psi}_i \ d\hat{\boldsymbol{x}} \\
&= \ \int_{\hat{\Omega}} \left(\left| \frac{\partial z}{\partial \hat{z}} \right|^{n+1} - \left| \frac{\partial z}{\partial \hat{z}} \right|^{n} \right) \ \hat{\psi}_i \ d\hat{\boldsymbol{x}}\,.
\end{aligned}$$

Using relation (3.26) we deduce that

$$\int_{\Omega_h^{n+1}} \psi_i^{n+1}\, d\boldsymbol{x} \;-\; \int_{\Omega_h^{n+1}} \psi_i^{n}\, d\boldsymbol{x} \;=\; \Delta t \int_{\hat\Omega} \frac{\partial\, \hat c_h^{n,n+1}}{\partial \hat z}\, \hat\psi_i\, d\hat{\boldsymbol{x}}$$
$$= \Delta t \int_{\hat\Omega} \hat\psi_i\ \mathrm{div}\, \hat{\boldsymbol{c}}_h^{n,n+1}\ d\hat{\boldsymbol{x}}\,.$$

$\diamondsuit$

Lemma 3.3 *If the approximated domain velocity within time step* $[t^n, t^{n+1}]$ *is constant and of the form* $(0,0,c_h^{n,n+1})^T$ *then*

$$\int_{\Omega_h^c} \psi_i^c\ \ div\, \boldsymbol{c}_h^c\ d\boldsymbol{x} \;=\; \int_{\hat\Omega} \hat\psi_i\ \ div\, \hat{\boldsymbol{c}}_h^{n,n+1}\ d\hat{\boldsymbol{x}} \tag{3.34}$$

for any time t^c *in* $[t^n, t^{n+1}]$.

Proof. Writing again that $t^c = \alpha\, t^{n+1} + (1-\alpha)\, t^n$ with $0 \le \alpha \le 1$, we obtain that

$$\int_{\Omega_h^c} \psi_i^c\ \mathrm{div}\, \boldsymbol{c}_h^c\ d\boldsymbol{x} \;=\; \alpha \int_{\Omega_h^{n+1}} \psi_i^{n+1}\ \frac{\partial\, \boldsymbol{c}_h^{n,n+1}}{\partial z}\ d\boldsymbol{x}$$
$$+\ (1-\alpha) \int_{\Omega_h^n} \psi_i^n\ \frac{\partial\, c_h^{n,n+1}}{\partial z}\ d\boldsymbol{x}\,.$$

Moreover, it is easy to establish that for any time t and any discrete function $f_h : \Omega_{h,t} \longrightarrow \mathbb{R}$,

$$\frac{\partial f_h}{\partial z} = \frac{\partial \hat f_h}{\partial \hat z}\ \frac{\partial \hat z}{\partial z} \quad \text{in } \Omega_{h,t}\,, \tag{3.35}$$

where $\hat f_h$ is the ALE counterpart of f_h defined by $\hat f_h(\hat{\boldsymbol{x}}) = f_h \circ \hat{\mathcal{A}}_{h,t}(\hat{\boldsymbol{x}})$ on $\hat\Omega$. Taking $f_h(x,y,z) = z$ we deduce :

$$1 \;=\; \frac{\partial z}{\partial \hat z}\ \frac{\partial \hat z}{\partial z} \quad \text{and thus} \quad \hat J_{h,t}\ \frac{\partial \hat z}{\partial z} \;=\; 1\,. \tag{3.36}$$

Using relations (3.35)-(3.36) we finally obtain :

$$\int_{\Omega_h^c} \psi_i^c\ \mathrm{div}\, \boldsymbol{c}_h^c\ d\boldsymbol{x} \;=\; \int_{\hat\Omega} \hat\psi_i\ \frac{\partial\, \hat c_h^{n,n+1}}{\partial \hat z}\ d\hat{\boldsymbol{x}} \;=\; \int_{\hat\Omega} \hat\psi_i\ \mathrm{div}\, \hat{\boldsymbol{c}}_h^{n,n+1}\ d\hat{\boldsymbol{x}}\ .$$

$\diamondsuit$

The following proposition can now be stated.

Proposition 3.10 *If the local fluxes (3.7) are computed exactly and the advection velocity is divergence-free in the sense of (3.30), and if the mesh moves in the vertical direction only and its velocity is considered constant within each time step, then the advection ALE-MURD scheme (3.28) is conservative in the sense of (3.19).*

The proof of this proposition becomes trivial by using Proposition 3.9 as well as lemma 3.2 and 3.3.

From this proposition we can conclude that the advection ALE-MURD scheme (3.28) is conservative in the particular framework presented in this section, provided the approximated domain Ω_h^c and the advection velocity $\boldsymbol{\beta}_h^c$ are defined by (3.29).

3.8.3 Numerical illustration

In this section we aim to evaluate the conservation property of the ALE-MURD schemes introduced in this chapter, when used in the particular framework just described. Especially, we want to prove numerically the validity of Proposition 3.10. For this purpose, we will make use of the PSI scheme implemented in the Telemac-3D system for the linear advection of a scalar in the framework of three-dimensional free surface flows. We recall that the implementation of this scheme was initially made by Janin [74] for a particular ALE mapping – corresponding to the classical sigma transformation we will describe in the next chapter. Using the generic formulation of the ALE-MURD schemes derived in section 3.3.3 for the linear advection problem in non-conservative form , we have extended the implementation of the scheme to a more general form of ALE mappings.

Before introducing the test case, let us describe some characteristics of the Telemac-3D system which must be considered when using the advection schemes. This system solves the Navier-Stokes equations describing the motion of a free-surface flow in a three-dimensional domain moving in the vertical direction only. At each time step, the hydrodynamic problem is solved, providing the velocities of the fluid and the free-surface profile. Then, the equations describing the motion of the tracers contained in the fluid are solved in the newly defined domain. The ALE-MURD sheme can be used for the advection of the fluid velocity components as well as for the tracer advection.

However, Proposition 3.10, stating the conservation under certain assumptions of the ALE-MURD schemes set on a three-dimensional domain moving in the vertical direction only, cannot be applied in the particular framework of the standard Telemac-3D system. Indeed, the assumption that the advection velocity is divergence free in the sense of (3.30) is not verified.

In fact, a particular technique is used, consisting in computing the vertical component of the ALE advection velocity $(\boldsymbol{\beta}_h - \boldsymbol{c}_h)$ so that the conservation constraint of the scheme is verified. Note that this technique will be described in the second part of this work, devoted more particularly to the three-dimensional hydrostatic free surface flows – see section in particular section 7.5. Note also that, for the test presented in this section, we have supressed the module applying this technique in the Telemac-3D system.

In order to get rid of this difficulty and satisfy the assumptions of Proposition 3.10, we won't use the entire solution of the hydrodynamic problem provided by the Telemac-3D system. We will rather choose a test case whose analytical solution is known, and we will impose the analytical free-surface profile and fluid velocity at each time. Only the scalar advection module of the system will then be used to simulate the advection of a passive tracer – *i.e.* a tracer which has no influence on the hydrodynamics of the fluid.

We have chosen the test case of the oscillating wave in a closed basin. A sloped free surface profile and zero velocities are imposed at the initial time, and the fluid is assumed to be not viscous : the total energy is therefore preserved and a continuous exchange of potential and kinetic energy takes place. Note that when the free surface is horizontal, the whole potential energy has been exchanged into kinetic one. On the contrary, when the highest elevation is reached at one of the basin boundaries, the maximum potential and minimum kinetic energy state is achieved, just as at the beginning of the simulation.

A closed square basin with side L and with a constant bottom at $z = -H$ is considered, where H is the equilibrium water depth. Since there is no active tracer the fluid density is constant ; it will be denoted by ρ_0. We assign the following initial conditions, represented in figure 3.2 :

$$\begin{cases} \eta(x,y) = \eta_0 \cos(kx), \\ u = v = w = 0, \\ p_d(x,y,z) = -\rho_0 g \eta + \rho_0 g \eta_0 \dfrac{\cosh(k(z+H))}{\cosh(kH)} \cos(kx), \end{cases} \tag{3.37}$$

where p_d denotes the hydrodynamic pressure correction and where

$$\eta_0 = 0.1\,\mathrm{m}, k = \frac{\Pi}{L},\ L = 10\,\mathrm{m} \text{ and } H = 10\,\mathrm{m}.$$

Note that since the initial free surface slope is parallel to the x-axis, no motion takes place in the y-axis direction. The following analytical solution is provided by the small amplitude wave theory for the linear equations – see

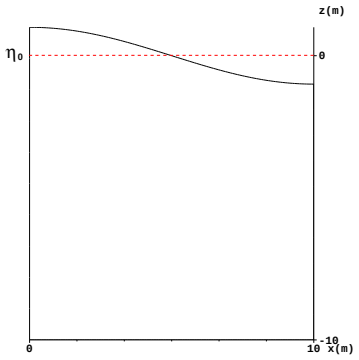


FIG. 3.2 – Initial test setting for the standing wave in a closed basin.

for instance [75] :

$$\begin{cases} \eta &= \eta_0 \, \cos(kx) \, \cos(\omega t) \,, \\ u &= \omega \, \eta_0 \, \dfrac{\cosh(k(z+H))}{\sinh(kH)} \, \sin(kx) \, \sin(\omega t) \,, \\ v &= 0 \,, \\ w &= -\omega \, \eta_0 \, \dfrac{\sinh(k(z+H))}{\sinh(kH)} \, \cos(kx) \, \sin(\omega t) \,, \end{cases} \tag{3.38}$$

with $\omega^2 = g\,k\tanh(kH)$. The analytical water height profile at $x = 0$ m is represented in Figure 3.3, and the analytical velocity at times $t = 14\,\mathrm{s}$ and $t = 30\,\mathrm{s}$ in figures 3.4 and 3.5 respectively.

A passive tracer T is advected by the flow just described. Its initial distribution in the domain is the following : T is equal to 25000 everywhere except inside a ball of radius 2 m whose center is located at $\boldsymbol{x}_0 = (5, 5, -5)$, where T is equal to 50000.

The simulation will be performed with a time step length of 0.1 seconds for a time interval of 30 seconds, and with 11 levels along the vertical. The initial three-dimensional mesh is shown in Figure 3.6. It has been obtain by piling up the two-dimensional grid presented in Figure 3.7 along the vertical in the discrete domain. For more visibility, the vertical cross-section of the three-dimensional mesh at $y = 5$ m is shown in Figure 3.8.

Let us now describe how we will apply the ALE-MURD advection scheme of Telemac-3D to the advection of T in this test. At each time step, the discrete advection velocity $\boldsymbol{\beta}_h$ is determined by interpolating the analytical fluid velocity on the three-dimensional domain. The mesh velocity $\boldsymbol{c}_h$ is then

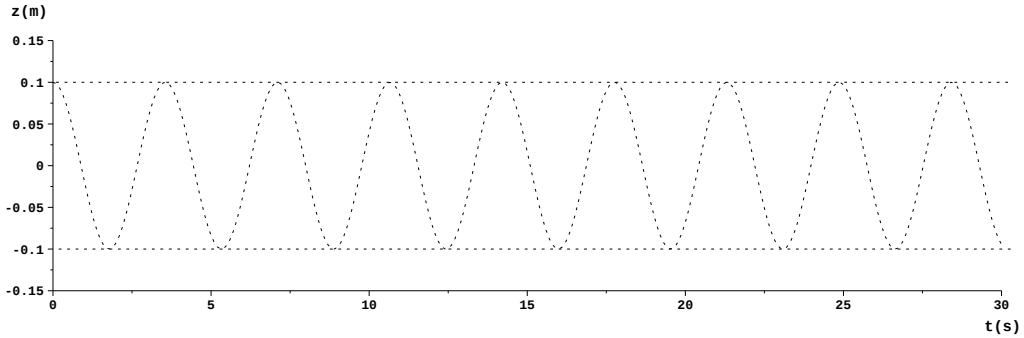


FIG. 3.3 – Temporal profile of the exact water height at $x = 0$ m, $y = 5$ m.

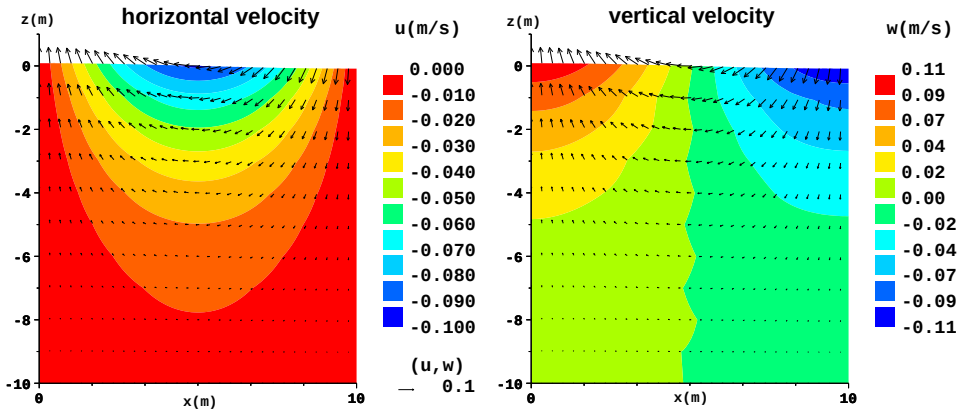


FIG. 3.4 – Analytical velocities at t=14s. Vertical cross-section at y=5m.

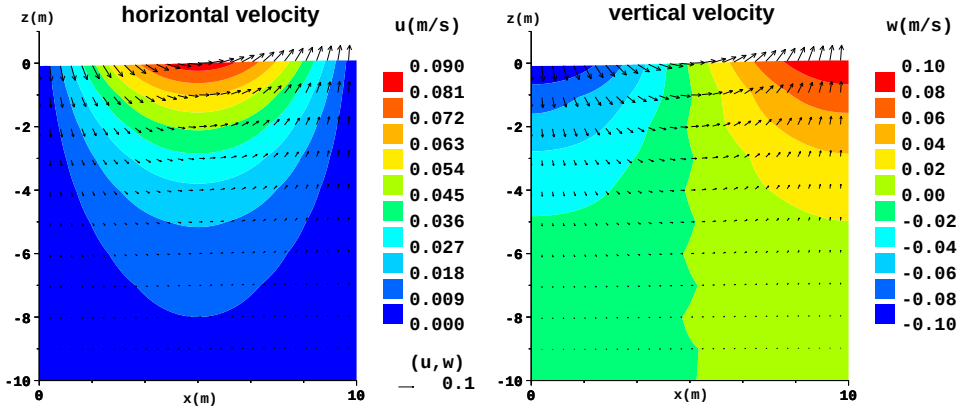


FIG. 3.5 – Analytical velocities at t=30s. Vertical cross-section at y=5m.

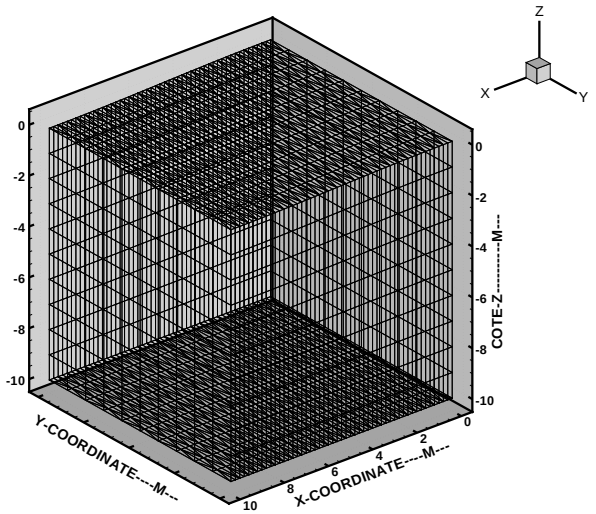


FIG. 3.6 – The initial 3D mesh.

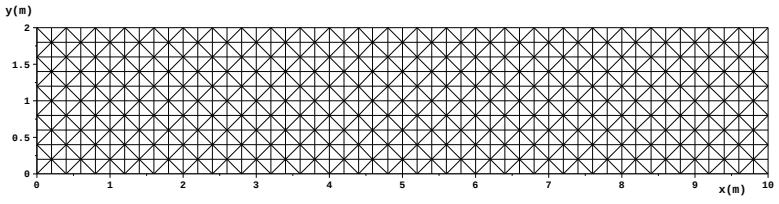


FIG. 3.7 – The horizontal 2D mesh.

evaluated, so that the ALE advection velocity $(\boldsymbol{\beta}_h - \boldsymbol{c}_h)$ can be determined. Finally, the local fluxes (3.7) are computed exactly. We recall that the discrete problem solved by the scheme is (3.28). Here we choose $t^c = t^{n+1}$: the problem therefore writes, given the initial condition T_0,

$$\text{for each} \quad n = 0, \ldots, N_t - 1 \quad \text{find} \quad T_h^{n+1} \in \mathcal{X}_h^{n+1} \quad \text{such that}$$

$$\begin{aligned} &\int_{\Omega_h^{n+1}} \left(T_h^{n+1} - T_h^n \circ \mathcal{A}_{n+1,n}\right) \, d\boldsymbol{x} \\ &\quad + \quad \Delta t \int_{\Omega_h^{n+1}} \left(\boldsymbol{\beta}_h^{n+1} - \boldsymbol{c}_h^{n,n+1}\right) \cdot \boldsymbol{\nabla} \left(T_h^n \circ \mathcal{A}_{n,n+1}\right) \, d\boldsymbol{x} \;=\; 0\,. \end{aligned} \tag{3.39}$$

Since the domain considered in this test case has no liquid boundaries, the flux through the boundaries is nil and therefore the loss of tracer quantity

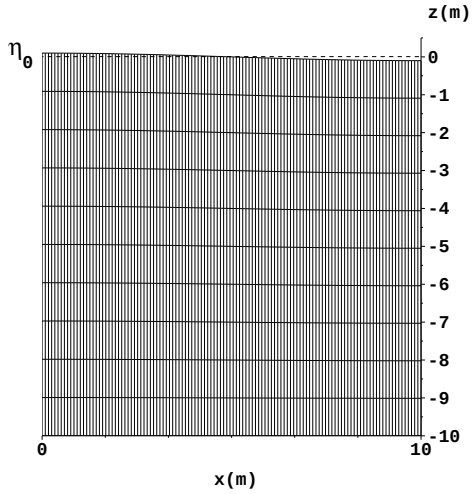


FIG. 3.8 – Vertical cross-section of the 3D mesh at $y = 5$ m.

at each time step $[t^n, t^{n+1}]$ – denoted by $\Delta M_T^{n,n+1}$ – should be zero :

$$\Delta M_T^{n,n+1} \;=\; \int_{\Omega_h^{n+1}} T_h^{n+1}\, d\boldsymbol{x} \;-\; \int_{\Omega_h^n} T_h^n\, d\boldsymbol{x} \;= 0\,. \tag{3.40}$$

Following Proposition 3.9, since the local fluxes (3.7) are computed exactly, if the fluid velocity is divergence free in the sense of (3.30) then the scheme is conservative in the sense of (3.40) if relation (3.31) holds true. In addition, following Proposition 3.10, this relation is verified if the mesh velocity is considered constant at each time step.

We would like to evaluate numerically the validity of this statement. However, the interpolation $\boldsymbol{\beta}_h$ of the analytical fluid velocity on the three-dimensional mesh may not be locally divergence free. Especially, relation (3.30) may not be satisfied exactly. This interpolation error deteriorates the conservation property of the scheme and must therefore be taken into account when evaluating the error on the conservation constraint. We will refer to this error as the *divergence error*, and denote its value at each time step $[t^n, t^{n+1}]$ by :

$$\epsilon_{Div}^{n,n+1} \;=\; \Delta t \int_{\Omega_h^{n+1}} T_h^n \;\mathrm{div}\, \boldsymbol{\beta}_h^{n+1}\, d\boldsymbol{x}\,. \tag{3.41}$$

We will also evaluate the error on relation (3.31), which will be referred

to as the *mesh velocity error* and denoted at each time step $[t^n, t^{n+1}]$ by :

$$\begin{aligned} \epsilon_{Mesh}^{n,n+1} &= \int_{\Omega_h^{n+1}} T_{h,n+1}^n \, d\boldsymbol{x} - \int_{\Omega_h^{n+1}} T_h^n \, d\boldsymbol{x} \\ &- \Delta t \int_{\Omega_h^{n+1}} T_{h,n+1}^n \, \mathrm{div}\, \boldsymbol{c}_h^{n,n+1} \, d\boldsymbol{x} \, . \end{aligned} \tag{3.42}$$

For the reasons exposed previously, the conservation constraint of the sheme – *i.e.* the condition for the scheme to satisfy the required discrete conservation property (3.40) – is that relations (3.30) and (3.31) be satisfied. Therefore, each time step $[t^n, t^{n+1}]$, the error on the conservation constraint will be evaluated as the addition of the divergence error and the mesh velocity error :

$$\epsilon_{Cons}^{n,n+1} = \epsilon_{Div}^{n+1} + \epsilon_{Mesh}^{n+1} \, . \tag{3.43}$$

Note that *the error on the conservation constraint increases with the length of the time step as well as with the value of the advected quantity.*

After each time step of the simulation, we will compute the different errors and make sure that $\Delta M_T^{n+1} = \epsilon_{Cons}^{n+1}$. The test will be performed using two different approximations of the mesh velocity. First, the mesh velocity is considered constant within each time step. It is computed as follows :

$$\boldsymbol{c}_h^{n,n+1} = (0, 0, c_h^{n,n+1}) \quad \text{with} \quad c_h^{n,n+1} = \frac{z^{n+1} - z^n}{\Delta t} \, . \tag{3.44}$$

Then, a second order approximation of the mesh velocity will be used, that is

$$\boldsymbol{c}_h^{n,n+1} = (0, 0, c_h^{n,n+1}) \quad \text{with} \quad c_h^{n,n+1} = \frac{4\, z^{n+1} - 3\, z^n + z^{n+1}}{2\, \Delta t} \, . \tag{3.45}$$

The evolution of the tracer distribution is shown in Figure 3.9. The profile is the same for both simulations.

However, when looking more carefully at the conservation of the tracer, we observe a relevant difference between these simulations. Figure 3.10 shows the values of the loss of T and the different errors at each time step, for the case where the mesh velocity is considered constant within each time step. We can see that *the mesh velocity error* ϵ_{Mesh} *is nil* during the entire simulation, so that the loss of tracer quantity is only due to the interpolation error

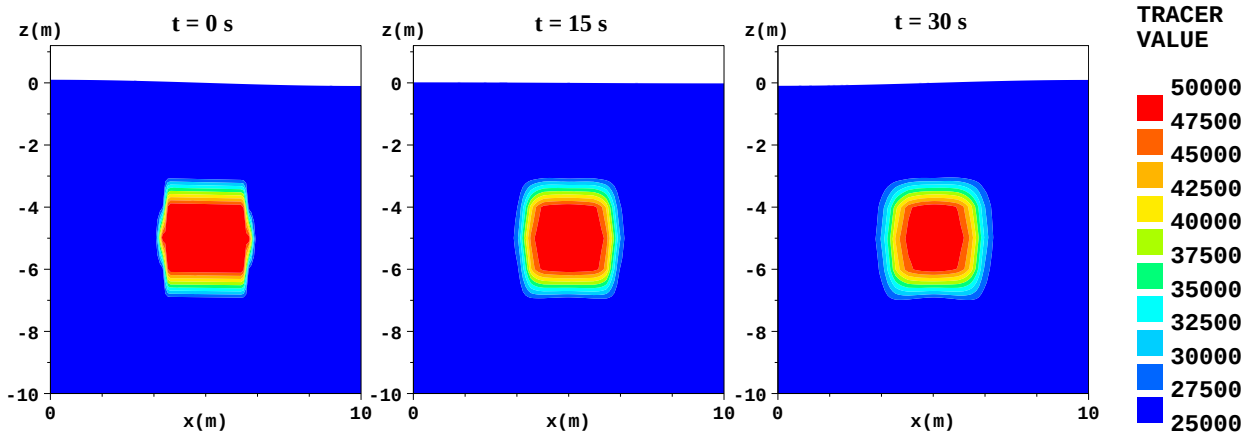


FIG. 3.9 – Tracer profile at $t = 0$ s, $t = 15$ s and $t = 30$ s. Vertical cross-section at y=5m.

ϵ_{Div} on the fluid velocity. Indeed, we observe that, at each time step $[t^n, t^{n+1}]$,

$$\Delta M_T^{n,n+1} \;=\; \epsilon_{Cons}^{n,n+1} \;=\; \epsilon_{Div}^{n,n+1}\,.$$

On the contrary, when using a second order approximation of the mesh velocity, *the mesh velocity error* ϵ_{Mesh} is not nil. Indeed, Figure 3.11 shows that ϵ_{Mesh} oscillates during the simulation between 0 and around 0.8.
In this case we observe that

$$\Delta M_T^{n,n+1} \;=\; \epsilon_{Cons}^{n,n+1} \;=\; \epsilon_{Mesh}^{n,n+1} \;+\; \epsilon_{Div}^{n,n+1}$$

at each time step $[t^n, t^{n+1}]$. Note that the loss of tracer quantity is then essentially due to the *mesh velocity error* ϵ_{Mesh} – the *interpolation error* ϵ_{Div} on the fluid velocity being ten times smaller.

These differences are better shown in Figure 3.12, where the loss of T, the *divergence error* and the *mesh velocity error* of the scheme for each simulation are compared. In addition, Figure 3.13 shows the evolution of the relative loss of tracer quantity – *i.e.*

$$(\Delta M_T / M_T^0)^n \;=\; \frac{\displaystyle\int_{\Omega_h^n} T_h^n \; d\boldsymbol{x} \;-\; \int_{\Omega_{h,0}} T_{h,0} \; d\boldsymbol{x}}{\displaystyle\int_{\Omega_{h,0}} T_{h,0} \; d\boldsymbol{x}}$$

– during 1000 seconds of simulation. We observe that using a second order approximation of the mesh velocity, the relative loss of T is 10 times larger

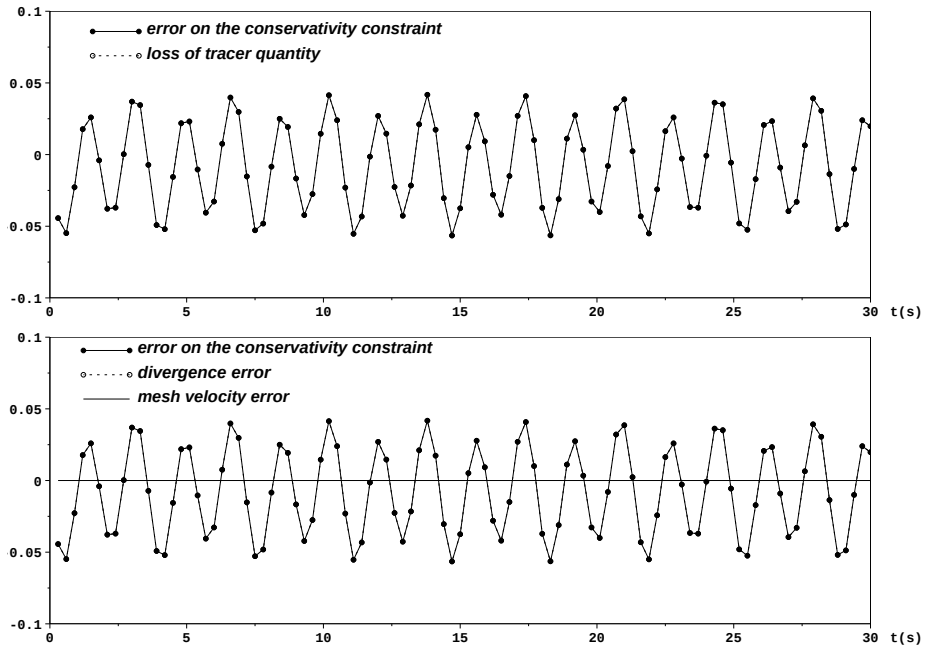


FIG. 3.10 – Values for a simulation using a constant mesh velocity. Top figure : at each time step, the loss of tracer quantity is equal to the *error on the conservativity constraint*. Bottom figure : the *error on the conservativity constraint* (solid line with plane circles) vs. the *divergence error* (dotted line with circles) and the *mesh velocity error* at each time step (solid line).

than considering a constant mesh velocity.

We must emphasize that the relative error is very low. Nevertheless, it may be significant in long time simulations, or with larger surface movements. Our purpose here was only to show that these results confirm the statements made in this chapter, and especially Proposition 3.10.

In order to show that the *mesh velocity error* increases with the length of the time step as well as with the amplitude of the mesh movement, we have performed two more simulations. Figure 3.14 shows the evolution of this error when a time step length of 1 second is used, that is ten times larger than in the previous simulations. We observe that the mesh velocity error reaches a maximum of around 50 when using a second order approximation.

In Figure 3.15 the same values are compared for a simulation in which the free surface oscillation has been amplified taking :

$$\eta_0 = 0.2 \quad \text{and} \quad k = \frac{2\,\Pi}{L}\,.$$

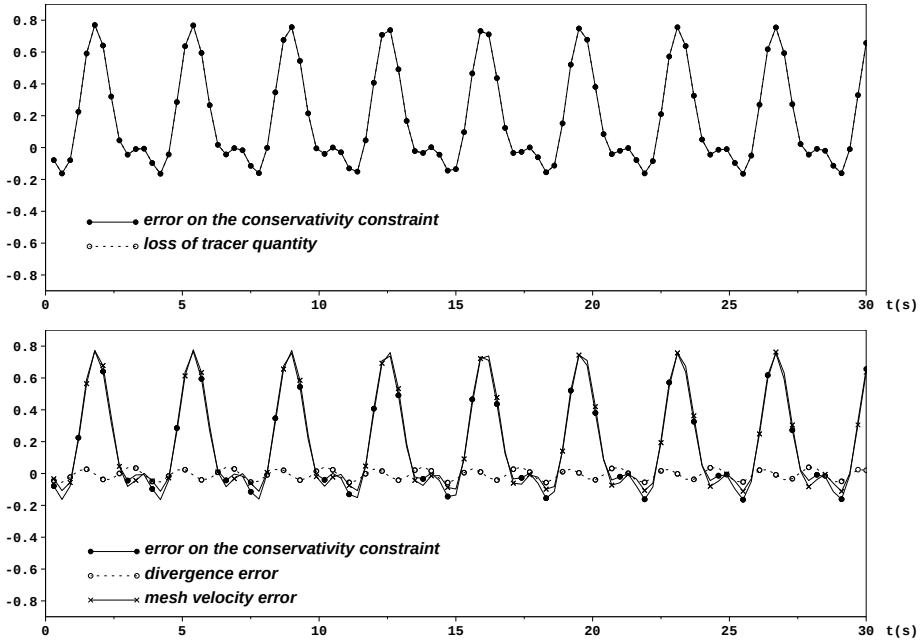


FIG. 3.11 – Values for a simulation using a second order approximation of the mesh velocity. Top figure : at each time step, the loss of tracer quantity is equal to the *error on the conservativity constraint*. Bottom figure : the *error on the conservativity constraint* (solid line with plane circles) vs. the *divergence error* (dotted line with circles) and the *mesh velocity error* at each time step (solid line).

In this case the mesh velocity error reaches a maximum of around 200 when using a second order approximation.

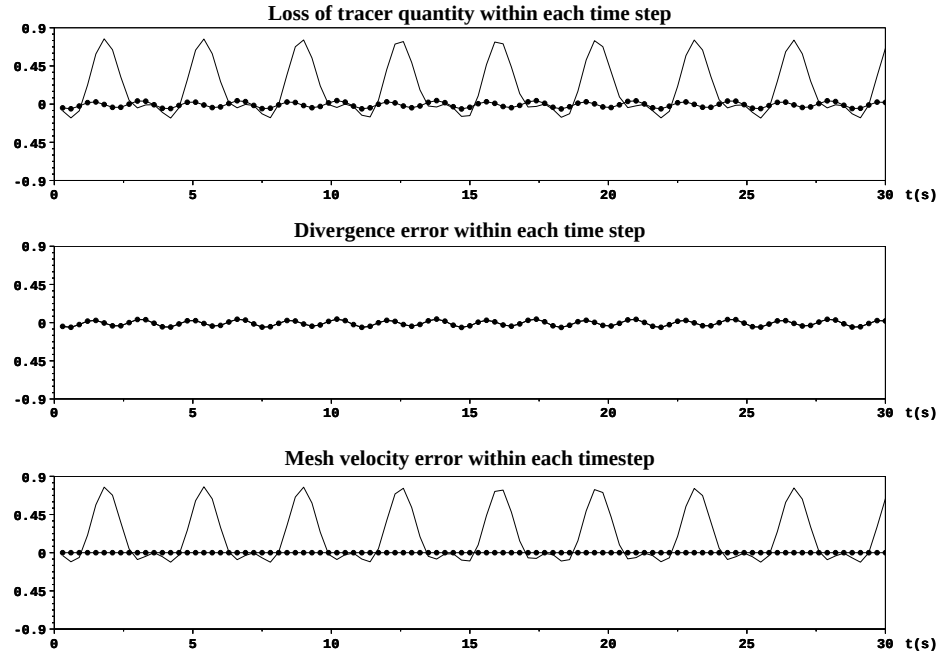


FIG. 3.12 – Comparison of the tracer quantity lost at each time step and the values of the different errors using a constant mesh velocity (solid line with plain circles) vs. a second order approximation of the mesh velocity (solid line) (with $\Delta t = 0.1$ s and $k = \Pi/L$).

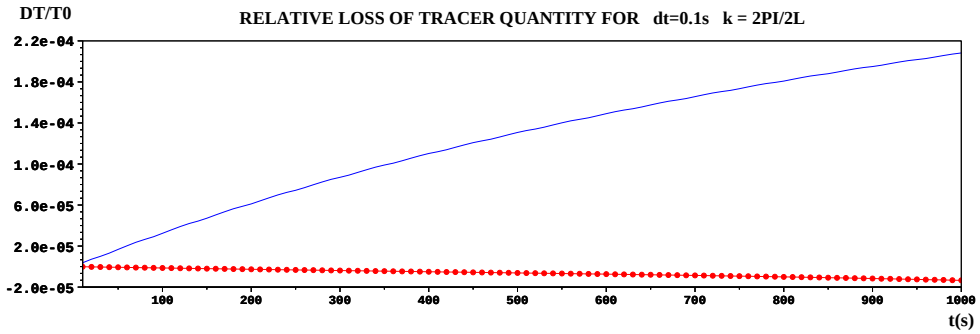


FIG. 3.13 – Relative loss of tracer quantity after 1000 seconds using a constant mesh velocity (solid line with plain circles) or a second order approximation of the mesh velocity (solid line) (with $\Delta t = 0.1$ s and $k = \Pi/L$).

FIG. 3.14 – Comparison of the value of the *mesh velocity error* at each time step using a constant mesh velocity (solid line with plain circles) vs. a second order approximation of the mesh velocity (solid line), with a time step length of $\Delta t = 1$ second (and with $k = \Pi/L$).

FIG. 3.15 – Comparison of the value of the *mesh velocity error* at each time step using a constant mesh velocity (line with plain circles) vs. a second order approximation of the mesh velocity (solid line), with $k = 2\Pi/L$ (and $\Delta t = 1$ s).

Chapitre 4

ALE interpretation of the sigma transformation for the three-dimensional free surface flow problem and generalization

4.1 Introduction

The three-dimensional free surface flow problem has been widely studied and many modelling systems have been developed since the early seventies, motivated by the study of different phenomena, from ocean and coastal currents to flows in estuaries, rivers, and lakes. In this chapter we focus on a particular way of treating the domain movement, which is one of the critical points of this problem.

Several techniques exist to deal with the moving surface of a three dimensional domain. Among them, the ALE approach has been adopted by many authors because of its great adaptability – see for instance Huerta and Liu [70], Maury [89], Soulaïmani *et al.* [129] and Cairncross *et al.* [23]. But another technique has been widely used, especially in the atmospheric and oceanographic communities, referred to as the sigma transformation or topography-following coordinate system. It consists in performing a transformation of the vertical coordinate – the z-coordinate – allowing to adopt a vertical discretization of the domain which at each time step follows the bathymetry and the free surface. Different systems have been developed on this approach, characterized by the choice of the discrete transformation function.

The link between this technique and the ALE approach is still not clear.

In the works describing the sigma transformation and its different variants, no mention is made of the ALE approach. Moreover, in [144] Zhou and Standsby present an ALE formulation based on a particular sigma transformation – the classical sigma transformation described further – for a shallow water flow problem. They claim that the conventional sigma models are based on flow equations without accounting for the effect of moving grids in the vertical direction. Nevertheless, we state the contrary and our purpose is to show that the sigma transformation is equivalent to the ALE approach for a particular type of ALE mappings. Therefore, we introduce the ALE-Sigma (ALES) approach, which consists in an ALE interpretation of the sigma transformation.

The sigma transformation technique presents many advantages ; essentially, the accurate assimilation of the bed and surface boundaries and the possibility of easily incorporating boundary layers. But the classical variants of this technique – which are still widely used – present important drawbacks. On one hand, they limit the possibilities of vertical discretization and therefore prevent from correctly adapting the mesh to the particular needs of the simulation. Especially, in some cases, they prevent from aplying a uniform vertical resolution to the boundary layers. On the other hand, they induce important errors in the computation of the horizontal pressure gradients in the presence of steep bottom, leading to an inaccurate representation of possible density-stratifications. Taking advantage of our reinterpretation of the sigma transformation as a particular ALE formulation, we propose a very general z-coordinate transformation in the ALES approach. This new point of view allows for a great adaptability of the vertical discretization and therefore overcomes some drawbacks of the classical sigma transformation.

After describing more precisely the sigma transformation technique, we will define the equivalent ALES approach. A new formulation of the problem using this approach will be proposed. Then, the source of the horizontal pressure gradient error will be explained and the different systems developed to overcome this problem presented. Finally, we will propose a generalization of the sigma transformation formulated through the ALES approach.

4.2 The sigma transformation approach

Let us first point out that this technique can only be applied to two- or three-dimensional domains in which *the lateral boundaries are perfectly vertical* – in the sequel we will only consider three-dimensional domains. This configuration is widely used for ocean and coastal applications, but also in other shallow areas. In this case, only the surface and bottom boundaries move, and they are described by two functions defined on the same two-dimensional domain. It is then possible to perform a variable change which

transforms the vertical coordinate of each point in the real domain into a value contained in a fixed interval – for instance $[0,1]$ – depending on the time and on the position of this point in the two-dimensional domain. In fact, this comes to transform the irregular physical domain into a fixed regular reference configuration.

Before going further, let us describe more explicitly the considered configuration. Let ω be the invariant two-dimensional domain, whereas Ω_t denotes the three-dimensional domain at each time $t \in I$. The functions describing the free surface and the bottom of Ω_t are, respectively, $\eta(x,y,t)$ and $b(x,y,t)$, whilst $h\,(x,y,t) = \eta\,(x,y,t) - b\,(x,y,t)$ denotes the water height. The current domain is then defined by

$$\Omega_t = \{\ \boldsymbol{x} = (x,y,z)\ \ /\ \ (x,y) \in \omega \quad \text{and} \quad b\,(x,y) \leq z \leq \eta\,(x,y,t)\,\}. \tag{4.1}$$

The sigma transformation approach consists in transforming the Cartesian coordinate system (x,y,z,t) into a topography-following coordinate system $(\hat{x},\hat{y},\hat{z},\hat{t})$, called σ-coordinate system and defined by

$$\hat{t} = t, \quad \hat{x} = x, \quad \hat{y} = y \quad \text{and} \quad \hat{z} = \sigma\,(x,y,z,t)\,, \tag{4.2}$$

where $\sigma\,(x,y,\hat{z},t)$ is an arbitrary continuous and monotonic $(\partial\sigma/\partial z \geq 0)$ function satisfying

$$\left\{ \begin{array}{rcl} \sigma(x,\ y,\ \eta(x,y,t),\ t) & = & 1 \\ \sigma(x,\ y,\ b(x,y,t),\ t) & = & 0 \end{array} \right. \qquad \forall\,(x,y,t) \in \omega \times I. \tag{4.3}$$

The reference configuration is therefore :

$$\hat{\Omega} = \{\ \hat{\boldsymbol{x}} = (\hat{x},\hat{y},\hat{z})\ \ /\ \ (\hat{x},\hat{y}) \in \omega \quad \text{and} \quad 0 \leq \hat{z} \leq 1\,\}. \tag{4.4}$$

An example of such a transformation is shown in Figure 4.1.

A particular transformation function $\sigma\,(x,y,\hat{z},t)$ is chosen. Then, a triangulation of the invariant two-dimensional domain ω is defined and $\sigma\,(x,y,\hat{z},t)$ is discretized. We indicate that in some cases the transformation function is directly defined at the discrete level. The three-dimensional mesh $\hat{\mathcal{T}}_h$ is defined on the reference configuration $\hat{\Omega}$ by arbitrarily distributing the two-dimensional mesh as horizontal levels along the vertical. The corresponding mesh on the real domain is then defined at each time step through the discrete transformation function $\sigma_h\,(x,y,\hat{z},t)$. Thus, *the real mesh only moves in the vertical direction*. For the sake of clarity, we indicate that, if i denotes a particular node and t^n denotes a particular time, then the coordinates of i in the real domain at $t = t^n$ are

$$x_i^n = \hat{x}_i, \quad y_i^n = \hat{y}_i \quad \text{and} \quad z_i^n = \sigma_h(x_i,y_i,\hat{z}_i,t^n), \tag{4.5}$$

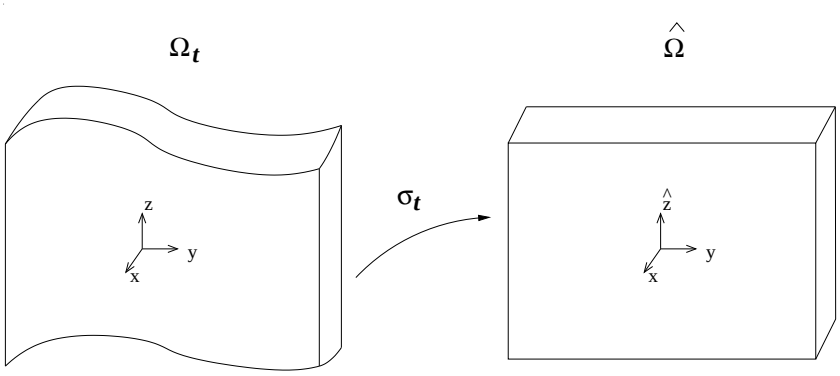


FIG. 4.1 – Example of a sigma transformation

where $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ are the constant coordinates of i in the reference configuration.

The equations describing the flow are then completely written in the reference domain ; more precisely all the partial derivatives are transformed into derivatives with respect to the σ-coordinates $(x, y, \hat{z})$.

The adoption of a σ-coordinate system allows an accurate and smooth definition of the bed and surface boundaries at each time. Indeed, the values of $\hat{z}$ at the bottom and the free surface are constant in time and space ; that is respectively $\hat{z} = 0$ and $\hat{z} = 1$. In addition, one can easily incorporate bottom and surface boundary layers, and increase their resolution without having to increase the global number of layers. This is not possible in the commonly used z-coordinate system – see for instance [21] –, in which the bottom profile is approximated by a staircase form. Such a step-discretization provoques a loss of accuracy in the assimilation of the bathymetry, as well as spurious errors. Many comparisons have been made between topography-following systems and the z-coordinate system (see for instance [54] and [93]), and these have shown that inferior numerical results are obtained with the second one, especially in the simulation of overflow processes and bottom layer dynamics.

Another essential advantage of the sigma transformation approach is the fact that it simplifies the kinematic boundary condition considerably both at the bottom and the free surface.

For all these reasons, it has been gaining popularity and is used for a very large range of ocean, coastal and shallow areas applications. Indeed, numerous free surface ocean circulation models have been developed on this approach, including the Princeton Ocean Model (POM) [19], the S-coordinate Rutgers University Model (SCRUM) [127], and the more recent Regional Ocean Model System (ROMS) [123]. Originally used mostly for re-

gional coastal simulations, they are now also applied to basin-scale problems [66, 144, 77, 141]. The finite element based Telemac-3D system [66] also uses the sigma transformation. It is mostly applied to coastal, estuarine and river flows, including transport and non-hydrostatic phenomena.

4.3 The sigma transformation : a particular ALE formulation

Our purpose here is to show that the application of the sigma transformation to the free surface flow problem can in fact be interpreted as the adoption of a particular ALE mapping and a special, unusual ALE formulation.

On one hand, the vertical-coordinate transformation (4.2)-(4.3) can be performed using the following ALE mapping between the reference configuration $\hat{\Omega}$ and the real domain Ω_t :

$$\hat{\mathcal{A}}_t \ : \ \hat{\Omega} \longrightarrow \Omega_t, \qquad \boldsymbol{x}(\hat{\boldsymbol{x}}, t) = \hat{\mathcal{A}}_t(\hat{\boldsymbol{x}})$$

$$\text{with} \quad x = \hat{x}, \quad y = \hat{y} \quad \text{and} \quad z = Z\left(x, y, \hat{z}, t\right), \tag{4.6}$$

where $Z\left(x, y, \hat{z}, t\right)$ is an arbitrary continuous and monotonic ($\partial Z / \partial \hat{z} \geq 0$) function satisfying

$$\begin{cases} Z(x, y, 0, t) = \eta(x, y, t), & \forall\, (x, y, t) \in \omega \times I, \\ Z(x, y, 1, t) = b(x, y, t). \end{cases} \tag{4.7}$$

Actually, this defines the ALE mappings for which the domain only moves in the vertical direction. Note that consequently the domain velocity has only a vertical non zero component, that is $\boldsymbol{c} = (0, 0, c)^T$.

On the other hand, as already mentioned, the sigma transformation approach states that the equations be completely transformed into the σ-coordinate system. Interpreting the sigma transformation as an ALE mapping makes clear that the transformation of the equations consists in a reformulation of the whole derivatives, including the spatial partial derivatives, into the ALE frame. Especially, the time derivatives are reformulated into ALE time derivatives – see section 2.1.3. Thus, the effects of the moving domain are taken into account.

Throughout, we will refer to the ALE interpretation of the sigma transformation as the ALE-Sigma (ALES) approach. We propose next a particular formulation of the free surface flow problem using this approach.

4.4 ALES formulation of the three-dimensional free surface flow problem

We consider the so-called Reynolds-averaged Navier-Stokes equations [104, 133, 118] describing the mean motion of an incompressible Newtonian fluid in a three-dimensional moving domain Ω_t of the type defined by (4.1). Let us indicate that most of the models using topography-following coordinates apply the hydrostatic approximation to the equations, because they deal with oceanic or atmospheric flows for which the shallow water condition is fulfilled. Never-theless, in order to be as general as possible, we will here stick to the non-hydrostatic equations.

4.4.1 Formulation in the Cartesian frame

In the Cartesian frame, these equations are the following, valid for all $t \in I$ and for all $\boldsymbol{x} \in \Omega_t$:

$$\frac{d\boldsymbol{u}}{dt} - \text{div}\,(\boldsymbol{\nu}\boldsymbol{\nabla}\boldsymbol{u}) + \frac{1}{\rho}\,\boldsymbol{\nabla}_2 p = \boldsymbol{f} \tag{4.8}$$

$$\frac{dw}{dt} - \text{div}\,(\boldsymbol{\nu}\boldsymbol{\nabla} w) + \frac{1}{\rho}\,\frac{\partial p}{\partial z} = -g \tag{4.9}$$

$$\text{div}\;\boldsymbol{U} = 0 \tag{4.10}$$

where $\boldsymbol{u}$ and w are respectively the horizontal and vertical velocities, $\boldsymbol{U} = (\boldsymbol{u}, w)^T$ is the total velocity, and $\boldsymbol{\nu} = (\nu_h, \nu_h, \nu_v)^T$ is the turbulent kinematic viscosity coefficient of the fluid. The horizontal source term is denoted by $\boldsymbol{f}$ and the gravity acceleration by g.

This system is closed by convenient boundary conditions. In particular, the kinematic condition on the bottom of the domain Γ_b expressing impermeability is

$$\boldsymbol{U}\cdot\boldsymbol{n} = 0 \qquad \text{on} \quad \Gamma_b \cup \Gamma_b\,, \tag{4.11}$$

where $\boldsymbol{n}$ is the outward normal to the current domain Ω_t. At the free surface $\Gamma_{s,t}$, the velocity of the fluid is equal to the velocity of the free surface itself, which is expressed by the following kinematic condition :

$$\frac{\partial \eta}{\partial t} - \boldsymbol{U}\cdot\boldsymbol{n} = 0 \quad \text{on} \quad \Gamma_{s,t}\,. \tag{4.12}$$

A complete set of boundary conditions closing the system in a very general case is given in the second part of this work, devoted more particularly to the three-dimensional free surface flow problem (see section 5.4).

The free surface hydrodynamic problem also involves the motion of the tracers, which represent a temperature or any physical quantity contained

in the water. The non-conservative form of the tracer equation is

$$\frac{\partial T}{\partial t} \; + \; \boldsymbol{U} \cdot \boldsymbol{\nabla} T \; - \; \operatorname{div}\left(\nu_T \boldsymbol{\nabla} T\right) \; = \; f_T \qquad \text{on} \quad \Omega_t \; , \tag{4.13}$$

where f_T is the source term and ν_T the tracer's constant diffusion coefficient.

4.4.2 ALES formulation

The generic form (4.6) of an ALE mapping leading to a σ-coordinate system is considered. We intend to write a particular formulation of the problem in the ALE frame which differs as well from the usual ALE formulation as from the classical formulation used in the sigma transformation approach. On one hand, we will not only reformulate the partial time derivatives into ALE time derivatives, but we will also transform the spatial partial derivatives present in the continuity equation as well as in the advection terms. On the other hand, we will not transform the second order derivatives in the diffusion terms, because this would lead to a very long and complex expression. In addition, we will see in section 4.6 that a transformation of the horizontal pressure gradient into a topography-following coordinate system can increase the computational error of this term. We therefore also keep the partial derivatives in this terms in the Cartesian frame.

Let us now introduce some relations involving the Jacobian determinant $\hat{J}_t$ of the mapping as well as several derivation formulas that will be needed to write the equations in the reference configuration $\hat{\Omega}$.

Since the mapping only transforms the vertical coordinates of the system, the continuous expression of the Jacobian determinant is

$$\hat{J}_t \left(x, y, \hat{z}, t\right) \; = \; \frac{\partial Z}{\partial \hat{z}} \left(x, y, \hat{z}, t\right) .$$

It can easily be seen from (4.3) that

$$\int_0^1 \hat{J}_t \left(x, y, \hat{z}\right) \; d\hat{z} \; = \; h \left(x, y, t\right) \qquad \forall \left(x, y, t\right) \in \omega \times I. \tag{4.14}$$

Moreover, for any function $\psi : \Omega_t \longrightarrow \mathbb{R}$, its integral on the real domain Ω_t is transformed as follows :

$$\int_{\Omega_{h,t}} \psi \; d\boldsymbol{x} \; = \; \int_{\hat{\Omega}} \hat{J}_{h,t} \; \hat{\psi} \; d\hat{\boldsymbol{x}}. \tag{4.15}$$

with $d\boldsymbol{x} = dx \; dy \; dz$ and $d\hat{\boldsymbol{x}} = dx \; dy \; d\hat{z}$.

Throughout, the ALE space derivatives will be denoted by

$$\frac{\partial}{\partial x}\Big|_{\hat{\boldsymbol{x}}}, \quad \frac{\partial}{\partial y}\Big|_{\hat{\boldsymbol{x}}} \quad \text{and} \quad \frac{\partial}{\partial \hat{z}},$$

where we recall that $\hat{\boldsymbol{x}} = (x, y, \hat{z})$, while the Eulerian space derivatives will be denoted as usual by

$$\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y} \quad \text{and} \quad \frac{\partial}{\partial z}.$$

An alternative to (2.10) relating the ALE time derivative of a function $\psi : \Omega_t \longrightarrow \mathbb{R}$ to its Eulerian time derivative is

$$\frac{\partial \psi}{\partial t} \;=\; \frac{\partial \psi}{\partial t}\Big|_{\hat{\boldsymbol{x}}} \;+\; \frac{\partial \hat{z}}{\partial t}\,\frac{\partial \psi}{\partial \hat{z}}. \tag{4.16}$$

The following formulas relate the different space partial derivatives of ψ :

$$\frac{\partial \psi}{\partial x} \;=\; \frac{\partial \psi}{\partial x}\Big|_{\hat{\boldsymbol{x}}} \;+\; \frac{\partial \hat{z}}{\partial x}\,\frac{\partial \psi}{\partial \hat{z}}, \tag{4.17}$$

$$\frac{\partial \psi}{\partial y} \;=\; \frac{\partial \psi}{\partial y}\Big|_{\hat{\boldsymbol{x}}} \;+\; \frac{\partial \hat{z}}{\partial y}\,\frac{\partial \psi}{\partial \hat{z}}, \tag{4.18}$$

$$\frac{\partial \psi}{\partial z} \;=\; \frac{\partial \hat{z}}{\partial z}\,\frac{\partial \psi}{\partial \hat{z}}. \tag{4.19}$$

Taking $\psi(x, y, z, t) \;=\; z$, essential relations can be derived :

$$\frac{\partial \hat{z}}{\partial t} \;=\; -\hat{J}_t^{-1}\,\frac{\partial z}{\partial t}\Big|_{\hat{\boldsymbol{x}}}, \qquad \frac{\partial \hat{z}}{\partial x} \;=\; -\hat{J}_t^{-1}\,\frac{\partial z}{\partial x}\Big|_{\hat{\boldsymbol{x}}}, \tag{4.20}$$

$$\frac{\partial \hat{z}}{\partial y} \;=\; -\hat{J}_t^{-1}\,\frac{\partial z}{\partial y}\Big|_{\hat{\boldsymbol{x}}}, \qquad \frac{\partial \hat{z}}{\partial z} \;=\; \hat{J}_t^{-1}. \tag{4.21}$$

The definition of a new "pseudo vertical velocity" simplifies the transformation of the equations :

$$w^* \;=\; \frac{d\hat{z}}{dt} \;=\; \frac{\partial \hat{z}}{\partial t} \;+\; u\,\frac{\partial \hat{z}}{\partial x} \;+\; v\,\frac{\partial \hat{z}}{\partial y} \;+\; w\,\frac{\partial \hat{z}}{\partial z} \qquad \text{in} \quad \Omega_t, \tag{4.22}$$

where $\dfrac{d}{dt}$ denotes the total derivative in the real domain. We will see that it is more convenient to use :

$$\begin{aligned} \hat{J}_t\, w^* \;&=\; \hat{w} \;-\; \frac{\partial z}{\partial t}\Big|_{\hat{\boldsymbol{x}}} \;-\; \hat{u}\,\frac{\partial z}{\partial x}\Big|_{\hat{\boldsymbol{x}}} \;-\; \hat{v}\,\frac{\partial z}{\partial y}\Big|_{\hat{\boldsymbol{x}}} \\ &=\; \hat{w} \;-\; \hat{c} \;-\; \hat{u}\,\frac{\partial z}{\partial x}\Big|_{\hat{\boldsymbol{x}}} \;-\; \hat{v}\,\frac{\partial z}{\partial y}\Big|_{\hat{\boldsymbol{x}}} \qquad \text{in} \quad \hat{\Omega}. \end{aligned} \tag{4.23}$$

Remark 4.4.1 Note that this definition of w^*, which is standard in the framework of the sigma transformation, can be misleading in the ALE formulation. In particular, the velocity of the domain is contained in this "pseudo vertical velocity". This may explain some confusions found in the literature, and in particular the false idea that models using the sigma transformation do not take into account the movement of the domain.

The expression of the continuity equation (4.10) in the reference configuration is :

$$\frac{\partial(\hat{J}_t)}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \frac{\partial(\hat{J}_t\ u)}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \frac{\partial(\hat{J}_t\ v)}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \frac{\partial(\hat{J}_t\ w^*)}{\partial \hat{z}} = 0 \qquad \text{in} \quad \hat{\Omega}. \quad (4.24)$$

To obtain this expression, one starts from equation (4.10), multiply with the Jacobian determinant $\hat{J}_t$, and then use relations (4.17)-(4.21).
The momentum equations (4.8)-(4.9) can be written in the following *non-conservative* form :

$$\begin{cases} \dfrac{\partial \hat{\boldsymbol{u}}}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \hat{u}\ \dfrac{\partial \hat{\boldsymbol{u}}}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \hat{v}\ \dfrac{\partial \hat{\boldsymbol{u}}}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \hat{w}^*\ \dfrac{\partial \hat{\boldsymbol{u}}}{\partial \hat{z}} \\ \qquad - (\operatorname{div}(\boldsymbol{\nu}\boldsymbol{\nabla}\boldsymbol{u})) \circ \hat{\mathcal{A}}_t + \left(\dfrac{1}{\rho}\ \boldsymbol{\nabla}_2 p\right) \circ \hat{\mathcal{A}}_t = \hat{f} \qquad \text{in} \quad \hat{\Omega}\,, \\ \dfrac{\partial \hat{w}}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \hat{u}\ \dfrac{\partial \hat{w}}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \hat{v}\ \dfrac{\partial \hat{w}}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \hat{w}^*\ \dfrac{\partial \hat{w}}{\partial \hat{z}} \\ \qquad - (\operatorname{div}(\boldsymbol{\nu}\boldsymbol{\nabla} w)) \circ \hat{\mathcal{A}}_t + \hat{J}_t^{-1}\ \dfrac{1}{\rho}\ \dfrac{\partial p}{\partial \hat{z}} = -g \qquad \text{in} \quad \hat{\Omega}\,. \end{cases} \quad (4.25)$$

But using formulation (4.24) of the continuity equation, we deduce a *conser-*

vative form of the momentum equations :

$$\begin{cases} \dfrac{\partial(\hat{J}_t\ \hat{\boldsymbol{u}})}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ \hat{u}\ \hat{\boldsymbol{u}})}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ \hat{v}\ \hat{\boldsymbol{u}})}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ w^*\ \hat{\boldsymbol{u}})}{\partial \hat{z}} \\ \qquad - \hat{J}_t\ (\operatorname{div}(\boldsymbol{\nu}\boldsymbol{\nabla}\boldsymbol{u})) \circ \hat{\mathcal{A}}_t + \hat{J}_t \left(\dfrac{1}{\rho}\ \boldsymbol{\nabla}_2 p\right) \circ \hat{\mathcal{A}}_t = \hat{f} \quad \text{in} \quad \hat{\Omega}, \\ \dfrac{\partial(\hat{J}_t\ \hat{w})}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ \hat{u}\ \hat{w})}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ \hat{v}\ \hat{w})}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \dfrac{\partial(\hat{J}_t\ w^*\ \hat{w})}{\partial \hat{z}} \\ \qquad - \hat{J}_t\ (\operatorname{div}(\boldsymbol{\nu}\boldsymbol{\nabla} w)) \circ \hat{\mathcal{A}}_t + \dfrac{1}{\rho}\ \dfrac{\partial p}{\partial \hat{z}} = -g \quad \text{in} \quad \hat{\Omega}. \end{cases} \tag{4.26}$$

Finally, the tracer equation (4.13) can be written in the reference configuration either in non-conservative form :

$$\frac{\partial T}{\partial t} + \hat{u}\ \frac{\partial T}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \hat{v}\ \frac{\partial T}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \hat{w}^*\ \frac{\partial T}{\partial \hat{z}} = f_T \quad \text{in} \quad \hat{\Omega}, \tag{4.27}$$

or in conservative form :

$$\frac{\partial(\hat{J}_t\ T)}{\partial t} + \frac{\partial(\hat{J}_t\ \hat{u}\ T)}{\partial x}\Big|_{\hat{\boldsymbol{x}}} + \frac{\partial(\hat{J}_t\ \hat{v}\ T)}{\partial y}\Big|_{\hat{\boldsymbol{x}}} + \frac{\partial(\hat{J}_t\ w^*\ T)}{\partial \hat{z}} = f_T \quad \text{in } \hat{\Omega}. \tag{4.28}$$

Remark 4.4.2 Note that in the case where the hydrostatic assumption is applied, the system simplifies considerably. Indeed, the accelerations of the vertical velocity w are then neglected and w disappears completely from the transformed equations. In that case, only the newly-defined transformed vertical velocity w^* has to be evaluated. The physical vertical velocity can then be recalculated using relation (4.23).

As we can see, the expression of the advection terms in a σ-coordinate system is quite simple using the Jacobian determinant of the ALE mapping. Moreover, it perfectly suits to the formulation of the continuity equation in this system. On the other hand, the kinematic boundary conditions at the top and bottom surfaces of the domain – (4.11) and (4.12) – can be expressed in a simple form in the reference configuration ; that is

$$w^* = 0 \quad \text{at} \quad \hat{z} = 0 \quad \text{and} \quad \hat{z} = 1. \tag{4.29}$$

This shows a major advantage of using a σ-coordinate system.

In the following sections, we will be led to describe some of the different topography-following coordinate systems developed for three-dimensional free surface flow modelling, as well as their advantages and drawbacks. The first and most commonly used σ-coordinate system is the one obtained by the so-called classical sigma transformation, introduced by Phillips in [111].

4.5 The classical sigma transformation

Phillips [111] proposed a linear transformation of the vertical coordinate, which consists in adopting the following σ function :

$$\sigma(x,y,z,t) = \frac{z - b(x,y,t)}{h(x,y,t)}. \tag{4.30}$$

Note that the discrete counterpart of this transformation involves the discrete water height and bottom functions $h_h(x,y,t)$ and $b_h(x,y,t)$. This means that the z-coordinate of the mesh nodes in the current domain are incremented of $b_h(x,y,t)$ and multiplied by the coefficient $h_h(x,y,t)$ relatively to their value in the reference domain. Moreover the relative width of each layer – $\Delta z/h_h$, where Δz is the width of the layer – is equal to its corresponding width in the reference configuration, and thus it is constant over the time and over the whole two-dimensional domain ω. Figure 4.2 shows an exemple of such a mapping.

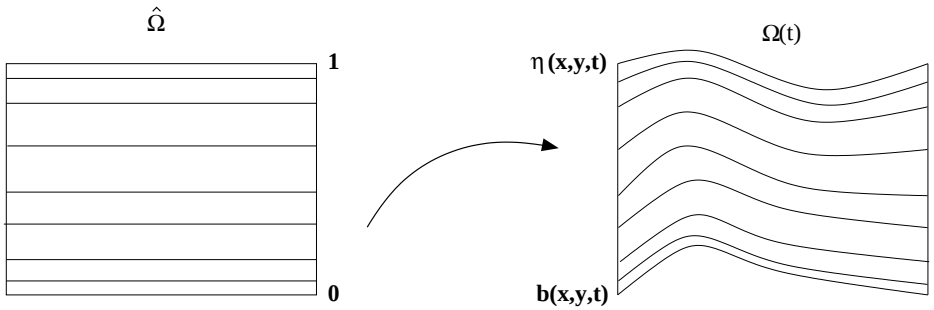


FIG. 4.2 – Example of a classical sigma-transformation

A clear advantage of this coordinate system is its simplicity. Indeed, not only the implementation of the transformation is easy, but the formulas for the variable change between the real and the reference framework are relatively simple. Note in particular that the determinant of the transformation

is equal to the water height function

$$\frac{\partial z}{\partial \hat{z}} = h(x, y, t). \tag{4.31}$$

However, the classical sigma-transformation also presents major drawbacks, due to the restriction it oppers on the vertical discretization. Indeed, this transformation implies that all the levels necessarily follow the bottom and free surface functions, and that the relative width of each layer is constant. As a consequence, it cannot maintain equally high resolution in particular zones – for instance near the surface and the bottom layers – independently of local water height. With a constant number of horizontal levels, deeper zones of the domain can be under-discretized and shallower zones over-discretized, and the latter may cause severe CFL restrictions. Many applications need a high resolution of the surface layer everywhere, for instance to represent accurately a surface mixing process arising from wind, thermal forcing and turbulence. In addition, the classical sigma-transformation prevents from fixing the mesh in a particular zone of the domain which shall be solved more accurately. Now the movement of the mesh can generate an amplification of the numerical diffusion, leading to a relevant perturbation of the solution. Finally, such a coordinate system can yield important errors, especially in the case of stratified flows over steep topography. This problem, known as the *sigma-coordinate pressure gradient error* [51, 76], is due to the topography-following shape of the mesh and has evoked considerable concern in the meteorological and oceanographic communities. To overcome these disadvantages, several alternative and more general coordinate systems were developed. Before reviewing them, let us describe more precisely the sigma-coordinate pressure gradient error.

4.6 The sigma-coordinate pressure gradient error

When the flow is controlled by topographic features a sigma-coordinate system is optimal. But when the pressure represents a major driving force it can produce ill-behaved results. In this case it becomes essential to compute the horizontal pressure gradient (HPG) in the momentum equation accurately.

Now, owing to the sigma-transformation, in the presence of important bottom or free surface gradients, the grid elements become non-orthogonal and may even deteriorate quite strongly in shallow areas – for instance near tidal flats. It has been recognized by several authors (see for instance the pioniers [51], [60] and [76]) that grids of this type cause problems when computing horizontal gradients.

4.6.1 Root of the error

This problem has essentially been analysed in the framework of hydrostatic flows, where the hydrostatic approximation reduces the momentum equation (4.9) on the vertical velocity to

$$\frac{\partial p}{\partial z} \;=\; -\,\rho g.$$

We have therefore focussed on the HPG error in this particular case.

Assuming a constant atmospheric pressure at the free surface and applying the Boussinesq approximation, the horizontal hydrostatic pressure gradient reduces to

$$\frac{1}{\rho}\,\boldsymbol{\nabla}_2 p \;=\; g\,\boldsymbol{\nabla}_2\eta \;-\; g\,\frac{\Delta\rho}{\rho_0}\,\boldsymbol{\nabla}_2\eta \;+\; g\,\boldsymbol{\nabla}_2\left(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz\right), \tag{4.32}$$

where $\Delta\rho$ denotes the fluctuation around an average value ρ_0 of the fluid density ρ, and $\boldsymbol{\nabla}_2$ denotes the horizontal gradient. Note that the terms involving $\Delta\rho$ form the so-called *buoyancy term*. We indicate that the hydrostatic approximation will be described more widely in the second main part of this work.

Clearly, in the hydrostatic case, the pressure gradient error concerns the buoyancy term, and more particularly its barotropic part, *i.e.*,

$$g\,\boldsymbol{\nabla}_2\Big(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz\Big),$$

since it involves the horizontal gradient of a z-dependent function. This term accounts for the accelerations arising from the slight density variations in the fluid.

The main difficulty in computing this term is attributed to the deterioration of the grid cells that can render the numerical scheme "hydrostatically inconsistent". This means that the difference between the discrete and the analytical version of the HPG force does not decrease when the resolution is increased, and that occurs when the so-called *hydrostatic consistency* condition (see [76] and [60]) is violated. Tipically this is the case when, in presence of very steep bottom or free surface gradients, one or more of the nodes at the bottom of a real mesh element are situated above some nodes at its top, as illustrated by Figure 4.3 for prismatic elements. Such a situation can sometimes be avoided already during mesh generation by finer horizontal discretization or lower vertical discretization, especially when it is only caused by the bottom gradient. Unfortunately, as pointed out by Mellor *et al.* in [91], the hydrostatic consistency condition is quite difficult to fulfil in oceanic

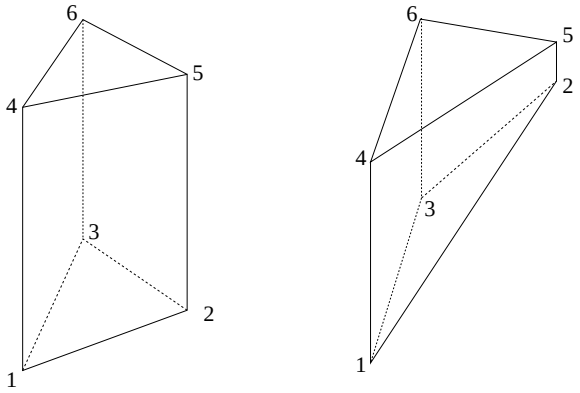


FIG. 4.3 – Example of hydrostatically consistent (left) and inconsistent (right) elements.

applications and especially for estuaries.

Such a situation is especially dangerous for the finite difference models because of the truncation errors inherent to the method. Besides, most of the authors have analysed this problem for the case of a finite difference discretization – see for instance [132], [125], [91]. They point out that the truncation errors in the computation of the gradient are increased by the transformation of the horizontal gradient in the sigma-coordinate system. For example, in the case of the classical sigma-transformation, the x-component of the horizontal gradient force is determined by the sum of two terms ; that is

$$\frac{\partial p}{\partial x} = \frac{\partial p}{\partial x}\Big|_{\hat{\boldsymbol{x}}} - \frac{1}{h}\left(\frac{\partial b}{\partial x} + \hat{z}\,\frac{\partial h}{\partial x}\right)\frac{\partial p}{\partial \hat{z}}\,. \tag{4.33}$$

The first term involves the variation of pressure along the grid line in $\hat{\Omega}$ and the second one involves the usual vertical variation of pressure. Now, near steep bottom these terms are large, of opposite sign and approximately the same size. Analytically, these terms compensate to a relatively small quantity, but the sum of the truncation error of each term may be large.

In the case of a finite element method, the errors are of smaller magnitude because the pressure gradient can be computed to a higher order. Even so, the HPG error exists and can be relevant. It arises essentially from the interpolation error on the relative density field, which can be amplified when computing its gradient. It is thus difficult to catch a stratification in strongly deformed elements when the interpolation functions are not of high order.

4.6.2 Effects of the error

An error in the computation of the horizontal pressure gradient can clearly be detected in the case of an initial density field varying only in the vertical direction – due for instance to a horizontal stratification of the temperature or the water salinity. In this case, the buoyancy source term is nil :

$$- g \frac{\Delta\rho\,(z)}{\rho_0} \boldsymbol{\nabla}_2\eta \;+\; g\,\boldsymbol{\nabla}_2\Big(\int_z^\eta \frac{\Delta\rho\,(z')}{\rho_0}\,dz'\Big) \;=\; 0\,.$$

In absence of exterior source terms, the configuration should theoretically produce zero velocities : this makes spurious velocities easy to detect. Moreover, the density field should remain unchanged, a disturbation of the stratification can therefore illustrate the error.

Other simple test cases of this kind can be performed in order to evaluate the accuracy of the computation of horizontal gradients. Many models using the classical sigma-transformation have been tested on such cases and they produce artificial flows, sometimes even degrading the solution beyond an acceptable limit (see in particular the results obtained on the idealized seamount test case [17], [92]). It has been shown that the pressure gradient error depends on the coordinate slope, the slope of the isopycnals and the polynomial degree of the vertical buoyancy distribution. Larger errors are found in the areas where the pycnoclines intersect the bottom slope, and especially in shallower zones.

Simulations performed with the finite element Telemac-3D system in regions with non-horizontal bottom have also shown important errors in cases where strong stratifications are present [83]. Moreover, they appear even when the hydrostatic consistency condition is fulfilled. Simple test cases have shown that these errors arise from the interpolation errors on the buoyancy term [83, 85] inherent to the finite element method.

In order to clarify this matter, let us develop the case where the density field varies quadratically in the vertical direction only, for instance

$$\Delta\rho\,(x,y,z,t) \;=\; a\,z^2 + b \quad \text{in} \quad \Omega_t \times I,$$

where a and b are two constants. At the continuous level we have that

$$\boldsymbol{\nabla}_2\,\Big(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz\Big) \;=\; 0\;.$$

Assume now that the density variations are interpolated by finite element functions which are piecewise linear in the vertical direction. At the discrete level, using the trapezium rule for the vertical integration – which is exact for linear functions–, and denoting by N_l the number of levels in the mesh,

we then have that

$$\begin{aligned}
\boldsymbol{\nabla}_2 \left(\int_z^{\eta_h} \frac{\Delta\rho_h}{\rho_0}\, dz \right) &= \frac{1}{\rho_0}\, \boldsymbol{\nabla}_2 \left(\sum_{k=1}^{N_l-1} \Delta z_{k+1/2}\, \frac{\Delta\rho\,(z_k) + \Delta\rho\,(z_{k+1})}{2} \right) \\
&= \frac{1}{\rho_0}\, \boldsymbol{\nabla}_2 \left(\sum_{k=1}^{N_l-1} \Delta z_{k+1/2}\, \frac{a\,(z_k^2 + z_{k+1}^2) + 2\,b}{2} \right) \\
&= \frac{1}{\rho_0} \sum_{k=1}^{N_l-1} \left(\frac{a\,(z_k^2 + z_{k+1}^2) + 2\,b}{2}\, \boldsymbol{\nabla}_2\, \Delta z_{k+1/2} \right).
\end{aligned}$$

where z_k is the height of the k-th level and $\Delta z_{k+1/2} = z_{k+1} - z_k$, for $k = 1, \ldots, N_l$. Now in zones where the grid elements are non-orthogonal, the horizontal gradient of the depth $\Delta z_{k+1/2}$ between the layers is not zero. In addition, because of the interpolation error on $\Delta\rho$, the terms do not compensate. Therefore, the value of the buoyancy term is non-zero, which causes spurious numerical density currents. This acts like an articial source term, creating spurious velocities, and if the error is important, like in shallow areas with steep bottom slopes, it can disturb the stratifications and even the free surface.

In order to illustrate this numerical phenomenon, let us show some results obtained on the simple test case of a standing basin using the Telemac-3d system. We consider a closed rectangular basin with a length of 500 m and a width of 100 m, and with a non-horizontal bottom varying quadratically between -25 m and -50 m. The three-dimensional mesh is represented in Figure 4.4. It has been obtained using the classical sigma transformation and piling up the two-dimensional grid shown in Figure 4.6 at 11 homogeneously distributed levels on the vertical. Its vertical cross-section at $y = 50$ m is shown in Figure 4.5.

The basin is filled with a fluid containing an active tracer T, acting on the density ρ following the linear state equation :

$$\rho \;=\; \rho_0 \;+\; 0.749979\; c,$$

where c is the concentration of the tracer T. The initial density field can therefore be imposed by setting a particular initial distribution of the tracer. We will make the relative density $\Delta\rho/\rho_0$ vary in the vertical direction only. At the initial time, the free surface is constant and the velocity of the fluid nil. In addition, the diffusivity of the tracer is neglected and no exterior source term is applied. We indicate that the results are independent of the time step length chosen.

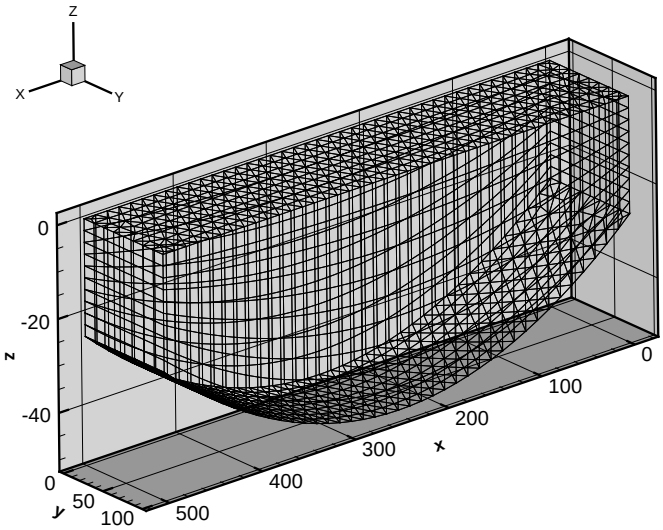


FIG. 4.4 – The initial 3D mesh.

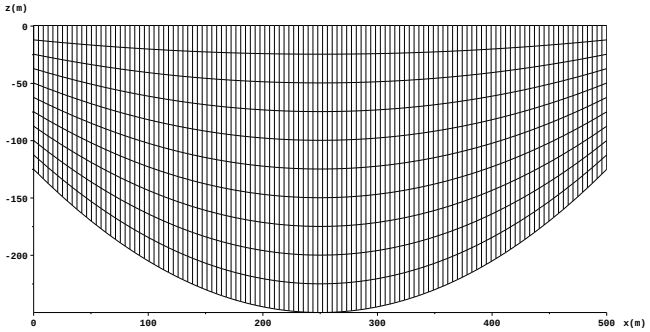


FIG. 4.5 – Vertical cross-section of the initial 3D mesh at $y = 50$m.

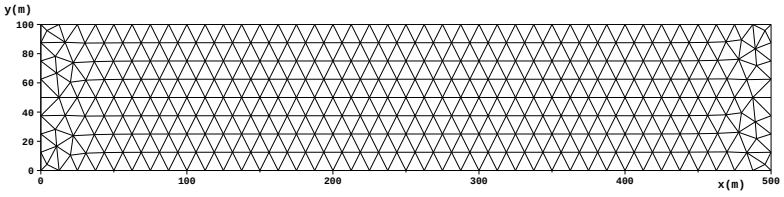


FIG. 4.6 – Horizontal 2D mesh.

First, we set a density field varying linearly in the vertical direction only. The results are presented in Figure 4.7 after 100 seconds. As expected, the system remains unchanged and no movement is detected.

The initial tracer distribution is then modified so that the density field varies quadratically in the vertical direction. The results in Figure 4.8 show spurious velocities, whose amplitude remain small : the maximal value after 100 seconds is obtained at the bottom, where the horizontal velocity $\boldsymbol{u}$ reaches $2.21\ 10^{-3}\,\mathrm{m/s}$ and the vertical velocity w reaches $2.9\ 10^{-4}\,\mathrm{m/s}$.

The spurious velocities increase significantly when a stratified density field is set. In the next test, ρ is piecewise constant and discontinuous at a horizontal interface – as it happens for instance in presence of a horizontal temperature or salinity stratification. The spurious velocities after 100 seconds are shown in figure 4.9. A maximal value of $0,12\,\mathrm{m/s}$ is obtained for the horizontal velocity at the surface, while the vertical velocity w reaches $0.041\,\mathrm{m/s}$. We also notice a disturbation of the density stratification.

This test cases clearly reveal the appearance of an *HPG error* when the density field is not interpolated exactly in the finite element method. In addition, they show that this error increases with the value of the initial vertical density gradient. Indeed, the maximal spurious velocities are obtained when the density field is discontinous, thus when its vertical gradient is infinite.

However, in this particular cases, the HPG error decreases with time. Indeed, the system tends to a new steady state, in which the tracer distribution and thus the density profile become linear, so that the error decays. This is illustrated in Figure 4.10, where we show the results obtained after 1000 seconds of simulation with a density stratification. But the error does not cancel : it seems to converge to a non zero value, as shown by Figure 4.11, where the time evolution of the spurious velocity's maximal value and norm are represented.

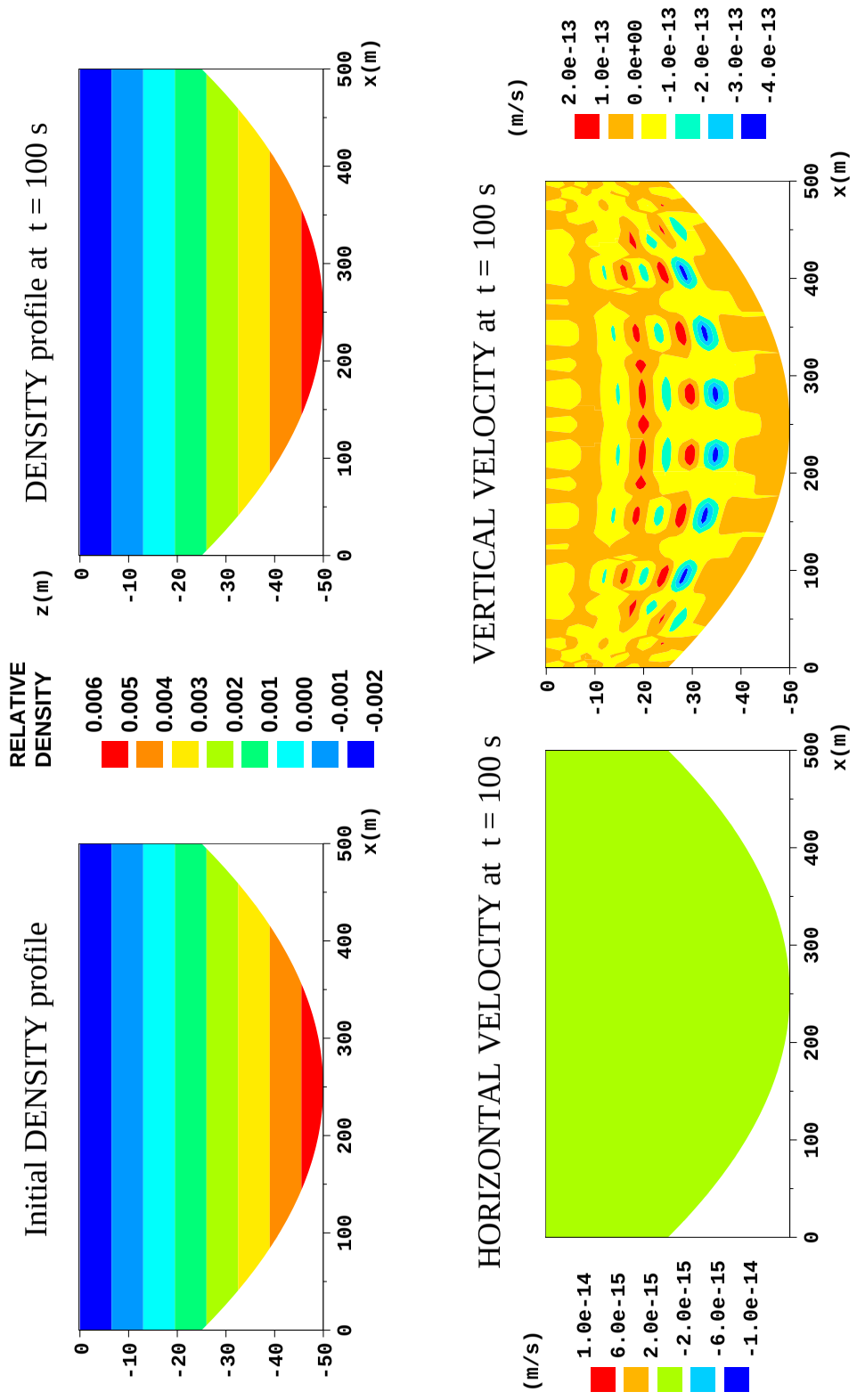


FIG. 4.7 – Results obtained at $t = 100$ s with a linear initial density profile. Vertical cross-section at $y = 50$ m.

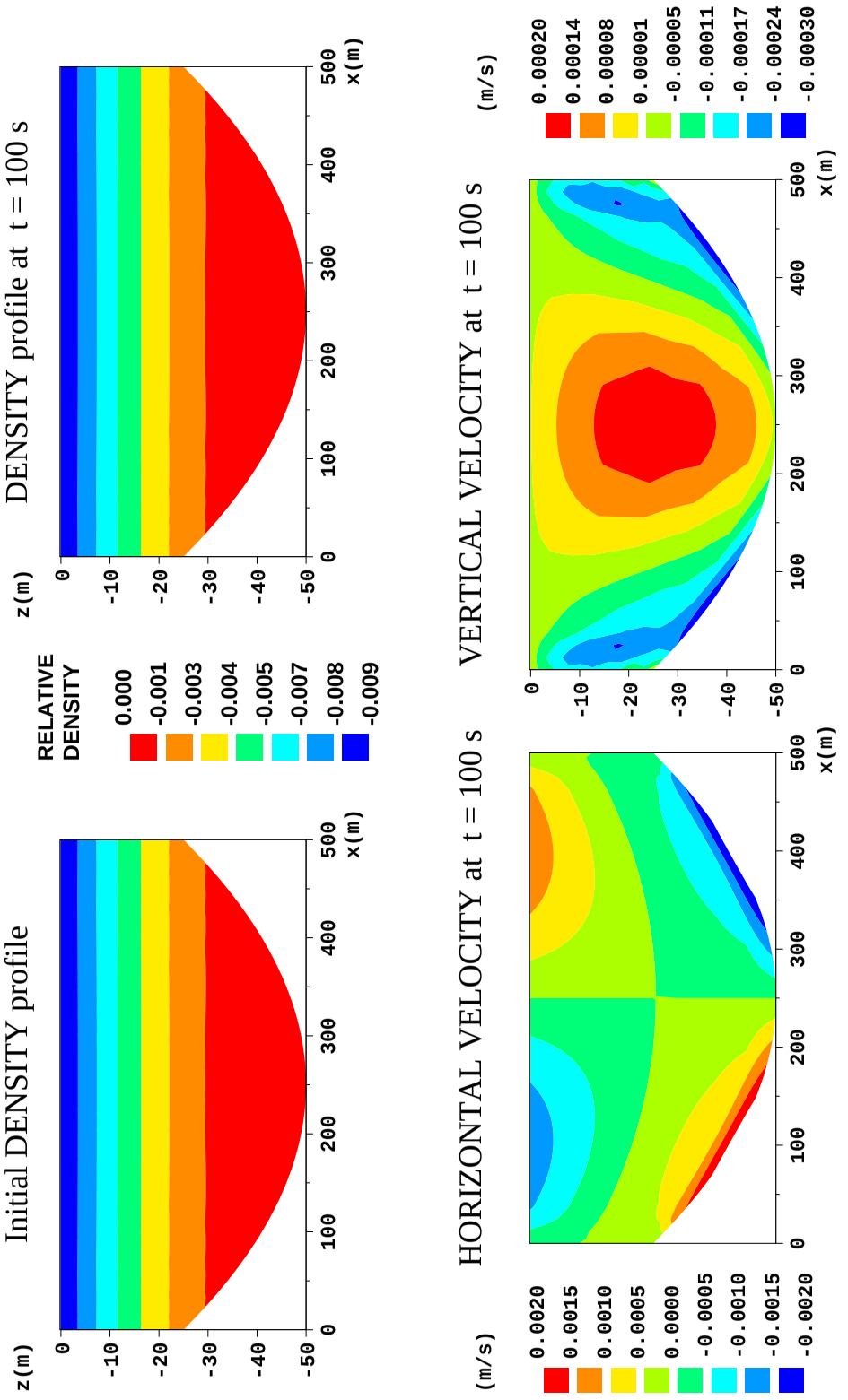


FIG. 4.8 – Results obtained at $t = 100$ s with a quadratic initial density profile. Vertical cross-section at $y = 50$ m.

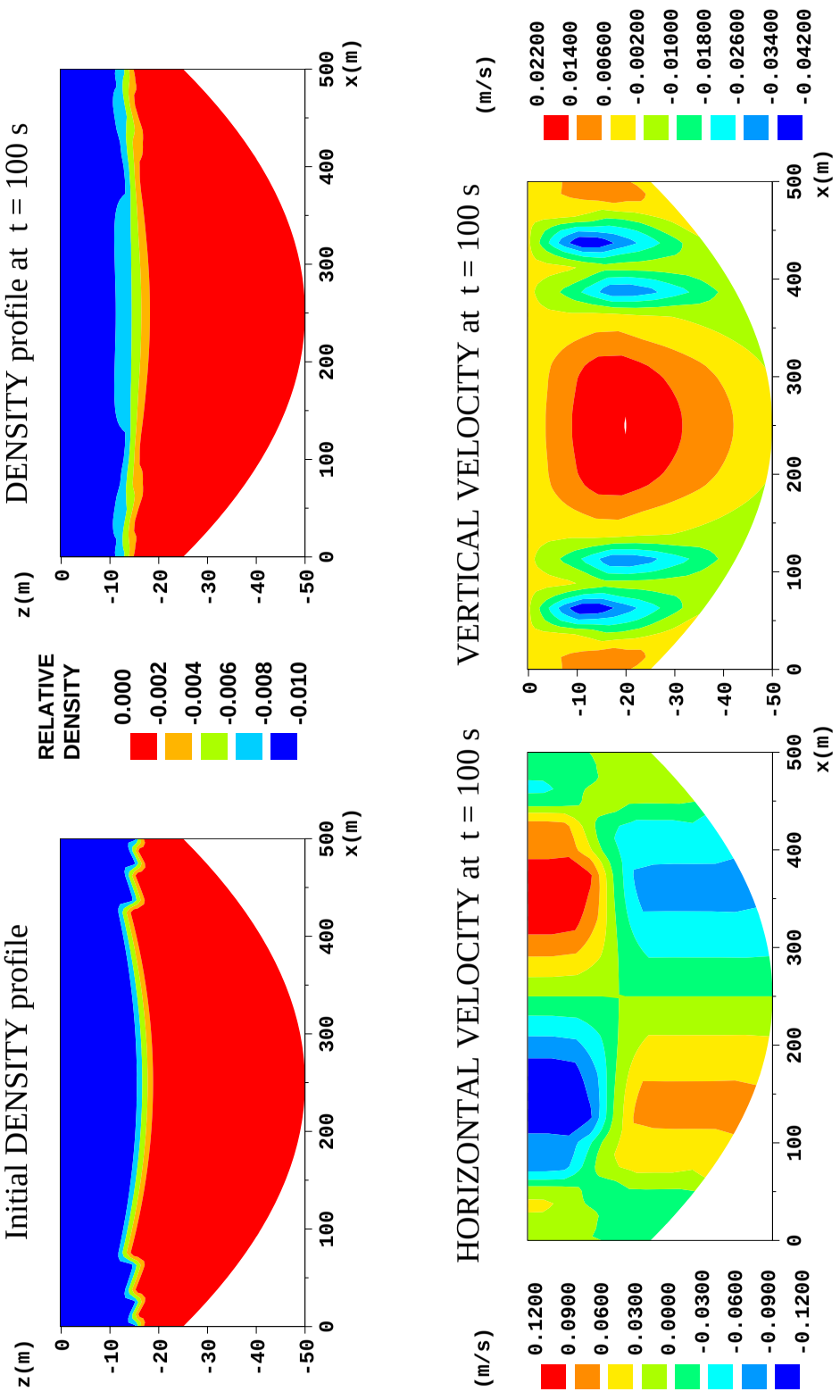


FIG. 4.9 – Results obtained at $t = 100$ s with a horizontal initial density stratification. Vertical cross-section at $y = 50$ m.

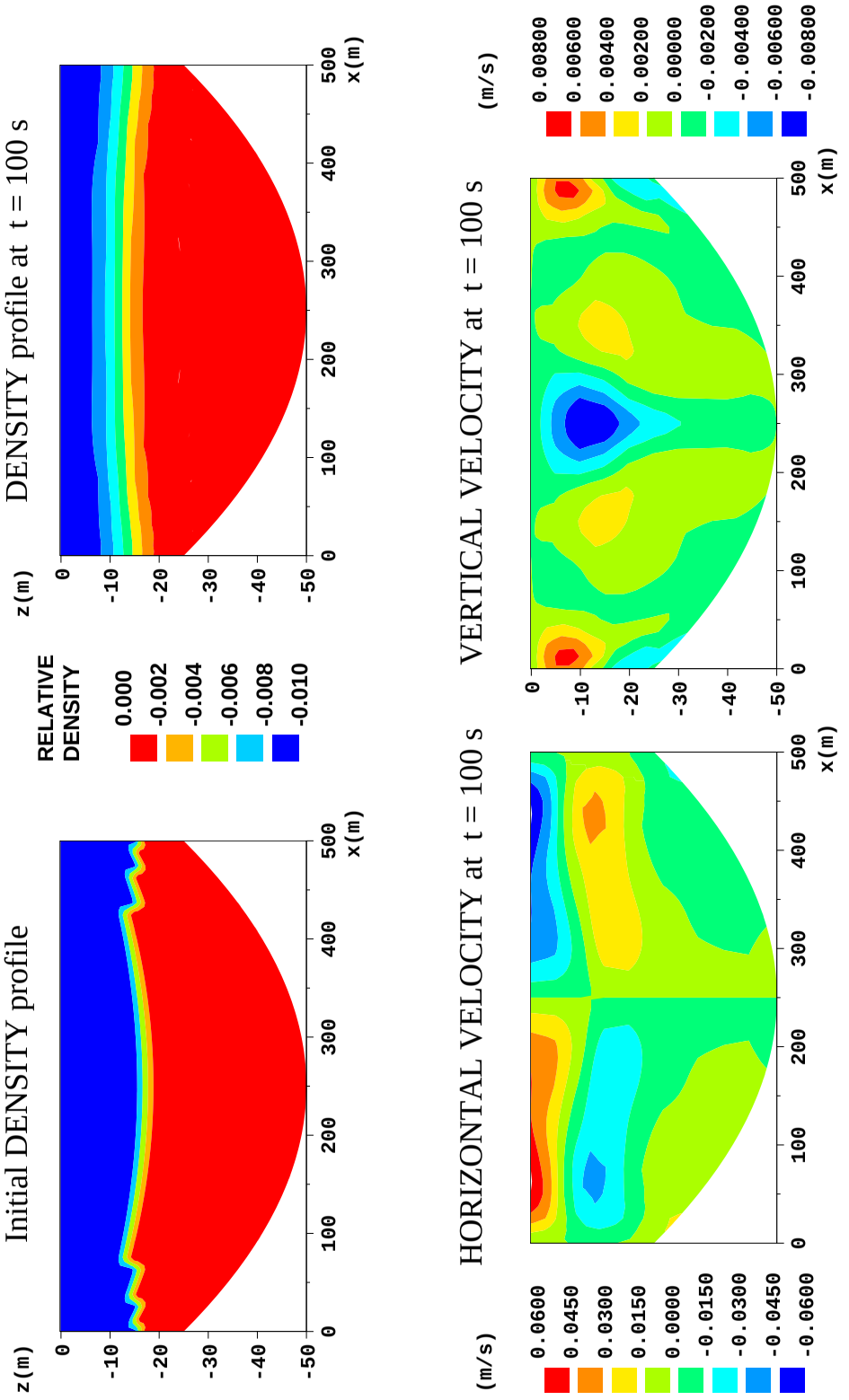


FIG. 4.10 – Results obtained at $t = 1000$ s with a horizontal initial density stratification. Vertical cross-section at $y = 50$ m.

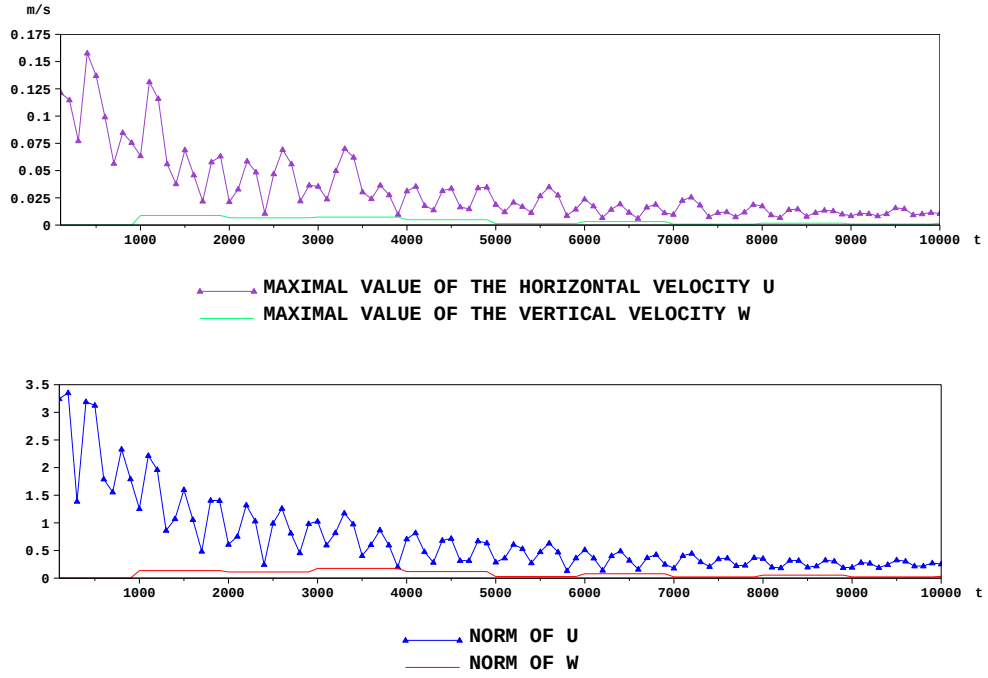


FIG. 4.11 – Time evolution of the spurious velocities with a horizontal initial density stratification.

4.6.3 Solutions

Since the stratifications are nearly horizontal in most situations, a z-coordinate system might be prefered, at least locally. But for applications where both the topography and the density are essential, this wouldn't be a valid option. Another system avoiding the pressure gradient problem is the so-called isopycnic coordinate system, which uses the density as the vertical coordinate. As a consequence the mesh is aligned with the density lines and the first part of the pressure term (4.33) vanishes. But this system is not adequate to study barotropic flows or whenever topography plays a major role.

Therefore, important efforts have been invested in the development of numerical methods improving the approximation of horizontal gradients. A simple and significant improvement is obtained by substracting the horizontal mean value from the actual density field before computing the HPG force [91].

A further improvement results from the use of the so-called density-

Jacobian formulation of the pressure gradient (see [19] and [127]). It is obtained by applying the Leibnitz rule to the baroclinic part of the buoyancy term in (4.32) :

$$g \boldsymbol{\nabla} \left(\int_z^\eta \frac{\Delta\rho}{\rho_0} \, dz \right) \tag{4.34}$$

$$= \; g \int_z^\eta \boldsymbol{\nabla} \left(\frac{\Delta\rho}{\rho_0} \right) dz \; + \; g \; \frac{\Delta\rho}{\rho_0} \, \boldsymbol{\nabla}\eta \tag{4.35}$$

The density fluctuation term is thus differenciated first and integrated after, what allows a more accurate approximation. For instance, if the density field varies linearly, its vertical integration results in a second-order term. Thus, a second-order accurate differencing method has to be used to achieve an exact computation of the gradient using expression (4.34). On the contrary, if expression (4.35) is used, basic first-order derivation and second-order integration methods can be used. This technique has been convincing for the finite element system Telemac-3D – see [75] – in the case of a linear density fluctuation $\Delta\rho$. However, for non-linear or discontinuous density fields, errors still appear.

Other techniques consist in implementing higher-order methods to estimate the pressure gradient term (see [17] and [90]), but they fail in some cases such as in presence of strongly-stratified flows.

Stelling and Van Kester [132] suggested an alternative method, which basically transforms the sigma-grid back to a Cartesian system before computing the HPG force, resulting in substancial reduction of the errors. A systematic underestimation of the calculated gradient was revealed in [125], leading in some cases to errors even larger than before. Thus, the author proposed a slight modification of this method, resulting in an important improvement of its efficiency. However, this method is very costly, especially in so-called prognostic simulations, i.e. simulations where the density field varies in time and has to be recomputed at each time-step. Moreover, substancial errors remain when there are strong vertical gradients in the density field.

Another way to overcome the problem is to design more general coordinate systems, combining the advantages of both the topography-following and the z-coordinate systems. Such a system should allow a transition from sigma- to z-coordinate, such that the top-most grid lines are almost flat, while bottom-most are still aligned with topography. In the next section, we present a brief review of the different topography-following coordinate systems developed since Phillips's classical sigma-transformation.

4.7 General topography-following coordinate systems

General vertical coordinate systems have been considered by meteorologists and oceanographers since the disadvantages of the classical sigma-transformation were detected. They consist in using more general forms of the transformation $z = Z(x, y, \hat{z}, t)$ – see (4.3).

The first generalizations of the sigma-transformation are classified as *separable*, because they use a non-linear function $S(\hat{z})$ that does not depend on the horizontal coordinates :

$$Z(x, y, \hat{z}, t) \;=\; S(\hat{z})\; h(x, y, t) \tag{4.36}$$

Song and Haidvogel introduced in [127] the so-called *general s-coordinate* system, developed in order to allow uniformly high resolution near the surface while maintaining the topography-following properties of the classical sigma-transformation. This system uses the following non-linear, *quasi-separable* transformation function :

$$Z(x, y, \hat{z}, t) \;=\; (1-\hat{z})\; \eta(x, y, t) \;-\; \hat{z}\; h_{min} \;+\; C(\hat{z})\; (b(x, y, t) - h_{min}) \tag{4.37}$$

where

$$C(\hat{z}) \;=\; (1 - \theta_b)\frac{sinh(-\theta\,\hat{z})}{sinh(-\hat{z})} \;+\; \theta_b \left(\frac{tanh[\theta\,(-\hat{z} + 1/2)]}{tanh(\theta/2)} - 1 \right), \tag{4.38}$$

h_{min} is a constant chosen to be the minimum water height or a width of the surface or bottom boundary layer, and θ, θ_b are two parameters controlling the vertical coordinate stretching. Letting θ approach zero the classical sigma-transformation is obtained. But the parameter θ can be chosen so that the highest resolution is achieved near the surface layer independently of the bottom bathymetry. In fact, the first term in (4.37) is used to follow the free surface, the second to prevent possible linear instabilities, and the third is designed to stretch the interior coordinate lines and follow the bottom. The system has achieved satisfactory results in accurately representing surface mixing dynamics. It has also proved to be more accurate and stable in the case of a strong stratification, because it allows a higher concentration of the vertical resolution within the isopycnals. However, note that the Jacobian determinant of this transformation is quite complex, it is thus difficult to implement.

Other authors have focused on systems that would be less sensitive to the pressure gradient error. They have worked on hybridization methods between *sigma*- and z-level coordinates, allowing a maximum of horizontal

levels in the real mesh. Such general systems are no longer separable ; they involve full three-dimensional transformations.

In [36] Deleersnijder and Beckers proposed to define a fixed horizontal level at a constant height $z = z_c$, splitting the domain into an upper and a lower region, and apply the classical sigma-transformation to both subdomains. In the case of a slight surface slope, this system leads to almost horizontal levels in the upper region. It is referenced to as the "double sigma transformation" and can be defined through the following discrete transformation function :

$$Z_h\left(x,y,\hat{z},t\right) \;=\; \begin{cases} \left(\eta_h\left(x,y,t\right)-z_c\right)\;\hat{z}\;+\;z_c & \text{if } \hat{z}\geq\hat{z}_c \\ \\ \left(z_c-b_h\left(x,y,t\right)\right)\;\hat{z}\;+\;b_h\left(x,y,t\right) & \text{if } \hat{z}\leq\hat{z}_c \end{cases}, \quad (4.39)$$

where $\hat{z}_c$ is the height of the fixed level in the reference configuration. Figure 4.12 shows a particular vertical discretization of an estuary obtained with this transformation.

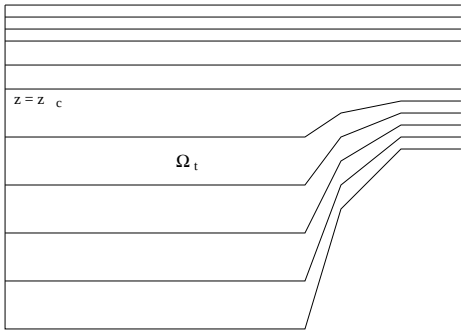


FIG. 4.12 – Example of a grid obtained with the double sigma transformation

A more general approach was introduced by Gerdes [54] and consists in assuming a z-level distribution on the whole domain and "correcting" locally the height of a level whenever its distance from the bottom or the surface is smaller than a minimal width h_{min} chosen arbitrarily. In such a case the classical sigma-transformation determines the position of the level. It is also classified as an s-coordinate system and is obtained using a particular discrete transformation function $Z_h\left(x,y,\hat{z},t\right)$ based on a Min-Mod formulation. It leads to a grid whose levels are horizontal to a certain distance of the bottom and surface boundaries.

However, in real domains the relative importance of topographical effect, density forcing, inertia and diffusion varies from point to point as well as

in time, and a models should thus allow to use different types of grids for different regions. Therefore, a generic discrete vertical transformation function Z_h should be implemented, allowing for arbitrary level distribution at each time step. This idea was first introduced by Deleersnijder and Ruddick in [37], where they applied it to a marine model. The Jacobian determinant of the mapping plays a central role in a system of this type. Indeed, it characterizes the particular transformation defined in each grid element and at each time step, and appears in the expression of the equations driving the flow written in the reference configuration.

Several systems have adopted such a general coordinate system, providing the user with more flexibility in the choice of the optimal model grid for a particular application. Indeed, the recently developed ROMS model [123] follows this approach. Moreover, in [93], Mellor *et al.* introduce a general coordinate system for the POM model. They propose a wide variety of particular vertical coordinate systems, including the z-level, the classical sigma transformation and Gerdes' s-coordinate systems.

A considerable amount of intercomparisons of the different vertical grids have been performed with such models. They have confirmed previous results showing the superiority of terrain-following systems with respect to the z-level model when the topography effects are important. Nevertheless, new information is provided ; in [93] it is shown that sigma-like models can handle much lower diffusivities than z-level systems. In [88], Martins *et al.* apply different vertical discretizations to the simulation of the flow in the Sado estuary. A region of steep bathymetry makes the transition between the estuary and the coastal waters, and a clear vertical recirculation has been identified in this zone during the ebb flow, whose intensity is inversely proportional to the diffusion. Now, the authors show that sigma-like grids provide similar and accurate results, whereas a bigger number of layers is necessary to obtain the recirculation with a z-level system. In [22] a model is presented that uses a generic vertical transformation and solves the equations integrated on each grid layer, so that layer-averaged quantities are computed. The authors show that general coordinate systems can be advantageous compared with classical σ- or z-coordinate systems. Anyway, important research is still necessary to understand the effects of the different sigma-like model grids on the simulated results.

As underscored in the previous section, important errors exist in the computation of the internal horizontal pressure gradient (HPG) in the Telemac-3D system, which we recall uses the classical sigma transformation. These errors subsist in spite of hydrostatic consistency and the adoption of the density-jacobian formulation of the pressure gradient. Moreover, the classical sigma transformation severely restricts the possibilities of vertical discretization. Therefore, we decided to implement a generalization of this trans-

formation allowing for an arbitrary distribution of the layers at each time step. Especially, this should allow to impose fixed levels in particular zones of the domain, for instance to capture stratifications. Next, we describe the implemented coordinate system.

4.8 A generalized σ-coordinate system in the ALES approach

4.8.1 Principle

Let us consider again the general form (4.6) of an ALE mapping between the reference configuration $\hat{\Omega}$ and the real domain Ω_t defined in section 4.3. We recall that it is characterized by a continuous and monotonic vertical transformation function $Z : \hat{\Omega} \times I \longrightarrow \Omega_t$ satisfying

$$\begin{cases} Z(x,y,0,t) = \eta(x,y,t), \\ Z(x,y,1,t) = b(x,y,t), \qquad \forall\,(x,y,t) \in \omega \times I. \end{cases} \tag{4.40}$$

The reference configuration $\hat{\Omega}$ is first discretized by N_l horizontal levels, distributed along the vertical following an arbitrary partition of the interval [0,1] :

$$0 = \hat{z}_1 < \ldots < \hat{z}_k < \ldots < \hat{z}_{N_l} = 1.$$

The height of the grid nodes located on the k-th level of $\hat{\mathcal{T}}_h$ is therefore $\hat{z}_k$, which is constant in space and time. We will denote by $l_{k+1/2}$ the layer delimited by levels k and $k+1$, for $k = 1, \ldots, (N_l - 1)$, and by

$$\Delta\hat{z}_{k+1/2} = \hat{z}_{k+1} - \hat{z}_k$$

its width.

The transformation function $Z\,(x,y,\hat{z},t)$ is then discretized by means of piecewise linear functions in the vertical direction. Inside each layer, this transformation can be expressed as a classical sigma transformation, characterized by the heights – in the real domain – of the levels delimiting the layer. Indeed, the restriction of the discrete transformation on each layer $l_{k+1/2}$ can be defined as follows on $\omega \times [\hat{z}_k, \hat{z}_{k+1}] \times I$:

$$Z_{k+1/2}\,(x,y,\hat{z},t) \;=\; \frac{\Delta z_{k+1/2}(x,y,t)}{\Delta\hat{z}_{k+1/2}}\;(\hat{z} - \hat{z}_k) \;+\; z_k\,(x,y,t)\,, \tag{4.41}$$

where $z_k\,(x,y,t)$ is an arbitrary discrete function describing the current height of the grid level k at each two-dimensional node and at each time

step. Then $\Delta z_{k+1/2}(x,y,t) = z_{k+1}(x,y,t) - z_k(x,y,t)$ denotes the width of the layer $l_{k+1/2}$.

Obviously, it is the choice of the set of discrete functions

$$\{z_k(x,y,t) \ / \ k = 1, \ldots, N_l\}$$

which determines the discrete mapping to use for a particular simulation. We point out that they are defined by their value at each node on the two-dimensional mesh and at each time step. Moreover, they must satisfy the following conditions for all $(x,y,t) \in \omega \times I$:

$$\left\{ \begin{array}{rcll} z_1(x,y,t) & = & b_h(x,y,t), & \\ z_{N_l}(x,y,t) & = & h_h(x,y,t), & \\ z_{k+1}(x,y,t) - z_k(x,y,t) & \geq & h_{min} > 0 \qquad & \forall k \in [1, N_l - 1], \end{array} \right. \tag{4.42}$$

where h_{min} is a constant minimal width chosen arbitrarily.

The discrete current domain $\Omega_{h,t}$ and its triangularisation $\mathcal{T}_{h,t}$ are then defined as in section 2.1.6, and the discrete ALE mapping is defined as follows :

$$\hat{\mathcal{A}}_{h,t} : \hat{\Omega} \longrightarrow \Omega_{h,t}, \qquad \boldsymbol{x}(\hat{\boldsymbol{x}}, t) = \hat{\mathcal{A}}_{h,t}(\hat{\boldsymbol{x}})$$

$$\text{with} \qquad x = \hat{x}, \quad y = \hat{y} \quad \text{and} \quad z = Z_h(x,y,\hat{z},t), \tag{4.43}$$

where $Z_h(x,y,\hat{z},t)$ is a continuous function, linear on each layer $l_{k+1/2}$ and defined by (4.41).

This expression of the discrete mapping is very general and allows a wide choice of layer distributions. Indeed, by imposing for $k = 1, \ldots, (N_l - 1)$

$$z_k(x,y,t) \ = \ h_h(x,y,t) \ \hat{z}_k \ + \ b_h(x,y,t) \tag{4.44}$$

the classical sigma transformation is recovered.

One can also impose some levels to be horizontal and fixed in time at some constant arbitrary height and define classical sigma transformations in between each one of these levels. This can be very useful to catch horizontal stratifications. Such a grid is illustrated in Figure 4.13. It counts 7 levels – of which 2 are horizontal and fixed in time – and it results from a discrete transformation function $Z_h(x,y,t)$ defined from the following set of real level-height functions :

$$\begin{cases} z_5(x,y,t) &= c_5, \\ z_4(x,y,t) &= c_4, \\ z_k(x,y,t) &= \left(\dfrac{\eta_h(x,y,t)-z_5}{1-\hat{z}_5}\right)(\hat{z}_k-\hat{z}_5) \;+\; z_5 \qquad \text{for } 5<k\leq 7, \\ z_k(x,y,t) &= \left(\dfrac{z_4-b_h(x,y,t)}{\hat{z}_4}\right)\hat{z}_k \;+b_h(x,y,t) \qquad \text{for } 1\leq k<4. \end{cases}$$

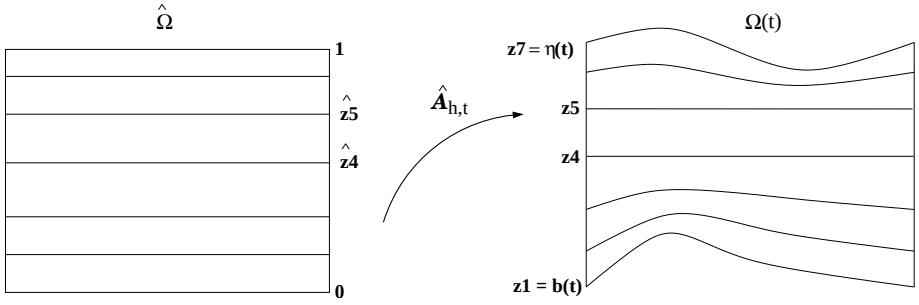


FIG. 4.13 – Example of a grid including one fixed layer obtained with the generalized sigma transformation

But one could as well imagine to impose fixed but non-horizontal levels for the case of steep stratifications.

More generally, the real height of each grid node can be imposed at each time step, defining the discrete function $z_k(x,y,t)$ for $k=1,\ldots,(N_l-1)$. The evolution of the grid is then controlled during the whole simulation.

We indicate that a test is made at each time step to make sure that condition (4.42) is satisfied. In the case the height of a level violates this condition at some node, its value is corrected using a Min-Mod formula.

4.8.2 Jacobian determinant of the discrete mapping

The model must be able to solve the equations for any particular mapping $\hat{\mathcal{A}}_{h,t}$ of the form described previously. Therefore, the Jacobian determinant $\hat{J}_{h,t}$ of the discrete mapping must be expressed in a *generic* form.

We recall that the Jacobian determinant of the classical sigma transformation is constant along the vertical and equal to the water depth. This is

no longer true when the transformation is generalized as it is here. Indeed, in this case, $\hat{J}_{h,t}$ varies along the vertical direction but it remains *constant on each grid layer* in the vertical direction. Its expression on each layer $l_{k+1/2}$, for $k = 1, \ldots, (N_l - 1)$ is

$$\hat{J}_h\,(x, y, \hat{z}, t) \;=\; \frac{\partial Z_{k+1/2}}{\partial \hat{z}}\,(x, y, \hat{z}, t) \;=\; \frac{\Delta z_{k+1/2}\,(x, y, t)}{\Delta \hat{z}_{k+1/2}} \qquad (4.45)$$

$$\text{in} \quad \omega \times [\hat{z}_k, \hat{z}_{k+1}], \quad \forall\, t \in I\,.$$

In the following, we will denote by $\hat{J}_{k+1/2}$ the restriction of $\hat{J}_h$ to layer $l_{k+1/2}$, where it is independent of the $\hat{z}$-coordinate :

$$\hat{J}_{k+1/2}\,(x, y, t) \;=\; \frac{\Delta z_{k+1/2}\,(x, y, t)}{\Delta \hat{z}_{k+1/2}} \qquad \text{in} \quad \omega, \quad \forall t \in I\,. \qquad (4.46)$$

The integral of any discrete function $\psi_h : \Omega_{h,t} \longrightarrow \mathbb{R}$ on the discrete domain $\Omega_{h,t}$ can thus be transformed and decomposed as follows :

$$\begin{aligned}
\int_{\Omega_{h,t}} \psi_h \; d\boldsymbol{x} \;&=\; \int_{\hat{\Omega}} \hat{J}_{h,t} \; \hat{\psi}_h \; d\hat{\boldsymbol{x}} \\
&=\; \sum_{k=1,\ldots,N_l-1} \quad \int_{\hat{l}_{k+1/2}} \hat{J}_{k+1/2}\,(x, y, t) \;\; \hat{\psi}_h\,(x, y, \hat{z}) \;\; d\hat{\boldsymbol{x}} \\
&=\; \sum_{k=1,\ldots,N_l-1} \quad \int_{\omega} \hat{J}_{k+1/2}\,(x, y, t) \left(\int_{\hat{z}_k}^{\hat{z}_{k+1}} \hat{\psi}_h\,(x, y, \hat{z}, t) \;\; d\hat{z} \right) \; d\omega\,.
\end{aligned}$$

The value of this discrete Jacobian determinant characterizes the particular transformation defined in each grid element and at each time step. It appears in the discrete counterpart of the equations driving the flow when written in the reference configuration.

4.8.3 Application to the Telemac-3D system

The general σ-coordinate system proposed in this work has been implemented in the Telemac-3D system in order to allow more flexibility in the choice of the optimal vertical domain discretization. The generalization of the mapping has required an adaptation of the whole system. On one hand, the update of the grid node coordinates in the real domain at each time step has considerably changed. On the other hand, the equations written in the reference configuration were up to now implemented using the value of the Jacobian determinant corresponding to the particular classical sigma transformation. Therefore, these equations had to be re-written in the general vertical coordinate system and implemented in this form.

We indicate that the Telemac-3D system solves the hydrostatic and non-hydrostatic Navier-Stokes equations using the formulation presented in section 4.4 – in the non-conservative form – for the particular case of the classical sigma transformation. We have extended the resolution to the general σ-coordinate system using the ALES formulation.

4.8.4 Results with the generalized sigma transformation

We will now show numerically some advantages of the general σ-coordinate system just described. For this purpose we present two particular tests cases, for which we compare the results obtained with the Telemac-3D system using different grids. Some of those grids can be defined using the classical sigma transformation, others must be defined using the generalized version.

Test case 1 : Advection of a passive tracer in a 3D closed basin with a standing wave. In the first test case, we aim to reveal the amplification of the numerical diffusion due to the movement of the mesh, and show that it is possible to minimize this phenomenon by fixing some particular zones of the mesh. For this purpose we consider again the case of the standing wave in a closed basin introduced in the previous chapter – see section 3.8.3. This time however we highlight the mesh movement by increasing the wave amplitude and decreasing the wave length ; indeed we choose

$$\eta_0 \;=\; 0.2\,\mathrm{m} \quad \text{and} \quad k = \frac{2\,\Pi}{L}\,.$$

The analytical solution is provided by the small amplitude wave theory. The temporal profile of the analytical free surface at $x = 0$ m is shown in Figure 4.14, while the fluid velocity is represented at times $t = 14\,\mathrm{s}$ and $t = 30\,\mathrm{s}$ in Figure 4.15.

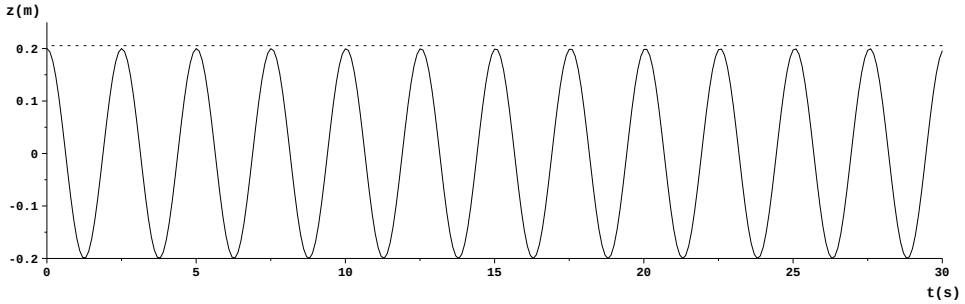


FIG. 4.14 – Temporal profile of the exact water height at $x = 0$ m and $y = 5$ m.

In order to get rid of the numerical errors coming from the hydrodynamic computation in the Telemac-3D system, we impose the analytical free

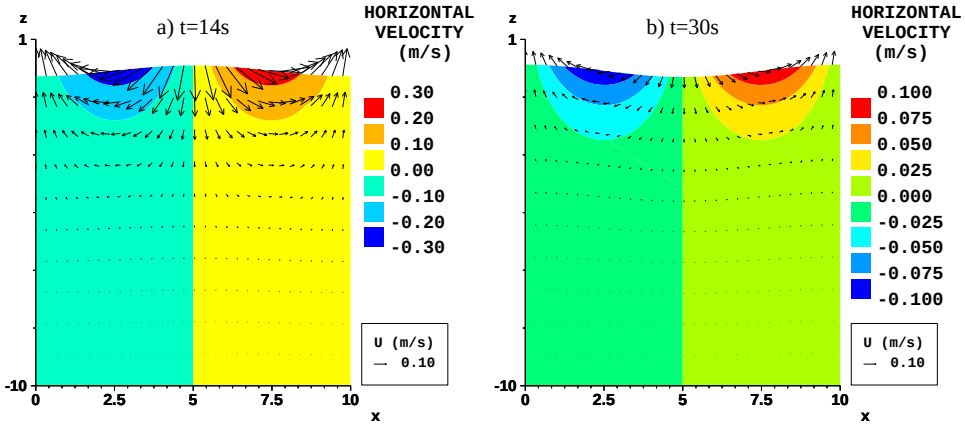


FIG. 4.15 – Analytical velocities at $t = 14\ s$ and $t = 30\ s$. Vertical cross-section at y=5m.

surface and velocities of the fluid at each time, and we only simulate the advection of a passive tracer in the basin. At the initial time, T is stratified along the vertical : its concentration c is equal to 4 if $z \geq -4.5\,\mathrm{m}$, and it is equal to 2 if $z \leq -4.5\,\mathrm{m}$.

A first simulation is performed on a three-dimensional grid obtained using the classical sigma transformation and made of 11 homogeneously distributed levels on the vertical. Its vertical cross-section at $z = 5$ m is shown in Figure 4.16.

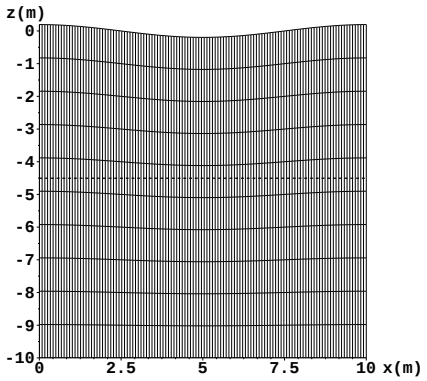


FIG. 4.16 – Mesh 1 obtained using the classical sigma transformation. Vertical cross-section at $y = 5$ m.

A time step length of 0.1 seconds is chosen. The initial state and the re-

sult obtained after 300 time steps are presented in Figure 4.17.

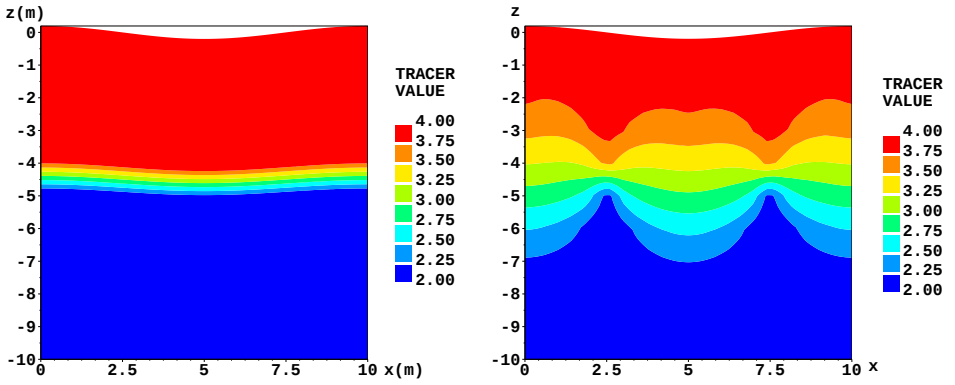


FIG. 4.17 – Initial tracer profile on mesh 1 (left) and tracer profile obtained on mesh 1 at $t = 30$ s (right). Vertical cross-section at $y = 5$ m.

We observe a huge diffusion of the tracer. In order to improve the precision of the simulation a second grid is used, represented in Figure 4.18 and such that the levels are more concentrated in the stratified zone.

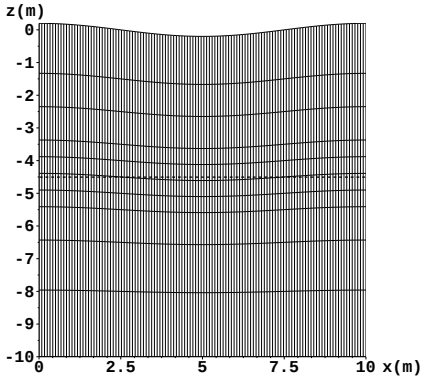


FIG. 4.18 – Mesh 2 obtained using the classical sigma transformation. Vertical cross-section at $y = 5$ m.

The tracer distribution obtained after 30 seconds is shown in Figure 4.19 ; it is slightly less diffused.

In order to evaluate the precision of these results, we have performed the same simulation whith a great number of levels : 55, 105 and finally 205 levels. Figure 4.20 presents the results of these simulations after 30 seconds.

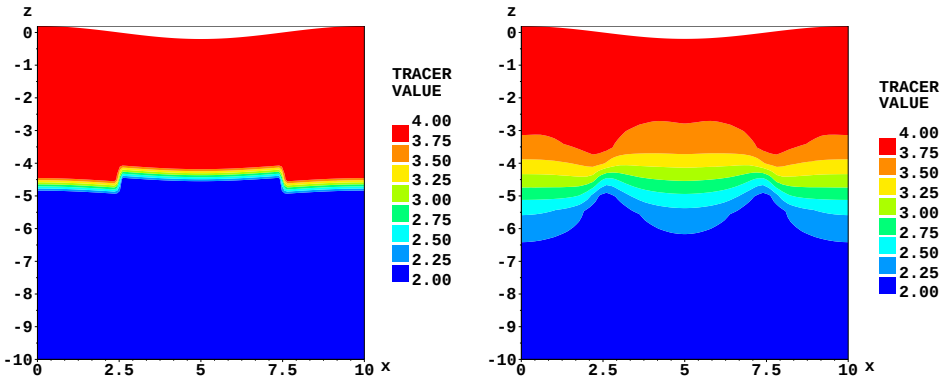


FIG. 4.19 – Initial tracer profile on mesh 2 (left) and tracer profile obtained on mesh 2 at $t = 30$ s (right). Vertical cross-section at $y = 5$ m.

They show that the discrete solution converges to a solution in which the stratification is almost undisturbed ; more precisely, a solution in which the diffusion is minimal.

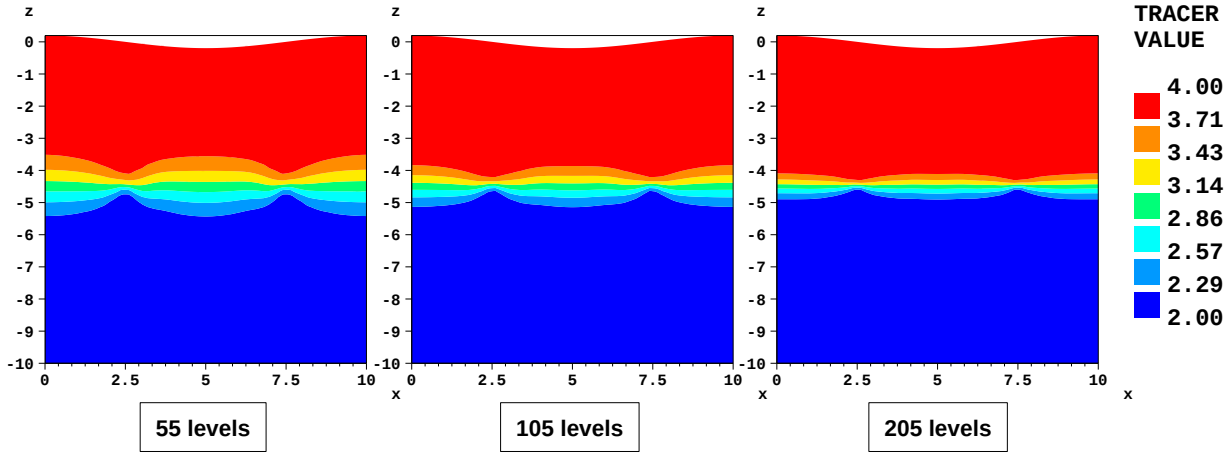


FIG. 4.20 – Tracer profile obtained at $t = 30$ s increasing the number of levels.

The same simulation is then performed on a third grid, shown in Figure 4.21. It also counts 11 levels, but it has been obtained using the generalized sigma transformation implemented in the Telemac-3d system. A horizontal plane is fixed at height $z = -3.4$ m, so that the inferior part of the mesh does not move. Note that the planes are distributed such that the discretization of the stratified zone is as fine as in the second grid.

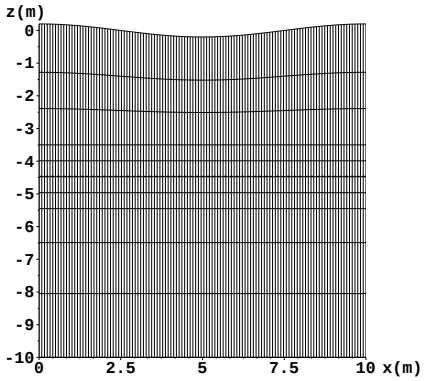


FIG. 4.21 – Mesh 3 obtained using the generalized sigma transformation. Vertical cross-section at $y = 5$ m.

The results obtained on this grid are presented in figure 4.22. The fixed plane has clearly limited the numerical diffusion in the simulation. In conclusion,

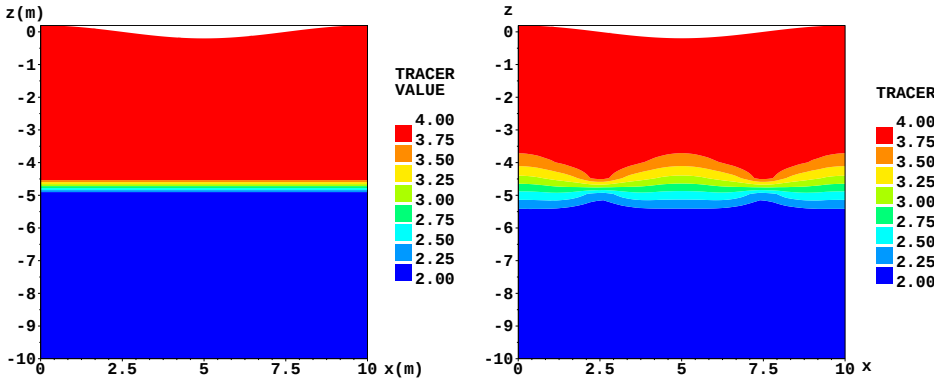


FIG. 4.22 – Initial tracer profile on mesh 3 (left) and tracer profile obtained on mesh 1 at $t = 30$ s (right). Vertical cross-section at $y = 5$ m.

the ALES formulation allows to recover with only 11 layers a better result than the classical sigma transformation with 55 layers.

Test case 2 : Density stratification in a standing basin. Let us now consider again the test case of a horizontal density stratification in a standing basin presented in section 4.6. The results of the simulation performed on a classical sigma grid clearly revealed the existence of the so-called horizontal pressure gradient (HPG) error.

This time, we perform the simulation on a grid obtained using the generalized sigma transformation. It still counts the same number of levels, 11, but they are distributed differently : a fixed plane is imposed exactly at the height of the stratification, and a homogeneous classical sigma transformation is opered in both the lower and the upper parts of the basin. The resulting mesh is shown in Figure 4.23, and the corresponding result after 300 seconds in figure 4.24.

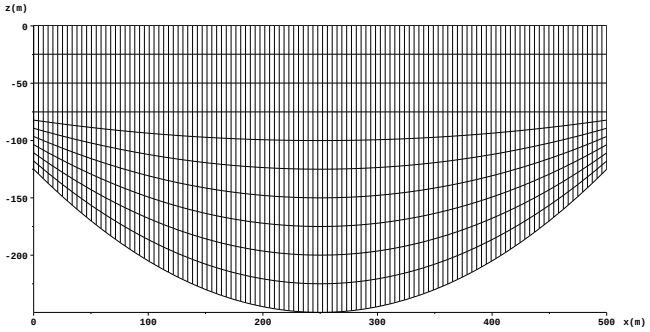


FIG. 4.23 – Mesh 2 obtained using the generalized sigma transformation. Vertical cross-section of the initial 3D mesh at $y = 50$m.

We observe that no spurious velocities have appeared, and that the stratification is undisturbed. The equilibrium has been completely preserved. This shows that fixed planes can help to reproduce accurately density stratifications.

Of course this solution is *ad hoc* since the mesh fits exactely to the shape of the stratification. Nevertheless, in some practical applications, this shape and its time evolution may be known, at least approximately. In such cases, the ALES formulation allows to adapt the mesh in order to reproduce accurately the stratification, getting rid of strong spurious diffusion obtained with the classical sigma transformation technique.

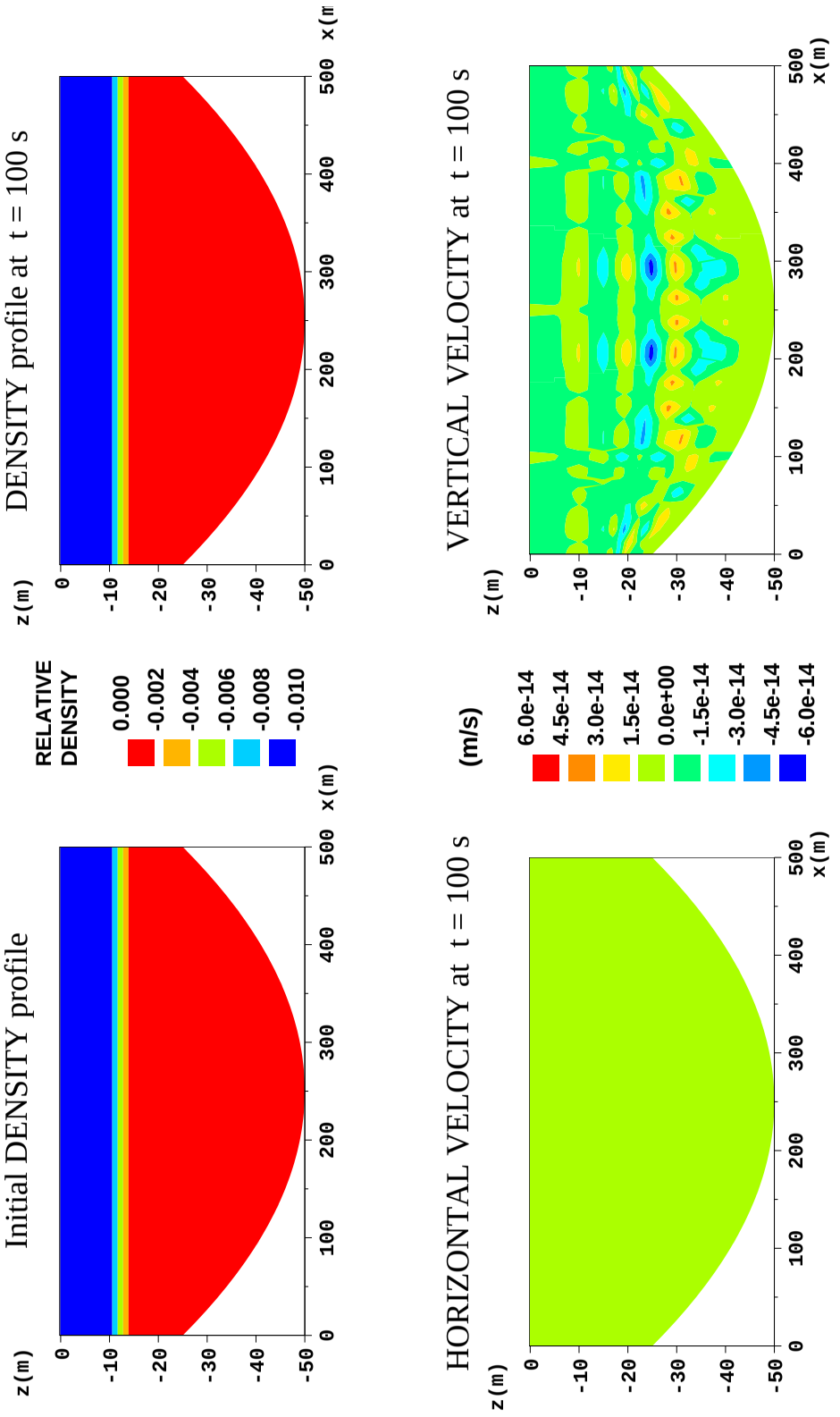


FIG. 4.24 – Results obtained at $t = 100$ s on mesh 2 with a horizontal initial density stratification. The maximal spurious velocities are of about 6.10^{-14} m/s.

Deuxième partie

Three Dimensional Hydrostatic Model for Free Surface Flows

Chapitre 5

Three-dimensional hydrostatic and non-hydrostatic models for incompressible free surface flows

5.1 Introduction

In this chapter we aim to introduce the three-dimensional (3D) free surface problem posed on domains whose lateral boundaries are perfectly vertical. Note that the free surface of these domains is assumed to be a single-valued function.

After describing the domain and its boundaries, we formulate the three-dimensional Navier-Stokes equations describing the motion of the fluid. The Boussinesq approximation is made, allowing to consider an incompressible flow. In order to close the system, we propose a list of convenient boundary conditions for the problem. Since the free surface is assumed to be a singled-valued function, the continuity equation expressing the mass conservation of the fluid can be depth-averaged. This allows to obtain an equation on the free surface. The entire problem is then formulated in the ALE framework.

Once the hydrostatic approximation has been described, we introduce the 3D hydrostatic model, whose study will be adressed in details in Chapters 6 and 7. Then we present a particular 3D model for the non-hydrostatic problem, based on the decomposition of the pressure into hydrostatic pressure and hydrodynamic pressure correction. This decomposition will allow to split the time-discretized equations into a hydrostatic part – which corresponds almost exactly to the hydrostatic model – and a hydrodynamic part. The 3D hydrostatic model can therefore also be viewed as an intermediate step in the solution of the full 3D Navier-Stokes equations.

Finally, a weak formulation of the three-dimensional hydrostatic problem

is proposed. After time-discretization and linearization of the problem, the equations are splitted, leading to three separate sub-problems : the advection of the horizontal velocity ; a system coupling the horizontal velocity and the free surface function, referenced to as the $\boldsymbol{u} - \eta$ *problem* ; and the continuity equation allowing to retrieve the vertical velocity. We point out that the $\boldsymbol{u} - \eta$ problem is linear, symmetric, and that it is based on the physical free surface equation. Chapter 6 will be dedicated to its mathematical analysis.

5.2 Description of the domain

We consider a three-dimensional domain in which *the lateral boundaries are perfectly vertical*. In order to simplify the problem, we assume the bottom to be fixed and impervious.

Let ω be a fixed bounded region of $\mathbb{R}^2$. For any time $t \in I$ the domain Ω_t is defined as follows :

$$\Omega_t = \{\boldsymbol{x} = (x, y, z) \ / \ (x, y) \in \omega \quad \text{and} \quad b\,(x, y) \leq z \leq \eta\,(x, y, t)\,\}$$

where η and b are two functions defined on ω describing respectively the free surface and the bottom (see Figure 5.1). We will denote by h the water height :

$$h\,(x, y, t) \ = \ \eta\,(x, y, t) - b\,(x, y)\,, \quad (x, y) \in \omega, \quad t \in I.$$

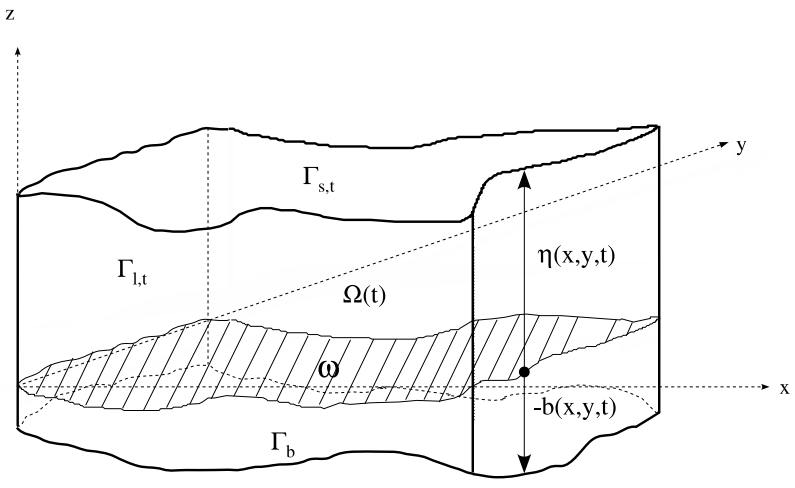


FIG. 5.1 – The three-dimensional domain and its boundaries.

Any integral over the 3D domain Ω_t can then be written as

$$\int_\Omega (\cdots)\ d\boldsymbol{x} = \int_\omega \left(\int_b^\eta (\cdots)\ dz \right)\ d\omega.$$

Note that the free surface as well as the lateral boundaries *only move in the vertical direction.*

The boundary of the current domain Ω_t is denoted by $\partial\Omega_t$ and can be decomposed into three separate parts, as illustrated in Figure 5.1 :

1. the free surface $\Gamma_{s,t}$,
2. the bottom surface Γ_b and
3. the lateral boundaries $\Gamma_{l,t}$. These can include :
   (a) solid lateral boundaries $\Gamma_{sol,t}$ and
   (b) fictitious liquid boundaries $\Gamma_{liq,t}$, composed of the *inflow* boundary $\Gamma_{in,t}$ and the *outflow* boundary $\Gamma_{out,t}$.

We will use the three-dimensional gradient and divergence operators – respectively $\boldsymbol{\nabla}$ and div –, as well as the corresponding horizontal operators, that are :

$$\boldsymbol{\nabla}_2 \boldsymbol{f} \ = \ \begin{pmatrix} \dfrac{\partial \boldsymbol{f}}{\partial x} \\ \\ \dfrac{\partial \boldsymbol{f}}{\partial y} \end{pmatrix} \quad \text{and} \quad \operatorname{div}_2 \boldsymbol{f} \ = \ \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \ .$$

5.3 The three-dimensional free surface Navier-Stokes equations

The equations describing the motion of the fluid are the well known Navier-Stokes equations. They express the mass conservation of the fluid as well as the conservation of the momentum. These equations, valid $\forall t \in I$ and $\forall \boldsymbol{x} \in \Omega_t$, can be written as :

$$\left\{ \begin{array}{rcl} \dfrac{d(\rho \boldsymbol{U})}{dt} - \operatorname{div}(\boldsymbol{\sigma_T}) & = & \rho \boldsymbol{f} + \rho \boldsymbol{g}, \\ \dfrac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{U}) & = & 0, \end{array} \right. \tag{5.1}$$

where $\boldsymbol{U} = (u, v, w)$ is the total velocity of the fluid and ρ its density, $\boldsymbol{\sigma_T}$ is the physical stress tensor, $\boldsymbol{f} = (f_x, f_y, f_z)^T$ the sum of the external forces applied on the fluid, and $\boldsymbol{g} = (0, 0, -g)$ the gravity acceleration.

Note that using the mass conservation equation we obtain the so-called *non-conservative* formulation of the momentum equation :

$$\frac{d\boldsymbol{U}}{dt} \; + \; \frac{1}{\rho} \; \text{div} \; (\, \boldsymbol{\sigma_T} \,) \; = \; \boldsymbol{f} \; + \; \boldsymbol{g}. \tag{5.2}$$

Now considering only Newtonian fluids, the symmetric stress tensor is obtained adding the pressure p to the viscosity tensor $\boldsymbol{\sigma}$:

$$\boldsymbol{\sigma_T} \; = \; -pId \; + \; \boldsymbol{\sigma}, \tag{5.3}$$

with

$$\boldsymbol{\sigma} \; = \; (\sigma_{ij}) \; = \; \mu \; \begin{bmatrix} \mathbb{D}_{11} & \mathbb{D}_{12} & \mathbb{D}_{13} \\ \mathbb{D}_{21} & \mathbb{D}_{22} & \mathbb{D}_{23} \\ \mathbb{D}_{31} & \mathbb{D}_{32} & \mathbb{D}_{33} \end{bmatrix}, \tag{5.4}$$

where $\mathbb{D} \; = \; \boldsymbol{\nabla U} + (\boldsymbol{\nabla U})^T$, and μ is the dynamic viscosity coefficient.

5.3.1 Variable density and Boussinesq approximation

In natural water bodies the density variations are relatively small – in the oceans for instance the variations of the relative density are less than 3% of the average density. However, for many applications, these variations cannot be neglected. In those cases, a state equation relates the fluid density to the concentration of the tracers interacting with the flow – that can be temperature, salinity or sediments. In estuaries for instance, it is essential to consider the fluctuations of the density due to the different salinity concentrations.

We suppose that the density can be written as the sum of a reference value ρ_0 and a fluctuation around this value $\Delta\rho$ such that $\Delta\rho/\rho_0 << 1$. Note that the state equation has the form

$$\frac{\Delta\rho}{\rho_0} \; = \; f(T_1, T_2, \ldots, T_n),$$

where T_i is the i-th active tracer transported.

The Boussinesq approximation – see [130] – states that small density variations must be taken into account only in those terms of the momentum equation where they give rise to buoyancy forces. Therefore the density is supposed constant in the continuity equation, but its variations are considered in the pressure gradient term of the momentum equation. This implies that the variations of fluid mass are neglected, but their effect on fluid weight is retained. In addition, the gradient pression is developed to the first order

in $\Delta\rho$:

$$\frac{1}{\rho}\,\boldsymbol{\nabla} p \;=\; \frac{1}{\rho_0}\,\frac{1}{(1+\frac{\Delta\rho}{\rho_0})}\,\boldsymbol{\nabla} p \;\approx\; \frac{1}{\rho_0}\,(1-\frac{\Delta\rho}{\rho_0})\,\boldsymbol{\nabla} p\,. \tag{5.5}$$

5.3.2 Incompressible fluid

According to the Boussinesq approximation the continuity equation writes

$$\text{div}\;\boldsymbol{U} \;=\; 0 \quad \text{in} \quad \Omega_t, \quad t\in I, \tag{5.6}$$

the flow is thus considered incompressible.

Using the continuity equation (5.6) the divergence of the viscosity tensor can be simplified as follows

$$\text{div}\,(\boldsymbol{\sigma}) \;=\; \text{div}\,(\,\mu\,\boldsymbol{\nabla U}). \tag{5.7}$$

This yields the following non-conservative expression of the three-dimensional free surface Navier-Stokes equations :

$$\left\{\begin{array}{rcl}
\dfrac{du}{dt} \;-\; \text{div}\,(\nu\boldsymbol{\nabla} u) \;+\; \dfrac{1}{\rho}\,\dfrac{\partial p}{\partial x} &=& f_x,\\[2ex]
\dfrac{dv}{dt} \;-\; \text{div}\,(\nu\boldsymbol{\nabla} v) \;+\; \dfrac{1}{\rho}\,\dfrac{\partial p}{\partial y} &=& f_y,\\[2ex]
\dfrac{dw}{dt} \;-\; \text{div}\,(\nu\boldsymbol{\nabla} w) \;+\; \dfrac{1}{\rho}\,\dfrac{\partial p}{\partial z} &=& -\,g \;+\; f_z,\\[2ex]
\dfrac{\partial u}{\partial x} \;+\; \dfrac{\partial v}{\partial y} \;+\; \dfrac{\partial w}{\partial z} &=& 0,
\end{array}\right. \tag{5.8}$$

where $\nu = \mu/\rho_0$ is the kinematic viscosity coefficient of the fluid, and where ρ is the real variable density .

Remark 5.3.1 In order to circumvent the problem of turbulence, the equations are averaged in time, leading to the so-called Reynolds-Averaged Navier-Stokes equations (RANS) [118] that describe the mean flow. The effects of the turbulent fluctuations on the mean flow are then taken care of by means of turbulence models (see for instance [140]). We won't elaborate on this matter. Note however that we have made use of some turbulence models implemented in the Telemac-3D system [66] for the simulations presented in this work. The eddy viscosity coefficients for the horizontal and the vertical diffusion can then be different, and they can even vary in space and time.

The system is closed by suitable initial and boundary conditions.

5.4 Boundary conditions

The boundary conditions prescribed here have been chosen in order to simplify the problem. However, the analysis can be extended to other types of conditions. They are valid for any $t > 0$.

Let $\boldsymbol{n}$ be the outward normal to the current domain Ω_t. We will denote by $\boldsymbol{n}_{xy}$ the horizontal component of $\boldsymbol{n}$ and by n_z its vertical component, so that $\boldsymbol{n} = (\boldsymbol{n}_{xy}, n_z)^T$. We will also make use of $\boldsymbol{n}_\omega$, the outward normal to the two-dimensional domain ω.

Remark 5.4.1 In the general case, $\boldsymbol{n}_{xy} \neq \boldsymbol{n}_\omega$, but on the vertical lateral surfaces,

$$n_z = 0 \quad \text{and} \quad \boldsymbol{n}_{xy} = \boldsymbol{n}_\omega \quad \text{on} \quad \Gamma_{l,t}.$$

On the solid boundaries, the kinematic condition expressing imperviousness is :

$$\boldsymbol{U} \cdot \boldsymbol{n} = 0 \qquad \text{on} \quad \Gamma_b \cup \Gamma_{sol,t}. \tag{5.9}$$

On the bottom, this is equivalent to

$$-\, \boldsymbol{u}_{|z=b} \cdot \boldsymbol{\nabla} b = w_{|z=b} \qquad \text{in} \quad \omega, \tag{5.10}$$

and on the solid lateral boundaries to

$$\boldsymbol{u} \cdot \boldsymbol{n}_{xy} = 0 \quad \text{on} \quad \Gamma_{sol,t}. \tag{5.11}$$

The dynamic condition at the solid boundaries accounts for the shear stress $\boldsymbol{\tau}$ acting on the fluid. This stress is opposed to the velocity – it acts as a brake –, and is equivalent to the tangential stress exerced by the fluid, *i.e.*

$$\boldsymbol{\tau} = -\, \mu \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \qquad \text{on} \qquad \Gamma_b \cup \Gamma_{sol,t}.$$

Since it depends on the flow, the stress $\boldsymbol{\tau}$ is unknown and must be modeled. At the bottom, dimensional analysis shows that, supposing $\boldsymbol{u}$ sufficiently far from the wall, $\boldsymbol{\tau}$ has the form

$$\boldsymbol{\tau} = -\, \frac{1}{2}\, \rho\, C_f\, |\boldsymbol{u}|\, \boldsymbol{u},$$

where C_f is a dimensionless positive friction coefficient. Several formulae, determined empirically, exist to compute this coefficient. We can use for instance the Strickler formula, where C_f depends on the water height and the

bottom roughness. Therefore, we will consider a dynamic condition at the solid boundaries of the form

$$\nu \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} = -\kappa(\boldsymbol{u}) \qquad \text{on} \qquad \Gamma_b \cup \Gamma_{sol,t}, \tag{5.12}$$

where $\kappa : \mathbb{R} \longrightarrow \mathbb{R}$ is a given function.

At the free surface, the velocity of the fluid is equal to the velocity of the free surface itself. This is expressed by the following kinematic condition :

$$\frac{\partial \eta}{\partial t} - \boldsymbol{U} \cdot \boldsymbol{n} = 0 \qquad \text{on} \quad \Gamma_{s,t}, \tag{5.13}$$

which is equivalent to

$$\frac{\partial \eta}{\partial t} + \boldsymbol{u}_{|z=\eta} \cdot \boldsymbol{\nabla} \eta = w_{|z=\eta} \qquad \text{in} \quad \omega. \tag{5.14}$$

The stress vector is continuous at the air-fluid interface ; that is

$$(\boldsymbol{\sigma_T} \cdot \boldsymbol{n})_{fluid} = (\boldsymbol{\sigma_T} \cdot \boldsymbol{n})_{air}.$$

We assume the pressure equal to the atmospherical pressure p_0 and the viscous air stress negligeable. Therefore, the physical dynamic boundary condition prescribed at the free surface is :

$$p = p_0 \quad \text{and} \quad \boldsymbol{\sigma} \cdot \boldsymbol{n} = 0 \quad \text{on} \quad \Gamma_{s,t}.$$

However, in view of the variational formulation, it is convenient to adapt this condition to the particular model used here, in which we assume that $\text{div}\,(\boldsymbol{\sigma}) = \text{div}\,(\mu \boldsymbol{\nabla U})$. Indeed, the following dynamic boundary condition should be prescribed at the free surface :

$$p = p_0 \quad \text{and} \quad \nu \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} = 0 \quad \text{on} \quad \Gamma_{s,t}. \tag{5.15}$$

On the fictitious liquid boundaries $\Gamma_{liq,t}$, the effects of the outside are taken into account. On one hand, the total velocity is prescribed at the inflow :

$$\boldsymbol{U}(x,y,z,t) = \boldsymbol{U}_{in}(x,y,z,t) \qquad \text{on} \quad \Gamma_{in,t}. \tag{5.16}$$

On the other hand, the viscous stress is considered nil at the outflow. The following dynamic condition is therefore prescribed :

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = 0 \quad \text{on} \quad \Gamma_{out,t}. \tag{5.17}$$

Note that this condition is not adapted to the variational formulation of the hydrostatic model we will introduce further in this section. We will see that a stronger condition is needed at the outflow, that is the prescription of the free surface and the assumption that the viscous stress is nil :

$$\eta(x,y,t) \;=\; \eta_{out}(x,y,t) \quad \text{and} \quad \nu\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \;=\; 0 \qquad \text{on} \quad \Gamma_{out,t}. \tag{5.18}$$

5.5 The free surface equation

Proposition 5.1 *If the continuity equation (5.6) as well as the kinematic condition on the bottom (5.10) are satisfied, then*

(i) the kinematic condition at the free surface (5.13)

$$\frac{\partial \eta}{\partial t} \;+\; \boldsymbol{u}_{|z=\eta}\cdot\boldsymbol{\nabla}\eta \;=\; w_{|z=\eta} \qquad \textit{in} \quad \omega$$

and

(ii) the following equation on the free surface

$$\frac{\partial \eta}{\partial t} \;+\; \textit{div}_2\,(\int_b^\eta \boldsymbol{u}\;dz) \;=\; 0 \quad \textit{in} \quad \omega \tag{5.19}$$

are equivalent.

Proof. Let us assume that the continuity equation (5.6) and the kinematic condition on the bottom (5.10) are satisfied.

(i) $\Rightarrow$ (ii) Assume first that the kinematic condition at the free surface (5.13) is satisfied. By depth-integrating the continuity equation (5.6) we obtain

$$\int_b^\eta \mathrm{div}_2\,\boldsymbol{u}\;\;dz \;+\; w_{|z=\eta} - w_{|z=b} \;=\; 0.$$

Applying the Leibnitz rule

$$\int_b^\eta \mathrm{div}_2\,\boldsymbol{u}\;\;dz \;=\; \mathrm{div}_2\,(\int_b^\eta \boldsymbol{u}\;dz) \;-\; \boldsymbol{u}_{|z=\eta}\cdot\boldsymbol{\nabla}\eta \;+\; \boldsymbol{u}_{|z=b}\cdot\boldsymbol{\nabla}b$$

and thanks to the kinematic conditions at the free surface (5.13) and on the bottom (5.10)

$$\int_b^\eta \mathrm{div}_2\,\boldsymbol{u}\;\;dz \;=\; \mathrm{div}_2\,(\int_b^\eta \boldsymbol{u}\;dz) \;+\; \frac{\partial \eta}{\partial t} \;-\; w_{|z=\eta} \;+\; w_{|z=b}.$$

Finally

$$\frac{\partial \eta}{\partial t} \;+\; \mathrm{div}_2\,(\int_b^\eta \boldsymbol{u}\;dz) \;=\; 0.$$

(ii) $\Rightarrow$ (i) Assume now that the free surface equation is satisfied :

$$\frac{\partial \eta}{\partial t} \; + \; \mathrm{div}_2 \, (\int_b^\eta \boldsymbol{u} \; dz) \; = \; 0.$$

The Leibnitz rule states that

$$\mathrm{div}_2 \, (\int_b^\eta \boldsymbol{u} \; dz) \; = \; \int_b^\eta \mathrm{div}_2 \, \boldsymbol{u} \; \; dz \; + \; \boldsymbol{u}_{|z=\eta} \cdot \boldsymbol{\nabla} \eta \; - \; \boldsymbol{u}_{|z=b} \cdot \boldsymbol{\nabla} b$$

The continuity equation yields

$$\int_b^\eta \mathrm{div}_2 \, \boldsymbol{u} \; \; dz \; = \; - \int_b^\eta \frac{\partial w}{\partial z} \; dz \; = \; - w_{|z=\eta} \; + \; w_{|z=b}$$

and the kinematic condition on the bottom (5.10) reads

$$- \boldsymbol{u}_{|z=b} \cdot \boldsymbol{\nabla} b \; = \; - w_{|z=b}.$$

Therefore

$$\frac{\partial \eta}{\partial t} \; + \; \boldsymbol{u}_{|z=\eta} \cdot \boldsymbol{\nabla} \eta \; = \; w_{|z=\eta}$$

$\diamondsuit$

This property allows us to replace the kinematic condition (5.13) by the free surface equation (5.19), which leads to a more convenient formulation.

5.6 The three-dimensional tracer equation

The free surface hydrodynamic problem also involves the motion of the tracers, which represent a temperature or any physical quantity contained in the water. If a tracer interacts with the flow, as for instance the temperature and the salinity, it is called active ; if not it is passive. The time evolution of a tracer depends on the velocity of the flow that drives it, the diffusion and the source sink terms.

The tracer equation in conservative form is

$$\frac{\partial (\rho_0 \, T \,)}{\partial t} \; + \; \mathrm{div} \, (\, \rho_0 \, T \, \boldsymbol{U} \; - \; \rho_0 \, \nu_T \boldsymbol{\nabla} T \,) \; = \; f_T \qquad \text{on} \quad \Omega_t,$$

where f_T is the source term and ν_T the tracer's constant diffusion coefficient. We recall that $\boldsymbol{U}$ is the divergence-free velocity of the flow and ρ_0 its average density.

Using the continuity equation of the fluid we obtain the following *non-conservative* form of this equation :

$$\frac{\partial T}{\partial t} \; + \; \boldsymbol{U} \cdot \boldsymbol{\nabla} T \; - \; \mathrm{div} \, (\nu_T \boldsymbol{\nabla} T) \; = \; f_T \qquad \text{on} \quad \Omega_t. \qquad (5.20)$$

This system is closed by suitable boundary conditions. The quantity T is prescribed at the inflow boundaries $\Gamma_{in,t}$ of the domain, *i.e.*,

$$T(\boldsymbol{x},t) \;=\; T_{in}(\boldsymbol{x},t) \qquad \boldsymbol{x} \in \Gamma_{in,t}, \quad t \in I, \tag{5.21}$$

where $\Gamma_{in,t}$ is defined by

$$\Gamma_{in,t} \;=\; \{\; \boldsymbol{x} \in \Gamma_{l,t} \;/\; \boldsymbol{u} \cdot \boldsymbol{n}_{xy} < 0 \;\}\,. \tag{5.22}$$

In addition, we prescribe a flux law at the free surface of the type :

$$\nu_T \boldsymbol{\nabla} T \cdot \boldsymbol{n} \;=\; a\,T \;+\; b. \tag{5.23}$$

We also assume the flux of the tracer to be nil on the solid boundaries of the domain ; that is, $\forall t \in I$,

$$\nu_T \boldsymbol{\nabla} T \cdot \boldsymbol{n} \;=\; 0 \quad \text{on} \quad \partial\Gamma_b \cup \Gamma_{sol,t}. \tag{5.24}$$

5.7 ALE formulation

The problem is now reformulated in the Arbitrary Lagrangian Eulerian (ALE) approach. We consider the reference configuration

$$\hat{\Omega} = \{\; \hat{\boldsymbol{x}} = (\hat{x},\hat{y},\hat{z}) \;/\; (\hat{x},\hat{y}) \in \omega \quad \text{and} \quad 0 \le \hat{z} \le 1 \;\}$$

and an ALE mapping of the ALE-Sigma approach form – see Part I, section 4.3 –, that is

$$\begin{gathered} \hat{\mathcal{A}}_t \;:\; \hat{\Omega} \longrightarrow \Omega_t, \qquad \boldsymbol{x}(\hat{\boldsymbol{x}},t) = \hat{\mathcal{A}}_t(\hat{\boldsymbol{x}}) \\ \text{with} \quad x = \hat{x}, \quad y = \hat{y} \quad \text{and} \quad z = Z\,(x,y,\hat{z},t)\,, \end{gathered} \tag{5.25}$$

where $Z\,(x,y,\hat{z},t)$ is an arbitrary continuous and monotonic ($\partial Z/\partial\hat{z} \ge 0$) function satisfying

$$\left\{ \begin{array}{ll} Z(x,y,0,t) = \eta(x,y,t) & \forall\,(x,y,t) \in \omega \times I \\ Z(x,y,1,t) = b(x,y,t). & \end{array} \right. \tag{5.26}$$

Since *the domain moves in the vertical direction only*, its velocity $\boldsymbol{c}$ has only a vertical non zero component :

$$\boldsymbol{c} \;=\; \frac{\partial \hat{\mathcal{A}}_t}{\partial t} \;=\; (0,0,c)^T. \tag{5.27}$$

We can now write the non-conservative ALE formulation of the considered

hydrodynamic problem, valid $\forall t \in I$ and $\forall \boldsymbol{x} \in \Omega_t$:

$$\frac{\partial \boldsymbol{U}}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + (\boldsymbol{U} - \boldsymbol{c}) \cdot \boldsymbol{\nabla} \boldsymbol{U} - \text{div}\,(\nu \boldsymbol{\nabla} \boldsymbol{U}) + \frac{1}{\rho} \boldsymbol{\nabla} p = \boldsymbol{f} + \boldsymbol{g} \quad (5.28)$$

$$\frac{\partial \eta}{\partial t} + \text{div}_2 \left(\int_b^\eta \boldsymbol{u}\, dz \right) = 0 \quad (5.29)$$

$$\text{div}\ \boldsymbol{U} = 0 \quad (5.30)$$

$$\frac{\partial T}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + (\boldsymbol{U} - \boldsymbol{c}) \cdot \boldsymbol{\nabla} T - \text{div}\,(\nu_T \boldsymbol{\nabla} T) = f_T. \quad (5.31)$$

It is completed with the folllowing boundary conditions on $\boldsymbol{U}$ and η :

$$\begin{cases} \boldsymbol{U} \cdot \boldsymbol{n} = 0 \quad \text{on} \quad \Gamma_b, \\ \boldsymbol{u} \cdot \boldsymbol{n}_{xy} = 0 \quad \text{on} \quad \Gamma_{sol,t}, \\ \nu \dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} = -\kappa(\boldsymbol{u}) \quad \text{on} \quad \Gamma_b \cup \Gamma_{sol,t}, \\ \nu \dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} = 0 \quad \text{on} \quad \Gamma_{s,t}, \\ \boldsymbol{U} = \boldsymbol{U}_{in} \quad \text{on} \quad \Gamma_{in,t}, \\ \eta = \eta_{out} \quad \text{on} \quad \Gamma_{out,t}, \\ \nu \dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} = 0 \quad \text{on} \quad \Gamma_{out,t}, \end{cases} \quad (5.32)$$

and on T :

$$\begin{cases} T = T_{in} \quad \text{on} \quad \Gamma_{in,t}, \quad t \in I, \\ \nu_T \boldsymbol{\nabla} T \cdot \boldsymbol{n} = a\, T + b \quad \text{on} \quad \Gamma_{s,t}, \\ \nu_T \boldsymbol{\nabla} T \cdot \boldsymbol{n} = 0 \quad \text{on} \quad \partial\Gamma_b \cup \Gamma_{sol,t}. \end{cases} \quad (5.33)$$

Remark 5.7.1 Note that the time derivative on the free surface function η is not turned into an ALE time derivative, because this variable is defined on the fixed two-dimensional domain ω.

5.8 The hydrostatic assumption

The hydrostatic approximation consists in neglecting the vertical acceleration, diffusion and source term, so that the momentum equation on the vertical velocity in (5.8) reduces to :

$$\frac{\partial p}{\partial z} \;=\; -\,\rho g. \tag{5.34}$$

Rescaling the continuity equation (5.30) and the momentum equation (5.28) using U, W, L and H, the scales for the horizontal and vertical velocities, for a characteristic horizontal length and for the depth, it can be shown that :

1. the vertical acceleration is negligible if the horizontal scale of motion dominates the vertical one, that is if $H << L$ or if $W << U$, which is usually the case in shallow water areas and in flows with long waves,
2. the diffusion on the vertical velocity can be neglected if the Reynolds number
$$R_e \;=\; \frac{L\,U}{\nu}$$
is large, which is the case for turbulent flows with dominating horizontal motions.

For the derivation of the rescaled system see for instance [110].

The hydrostatic approximation can therefore be applied if conditions 1 and 2 are fullfilled. In that case, since the pressure at the free surface is assumed to be equal to the atmospherical pressure p_0 – considered constant and known – we have that

$$p(x,y,z,t) \;=\; \rho_0 g\,(\eta\,(x,y,t)-z) \;+\; p_0 \;+\; g\rho_0\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz. \tag{5.35}$$

The horizontal pressure gradient term in the momentum equation is then :

$$\frac{1}{\rho}\,\boldsymbol{\nabla}_2 p \;=\; \frac{1}{\rho}\,\left(g\,\boldsymbol{\nabla}_2\eta \;+\; g\,\boldsymbol{\nabla}_2(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz)\right).$$

According to the Boussinesq approximation the second order terms in $\Delta\rho/\rho_0$ can be neglected, yielding

$$\frac{1}{\rho}\,\boldsymbol{\nabla}_2 p \;\approx\; g\,\boldsymbol{\nabla}_2\eta \;-\; g\,\frac{\Delta\rho}{\rho_0}\,\boldsymbol{\nabla}_2\eta \;+\; g\,\boldsymbol{\nabla}_2\left(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz\right). \tag{5.36}$$

The terms accounting for the effects of the density fluctuation are referenced to as the *buoyancy terms*. We will denote them by

$$\boldsymbol{f}^{buoy} \;=\; -\,g\,\frac{\Delta\rho}{\rho_0}\,\boldsymbol{\nabla}_2\eta \;+\; g\,\boldsymbol{\nabla}_2\left(\int_z^\eta \frac{\Delta\rho}{\rho_0}\,dz\right). \tag{5.37}$$

5.9 The three-dimensional hydrostatic model

According to (5.34) and (5.36), the ALE formulation of the equations describing a three-dimensional hydrostatic free surface flow with divergence free velocity is :

$$\frac{\partial \boldsymbol{u}}{\partial t}\Big|_{\hat{\boldsymbol{x}}} \;+\; (\boldsymbol{U}-\boldsymbol{c})\cdot\boldsymbol{\nabla}\boldsymbol{u} \;-\; \mathrm{div}\,(\nu\boldsymbol{\nabla}\boldsymbol{u}) \;+\; g\,\boldsymbol{\nabla}_2\eta \;=\; \boldsymbol{f}^{hor}, \tag{5.38}$$

$$\frac{\partial \eta}{\partial t} \;+\; \mathrm{div}_2\,(\int_b^\eta \boldsymbol{u}\,dz) \;=\; 0, \tag{5.39}$$

$$\mathrm{div}_2\;\boldsymbol{u} \;+\; \frac{\partial w}{\partial z} \;=\; 0, \tag{5.40}$$

$$\tag{5.41}$$

where $\boldsymbol{f}^{hor} = (f_x, f_y)^T + \boldsymbol{f}^{buoy}$. This formulation is valid for any $t \in I$ and for all $\boldsymbol{x} \in \Omega_t$.

Let us explicit the system for more clarity :

$$\begin{cases}
\dfrac{\partial u}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} + (w-c)\dfrac{\partial u}{\partial z} - \mathrm{div}\,(\nu\boldsymbol{\nabla}u) + g\dfrac{\partial \eta}{\partial x} = f_x^{hor}, \\[2ex]
\dfrac{\partial v}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} + (w-c)\dfrac{\partial v}{\partial z} - \mathrm{div}\,(\nu\boldsymbol{\nabla}v) + g\dfrac{\partial \eta}{\partial y} = f_y^{hor}, \\[2ex]
\dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}(\int_b^\eta u\,dz) + \dfrac{\partial}{\partial y}(\int_b^\eta v\,dz) = 0, \\[2ex]
\dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z} = 0,
\end{cases}$$

This problem is also referred to as the *three-dimensional shallow water equations.*

We observe that the hydrostatic model does not involve the time-derivative of w. Therefore, no initial condition is needed on the vertical velocity of the fluid. Moreover, in the momentum equation (5.38) on $\boldsymbol{u}$, w only appears as part of $(w-c)$, which we call the vertical ALE advective velocity and denote

by

$$\bar{w} \; = w - c \, . \tag{5.42}$$

In fact, the vertical velocity w is not needed separately and we can determine directly $\bar{w}$.

Following this remark, we choose to reformulate the continuity equation in order to make appear the vertical ALE advective velocity $\bar{w}$.
Let us multiply equation (5.40) by the determinant of the ALE mapping. Then, using the Euler expansion formula (2.11), we obtain the following ALE formulation of the continuity equation :

$$\frac{\partial \hat{J}_t}{\partial t} \, (\, \mathcal{A}_t^{-1}(\boldsymbol{x}), \, t \,) \; + \; \hat{J}_t \, (\, \mathcal{A}_t^{-1}(\boldsymbol{x}), t \,) \; \text{div} \, (\boldsymbol{U} - \boldsymbol{c}) \; = \; 0 \qquad \text{on} \quad \Omega_t. \tag{5.43}$$

Denoting the three-dimensional velocity $(\boldsymbol{U} - \boldsymbol{c})$ by

$$\bar{\boldsymbol{U}} \; = \; \boldsymbol{U} - \boldsymbol{c} \; = \; (\boldsymbol{u}, \bar{w})^T \, , \tag{5.44}$$

the continuity equation can be written in the ALE frame as follows :

$$\frac{\partial \hat{J}_t}{\partial t} \; + \; \hat{J}_t \; \left(\text{div} \, \bar{\boldsymbol{U}} \circ \hat{\mathcal{A}}_t \right) \; = \; 0 \qquad \text{on} \quad \hat{\Omega}. \tag{5.45}$$

This formulation of the continuity equation will only be used for the hydrostatic model. It simplifies system (5.38)-(5.40) since the only variables involved are now the horizontal velocity $\boldsymbol{u}$, the vertical ALE advective velocity $\bar{w}$ and the free surface function η.

5.10 The three-dimensional non-hydrostatic model

When the pressure cannot be assumed to be hydrostatic or when the vertical accelerations are important, the non-hydrostatic model must be used. This is the case, for instance,

- in deep waters with strong bathymetry gradients, where the influence of the steep bottom slopes on the velocity field is not damped like in shallow waters and results in changing the flow direction ;
- in presence of short waves, where there are important vertical orbital movements of the fluid particles in the wave ;
- in flows around structures, where the pressure field before the obstacle cannot be considered hydrostatic.

The formulation of the three-dimensional non-hydrostatic model presented next is based on the pressure decomposition into hydrostatic and hydrodynamic parts. This idea was applied by Casulli and Stelling to geophysical

flows in a finite difference implementation presented in [27]. Saleri *et al.* also proceed to this pressure decomposition in [97], where they propose a finite element approximation of the 3D free surface Navier Stokes equations.

The ALE formulation of the non-hydrostatic equations, valid for any $t \in I$ and for all $\boldsymbol{x} \in \Omega_t$, is then :

$$\frac{\partial \boldsymbol{u}}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \bar{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \text{div}\,(\nu \boldsymbol{\nabla} \boldsymbol{u}) + g\,\boldsymbol{\nabla}_2 \eta + \frac{1}{\rho}\boldsymbol{\nabla}_2 p_d = \boldsymbol{f}^{hor} \quad (5.46)$$

$$\frac{\partial w}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + \bar{\boldsymbol{U}} \cdot \boldsymbol{\nabla} w - \text{div}\,(\nu \boldsymbol{\nabla} w) + \frac{1}{\rho}\frac{\partial p_d}{\partial z} = f_z \quad (5.47)$$

$$\frac{\partial \eta}{\partial t} + \text{div}_2\,(\int_b^\eta \boldsymbol{u}\, dz) = 0 \quad (5.48)$$

$$\text{div}_2\ \boldsymbol{u} + \frac{\partial w}{\partial z} = 0, \quad (5.49)$$

where p_d denotes the hydrodynamic pressure correction.

Here again we explicit the system for more clarity :

$$\begin{cases} \dfrac{\partial u}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial x} + (w-c)\dfrac{\partial u}{\partial z} - \text{div}\,(\nu \boldsymbol{\nabla} u) + g\dfrac{\partial \eta}{\partial x} + \dfrac{\partial p_d}{\partial x} = f_x^{hor}, \\ \dfrac{\partial v}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} + (w-c)\dfrac{\partial v}{\partial z} - \text{div}\,(\nu \boldsymbol{\nabla} v) + g\dfrac{\partial \eta}{\partial y} + \dfrac{\partial p_d}{\partial y} = f_y^{hor}, \\ \dfrac{\partial w}{\partial t}\Big|_{\hat{\boldsymbol{x}}} + u\dfrac{\partial w}{\partial x} + v\dfrac{\partial w}{\partial y} + (w-c)\dfrac{\partial w}{\partial z} - \text{div}\,(\nu \boldsymbol{\nabla} w) + \dfrac{\partial p_d}{\partial z} = f_z, \\ \dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}(\int_b^\eta u\, dz) + \dfrac{\partial}{\partial y}(\int_b^\eta v\, dz) = 0, \\ \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z} = 0, \end{cases}$$

The non-hydrostatic model is more complex but it allows an accurate computation of the vertical velocity, which is now coupled with the whole system.

The decomposition of the pressure p as $p = p_d + p_h$, where p_h is the hydrostatic pressure given by (5.35), simplifies the resolution of the problem.

Indeed, it leads to the splitting of the time-discrete equations into :

1. a hydrostatic part, which corresponds almost exactly to the hydrostatic problem (5.38)-(5.39), with the exception that the vertical velocity is also advected and diffused. This step provides the free surface function.
2. a hydrodynamic part, in which the total velocity $\boldsymbol{U} = (u, v, w)$ is corrected by the hydrodynamic pressure term p_d, ensuring the continuity equation.

The 3D hydrostatic model can therefore either be viewed as an independent model or as an intermediate step in the solution of the full 3D Navier Stokes equations with free surface. The study of the 3D hydrostatic model is one of the focus of this thesis and is addressed in details in Chapters 6 and 7.

5.11 Weak formulation of the three-dimensional hydrostatic problem

For the sake of simplicity, we will assume the following regularity properties on the domain at any time $t \in I$:

$$\eta(t) \in W^{1,\infty}(\omega), \quad b \in W^{1,\infty}(\omega), \tag{5.50}$$

and

$$\exists\, h_0 > 0 \quad \text{such that} \quad h(t) \geq h_0 \quad \text{on } \omega. \tag{5.51}$$

In addition, we will assume that the Dirichlet condition (5.16) on the velocity at the inflow is homogeneous, *i.e.* that

$$\boldsymbol{U}(x, y, z, t) = 0 \qquad \text{on} \quad \Gamma_{in,t}. \tag{5.52}$$

Let us consider the following functional spaces defined on the reference configuration $\hat{\Omega}$:

$$\hat{\mathcal{X}} = H^1(\hat{\Omega})^2, \quad \hat{\mathcal{V}} = L^2(\hat{\Omega}) \quad \text{and} \quad \hat{\mathcal{W}} = H^1(\hat{\Omega}), \tag{5.53}$$

as well as the space

$$\mathcal{M} = L^2(\omega) \tag{5.54}$$

defined on the two-dimensional domain. The ALE mapping (5.25) then determines the corresponding spaces $\mathcal{X}(\Omega_t)$ and $\mathcal{X}(\Omega_t, t)$, as well as $\mathcal{W}(\Omega_t)$ and $\mathcal{V}(\Omega_t, t)$, defined on the current domain – see (2.13) and (2.15) in section 2.1.5.

In addition, we define

$$\begin{aligned}\mathcal{X}_0(\Omega_t) \;=\; \{\; \boldsymbol{\psi} \in \mathcal{X}(\Omega_t) \;\;/\;\; \boldsymbol{\psi}(t)\cdot \boldsymbol{n}_{xy} \;=\; 0 \text{ on } \Gamma_{sol,t} \;\text{ and}\\ \boldsymbol{\psi}(t) \;=\; 0 \text{ on } \Gamma_{in,t}\,,\; t\in I \;\}\end{aligned} \tag{5.55}$$

and

$$\begin{aligned}\mathcal{X}_0(\Omega_t,t) \;=\; \{\; \boldsymbol{\psi} \in \mathcal{X}(\Omega_t,t) \;\;/\;\; \boldsymbol{\psi}(t)\cdot \boldsymbol{n}_{xy} \;=\; 0 \text{ on } \Gamma_{sol,t} \;\text{ and}\\ \boldsymbol{\psi}(t) \;=\; 0 \text{ on } \Gamma_{in,t}\,,\; t\in I \;\}\end{aligned} \tag{5.56}$$

We will prove later on that these spaces are regular enough for the following weak formulation of the hydrostatic system (5.38)-(5.40) to be well-defined. For now, let us state the following proposition :

Proposition 5.2 *A weak ALE formulation of the 3D hydrostatic free surface flow problem is, given the initial conditions* $\boldsymbol{u}(0)$ *and* $\eta(0)$*,*

$$\begin{aligned}&\textit{find } \boldsymbol{u}\in\mathcal{X}_0(\Omega_t,t),\; \bar{w}\in\mathcal{V}(\Omega_t,t) \textit{ and } \eta(t)\in\mathcal{M} \textit{ such that}\\ &\forall\,\boldsymbol{\psi}\in\mathcal{X}_0(\Omega_t),\; \forall\,\chi\in\mathcal{M},\; \forall\,\phi\in\mathcal{W}(\Omega_t) \textit{ and for each } t\in I,\end{aligned}$$

$$\left\{\begin{array}{l}\displaystyle\int_{\Omega_t}\left(\frac{\partial \boldsymbol{u}}{\partial t} \;+\; \bar{\boldsymbol{U}}\cdot\boldsymbol{\nabla}\boldsymbol{u}\right)\cdot\boldsymbol{\psi}\;d\boldsymbol{x} \;+\; \int_{\Omega_t}\nu\boldsymbol{\nabla}\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{\psi}\;d\boldsymbol{x}\\ \qquad\qquad - g\displaystyle\int_\omega \eta\; div_2\,(\int_b^\eta \boldsymbol{\psi}\;dz)\;d\omega \;=\; <\boldsymbol{F},\boldsymbol{\psi}>_{\mathcal{X}'\times\mathcal{X}}\,,\\[2ex] \qquad g\displaystyle\int_\omega \frac{\partial\eta}{\partial t}\;\chi\;d\omega \;+\; g\int_\omega \chi\; div_2\,(\int_b^\eta \boldsymbol{u}\;dz)\;d\omega \;=\; 0\,,\\[2ex] \displaystyle\int_{\Omega_t}\bar{\boldsymbol{U}}\cdot\boldsymbol{\nabla}\phi\;d\boldsymbol{x} \;-\; \int_{\Gamma_{liq,t}}\phi\;\boldsymbol{u}\cdot\boldsymbol{n}_{xy}\;d\Gamma \;=\; \frac{d}{dt}\int_{\Omega_t}\phi\;d\boldsymbol{x}\,,\end{array}\right. \tag{5.57}$$

where

$$\begin{aligned}<\boldsymbol{F},\boldsymbol{\psi}>_{\mathcal{X}'\times\mathcal{X}} \;=\; &\int_{\Omega_t}\boldsymbol{f}^{hor}\cdot\boldsymbol{\psi}\;d\boldsymbol{x} \;-\; \int_{\Gamma_b\cup\Gamma_{sol,t}}\kappa(\boldsymbol{u})\;\boldsymbol{\psi}\;d\Gamma\\ &+\int_{\Gamma_{out,t}}\eta_{out}\;\boldsymbol{\psi}\cdot\boldsymbol{n}_{xy}\;d\Gamma.\end{aligned} \tag{5.58}$$

Proof.

Momentum equation Let us consider a test function $\boldsymbol{\psi} = (\psi_u, \psi_v)$ in $\mathcal{X}_0(\Omega_t)$. We multiply equation (5.38) by ψ and we integrate on Ω_t. Then, classical integration by parts of the flux term gives

$$\int_{\Omega_t} \left(\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{U} - \boldsymbol{c}) \cdot \boldsymbol{\nabla} \boldsymbol{u} \right) \cdot \boldsymbol{\psi} \; d\boldsymbol{x} \; + \; \int_{\Omega_t} \nu \boldsymbol{\nabla} \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{\psi} \; d\boldsymbol{x}$$
$$+ \, g \int_{\Omega_t} \boldsymbol{\nabla}_2 \eta \boldsymbol{\psi} \; d\boldsymbol{x} \; = \; \int_{\partial \Omega_t} (\nu \boldsymbol{\nabla} \boldsymbol{u} \cdot \boldsymbol{n}) \cdot \boldsymbol{\psi} \; d\Gamma \; + \; \int_{\Omega_t} \boldsymbol{f}^{hor} \cdot \boldsymbol{\psi} \; d\boldsymbol{x}.$$

We impose strongly the homogeneous Dirichlet condition (5.52) on the horizontal velocity at the inflow boundary. In addition, we use the dynamic boundary conditions (5.12) at the bottom and the solid lateral boundaries, (5.15) at the free surface, and (5.18) at the outflow. We obtain :

$$\int_{\partial \Omega_t} (\nu \boldsymbol{\nabla} \boldsymbol{u} \cdot \boldsymbol{n}) \cdot \boldsymbol{\psi} \; d\Gamma \; = \; - \int_{\Gamma_b \cup \Gamma_{sol,t}} \kappa(\boldsymbol{u}) \; \boldsymbol{\psi} \; d\Gamma.$$

On the other hand, since η is defined on the two-dimensional domain ω, we can write

$$\int_{\Omega_t} \boldsymbol{\nabla}_2 \eta \cdot \boldsymbol{\psi} \; d\boldsymbol{x} \; = \; \int_{\omega} \boldsymbol{\nabla}_2 \eta \cdot \left(\int_b^{\eta} \boldsymbol{\psi} \; dz \right) \; d\omega,$$

and integrating by parts this term we get :

$$\int_{\omega} \boldsymbol{\nabla}_2 \eta \cdot \left(\int_b^{\eta} \boldsymbol{\psi} \; dz \right) \; d\omega$$
$$= - \int_{\omega} \eta \, \mathrm{div}_2 \, (\int_b^{\eta} \boldsymbol{\psi} \; dz) \; d\omega \; + \; \int_{\partial \omega} \eta \left(\int_b^{\eta} \boldsymbol{\psi} \; dz \right) \cdot \boldsymbol{n}_{\omega} \; d\gamma.$$

Boundary $\partial\omega$ corresponds to the two-dimensional basis of the lateral boundaries $\Gamma_{l,t}$. Moreover, according to remark (5.4.1), $\boldsymbol{n}_{\omega} = \boldsymbol{n}_{xy}$ on $\Gamma_{l,t}$. Therefore we have that

$$\int_{\partial \omega} \eta \left(\int_b^{\eta} \boldsymbol{\psi} \; dz \right) \cdot \boldsymbol{n}_{\omega} \; d\gamma \; = \; \int_{\partial \omega} \int_b^{\eta} \eta \; \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \; dz \; d\gamma \; = \; \int_{\Gamma_{l,t}} \eta \; \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \; d\Gamma.$$

We impose strongly the imperviousness condition on the solid lateral boundaries (5.11). Using in addition the prescribed free surface condition at the outflow we obtain :

$$\int_{\Gamma_{l,t}} \eta \; \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \; d\Gamma \; = \; \int_{\Gamma_{out,t}} \eta_{out} \; \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \; d\Gamma.$$

This leads to the weak formulation of the momentum equation in (5.57).

Free surface equation Let χ be a function of $\mathcal{M}$. Multiplying equation (5.39) by $g\ \chi$ and integrating on ω, we obtain the weak formulation of the free surface equation in (5.57).

Continuity equation We now multiply equation (5.45) by a test function $\hat{\phi}$ in $\hat{\mathcal{W}}$ and integrate on $\hat{\Omega}$:

$$\int_{\hat{\Omega}} \frac{\partial \hat{J}_t}{\partial t}\ \hat{\phi}\ d\hat{\boldsymbol{x}}\ +\ \int_{\hat{\Omega}} \hat{J}_t\ \left(\operatorname{div} \bar{\boldsymbol{U}} \circ \hat{\mathcal{A}}_t\right)\ \hat{\phi}\ d\hat{\boldsymbol{x}}\ =\ 0.$$

Since $\hat{\phi}$ and $\hat{\Omega}$ do not depend on time we can write that

$$\int_{\hat{\Omega}} \frac{\partial \hat{J}_t}{\partial t}\ \hat{\phi}\ d\hat{\boldsymbol{x}}\ =\ \frac{d}{dt}\int_{\hat{\Omega}} \hat{J}_t\ \hat{\phi}\ d\hat{\boldsymbol{x}}\ =\ \frac{d}{dt}\int_{\Omega_t} \phi\ d\boldsymbol{x}, \qquad \text{with} \quad \phi\ =\ \hat{\phi} \circ \mathcal{A}_t^{-1}.$$

The divergence term is then integrated by parts :

$$\int_{\Omega_t} \phi\ \operatorname{div} \bar{\boldsymbol{U}}\ d\boldsymbol{x}\ =\ -\int_{\Omega_t} \bar{\boldsymbol{U}} \cdot \boldsymbol{\nabla}\phi\ d\boldsymbol{x}\ +\ \int_{\partial\Omega_t} \phi\ \bar{\boldsymbol{U}} \cdot \boldsymbol{n}\ d\Gamma.$$

On the lateral boundaries, using Remark 5.4.1 and condition (5.11) imposed strongly :

$$\int_{\Gamma_{l,t}} \phi\ \bar{\boldsymbol{U}} \cdot \boldsymbol{n}\ d\Gamma\ =\ \int_{\Gamma_{l,t}} \phi\ \boldsymbol{u} \cdot \boldsymbol{n}_{xy}\ d\Gamma\ =\ \int_{\Gamma_{liq,t}} \phi\ \boldsymbol{u} \cdot \boldsymbol{n}_{xy}\ d\Gamma.$$

Moreover, we recall that the ALE mapping $\hat{\mathcal{A}}_t$ is such that the velocity of the domain satisfies

$$\bar{\boldsymbol{U}} \cdot \boldsymbol{n} =\ (\boldsymbol{U} - \boldsymbol{c}) \cdot \boldsymbol{n}\ =\ 0$$

at the impervious boundaries of the domain. Especially, this holds true at the free surface $\Gamma_{s,t}$ and at the bottom Γ_b.

Remark 5.11.1 Note that this comes to impose weakly the imperviousness conditions (5.13) at the free surface and (5.10) on the bottom.

The boundary term then reduces to

$$\int_{\partial\Omega_t} \phi\ \bar{\boldsymbol{U}} \cdot \boldsymbol{n}\ d\Gamma\ =\ \int_{\Gamma_{liq,t}} \phi\ \boldsymbol{u} \cdot \boldsymbol{n}_{xy}\ d\Gamma,$$

and we finally get the weak formulation of the continuity equation in (5.57). ♢

Let us now prove that the weak formulation is well-defined for the continuous functional spaces $\hat{\mathcal{X}}$, $\hat{\mathcal{V}}$, $\hat{\mathcal{W}}$ and $\mathcal{M}$ introduced in (5.53) and (5.54).

For that purpose, we state two functional analysis results, valid for any bounded moving domain $\Omega_t \cup \mathbb{R}^d$, $d \geq 2$, with Lipschitz boundary Γ_t.

Lemma 5.1 *Following the trace theorem, at any time $t \in I$, there is a strictly positive constant $c_{T,t} > 0$, depending only on the current domain Ω_t, such that*

$$\forall \boldsymbol{\psi} \in H^1(\Omega_t)^2 \qquad \|\boldsymbol{\gamma}_0(\boldsymbol{\psi})\|_{0,\partial\Omega_t} \leq c_{T,t} \|\boldsymbol{\psi}\|_{1,\Omega_t}, \tag{5.59}$$

where $\boldsymbol{\gamma}_0$ is the continuous application associating to each element $\boldsymbol{\psi} \in H^1(\Omega_t)^2$ its trace on $\partial\Omega_t$.

We introduce, for any non empty part $\Gamma_{1,t}$ of the domain's boundary Γ_t, the functional space

$$H_{\Gamma_1}(\Omega_t) = \{ v \in H^1(\Omega), \quad v = 0 \quad \text{on} \quad \Gamma_1 \}.$$

Lemma 5.2 *Following Poincaré's Inequality, at any time $t \in I$, there is a strictly positive constant $c_{P,t} > 0$, depending only on the current domain Ω_t, such that*

$$\forall \boldsymbol{\psi} \in H_{\Gamma_1}(\Omega_t)^2 \qquad \|\boldsymbol{\psi}\|_{0,\Omega_t} \leq c_{P,t} \|\boldsymbol{\nabla}\boldsymbol{\psi}\|_{0,\Omega_t}. \tag{5.60}$$

This shows that the semi-norm $|\cdot|_{1,\Omega_t}$ in $H^1(\Omega_t)^2$ is a norm in $H_0^1(\Omega_t)^2$.

There is also a strictly positive constant c_p, depending only on the two-dimensional domain ω, and such that

$$\forall \chi \in H_0^1(\omega), \quad \|\chi\|_{0,\omega} \leq c_p |\chi|_{1,\omega}. \tag{5.61}$$

Let us now prove the following proposition :

Proposition 5.3 *Assuming (5.50) and (5.51), for any $\boldsymbol{\psi}$ in $H^1(\Omega_t)^2$ and for all t in I,*

$$div_2 \left(\int_b^{\eta(t)} \boldsymbol{\psi} \, dz \right) \in L^2(\omega)^2$$

and there is a strictly positive constant $C(\Omega_t) > 0$, depending on the domain Ω_t at time t, such that

$$\| div_2 \left(\int_b^{\eta(t)} \boldsymbol{\psi} \, dz \right) \|_{0,\omega} \leq C(\Omega_t) \|\boldsymbol{\psi}\|_{1,\Omega_t}. \tag{5.62}$$

Proof.
Let $\boldsymbol{\psi}$ be a function in $H^1(\Omega_t)^2$:

$$\begin{aligned}
&\int_\omega |\ \mathrm{div}_2\ (\int_b^{\eta(t)} \boldsymbol{\psi}\ dz)\ |^2\ d\omega \\
=\ &\int_\omega |\int_b^{\eta(t)}\ \mathrm{div}_2\,\boldsymbol{\psi}\ dz\ +\ \boldsymbol{\nabla}_2\eta(t)\cdot\boldsymbol{\psi}_{|z=\eta}\ -\ \boldsymbol{\nabla}_2 b\cdot\boldsymbol{\psi}_{|z=b}\ |^2\ d\omega \\
\leq\ &2\int_\omega |\int_b^{\eta(t)}\ \mathrm{div}_2\,\boldsymbol{\psi}\ dz\,|^2\ d\omega\ +\ 2\int_\omega |\ \boldsymbol{\nabla}_2\eta(t)\cdot\boldsymbol{\psi}_{|z=\eta} \\
&\qquad\qquad - \boldsymbol{\nabla}_2 b\cdot\boldsymbol{\psi}_{|z=b}\ |^2\ d\omega \\
\leq\ &2\ \|h(t)\|_{L^\infty(\omega)}\ \|\,\mathrm{div}_2\ \boldsymbol{\psi}\ \|^2_{0,\Omega_t}\ +\ 4\ \|\boldsymbol{\nabla}_2\eta(t)\|^2_{L^\infty(\omega)}\ \|\boldsymbol{\psi}_{|z=\eta}\|^2_{0,\omega} \\
&+ 4\ \|\boldsymbol{\nabla}_2 b\|^2_{L^\infty(\omega)}\ \|\boldsymbol{\psi}_{|z=b}\|^2_{0,\omega}\,.
\end{aligned}$$

Now, according to lemma 5.1 we have that

$$\begin{aligned}
&\|\boldsymbol{\psi}_{|z=\eta}\|_{0,\omega} \leq c_{T,t}\ \|\boldsymbol{\psi}\|_{1,\Omega_t} \\
\text{and}\qquad &\|\boldsymbol{\psi}_{|z=b}\|_{0,\omega} \leq c_{T,t}\ \|\boldsymbol{\psi}\|_{1,\Omega_t}\,.
\end{aligned}$$

Finally, using that

$$\|\,\mathrm{div}_2\ \boldsymbol{\psi}\ \|_{0,\Omega_t}\ \leq\ 2\ \|\boldsymbol{\nabla}_2\boldsymbol{\psi}\|_{0,\Omega_t}\,,$$

we obtain

$$\begin{aligned}
\|\,\mathrm{div}_2\ (\int_b^{\eta(t)} \boldsymbol{\psi}\ dz)\ \|^2_{0,\omega}\ \leq\ 4\ \Big(\ &2\ \|h(t)\|_{L^\infty(\omega)} \\
&+ c^2_{T,t}\left(\|\boldsymbol{\nabla}_2\eta(t)\|^2_{L^\infty(\omega)}\ +\ \|\boldsymbol{\nabla}_2 b\|^2_{L^\infty(\omega)}\right)\Big)\ \|\boldsymbol{\psi}\|^2_{1,\Omega_t}\,.
\end{aligned}$$

$\diamondsuit$

Following Proposition 5.3, if at each time $t \in I$

$$\boldsymbol{u}(t) \in H^1(\Omega_t)^2, \qquad \eta(t) \in L^2(\omega) \qquad \text{and} \qquad \bar{w}(t) \in L^2(\Omega_t),$$

and

$$\boldsymbol{\psi}(t) \in H^1(\Omega_t)^2, \qquad \chi(t) \in L^2(\omega) \qquad \text{and} \qquad \phi(t) \in H^1(\Omega_t),$$

then the weak formulation (5.57) of the hydrostatic problem is well defined.

According to (2.19) and (2.20), the following properties hold :

$$\forall\, \boldsymbol{u} \in \mathcal{X}_0(\Omega_t, t), \quad \boldsymbol{u}(t) \in H^1(\Omega_t)^2 \quad \forall t \ \in I\,, \tag{5.63}$$

$$\forall\, \boldsymbol{\psi} \in \mathcal{X}_0(\Omega_t), \quad \boldsymbol{\psi}(t) \in H^1(\Omega_t)^2 \quad \forall t \ \in I\,, \tag{5.64}$$

$$\forall\, \bar{w} \in \mathcal{V}(\Omega_t, t), \quad \bar{w}(t) \in L^2(\Omega_t), \quad \forall t \ \in I\,, \tag{5.65}$$

$$\forall\, \phi \in \mathcal{W}(\Omega_t), \quad \phi(t) \in H^1(\Omega_t) \quad \forall t \ \in I\,. \tag{5.66}$$

5.12 Time discretization and linearization

Let us now introduce the time-advancing scheme chosen for the hydrostatic problem. A discussion on the proper choice of the terms to discretize implicitly rather than explicitly in a three-dimensional shallow water flow problem can be found in [26]. The authors show through numerical experiments that the semi-implicit discretization of the free surface gradient and velocity divergence terms, and the explicitation of the remaining terms is stable and efficient. We will follow this method. Note that the explicitation of the advection terms allows their treatment through the ALE-MURD schemes introduced in Chapter 3. The results can be improved by the use of the so-called θ-method (ref).

The time interval I is divided in N_t time steps of equal length Δt. We denote by Ω^n the approximation of the domain occupied by the fluid at time $t^n = n\,\Delta t$, for $n = 0, \ldots, N_t$, and by $(\,\boldsymbol{u}^n,\, \eta^n,\, \bar{w}^n)$ the approximation of the solution $(\,\boldsymbol{u}(t^n),\, \eta(t^n),\, \bar{w}(t^n)\,)$. Let us introduce the following application :

$$\mathcal{A}_{n,n+1} : \Omega^n \longrightarrow \Omega^{n+1}, \qquad \mathcal{A}_{n,n+1}(\boldsymbol{x}) \ = \ \hat{\mathcal{A}}_{n+1} \circ \mathcal{A}_n^{-1}(\boldsymbol{x})\,.$$

We introduce the following time-discrete functional spaces

$$\begin{aligned}
\mathcal{X}^n &= \{\, \boldsymbol{\psi} : \Omega^n \longrightarrow \mathbb{R}, \quad \boldsymbol{\psi} \ = \ \hat{\boldsymbol{\psi}} \circ \mathcal{A}_n^{-1}, \quad \hat{\boldsymbol{\psi}} \in \hat{\mathcal{X}} \,\}\,, \\
\mathcal{V}^n &= \{\phi : \Omega^n \longrightarrow \mathbb{R}, \quad \phi \ = \ \hat{\phi} \circ \mathcal{A}_n^{-1}, \quad \hat{\phi} \in \hat{\mathcal{V}} \,\}\,, \\
\mathcal{W}^n &= \{\phi : \Omega^n \longrightarrow \mathbb{R}, \quad \phi \ = \ \hat{\phi} \circ \mathcal{A}_n^{-1}, \quad \hat{\phi} \in \hat{\mathcal{W}} \,\}\,,
\end{aligned} \tag{5.67}$$

as well as

$$\mathcal{X}_0^n \ = \ \{\, \boldsymbol{\psi} \in \mathcal{X}^n \ / \ \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \ = \ 0 \text{ on } \Gamma_{sol}^n \quad \text{and} \quad \boldsymbol{\psi} \ = \ 0 \text{ on } \Gamma_{in}^n \,\}\,. \tag{5.68}$$

Applying Proposition 2.1, we can show that :

$$\mathcal{X}^n \;=\; H^1(\Omega^n)^2, \quad \mathcal{V}^n \;=\; L^2(\Omega^n) \quad \text{and} \quad \mathcal{W}^n \;=\; H^1(\Omega^n)\;. \quad (5.69)$$

Let us now define some notations we will use in order to simplify the time-discretized weak formulation of the problem. We will denote by $(.,.)_{\Omega^n}$ the scalar product in $L^2(\Omega^n)^2$:

$$(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,)_{\,\Omega^n} \;=\; \int_{\Omega^n} \boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_2 \; d\boldsymbol{x}\,,$$

and by $(.,.)_\omega$ the scalar product in $L^2(\omega)$:

$$(\,\chi_1\,,\,\chi_2\,)_{\,\omega} \;=\; \int_{\omega} \chi_1 \; \chi_2 \; d\omega.$$

Moreover, we will make use of the following shortcut, valid for any time-discrete function f such that, for each $n = 0, \ldots, N_t$, f^n is defined on Ω^n

$$\text{for any } (n_1, n_2) \in [1, N_t]^2, \quad f^{n_1}_{n_2} \;=\; f^{n_1} \circ \mathcal{A}_{n_2,n_1} \quad \text{in} \quad \Omega^{n_2}.$$

Finally, we will denote, for any time independent function $\hat{\boldsymbol{\psi}}$ defined on the reference configuration $\hat{\Omega}$:

$$\boldsymbol{\psi}^n \;=\; \hat{\boldsymbol{\psi}} \circ \hat{\mathcal{A}}_n \quad \text{in} \quad \Omega^n, \quad \text{for } n = 0, \ldots, N_t.$$

We point out that the time-discretization chosen leads to a linear system. Moreover, the explicit time-discretization of the advective term uncouples the continuity equation from the two other ones. Indeed, at each time t^n, the system can be solved in two steps : first, the sub-system formed by the momentum equation and the free surface equation is solved, determining the horizontal velocity $\boldsymbol{u}$ and the free surface function η at $t = t^{n+1}$. Then, the continuity equation is solved independently, by deriving the vertical ALE advective velocity $\bar{w}$ at t^{c_n} from the values $\boldsymbol{u}^{c_n}$ and $\eta^{\,c_n}$ just computed.

Given the initial conditions $\boldsymbol{u}_0$ and η_0, we consider the following semi-implicit time-advancing scheme for problem (5.57) :

$$\text{for each} \quad n = 0, \ldots, N_t - 1,$$

$$\text{find} \quad \boldsymbol{u}^{n+1} \in \mathcal{X}_0^{n+1}, \quad \eta^{n+1} \in \mathcal{M} \quad \text{and} \quad \bar{w}^{n+1} \in \mathcal{V}^{n+1} \quad \text{such that,}$$

$$\forall\, \hat{\boldsymbol{\psi}} \in \hat{\mathcal{X}}_0, \quad \forall\, \chi \in \mathcal{M} \quad \text{and} \quad \forall \hat{\phi} \in \hat{\mathcal{W}},$$

$$\left\{\begin{array}{l} \dfrac{1}{\Delta t}\ (\,\boldsymbol{u}_n^{n+1}\ -\ \boldsymbol{u}^n\,,\,\boldsymbol{\psi}^n\,)_{\Omega^n}\quad +\quad \displaystyle\int_{\Omega^{a_n}} (\bar{\boldsymbol{U}}^{a_n}\cdot\boldsymbol{\nabla}\boldsymbol{u}_{a_n}^n)\cdot\boldsymbol{\psi}^{a_n}\ d\boldsymbol{x} \\ \\ +\quad \nu\ (\,\boldsymbol{\nabla}\boldsymbol{u}_n^{n+1}\,,\,\boldsymbol{\nabla}\boldsymbol{\psi}^n\,)_{\Omega^n}\quad -\quad g\displaystyle\int_\omega \eta^{\theta_n}\ \mathrm{div}_2\ (\int_b^{\eta^n}\boldsymbol{\psi}^n\ dz)\ d\omega \\ \\ \qquad\qquad\qquad\qquad\qquad\qquad =\quad <\boldsymbol{F}^{n+1},\boldsymbol{\psi}^n>_{(\mathcal{X}^n)'\times\mathcal{X}^n}\,, \\ \\ \dfrac{g}{\Delta t}\ (\,\eta^{n+1}\ -\ \eta^n\,,\,\chi\,)_\omega\quad +\quad g\displaystyle\int_\omega \chi\ \mathrm{div}_2\ (\int_b^{\eta^n}\boldsymbol{u}_n^{\theta_n}\ dz)\ d\omega\ =\ 0\,, \end{array}\right. \tag{5.70}$$

$$\left\{\begin{array}{l} \displaystyle\int_{\Omega^{c_n}}\bar{w}^{c_n}\ \frac{\partial\phi^{c_n}}{\partial z}\ d\boldsymbol{x}\quad =\quad \frac{1}{\Delta t}\ (\int_{\Omega^{n+1}}\phi^{n+1}\ d\boldsymbol{x}\ -\ \int_{\Omega^n}\phi^n\ d\boldsymbol{x}\,) \\ \\ \qquad -\quad \displaystyle\int_{\Omega^{c_n}}\boldsymbol{u}^{c_n}\cdot\boldsymbol{\nabla}_2\phi^{c_n}\ d\boldsymbol{x}\quad +\quad \int_{\Gamma_{liq}^{c_n}}\phi^{c_n}\ \boldsymbol{u}^{c_n}\cdot\boldsymbol{n}_{xy}\ d\Gamma\,, \end{array}\right. \tag{5.71}$$

where t^{a_n} is a particular time in $[t^{n-1},t^n]$ and t^{θ_n}, t^{c_n} are two particular times in $[t^n,t^{n+1}]$, chosen in order to satisfy some properties. Time t^{θ_n} is determined by an implicitness parameter $0\le\theta\le 1$, defining

$$t^{\theta_n}=\theta\ t^{n+1}\ +\ (1-\theta)\ t^n$$

for each $n=0,\ldots,N_t-1$. The corresponding variables are then

$$\eta^{\theta_n}\ =\ \theta\ \eta^{n+1}\ +\ (1-\theta)\ \eta^n\qquad\text{and}\qquad \boldsymbol{u}^{\theta_n}\ =\ \theta\ \boldsymbol{u}^{n+1}\ +\ (1-\theta)\ \boldsymbol{u}^n.$$

Finally,

$$\begin{array}{rcl} <\boldsymbol{F}^{n+1},\boldsymbol{\psi}>_{(\mathcal{X}^n)'\times\mathcal{X}^n} & = & \displaystyle\int_{\Omega^n}(\boldsymbol{f}^{hor})_n^{n+1}\cdot\boldsymbol{\psi}\ d\boldsymbol{x}\ -\ \int_{\Gamma_b\cup\Gamma_{sol}^n}\kappa(\boldsymbol{u}^n)\ \boldsymbol{\psi}\ d\Gamma \\ & & \displaystyle -\int_{\Gamma_{out}^n}\eta_{out}^{n+1}\ \boldsymbol{\psi}\cdot\boldsymbol{n}_{xy}\ d\Gamma\,. \end{array} \tag{5.72}$$

Remark 5.12.1 The uncoupling of the vertical velocity from the horizontal velocity and the free surface function simplifies the solution, but it has an important drawback, which is inherent to the hydrostatic model : all inaccuracy in the computation of the horizontal velocity will result in errors on the vertical ALE advective velocity.

5.13 Fractional step method

In order to simplify the analysis and the approximation of the hydrostatic problem, the first oder in time fractional step method is applied : the momentum equation on the horizontal velocity $\boldsymbol{u}$ is splitted so that the ALE advective term is isolated from the rest. Consequently, system (5.70)-(5.71) is separated into three parts, which can be treated separately using the most appropriate numerical methods according to the mathematical properties of the operators.

First, $\boldsymbol{u}$ is advected, leading to an intermediate value $\tilde{\boldsymbol{u}}$:

$$(\tilde{\boldsymbol{u}} \; - \; \boldsymbol{u}^n \, , \, \boldsymbol{\psi}^n \,)_{\Omega^n} \; + \; \int_{\Omega^{\,a_n}} (\bar{\boldsymbol{U}}^{a_n} \cdot \boldsymbol{\nabla} \boldsymbol{u}^n_{a_n}) \cdot \boldsymbol{\psi}^{a_n} \; d\boldsymbol{x} \;\; = \;\; 0. \tag{5.73}$$

The hyperbolic nature of the advection equations implies that the space derivatives must be upwinded. Appropriate numerical methods for solving these equations in the finite element framework are for example the characteristic methods (see for instance [112]), the SUPG method [71] and the MURD schemes (see for instance [107]). Note that the latter are extended to moving domains in Chapter 3, leading to ALE-MURD schemes. The application of these schemes to the advection equations present in the 3D hydrostatic free surface flow problem will be discussed later on in this work – see section 7.5 in the next chapter.

Then, the rest of the momentum equation coupled with the free surface equation is solved, determining $\boldsymbol{u}^{n+1}$ and η^{n+1} :

$$\left\{ \begin{array}{l} \dfrac{1}{\Delta t} \;\; (\, \boldsymbol{u}^{n+1}_n \;\; - \;\; \tilde{\boldsymbol{u}} \, , \, \boldsymbol{\psi}^n \,)_{\Omega^n} \;\; + \;\; \nu \; (\, \boldsymbol{\nabla} \boldsymbol{u}^{n+1}_n \, , \, \boldsymbol{\nabla} \boldsymbol{\psi}^n \,)_{\Omega^n} \\ \\ - \; g \displaystyle\int_\omega \eta^{\theta_n} \; \mathrm{div}_2 \; (\int_b^{\eta^n} \boldsymbol{\psi}^n \; dz) \; d\omega \;\; = \;\; < \boldsymbol{F}^{n+1}, \boldsymbol{\psi}^n >_{(\mathcal{X}^n)' \times \mathcal{X}^n} , \\ \\ \dfrac{g}{\Delta t} \;\; (\, \eta^{n+1} \;\; - \;\; \eta^n \, , \, \chi \,)_\omega \;\; + \;\; g \displaystyle\int_\omega \chi \; \mathrm{div}_2 \; (\int_b^{\eta^n} \boldsymbol{u}^{\theta_n}_n \; dz) \; d\omega \;\; = \;\; 0 \, . \end{array} \right. \tag{5.74}$$

Finally, the continuity equation gives the value of $\bar{w}^{c_n}$:

$$\begin{array}{rcl} \displaystyle\int_{\Omega^{\,c_n}} \bar{w}^{c_n} \; \frac{\partial \phi^{c_n}}{\partial z} \; d\boldsymbol{x} & = & \dfrac{1}{\Delta t} \;\; (\displaystyle\int_{\Omega^{n+1}} \phi^{n+1} \; d\boldsymbol{x} \;\; - \;\; \int_{\Omega^n} \phi^n \; d\boldsymbol{x} \,) \\ & - & \displaystyle\int_{\Omega^{\,c_n}} \boldsymbol{u}^{c_n} \cdot \boldsymbol{\nabla}_2 \phi^{c_n} \; d\boldsymbol{x} \;\; + \;\; \int_{\Gamma^{\,c_n}_{liq}} \phi^{c_n} \; \boldsymbol{u}^{c_n} \cdot \boldsymbol{n}_{xy} \; d\Gamma \, . \end{array} \tag{5.75}$$

The semi-discrete problem (5.74) will throughout be referenced to as the $\boldsymbol{u} - \eta$ *problem.*

Note that Miglio and Saleri proposed in [98] a semi-discrete weak formulation of a linearized version of the hydrostatic free surface problem, in which the horizontal diffusion was neglected. By treating the advection terms explicitly, they achieve a problem coupling the horizontal velocity and the free surface which is very similar to the $\boldsymbol{u} - \eta$ problem presented here. However, the problem they obtain is non-symmetric. The authors then proceed to a modification of the free surface equation in order to achieve a symmetric system. We point out that our formulation (5.74) of the problem coupling the horizontal velocity and the free surface is symmetric and based on the physical free surface equation.

In the next chapter we will analyse the $\boldsymbol{u} - \eta$ problem.

Chapitre 6

The $\boldsymbol{u} - \eta$ problem.

6.1 Introduction

In this chapter we aim to analyse problem (5.74), arising from the time-discretization of the three-dimensional free surface flow problem introduced previously, and coupling the horizontal velocity and the free surface function. We call this problem the $\boldsymbol{u} - \eta$ *problem*.

We wish to point out that Miglio and Saleri have introduced in [98] a similar problem, obtained from a semi-discrete weak formulation of the hydrostatic free surface problem, in which the horizontal diffusion was neglected and the advection terms treated explicitly. This problem also couples the horizontal velocity and the free surface but it is non-symmetric. The authors analyse the problem after modifying the free surface equation in order to achieve a symmetric system. An existence and uniqueness result is then derived in the mixed problem framework for the modified symmetric semi-discrete problem.

We underline that the $\boldsymbol{u} - \eta$ problem analysed in this chapter is symmetric and based on the physical free surface equation. In addition, it includes the horizontal diffusion terms.

For the sake of clarity, we will drop the time subscripts of the variables and functions. Consequently, we will consider a continuous time-independent problem. Note that this problem can be viewed as part of the hydrostatic as well as the non-hydrostatic free surface flow models, but it may also arise from completely different problems.

We will first analyse the continuous problem in the Lax-Milgram framework, yielding an existence and uniqueness result. However, this analysis only holds in the case where a particular parameter α, corresponding to the viscosity coefficient of the fluid in the free surface flow problem, is not zero. In addition, when this parameter gets very small, the a priori estimation of the solution deteriorates. Therefore we will also analyse the continous $\boldsymbol{u} - \eta$

problem in the mixed problem framework. This will allow us to prove the existence and uniqueness of a solution independently of the value of α.

Then, we will analyse an approximation of the $\boldsymbol{u}-\eta$ problem through the Galerkin method. The analysis is made in the Lax-Milgram framework first, yielding an existence and uniqueness result as well as an estimation of the error between the solution of the continous problem and its approximation. The analysis is then made in the mixed problem framework, leading to the proof of existence and uniqueness of a solution to the discrete $\boldsymbol{u}-\eta$ problem provided a particular discrete *inf-sup* condition is satisfied. In this case, an error estimation is given, valid for any value of the parameter α. We also propose a pair of stable discrete spaces approximating the horizontal velocity and the free surface. That means that the discrete *inf-sup condition* required is satisfied for these spaces.

Finally, we introduce an alternative approximation of the continuous $\boldsymbol{u}-\eta$ problem, allowing the weak treatment of the imperviousness condition at the lateral boundaries. The existence and uniqueness of a solution to this discrete problem can be proven in the Lax-Milgram framework for the case where α is non zero. It can also be proven in the mixed problem framework, independently of α, provided a particular discrete *inf-sup* condition is satisfied. However, since this is not an internal approximation of the continuous weak formulation, we will not provide an estimation of the approximation error. A complex analysis using non-standard techniques would be needed for such an estimation.

6.2 The continuous $\boldsymbol{u}-\eta$ problem.

We consider in this chapter a generic form of the space-continuous weak $\boldsymbol{u}-\eta$ problem (5.74) arising from a time discretization of the three-dimensional hydrostatic model (5.38)–(5.40) – or from the non-hydrostatic model (5.46)–(5.40) – introduced in the previous chapter.

We recall that these models are posed on three-dimensional domains of the form

$$\Omega_t = \{\boldsymbol{x} = (x,y,z) \ / \ (x,y) \in \omega \quad \text{and} \quad b\,(x,y) \leq z \leq \eta\,(x,y,t)\,\}\,,$$

where ω is the two-dimensional domain – a fixed bounded region of $\mathbb{R}^2$ –, and where $\eta(x,y,t)$ and $b(x,y)$ are two functions defined on ω and describing respectively the free surface and the bottom (see Figure 6.1). We also recall that the boundary of this domain is denoted by $\partial\Omega_t$ and can be decomposed into three separate parts, as illustrated in Figure 6.1 : the free surface $\Gamma_{s,t}$, the bottom surface Γ_b and the lateral boundaries $\Gamma_{l,t}$. The latter can include solid lateral boundaries $\Gamma_{sol,t}$, an *inflow* boundary $\Gamma_{in,t}$ and an *outflow*

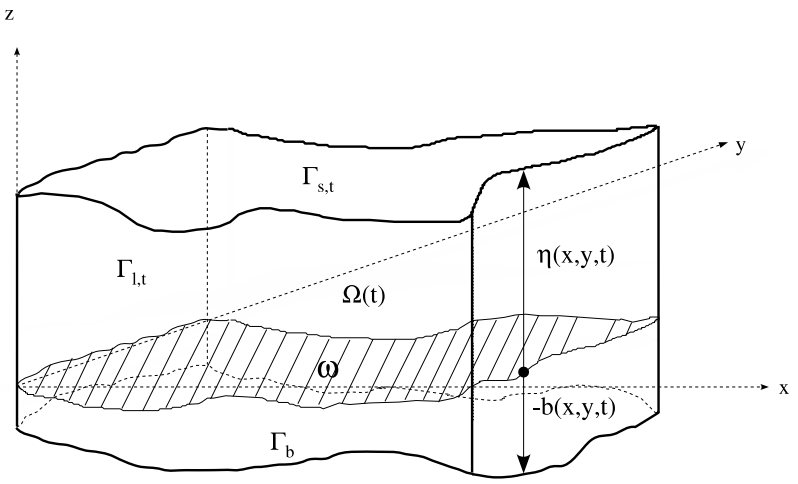


FIG. 6.1 – The three-dimensional domain and its boundaries.

boundary $\Gamma_{out,t}$.

For the sake of clarity, let us recall the ALE formulation of the equations describing the hydrostatic free surface flow model, valid for any $t \in I$ and for all $\boldsymbol{x}$ in the three-dimensional domain Ω_t :

$$\left\{ \begin{array}{rcl} \left.\dfrac{\partial \boldsymbol{u}}{\partial t}\right|_{\hat{\boldsymbol{x}}} + (\boldsymbol{U}-\boldsymbol{c})\cdot\boldsymbol{\nabla}\boldsymbol{u} - \operatorname{div}(\nu\boldsymbol{\nabla}\boldsymbol{u}) + g\,\boldsymbol{\nabla}_2\eta & = & \boldsymbol{f}^{hor}, \\ \dfrac{\partial \eta}{\partial t} + \operatorname{div}_2\left(\displaystyle\int_b^\eta \boldsymbol{u}\, dz\right) & = & 0, \\ \operatorname{div}_2\,\boldsymbol{u} + \dfrac{\partial w}{\partial z} & = & 0, \end{array} \right. \tag{6.1}$$

where $\boldsymbol{U}$ is the velocity of the fluid, $\boldsymbol{u}$ and w its horizontal and vertical components, η is the free surface variable and $\boldsymbol{c} = (O, O, c)^T$ the velocity of the domain. In addition, $\boldsymbol{f}^{hor}$ is the horizontal source term, comprising the external forces applied on the fluid as well as the bouyancy terms (5.37). This system is closed by the following boundary conditions – see section 5.4 :

$$\left\{\begin{array}{rcl}
\boldsymbol{U}\cdot\boldsymbol{n} &=& 0 \quad \text{on} \quad \Gamma_b\,,\\
\boldsymbol{u}\cdot\boldsymbol{n}_{xy} &=& 0 \quad \text{on} \quad \Gamma_{sol,t}\,,\\
\nu\,\dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} &=& -\,\kappa(\boldsymbol{u}) \quad \text{on} \quad \Gamma_b \cup \Gamma_{sol,t}\,,\\
\dfrac{\partial \eta}{\partial t} - \boldsymbol{U}\cdot\boldsymbol{n} &=& 0 \quad \text{on} \quad \Gamma_{s,t}\,,\\
\nu\dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} &=& 0 \quad \text{on} \quad \Gamma_{s,t}\,,\\
\boldsymbol{U}(x,y,z,t) &=& \boldsymbol{U}_{in}(x,y,z,t) \quad \text{on} \quad \Gamma_{in,t}\,,\\
\eta(x,y,t) &=& \eta_{out}(x,y,t) \quad \text{on} \quad \Gamma_{out,t}\,,\\
\nu\dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} &=& 0 \quad \text{on} \quad \Gamma_{out,t}\,.
\end{array}\right.$$

For the sake of simplicity, we will consider a *constant density*, *homogeneous boundary conditions*, *no external forces* and a *gravity acceleration of 1*, which means that *i.e.* $\rho = \rho_0$, $\boldsymbol{U}_{in} = 0$, $\eta_{out} = 0$, $\kappa = 0$, $f_{x,y} = 0$ and $g = 1$.

In order to derive a generic time-independent form of this problem, we consider a three-dimensional domain $\tilde{\Omega}$, defined through the two-dimensional domain ω, the bottom function $b(x,y)$ and a generic free surface function $\tilde{\eta}(x,y)$ as follows :

$$\tilde{\Omega} = \{\ \boldsymbol{x} = (x,y,z)\ /\ (x,y) \in \omega \quad \text{and} \quad \tilde{b}\,(x,y) \leq z \leq \tilde{\eta}\,(x,y,t)\ \}\ . \tag{6.2}$$

We will denote by $\tilde{\Gamma}_{sol}$ the solid lateral boundaries of the domain $\tilde{\Omega}$ and by $\tilde{\Gamma}_{in}$ its inflow boundary.

The following regularity assumptions are made on the domain $\tilde{\Omega}$:

$$\tilde{\eta} \ \in\ W^{1,\infty}(\omega), \quad \tilde{b} \ \in\ W^{1,\infty}(\omega), \tag{6.3}$$

$$\text{and} \qquad \exists\, h_0 \ >\ 0 \quad \text{such that} \quad \tilde{h} = \tilde{\eta} - \tilde{b} \ \geq\ h_0 \quad \text{on}\ \omega\,. \tag{6.4}$$

We will use the following functional spaces :

$$\begin{array}{rcl}
\mathcal{M} &=& L^2(\omega), \qquad \tilde{\mathcal{X}} \ =\ H^1(\tilde{\Omega}) \\
\tilde{\mathcal{X}}_0 &=& \{\ \boldsymbol{\psi} \in H^1(\tilde{\Omega})^2\ /\ \boldsymbol{\psi}\cdot\boldsymbol{n}_{xy} \ =\ 0 \text{ on } \tilde{\Gamma}_{sol} \quad \text{and} \quad \boldsymbol{\psi} \ =\ 0 \text{ on } \tilde{\Gamma}_{in}\ \}\,.
\end{array} \tag{6.5}$$

Let us introduce two space-independent bilinear forms :

$$\tilde{d}(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,) \;=\; \int_{\tilde{\Omega}} \boldsymbol{\nabla}\boldsymbol{\psi}_1 \cdot \boldsymbol{\nabla}\boldsymbol{\psi}_2 \; d\boldsymbol{x}, \tag{6.6}$$

$$\tilde{b}(\,\boldsymbol{\psi}\,,\,\chi\,) \;=\; -\int_{\omega} \chi \; \mathrm{div}_2 \; (\int_b^{\tilde{\eta}} \boldsymbol{\psi} \; dz) \; d\omega. \tag{6.7}$$

The bilinear form $\tilde{b}(\,\cdot\,,\,\cdot\,)$ defines the linear oberator $\mathcal{B} : \tilde{\mathcal{X}}_0 \longrightarrow \mathcal{M}'$ and its transpose $\mathcal{B}^T : \mathcal{M} \longmapsto (\tilde{\mathcal{X}}_0)'$ as

$$< \mathcal{B}\,\boldsymbol{\psi}, \chi >_{\mathcal{M}'\times\mathcal{M}} = < \boldsymbol{\psi}, \mathcal{B}^T\,\chi >_{(\tilde{\mathcal{X}})'\times\tilde{\mathcal{X}}} = \; \tilde{b}(\,\boldsymbol{\psi}\,,\,\chi\,). \tag{6.8}$$

We consider the following generic form of the continuous weak $\boldsymbol{u} - \eta$ problem :

$$\text{find} \quad \boldsymbol{u} \in \tilde{\mathcal{X}}_0 \quad \text{and} \quad \eta \in \mathcal{M} \quad \text{such that,}$$

$$\forall\, \boldsymbol{\psi} \in \tilde{\mathcal{X}}_0 \quad \text{and} \quad \forall\, \chi \in \mathcal{M},$$

$$\left\{ \begin{array}{lllll} \delta\;(\,\boldsymbol{u}\,,\,\boldsymbol{\psi}\,)_{\tilde{\Omega}} & + \quad \alpha\;\tilde{d}(\,\boldsymbol{u}\,,\,\boldsymbol{\psi}\,) & + \quad \tilde{b}(\eta\,,\,\boldsymbol{\psi}\,) & = & < \boldsymbol{G}, \boldsymbol{\psi} >_{\tilde{\mathcal{X}}'\times\tilde{\mathcal{X}}}\,, \\ & & & & \\ & \delta\;(\,\eta\,,\,\chi\,)_{\omega} & - \quad \tilde{b}(\chi\,,\,\boldsymbol{u}\,) & = & < H, \chi >_{\mathcal{M}'\times\mathcal{M}}\,, \end{array} \right. \tag{6.9}$$

where δ is a strictly positive and α a positive parameter, and

$$\begin{array}{rcl} < \boldsymbol{G}, \boldsymbol{\psi} >_{\tilde{\mathcal{X}}'\times\tilde{\mathcal{X}}} & = & \delta\;(\,\tilde{\boldsymbol{u}}\,,\,\boldsymbol{\psi}\,)_{\tilde{\Omega}}, \\ < H, \chi >_{\mathcal{M}'\times\mathcal{M}} & = & \delta\;(\,\tilde{\eta}\,,\,\chi\,)_{\omega}\,, \end{array}$$

with $\tilde{\boldsymbol{u}}$ a function in $L^2(\tilde{\Omega})^2$.

Remark 6.2.1 When the $\boldsymbol{u} - \eta$ problem arises from the time-discrete free surface flow problem, we have that

$$\delta \;=\; \frac{1}{\Delta t}\,, \quad \alpha \;=\; \nu,$$

where Δt is the time step length and ν the viscosity coefficient of the fluid. Note that many simulations are performed without diffusion, or with a very small viscosity coefficient. It is therefore worthwhile to note that α can be very small and even zero. This case will be adressed in Remark 6.2.4.

6.2.1 Analysis in the Lax-Milgram framework

The variational problem (6.9) can be written as follows. Given $\tilde{\boldsymbol{u}}$ and $\tilde{\eta}$, find

$$\begin{aligned} &(\boldsymbol{u}\,,\eta) \in \tilde{\mathcal{X}}_0 \times \mathcal{M} \quad \text{such that,} \quad \forall\,(\boldsymbol{\psi}\,,\chi) \in \tilde{\mathcal{X}}_0 \times \mathcal{M}, \\ &\qquad L\,((\boldsymbol{u}\,,\eta)\,,(\boldsymbol{\psi}\,,\chi)) = F((\boldsymbol{\psi}\,,\chi)), \end{aligned} \tag{6.10}$$

where

$$\begin{aligned} L\,((\boldsymbol{\psi}_1\,,\chi_1)\,,(\boldsymbol{\psi}_2\,,\chi_2)) &= \delta\,(\boldsymbol{\psi}_1\,,\boldsymbol{\psi}_2)_{\tilde{\Omega}} + \alpha\,\tilde{d}(\boldsymbol{\psi}_1\,,\boldsymbol{\psi}_2) \\ &\quad + \tilde{b}(\boldsymbol{\psi}_2\,,\chi_1) - \tilde{b}(\boldsymbol{\psi}_1\,,\chi_2) + \delta\,(\chi_1\,,\chi_2)_{\omega}, \end{aligned}$$

$$F((\boldsymbol{\psi}\,,\chi)) = <\boldsymbol{G},\boldsymbol{\psi}>_{\tilde{\mathcal{X}}'\times\tilde{\mathcal{X}}} + <H,\chi>_{\mathcal{M}'\times\mathcal{M}}\,.$$

Proposition 6.1 *Under the regularity assumptions (6.3) and (6.4) on the domain $\tilde{\Omega}$, the bilinear form $L(\cdot,\cdot)$ is symmetric and continuous on $\left(\tilde{\mathcal{X}}_0 \times \mathcal{M}\right)^2$.*

Proof. Let us consider $(\boldsymbol{\psi}_1,\chi_1)$ and $(\boldsymbol{\psi}_2,\chi_2)$ in $(\tilde{\mathcal{X}}_0 \times \mathcal{M})^2$.

$$\begin{aligned} |L\,((\boldsymbol{\psi}_1\,,\chi_1)\,,(\boldsymbol{\psi}_2\,,\chi_2))| &\leq \delta\,\|\boldsymbol{\psi}_1\|_{0,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{0,\tilde{\Omega}} \\ &+ \alpha\,\|\boldsymbol{\nabla}\boldsymbol{\psi}_1\|_{0,\tilde{\Omega}}\,\|\boldsymbol{\nabla}\boldsymbol{\psi}_2\|_{0,\tilde{\Omega}} + \delta\,\|\chi_1\|_{0,\omega}\,\|\chi_2\|_{0,\omega} \\ &+ \|\chi_2\|_{0,\omega}\,\|\operatorname{div}_2(\int_b^{\tilde{\eta}}\boldsymbol{\psi}_1\,dz)\|_{0,\omega} + \|\chi_1\|_{0,\omega}\,\|\operatorname{div}_2(\int_b^{\tilde{\eta}}\boldsymbol{\psi}_2\,dz)\|_{0,\omega} \end{aligned}$$

Using Proposition 5.3 :

$$\begin{aligned} &\leq \delta\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}} + \alpha\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}}\,. \\ &+ \delta\,\|\chi_1\|_{0,\omega}\,\|\chi_2\|_{0,\omega} + C(\tilde{\Omega})\,\left(\|\chi_2\|_{0,\omega}\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}} + \|\chi_1\|_{0,\omega}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}}\right) \\ &\leq \operatorname{Max}\left(\delta+\alpha\,,\delta\,,C(\tilde{\Omega})\right)\quad\Big(\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}} \\ &+ \|\chi_1\|_{0,\omega}\,\|\chi_2\|_{0,\omega} + \|\chi_2\|_{0,\omega}\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}} + \|\chi_1\|_{0,\omega}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}}\Big) \\ &\leq l(\alpha,\delta,\tilde{\Omega})\quad\|(\boldsymbol{\psi}_1,\chi_1)\|_{\tilde{\mathcal{X}}\times\mathcal{M}}\,\|(\boldsymbol{\psi}_2,\chi_2)\|_{\tilde{\mathcal{X}}\times\mathcal{M}} \end{aligned}$$

where

$$l(\alpha,\delta,\tilde{\Omega}) = 8\,\operatorname{Max}\left(\delta+\alpha\,,\delta\,,C(\tilde{\Omega})\right) \tag{6.11}$$

is the continuity constant of $L(\cdot,\cdot)$ and depends on α, δ, and on the domain $\tilde{\Omega}$. $\diamondsuit$

Proposition 6.2 *The bilinear form* $L(\cdot,\cdot)$ *is coercive on* $\left(\tilde{\mathcal{X}}_0 \times \mathcal{M}\right)^2$.

Proof. Let us consider $(\boldsymbol{\psi}, \chi)$ in $\tilde{\mathcal{X}}_0 \times \mathcal{M}$.

$$\begin{aligned} L\left((\boldsymbol{\psi}, \chi), (\boldsymbol{\psi}, \chi)\right) &= \delta \|\boldsymbol{\psi}\|^2_{0,\tilde{\Omega}} + \alpha \|\boldsymbol{\nabla}\boldsymbol{\psi}\|^2_{0,\tilde{\Omega}} + \delta \|\chi\|^2_{0,\omega} \\ &\geq \alpha |\boldsymbol{\psi}|^2_{1,\tilde{\Omega}} + \delta \|\chi\|^2_{0,\omega} \\ &\geq \mathrm{Min}(\alpha, \delta) \left(|\boldsymbol{\psi}|^2_{1,\tilde{\Omega}} + \|\chi\|^2_{0,\omega} \right). \end{aligned}$$

Following Poincaré's Lemma, the semi-norm $|\cdot|_{1,\tilde{\Omega}}$ is a norm in $\tilde{\mathcal{X}}_0$. Therefore

$$L\left((\boldsymbol{\psi}, \chi), (\boldsymbol{\psi}, \chi)\right) \geq c_L(\alpha,\delta) \|(\boldsymbol{\psi}, \chi)\|^2_{\tilde{\mathcal{X}}_0 \times \mathcal{M}},$$

where

$$c_L(\alpha, \delta,) = \mathrm{Min}(\alpha, \delta) > 0. \tag{6.12}$$

is the coercivity constant of $L(\cdot,\cdot)$ and depends on α and δ. $\diamondsuit$

Proposition 6.3 *Application* $F(\cdot)$ *is continuous on* $\tilde{\mathcal{X}} \times \mathcal{M}$.

Proof. Let us consider $(\boldsymbol{\psi}, \chi)$ in $\tilde{\mathcal{X}} \times \mathcal{M}$.

$$\begin{aligned} |F((\boldsymbol{\psi}, \chi))| &= |\delta \int_{\tilde{\Omega}} \tilde{\boldsymbol{u}} \cdot \boldsymbol{\psi}\, d\boldsymbol{x} + \delta \int_{\omega} \tilde{\eta}\chi\, d\omega| \\ &\leq \delta \|\tilde{\boldsymbol{u}}\|_{0,\tilde{\Omega}} \|\boldsymbol{\psi}\|_{0,\tilde{\Omega}} + \delta \|\tilde{\eta}\|_{0,\omega} \|\chi\|_{0,\omega} \\ &\leq f(\tilde{\boldsymbol{u}}, \tilde{\eta}, \delta) \|(\boldsymbol{\psi}, \chi)\|_{\tilde{\mathcal{X}} \times \mathcal{M}}, \end{aligned}$$

where

$$f(\tilde{\boldsymbol{u}}, \tilde{\eta}, \delta) = \sqrt{2}\,\delta\, \mathrm{Max}\{ \|\tilde{\boldsymbol{u}}\|_{0,\tilde{\Omega}} + \|\tilde{\eta}\|_{0,\omega} \} \tag{6.13}$$

is the continuity constant of $F(\cdot)$. $\diamondsuit$

We can now state the following existence and uniqueness result, whose proof is an immediate consequence of the Lax-Milgram Theorem.

Theorem 6.1
Under the regularity assumptions (5.50) and (5.51) on the domain, for every $\tilde{\boldsymbol{u}} \in \tilde{\mathcal{X}}$, *problem (6.10) has a solution* $(\boldsymbol{u}, \eta)$, *which is unique in* $\tilde{\mathcal{X}}_0 \times \mathcal{M}$ *and satisfies :*

$$\|\boldsymbol{u}\|^2_{1,\tilde{\Omega}} + \|\eta\|^2_{0,\omega} \leq \left(\frac{f(\tilde{\boldsymbol{u}}, \tilde{\eta}, \delta, \delta)}{c_L(\alpha, \delta)} \right)^2, \tag{6.14}$$

where $c_L(\alpha,\delta)$ *and* $f(\tilde{\boldsymbol{u}}, \tilde{\eta}, \delta)$ *are defined respectively by (6.12) and (6.13).*

As explained in Remark 6.2.1, α can be equal to zero. In that case, the analysis just presented is not valid because the bilinear form $L(\cdot,\cdot)$ is not coercive on $(\tilde{\mathcal{X}}_0 \times \mathcal{M})^2$ and the Lax-Milgram theorem cannot be applied.

In addition, for a very small value of α the estimation of the solution deteriorates : the majoration gets very large. This can be interpreted as the consequence of the fact that the coercivity of the bilinear form $L(\cdot,\cdot)$ is guaranteed by the presence of a term controlling the norm of the gradient of the solution $\boldsymbol{u}$. Since this term is proportional to α, the well-posedness of the problem can degenerate when α becomes too small.

For these reasons we will next analyse the problem in another framework.

6.2.2 Analysis in the mixed problem framework

Let us now introduce the following bilinear forms :

$$\tilde{a}(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2) \;=\; \delta\,(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2)_{\tilde{\Omega}} \;+\; \alpha\,\tilde{d}(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2),$$

$$m(\chi_1, \chi_2) \;=\; -\delta\,(\chi_1, \chi_2)_{\omega}\,.$$

The variational problem (6.9) can be reformulated into a mixed problem as follows. Given $\tilde{\boldsymbol{u}}$ and $\tilde{\eta}$,

$$\text{find} \quad (\boldsymbol{u}, \eta) \;\in\; \tilde{\mathcal{X}}_0 \times \mathcal{M} \quad \text{such that}$$

$$\forall\,(\boldsymbol{\psi}, \chi) \;\in\; \tilde{\mathcal{X}}_0 \times \mathcal{M}$$

$$\begin{array}{rcl} \tilde{a}(\boldsymbol{u}, \boldsymbol{\psi}) \;+\; \tilde{b}(\boldsymbol{\psi}, \eta) & = & < \boldsymbol{G}, \boldsymbol{\psi} >_{\tilde{\mathcal{X}}' \times \tilde{\mathcal{X}}}\,, \\ m(\eta, \chi) \;+\; \tilde{b}(\boldsymbol{u}, \chi) & = & - < H, \chi >_{\mathcal{M}' \times \mathcal{M}}\,. \end{array} \tag{6.15}$$

This formulation fits in with a very general form of mixed problems – more general than the one used for the Stokes problem where $m(\cdot\,,\cdot) = 0$. An existence and uniqueness result for this general form has been proven by Brezzi and Fortin in [20].

In order to apply the theorem proving the existence and uniqueness of a solution, we must first verify that the bilinear forms defining our problem satisfy some particular properties.

Proposition 6.4 *The bilinear form* $\tilde{a}(\cdot\,,\cdot)$ *is symmetric and continous on* $\tilde{\mathcal{X}}^2$.

Proof.
Let us consider $(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2)$ in $\tilde{\mathcal{X}}^2$.

$$\begin{aligned} |\,\tilde{a}(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,)\,| &\leq \delta\,\|\boldsymbol{\psi}_1\|_{0,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{0,\tilde{\Omega}} + \alpha\,|\boldsymbol{\psi}_1|_{1,\tilde{\Omega}}\,|\boldsymbol{\psi}_2|_{1,\tilde{\Omega}} \\ &\leq (\delta+\alpha)\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}} \\ &\leq a(\alpha,\delta)\,\|\boldsymbol{\psi}_1\|_{1,\tilde{\Omega}}\,\|\boldsymbol{\psi}_2\|_{1,\tilde{\Omega}} \end{aligned}$$

where

$$a(\alpha,\delta) \;=\; \delta \;+\; \alpha$$

is the continuity constant of $\tilde{a}(\,\cdot\,,\,\cdot\,)$, depending on δ and α.$\diamondsuit$

Proposition 6.5 *The bilinear form* $\tilde{a}(\,\cdot\,,\,\cdot\,)$ *is coercive on* $(\tilde{\mathcal{X}}_0)^2$.

Proof. Consider $\boldsymbol{\psi}$ in $H^1(\tilde{\Omega})$.

$$|\,\tilde{a}(\,\boldsymbol{\psi}\,,\,\boldsymbol{\psi}\,)\,| \;=\; \delta\,\|\boldsymbol{\psi}\|^2_{0,\tilde{\Omega}} \;+\; \alpha\,|\boldsymbol{\psi}|^2_{1,\tilde{\Omega}} \;\geq\; c_A(\alpha)\,\|\boldsymbol{\psi}\|_{\tilde{\mathcal{X}}_0}{}^2$$

where $c_A(\alpha) \;=\; \alpha \;>0$ is the coercivity constant of $\tilde{a}(\,\cdot\,,\,\cdot\,)$.$\diamondsuit$

Proposition 6.6 *Under the regularity assumptions (6.3) and (6.4) on the domain* $\tilde{\Omega}$*, the bilinear form* $\tilde{b}(\,\cdot\,,\,\cdot\,)$ *is continuous on* $\tilde{\mathcal{X}} \times \mathcal{M}$.

Proof. Let us consider $(\boldsymbol{\psi}, \chi)$ in $\tilde{\mathcal{X}} \times \mathcal{M}$.

$$\begin{aligned} |\,\tilde{b}(\,\boldsymbol{\psi}\,,\,\chi\,)\,| &\leq |\int_\omega \chi\,\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \boldsymbol{\psi}\;dz)\;d\omega| \\ &\leq \|\chi\|_{0,\omega}\;\|\,\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \boldsymbol{\psi}\;dz)\|_{0,\omega} \end{aligned}$$

Using Proposition 5.3 :

$$\begin{aligned} &\leq C(\tilde{\Omega})\,\|\chi\|_{0,\omega}\;\|\boldsymbol{\psi}\|_{1,\tilde{\Omega}} \\ &\leq b(\tilde{\Omega})\,\|\chi\|_{0,\omega}\;\|\boldsymbol{\psi}\|_{1,\tilde{\Omega}}\,, \end{aligned}$$

where $b(\tilde{\Omega}) = \; C(\tilde{\Omega})$ is the continuity constant of $\tilde{b}(\,\cdot\,,\,\cdot\,)$. $\diamondsuit$

Proposition 6.7 *Under the regularity assumptions (6.3) and (6.4) on the domain* $\tilde{\Omega}$*, the following* inf-sup *condition holds :*

$$\exists\;\tilde{\beta}>0 \quad \textit{such that,} \quad \forall\;\chi\in\mathcal{M}, \quad \exists\;\boldsymbol{\psi}\in\tilde{\mathcal{X}}_0 \quad \textit{such that}$$

$$\tilde{b}(\,\boldsymbol{\psi}\,,\,\chi\,) \;\geq\; \tilde{\beta}\,\|\boldsymbol{\psi}\|_{1,\tilde{\Omega}}\;\|\chi\|_{0,\omega}\,. \tag{6.16}$$

For the proof of this proposition we need to introduce the following two-dimensional functional space :

$$\mathcal{Y}_0 \;=\; \{\; \boldsymbol{v} \in H^1(\omega)^2 \;/\; \boldsymbol{v} \;=\; 0 \text{ on } \partial\omega_{in} \;\}\,, \tag{6.17}$$

where $\partial\omega_{in}$ denotes the boundary of the 2D domain ω corresponding to the inflow. In addition, we need the two following lemma.

Lemma 6.1 *Under the regularity assumptions (6.3) and (6.4) on the domain* $\tilde{\Omega}$*, for any* $\boldsymbol{v} \in \mathcal{Y}_0$ *:*

$$\tilde{h}\,\boldsymbol{v} \in \mathcal{Y}_0 \quad \textit{and} \quad \frac{\boldsymbol{v}}{\tilde{h}} \in \mathcal{Y}_0.$$

Moreover, the function $\tilde{\boldsymbol{v}} : (x,y,z) \longrightarrow \mathbb{R}$ *defined by* $\tilde{\boldsymbol{v}}(x,y,z) = \boldsymbol{v}(x,y)$ *belongs to the space* $\tilde{\mathcal{X}}_0$ *and the following inequality holds :*

$$\|\tilde{\boldsymbol{v}}\|^2_{1,\tilde{\Omega}} \;\leq\; \|\tilde{h}\|_{L^\infty(\omega)}\;\|\boldsymbol{v}\|^2_{1,\omega}\,. \tag{6.18}$$

Proof. Let us consider $\boldsymbol{v} \in \mathcal{Y}_0$.

$$\begin{aligned}
\int_\omega |\,\boldsymbol{\nabla}_2\left(\frac{\boldsymbol{v}}{\tilde{h}}\right)|^2\,d\omega \;&\leq\; \|\frac{1}{\tilde{h}}\|^4_{L^\infty(\omega)}\int_\omega |\,\tilde{h}\;\boldsymbol{\nabla}_2\boldsymbol{v} \;-\; \boldsymbol{v}\;\boldsymbol{\nabla}_2\tilde{h}\,|^2\,d\omega \\
&\leq\; 2\,\|\frac{1}{\tilde{h}}\|^4_{L^\infty(\omega)}\;\left(\int_\omega |\,\tilde{h}\;\boldsymbol{\nabla}_2\boldsymbol{v}\,|^2\,d\omega \;+\; \int_\omega |\,\boldsymbol{v}\;\boldsymbol{\nabla}_2\boldsymbol{v}\,|^2\,d\omega\right) \\
&\leq\; 2\,\|\frac{1}{\tilde{h}}\|^2_{L^\infty(\omega)}\;\left(\|\boldsymbol{\nabla}_2\boldsymbol{v}\|^2_{0,\omega} \;+\; \|\boldsymbol{\nabla}_2\tilde{h}\|_{L^\infty(\omega)}\;\|\boldsymbol{v}\|^2_{1,\omega}\right) \\
&\leq\; 2\,\|\frac{1}{\tilde{h}}\|^2_{L^\infty(\omega)}\;\left(1 + \|\boldsymbol{\nabla}_2\tilde{h}\|^2_{L^\infty(\omega)}\right)\;\|\boldsymbol{v}\|^2_{1,\omega} \\
&<\; \infty
\end{aligned}$$

$$\begin{aligned}
\int_\omega |\,\boldsymbol{\nabla}_2\left(\tilde{h}\,\boldsymbol{v}\right)|^2\,d\omega \;&\leq\; \int_\omega |\,\tilde{h}\,|^2\;|\,\boldsymbol{\nabla}_2\boldsymbol{v}\,|^2\,d\omega \;+\; \int_\omega |\,\boldsymbol{v}\,|^2\;|\,\boldsymbol{\nabla}_2\tilde{h}\,|^2\,d\omega \\
&\leq\; \|\tilde{h}\|^2_{L^\infty(\omega)}\;\|\boldsymbol{\nabla}_2\boldsymbol{v}\|^2_{0,\omega} \;+\; \|\boldsymbol{\nabla}_2\tilde{h}\|^2_{L^\infty(\omega)}\;\|\boldsymbol{v}\|^2_{0,\omega} \\
&\leq\; \left(\|\tilde{h}\|^2_{L^\infty(\omega)} + \|\boldsymbol{\nabla}_2\tilde{h}\|^2_{L^\infty(\omega)}\right)\;\|\boldsymbol{v}\|^2_{1,\omega} \\
&<\; \infty
\end{aligned}$$

$$\begin{aligned}
\|\tilde{\boldsymbol{v}}\|^2_{1,\tilde{\Omega}} &= \int_{\tilde{\Omega}} |\,\tilde{\boldsymbol{v}}\,|^2\, d\boldsymbol{x} \;+\; \int_{\tilde{\Omega}} |\,\boldsymbol{\nabla}\tilde{\boldsymbol{v}}\,|^2\, d\boldsymbol{x} \\
&= \int_{\omega} \tilde{h}\,|\,\boldsymbol{v}\,|^2\, d\omega \;+\; \int_{\omega} \tilde{h}\,|\,\boldsymbol{\nabla}_2\boldsymbol{v}\,|^2\, d\omega \\
&\leq \|\tilde{h}\|_{L^\infty(\omega)}\;\|\boldsymbol{v}\|^2_{1,\omega} \\
&< \infty
\end{aligned}$$

In addition, since the lateral boundaries are perfectly vertical, if $v = 0$ on $\partial\omega_{in}$, then $v = 0$ on the corresponding inflow boundary $\tilde{\Gamma}_{in}$ of the three dimensional domain. $\diamondsuit$

Let us now introduce the following bilinear form :

$$\tilde{b}^{2d} : \mathcal{Y}_0 \times L^2(\omega) \longrightarrow \mathbb{R}$$

$$\tilde{b}^{2d}(\,\boldsymbol{v}\,,\,\chi\,) \;=\; -\int_{\omega} \chi\,\mathrm{div}\,\left(\tilde{h}\,\boldsymbol{v}\right)\, d\omega. \tag{6.19}$$

Remark 6.2.2 We point out that $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$ is defined on $\mathcal{Y}_0 \times L^2(\omega)$, where $\mathcal{Y}_0$ is a space of functions defined in the two-dimensional domain ω, whereas $\tilde{b}(\,\cdot\,,\,\cdot\,)$ is defined on $H^1(\tilde{\Omega}) \times L^2(\omega)$, where $H^1(\tilde{\Omega})$ is a space of functions defined in the three-dimensional space $\tilde{\Omega}$.

Lemma 6.2 *Under the regularity assumptions (6.3) and (6.4) on the domain, the following* inf-sup *condition holds :*

$$\exists\ \tilde{\beta}^{2d} > 0 \quad \textit{such that,} \quad \forall\ \chi \in L^2(\omega), \quad \exists\ \boldsymbol{v} \in \mathcal{Y}_0 \quad \textit{such that}$$

$$\tilde{b}^{2d}(\,\boldsymbol{v}\,,\,\chi\,) \;\geq\; \tilde{\beta}^{2d}\;\|\boldsymbol{v}\|_{1,\omega}\;\|\chi\|_{0,\omega}\,. \tag{6.20}$$

Proof. The operator $\mathrm{div} : \mathcal{Y}_0 \longrightarrow L^2(\omega)'$ is surjective.
Moreover, lemma 6.1 allows us to state that application

$$\tilde{f} : \mathcal{Y}_0 \longrightarrow \mathcal{Y}_0, \quad \text{with} \quad \tilde{f}(\boldsymbol{v}) \;=\; \tilde{h}\boldsymbol{v}$$

is continuous and surjective.

Let $\tilde{B} : \mathcal{Y}_0 \longrightarrow L^2(\omega)'$ be the operator corresponding to the bilinear form $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$ – i.e. $\tilde{B} \;=\; div \circ \tilde{f}$. We deduce that $\tilde{B}$ is surjective, which is equivalent to stating that $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$ satisfies condition(6.20). $\diamondsuit$

We can now proceed with the proof of Proposition 6.7.
Proof. Let us consider χ in $L^2(\omega)$. We choose $\boldsymbol{v}$ in $\mathcal{Y}_0$ such that inequality

(6.20) is satisfied. Following lemma 6.1, function $\tilde{\boldsymbol{v}} : (x, y, z) \longrightarrow \mathbb{R}$ defined by $\tilde{\boldsymbol{v}}(x, y, z) = \boldsymbol{v}(x, y)$ belongs to $\tilde{\mathcal{X}}_0$ and

$$\|\tilde{\boldsymbol{v}}\|^2_{1,\tilde{\Omega}} \quad \leq \|\tilde{h}\|_{L^\infty(\omega)} \|\boldsymbol{v}\|^2_{1,\omega}.$$

Therefore

$$\begin{aligned} \tilde{b}(\tilde{\boldsymbol{v}}, \chi) = \tilde{b}^{2d}(\boldsymbol{v}, \chi) &\geq \tilde{\beta}^{2d} \|\boldsymbol{v}\|_{1,\omega} \|\chi\|_{0,\omega} \\ &\geq \frac{\tilde{\beta}^{2d}}{\left(\|\tilde{h}\|_{L^\infty(\omega)}\right)^{1/2}} \|\tilde{\boldsymbol{v}}\|_{1,\tilde{\Omega}} \|\chi\|_{0,\omega}. \end{aligned}$$

We have finally proven that, for any χ in $L^2(\omega)$, there is a function $\boldsymbol{\psi}$ in $\tilde{\mathcal{X}}_0$ such that

$$\tilde{b}(\boldsymbol{\psi}, \chi) \geq \tilde{\beta} \|\boldsymbol{\psi}\|_{1,\tilde{\Omega}} \|\chi\|_{0,\omega},$$

with

$$\tilde{\beta} = \frac{\tilde{\beta}^{2d}}{\left(\|\tilde{h}\|_{L^\infty(\omega)}\right)^{1/2}} > 0. \tag{6.21}$$

$\diamondsuit$

Proposition 6.8 *The bilinear form* $m(\cdot, \cdot)$ *is symmetric, continuous and coercive on* $\mathcal{M}^2$.

Proof. We have that

$$|m(\chi_1, \chi_2)| \leq m(\delta) \|\chi_1\|_{0,\omega} \|\chi_2\|_{0,\omega},$$

where $m(\delta) = \delta$ is the continuity constant of $m(\cdot, \cdot)$. Moreover, since

$$m(\chi, \chi) = \delta \|\chi\|^2_{0,\omega},$$

$m(\delta)$ is also its coercivity coefficient.$\diamondsuit$

We are now in position to state the following existence and uniquenes result, which is an immediate consequence of Theorem 1.2 in [20].

Theorem 6.2
Under the regularity assumptions (5.50) and (5.51) on the domain, for every $\tilde{\boldsymbol{u}} \in \tilde{\mathcal{X}}$, *the variational problem (6.15) has a solution* $(\boldsymbol{u}, \eta)$, *which is unique in* $\tilde{\mathcal{X}}_0 \times \mathcal{M}$ *and satisfies :*

$$\|\boldsymbol{u}\|_{1,\tilde{\Omega}} + \|\eta\|_{0,\omega} \leq K_1 \|\boldsymbol{G}\|_{\mathcal{X}'} + K_2 \|H\|_{\mathcal{M}'}, \tag{6.22}$$

where

$$K_1 = \frac{1}{c_A(\alpha)}\left(4+\frac{1}{c_A(\alpha)}\right)\left(1+2\left(\frac{a(\alpha,\delta)}{\tilde{\beta}}\right)^2\right) + \frac{2}{(\tilde{\beta})^2}$$

$$\text{and} \quad K_2 = \left(\frac{1}{c_A(\alpha)}+2\left(\frac{a(\alpha,\delta)}{c_A(\alpha)\,\tilde{\beta}}\right)^2\right)\left(1+2\left(\frac{a(\alpha,\delta)}{\tilde{\beta}}\right)^2\right). \tag{6.23}$$

We recall that

$$a(\alpha,\delta) = \delta+\alpha, \qquad c_A(\alpha) = \alpha\,,$$

$$\tilde{\beta} > 0 \quad \text{only depends on the domain } \tilde{\Omega}\,,$$

$$\|\boldsymbol{G}\|_{\mathcal{X}'} = \delta\,\|\tilde{\boldsymbol{u}}\|_{0,\tilde{\Omega}} \quad \text{and} \quad \|H\|_{\mathcal{M}'} = \delta\,\|\eta\|_{0,\omega}\,.$$

Remark 6.2.3 Two additional assumptions are necessary for the application of the theorem proving the existence and uniqueness of a solution to the general mixed problem. On one hand, we must have that Let us consider two assumptions, necessary for the application of the theorem prooving the existence and uniqueness of a solution to the general mixed problem. On one hand, we must have that

$$H \in Im\,\mathcal{B}, \tag{6.24}$$

and on the other hand that

$$\begin{gathered}\exists\, c_0 > 0 \quad \text{such that} \quad \forall \bar{\psi} \in \mathcal{K}er(\mathcal{B}^T)^{\perp} \quad \text{and} \quad \forall \varepsilon > 0,\\ \text{the solution } \psi_0 \in \mathcal{K}er(\mathcal{B}^T) \text{ of equation}\\ \varepsilon\,(\psi_0,\chi)_{0,\omega} + m(\psi_0,\chi) = -\,m(\bar{\psi},\chi) \quad \forall \chi \in \mathcal{K}er(\mathcal{B}^T)\\ \text{is bounded by} \quad c_0\,\|\psi_0\|_{0,\omega} \le \|\bar{\psi}\|_{0,\omega}\,.\end{gathered} \tag{6.25}$$

In the particular case treated here, these assumptions are trivially verified, since $Im(\mathcal{B}) = \mathcal{M}'$ and $\mathcal{K}er(\mathcal{B}^T) = \{0\}$, as can be deduced from Proposition 6.7.

Remark 6.2.4 In the case where $\nu = 0$, the existence and uniqueness of a solution to the variational problem (6.15) can still be proven in the mixed problem framework, but the proof is slightly different. In fact, the proof is then very similar to the proof made by Miglio and Saleri in [98] of the existence and uniqueness of a solution to the semi-discrete weak formulation

they propose of a linearized version of the hydrostatic free surface problem, in which the horizontal diffusion is neglected. Let us give the general idea of this proof.

If $\alpha = 0$, the definition space for the variable $\boldsymbol{u}$ can be enlarged to

$$\tilde{\mathcal{X}}_2 \;=\; \{ \quad \boldsymbol{\psi} \in L^2(\tilde{\Omega})^2 \quad / \quad (\,\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \boldsymbol{\psi}\; dz)) \in L^2(\omega)^2 \quad \}. \tag{6.26}$$

No Dirichlet boundary condition can then be imposed strongly at the inflow, therefore $\boldsymbol{u}$ must be sought in

$$\tilde{\mathcal{X}}_{2,z} \;=\; \{\, \boldsymbol{\psi} \in \tilde{\mathcal{X}} \;/\; \boldsymbol{\psi} \cdot \boldsymbol{n}_{xy} \;=\; 0 \text{ on } \tilde{\Gamma}_{sol} \,\}, \tag{6.27}$$

where we recall that $\tilde{\Gamma}_{sol}$ denotes the solid lateral boundaries of the domain $\tilde{\Omega}$.

Let us consider the bilinear form $\tilde{a}(\,\cdot\,,\,\cdot\,)$ with $\alpha = 0$. Clearly, it is symmetric and continous on $(\tilde{\mathcal{X}}_2)^2$, with a continuity constant equal to δ. In addition, defining the linear operator $\mathcal{B}_2 : \tilde{\mathcal{X}}_{2,z} \longrightarrow \mathcal{M}'$ and its transpose $\mathcal{B}^T : \mathcal{M} \longmapsto (\tilde{\mathcal{X}}_2)'$ by

$$< \mathcal{B}_2\, \boldsymbol{\psi}, \chi >_{\mathcal{M}' \times \mathcal{M}} = < \boldsymbol{\psi}, \mathcal{B}_2^T\, \chi >_{(\tilde{\mathcal{X}}_2)' \times \tilde{\mathcal{X}}_2} \;=\; \tilde{b}(\,\boldsymbol{\psi}\,,\,\chi\,)\;, \tag{6.28}$$

we have that

$$Ker(\mathcal{B}_2) \;=\; \{\, \boldsymbol{\psi} \in \tilde{\mathcal{X}}_{2,z} \;/\; (\,\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \boldsymbol{\psi}\; dz)) \;=\; 0 \text{ in } L^2(\omega)^2 \,\}\,.$$

Therefore, $\tilde{a}(\,\cdot\,,\,\cdot\,)$ with $\alpha = 0$ is coercive on $Ker(\mathcal{B}_2)$, the coercivity constant being equal to $\delta \;> 0$.

Finally, note that $\tilde{\mathcal{X}} \subset \tilde{\mathcal{X}}_2$ and, following Proposition 5.3,

$$\|\,\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \boldsymbol{\psi}\; dz)\,\|_{\,0,\omega} \;\leq\; C(\tilde{\Omega})\;\|\boldsymbol{\psi}\|_{\,1,\tilde{\Omega}}\,, \tag{6.29}$$

where $C(\tilde{\Omega})$ is a strictly positive constant depending only on the domain $\tilde{\Omega}$. Using these properties, we can easily extend Proposition 6.7 to the space $\tilde{\mathcal{X}}_2$.

Theorem 1.2 in [20] can therefore also be applied to problem (6.15) in the case where $\alpha = 0$, replacing $\tilde{\mathcal{X}}$ by $\tilde{\mathcal{X}}_{2,z}$: the existence and uniqueness result for the continuous $\boldsymbol{u} - \eta$ problem given by Theorem 6.2 holds true when $\alpha = 0$. Note that $a(\alpha, \delta)$ and $c_A(\alpha)$ must then be replaced by δ in 6.23.

6.3 Approximation of the $\boldsymbol{u} - \eta$ problem

We now analyse the finite element approximation of problem (6.9) through the Galerkin method. For this purpose we consider a triangulation $\mathcal{T}_h^{2d}$ of the two-dimensional domain ω as well as a triangulation $\mathcal{T}_h$ of the three-dimensional domain $\tilde{\Omega}$.

Remark 6.3.1 In order to simplify the analysis, we assume that the two-dimensional domain, as well as the three-dimensional real domain are polyhedral, so that they remain unchanged after triangulation : $\omega_h = \omega$ and $\tilde{\Omega}_h = \tilde{\Omega}$. Especially, this implies that $\tilde{\eta}_h = \tilde{\eta}$ and $b_h = b$.

On one hand, a Lagrangian finite element space on $\mathcal{T}_h^{2d}$ is chosen for the discrete space $\mathcal{M}_h$ approximating $\mathcal{M}(\omega)$. On the other hand, two degrees k and n satisfying $k \leq n$ are chosen in order to determine the finite element function space

$$\tilde{\mathcal{X}}_h = (\mathcal{F}_{n,k}(\mathcal{T}_h))^2 , \tag{6.30}$$

a Lagrangian finite element defined in (2.24) and approximating $\tilde{\mathcal{X}} = H^1(\tilde{\Omega})^2$. In addition, we define :

$$\tilde{\mathcal{X}}_{h,0} = \{ \boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_h \mid \boldsymbol{\psi}_h \cdot \boldsymbol{n}_{xy} = 0 \text{ on } \Gamma_{sol} \text{ and } \boldsymbol{\psi}_h = 0 \text{ on } \Gamma_{in} \}. \tag{6.31}$$

where Γ_{sol} denotes the solid lateral boundaries of the domain $\tilde{\Omega}$ and Γ_{in} its *inflow* boundary. Note that the discrete functional spaces verify :

$$\tilde{\mathcal{X}}_h \subset H^1(\tilde{\Omega})^2 \qquad \text{and} \qquad \mathcal{M}_h \subset L^2(\omega), \tag{6.32}$$

and thus the approximation is internal ; *i.e.*

$$\tilde{\mathcal{X}}_{h,0} \subset \tilde{\mathcal{X}}_0 \quad \text{and} \quad \mathcal{M}_h \subset \mathcal{M}.$$

We can now write the fully-discretized problem. Given $\tilde{\boldsymbol{u}}$ and $\tilde{\eta}$, find

$$\boldsymbol{u}_h \in \tilde{\mathcal{X}}_{h,0} \quad \text{and} \quad \eta_h \in \mathcal{M}_h \quad \text{such that,}$$

$$\forall\, \boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_{h,0} \quad \text{and} \quad \forall\, \chi \in \mathcal{M}_h,$$

$$\begin{cases} \delta\,(\,\boldsymbol{u}_h\,,\,\boldsymbol{\psi}_h\,)_{\tilde{\Omega}} + \alpha\,\tilde{d}(\,\boldsymbol{u}_h\,,\,\boldsymbol{\psi}_h\,) + \tilde{b}(\,\eta_h\,,\,\boldsymbol{\psi}_h\,) = < \boldsymbol{G}_h, \boldsymbol{\psi}_h >_{\tilde{\mathcal{X}}_h' \times \tilde{\mathcal{X}}_h} \\ \\ \qquad\qquad \delta\,(\,\eta_h\,,\,\chi_h\,)_{\omega} - \tilde{b}(\,\chi_h\,,\,\boldsymbol{u}_h\,) = < H_h, \chi_h >_{\mathcal{M}_h' \times \mathcal{M}_h}, \end{cases} \tag{6.33}$$

where

$$\begin{aligned} < \boldsymbol{G}_h, \boldsymbol{\psi}_h >_{\tilde{\mathcal{X}}_h' \times \tilde{\mathcal{X}}_h} &= \delta\,(\,\tilde{\boldsymbol{u}}_h\,,\,\boldsymbol{\psi}_h\,)_{\tilde{\Omega}}, \\ < H_h, \chi_h >_{\mathcal{M}_h' \times \mathcal{M}_h} &= \delta\,(\,\tilde{\eta}_h\,,\,\chi_h\,)_{\omega}. \end{aligned}$$

6.3.1 Analysis in the Lax-Milgram framework

The discrete variational problem (6.33) can be written as follows. Given $\tilde{\boldsymbol{u}}_h$ and $\tilde{\eta}$,

$$\text{find} \quad (\,\boldsymbol{u}_h\,,\,\eta_h\,) \;\in\; \tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h \quad \text{such that}$$

$$\forall\,(\,\boldsymbol{\psi}_h\,,\,\chi_h\,) \;\in\; \tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h,$$

$$L\,(\,(\,\boldsymbol{u}_h\,,\,\eta_h\,)\,,\,(\,\boldsymbol{\psi}_h\,,\,\chi_h\,)\,) \;=\; F(\,(\,\boldsymbol{\psi}_h\,,\,\chi_h\,)\,). \tag{6.34}$$

Since $\tilde{\mathcal{X}}_{h,0} \subset H^1(\tilde{\Omega})^2$ and $\mathcal{M}_h \subset L^2(\omega)$, the coercivity and continuity of the bilinear form $L(\cdot,\cdot)$ are still verified on $(\tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h)^2$ for the norm of $\tilde{\mathcal{X}} \times \mathcal{M}$, as well as the continuity of the linear application $F(\cdot)$. In addition, the continuity and coercivity coefficients remain unchanged. We can therefore apply the Lax-Milgram theorem and the Céa lemma in order to prove the following proposition :

Proposition 6.9 *Under the regularity assumptions (5.50) and (5.51) on the domain $\tilde{\Omega}$, for every $\tilde{\boldsymbol{u}}_h \in \tilde{\mathcal{X}}_h$, the discrete problem (6.34) has a solution $(\boldsymbol{u}_h, \eta_h)$, which is unique in $\tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h$ and satisfies the following a priori error estimation :*

$$\begin{aligned} &\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\tilde{\Omega}} \;+\; \|\eta - \eta_h\|_{0,\omega} \\ \leq\quad &\left(\frac{l(\alpha,\delta,\tilde{\Omega})}{c_L(\alpha,\delta)}\right)^2 \Big(\,\inf_{\boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_{h,0}} \|\boldsymbol{u} - \boldsymbol{\psi}_h\|_{1,\tilde{\Omega}} \;+\; \inf_{\chi_h \in \mathcal{M}_h} \|\eta_h - \chi_h\|_{0,\omega}\,\Big)\,. \end{aligned} \tag{6.35}$$

We recall that :

$$\begin{gathered} l(\alpha,\delta,\tilde{\Omega},) \;=\; 8\;\mathrm{Max}\Big(\delta+\alpha\,,\,\delta\,,\,C(\tilde{\Omega})\Big) \\ \text{and} \quad c_L(\alpha,\delta) \;=\; \mathrm{Min}\,(\alpha\,,\,\delta)\,, \end{gathered}$$

where $C(\tilde{\Omega})$ is a constant defined in Proposition 5.3 and depending on the domain $\tilde{\Omega}$ only.

Next, we analyse the discrete problem in the framework of the mixed problems.

6.3.2 Analysis in the mixed problem framework

The discrete variational problem (6.33) can be reformulated into a mixed problem as follows. Given $\tilde{\boldsymbol{u}}_h$ and $\tilde{\eta}$,

$$\text{find} \quad (\,\boldsymbol{u}_h\,,\,\eta_h\,) \;\in\; \tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h \quad \text{such that}$$

$$\forall\,(\,\boldsymbol{\psi}_h\,,\,\chi_h\,) \;\in\; \tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h$$

$$\begin{aligned} \tilde{a}(\,\boldsymbol{u}_h\,,\,\boldsymbol{\psi}_h\,) \;+\; \tilde{b}(\,\boldsymbol{\psi}_h\,,\,\eta_h\,) \;&=\; <\boldsymbol{G}_h, \boldsymbol{\psi}_h>_{\mathcal{X}_h' \times \mathcal{X}_h} \\ m(\,\eta_h\,,\,\chi_h\,) \;+\; \tilde{b}(\,\boldsymbol{u}_h\,,\,\chi_h\,) \;&=\; -\,<H_h, \chi_h>_{\mathcal{M}_h' \times \mathcal{M}_h}\,. \end{aligned} \tag{6.36}$$

Since $\tilde{\mathcal{X}}_{h,0} \subset H^1(\tilde{\Omega})^2$ and $\mathcal{M}_h \subset L^2(\omega)$, the continuity and coercivity of the bilinear form $a(\cdot,\cdot)$ are still verified on $(\tilde{\mathcal{X}}_{h,0})^2$ for the norm of $\mathcal{X}$. This also holds for $m(\cdot,\cdot)$ on $\mathcal{M}_h^2$ for the norm of $\mathcal{M}$. Applying Proposition 2.11 in [20] we can prove the following proposition :

Proposition 6.10 *Under the regularity assumptions (5.50) and (5.51) on the domain* $\tilde{\Omega}$*, if the discrete spaces* $\tilde{\mathcal{X}}_h$ *and* $\mathcal{M}_h$ *are chosen such that the following discrete* inf-sup *condition holds :*

$$\begin{aligned} &\exists\ \tilde{\beta}_h > 0 \quad \textit{such that} \quad \forall\ \chi_h \in \mathcal{M}_h \quad \exists\ \boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_{h,0} \quad \textit{such that} \\ &\qquad \tilde{b}(\,\boldsymbol{\psi}_h\,,\,\chi_h\,) \;\geq\; \tilde{\beta}_h\ \|\boldsymbol{\psi}_h\|_{1,\tilde{\Omega}}\ \|\chi_h\|_{0,\omega}\,, \end{aligned} \tag{6.37}$$

then, for every $\tilde{\boldsymbol{u}}_h \in \tilde{\mathcal{X}}_h$*, the discrete problem (6.36) has a solution* $(\boldsymbol{u}_h, \eta_h)$ *, which is unique in* $\tilde{\mathcal{X}}_{h,0} \times \mathcal{M}_h$ *and satisfies the following a priori error estimation :*

$$\begin{aligned} &\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\tilde{\Omega}} \;+\; \|\eta - \eta_h\|_{0,\omega} \\ \leq\quad & K_3\ \inf_{\boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_h}\ \|\boldsymbol{u} - \boldsymbol{\psi}_h\|_{1,\tilde{\Omega}} \;+\; K_4\ \inf_{\chi_h \in \mathcal{M}_h}\ \|\eta - \chi_h\|_{0,\omega}\,, \end{aligned} \tag{6.38}$$

where

$$K_4 \;=\; 1 \;+\; 2\ K_1\ b(\tilde{\Omega})^2 \;+\; 2\ K_2\ (m(\delta))^2,$$

$$K_3 \;=\; 1 \;+\; 2\ K_1\ (a(\alpha,\delta))^2 \;+\; 2\ K_2\ b(\tilde{\Omega})^2,$$

$$K_2 \;=\; \left(\frac{1}{c_A(\alpha)} \;+\; 2\left(\frac{a(\alpha,\delta)}{c_A(\alpha)\ \tilde{\beta}_h}\right)^2\right)\left(1 + 2\left(\frac{a(\alpha,\delta)}{\tilde{\beta}_h}\right)^2\right),$$

$$K_1 \;=\; \frac{1}{c_A(\alpha)}\left(4 + \frac{1}{c_A(\alpha)}\right)\left(1 + 2\left(\frac{a(\alpha,\delta)}{\tilde{\beta}_h}\right)^2\right) \;+\; \frac{2}{(\tilde{\beta}_h)^2}\,.$$

We recall that

$$a(\alpha, \delta) \;=\; \delta + \alpha, \qquad \alpha_a(\alpha) \;=\; \alpha\,,$$

$$b(\tilde{\Omega}) = g\; C(\tilde{\Omega}) \qquad \text{and} \qquad m(\delta) = \delta,$$

where $C(\tilde{\Omega})$ is a constant defined in Proposition 5.3 and depending on the domain $\tilde{\Omega}$ only.

In the next section we propose a pair of finite element spaces approximating the horizontal velocity and the free surface function, and such that the discrete *inf-sup condition* (6.37) is satisfied.

6.3.3 A pair of compatible finite element spaces

Let us first recall the following definition.

Definition 6.3.1 Consider two Hilbert spaces $\mathcal{X}$ and $\mathcal{M}$, and a continuous bilinear form $b(\cdot,\cdot)$ on $\mathcal{X} \times \mathcal{M}$. Two finite element spaces $\mathcal{X}_h$ and $\mathcal{M}_h$ approximating $\mathcal{X}$ and $\mathcal{M}$ are said to be compatible for the bilinear form b if the following discrete *inf-sup* condition holds :

$$\exists\, \beta_h > 0 \quad \text{such that} \quad \forall\, q_h \in \mathcal{M}_h \quad \exists\, \boldsymbol{v}_h \in \mathcal{X}_h \quad \text{such that}$$

$$b(\boldsymbol{v}_h, q_h) \;\geq\; \beta_h\; \|\boldsymbol{v}_h\|_X\; \|q_h\|_M. \tag{6.39}$$

Our aim in this section is to construct two discrete spaces $\tilde{\mathcal{X}}_h \subset H^1(\tilde{\Omega})^2$ and $\mathcal{M}_h \subset L^2(\omega)$, approximating respectively the horizontal velocity and the free surface and such that the corresponding finite element space $\tilde{\mathcal{X}}_{h,0}$ and $\mathcal{M}_h$ are compatible for the bilinear form $\tilde{b}(\cdot,\cdot)$ – *i.e.* the discrete inf-sup condition (6.37) holds.

For this purpose we consider a triangulation $\mathcal{T}_h^{2d}$ of the two-dimensional domain ω composed of triangles. By piling up this mesh along the vertical in $\tilde{\Omega}$ we obtain a prismatic mesh $\mathcal{T}_h$ of the three-dimensional domain, as illustrated in Figure 6.2.

We will denote by T_R the reference triangle for $\mathcal{T}_h^{2d}$ and by P_R the reference prism for $\mathcal{T}_h$. Note that the prism P_R is straight, its triangular basis coincides with the reference triangle T_R and its vertical is $[-1, 1]$. In addition, each triangle $T \in \mathcal{T}_h^{2d}$ and each prism $P \in \mathcal{T}_h$ are given by

$$T \;=\; \mathcal{M}_1^T(T_R) \qquad \text{and} \qquad P \;=\; \mathcal{M}^P(P_R),$$

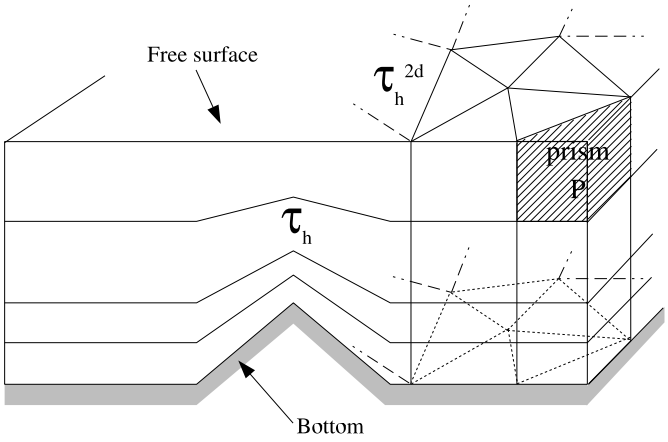


FIG. 6.2 – Example of a three-dimensional mesh $\mathcal{T}_h$ obtained by piling up the two-dimensional mesh $\mathcal{T}_h^{2d}$.

where $\mathcal{M}_1^T$ and $\mathcal{M}^P$ are two homeomorphic mappings from, respectively, T_R to T and P_R to P. Note that $\mathcal{M}_1^T$ is of degree 1. For each triangle $T \in \mathcal{T}_h^{2d}$ we will denote :

$$h_T \;=\; \mathrm{diam}(T) \qquad \text{and} \qquad \rho_T \;=\; \sup\,\{\; \mathrm{diam}(\mathcal{B}) \;/\quad \mathcal{B} \text{ ball} \subset T \;\}.$$

Let us now recall the definition of the bilinear form $\tilde{b}^{2d}(\cdot,\cdot)$ introduced in section 6.2.2 of this chapter :

$$\tilde{b}^{2d} : \mathcal{Y}_0 \times L^2(\omega) \longrightarrow \mathbb{R}$$

$$\tilde{b}^{2d}(\,\boldsymbol{v}\,,\,\chi\,) \;=\; -\int_\omega \chi\,\mathrm{div}\,\left(\tilde{h}\,\boldsymbol{v}\right)\,d\omega\,.$$

The idea is to choose two finite element spaces $\mathcal{Y}_h \subset H^1(\omega)^2$ and $\mathcal{M}_h \subset L^2(\omega)$ compatible for the bilinear form $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$, *i.e.*, such that the following discrete *inf-sup* condition holds :

$$\exists\;\tilde{\beta}_h^{2d} > 0 \quad \text{such that} \quad \forall\;\chi_h \in \mathcal{M}_h \quad \exists\;\boldsymbol{v}_h \in \mathcal{Y}_{h,0} \quad \text{such that}$$

$$\tilde{b}^{2d}(\,\boldsymbol{v}_h\,,\,\chi_h\,) \;\geq\; \tilde{\beta}_h^{2d}\;\|\boldsymbol{v}_h\|_{1,\omega}\,\|\chi_h\|_{0,\omega}\,, \tag{6.40}$$

where $\mathcal{Y}_{h,0} \;=\; \{\boldsymbol{v}_h \in \mathcal{Y}_h \quad | \quad \boldsymbol{v}_h \;=\; 0 \quad \text{on} \quad \partial\omega_{in}\;\}$.

Let $\tilde{\mathcal{Y}}_h$ be the three-dimensional function space defined from $\mathcal{Y}_h$ by

$$\tilde{\mathcal{Y}}_h = \{\;\tilde{\boldsymbol{v}} : (x,y,z) \longrightarrow \mathbb{R} \quad | \quad \tilde{\boldsymbol{v}}_h(x,y,z) \;=\; \boldsymbol{v}_h(x,y),\;\boldsymbol{v}_h \in \mathcal{Y}_h\;\}\,. \tag{6.41}$$

A sufficient condition for (6.37) to be verified is then to choose a finite element space $\tilde{\mathcal{X}}_h \subset H^1(\tilde{\Omega})^2$ such that $\tilde{\mathcal{Y}}_h \subset \tilde{\mathcal{X}}_h$. Indeed, the following proposition can be proven :

Proposition 6.11 *Let* $\mathcal{Y}_h \subset H^1(\omega)^2$ *and* $\mathcal{M}_h \subset L^2(\omega)$ *be two finite element spaces such that* $\mathcal{Y}_{h,0}$ *and* $\mathcal{M}_h$ *are compatible for the bilinear form* $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$. *In addition, let* $\tilde{\mathcal{X}}_h$ *be a finite element space approximating* $H^1(\tilde{\Omega})^2$ *and* $\tilde{\mathcal{Y}}_h$ *defined from* $\mathcal{Y}_h$ *by (6.41).*

If $\tilde{\mathcal{Y}}_h \subset \tilde{\mathcal{X}}_h$ *then the finite element spaces* $\tilde{\mathcal{X}}_{h,0}$ *and* $\mathcal{M}_h$ *are compatible for the bilinear form* $\tilde{b}(\,\cdot\,,\,\cdot\,)$, *ie, the discrete* inf-sup *condition (6.37) is verified.*

Proof. Consider $\chi_h \in \mathcal{M}_h$. We choose $\boldsymbol{v}_h \in \mathcal{Y}_{h,0}$ such that the inequality in the discrete *inf-sup* condition (6.40) holds. We define $\tilde{\boldsymbol{v}}_h$ in

$$\tilde{\mathcal{Y}}_{h,0} \;=\; \{\; \tilde{\boldsymbol{v}_h} \in \tilde{\mathcal{Y}}_h \quad | \quad \tilde{\boldsymbol{v}_h} \;=\; 0 \quad \text{on} \quad \tilde{\Gamma}_{in} \;\}$$

by $\tilde{\boldsymbol{v}}_h(x,y,z) = \boldsymbol{v}_h(x,y)$. Clearly, since $\tilde{\mathcal{Y}}_h \subset \tilde{\mathcal{X}}_h$, $\tilde{\boldsymbol{v}}_h \in \tilde{\mathcal{X}}_{h,0}$. Moreover, since $\tilde{\mathcal{X}}_h \subset H^1(\tilde{\Omega})^2$ – see (6.32) –, we can apply Lemma 6.1 and state that

$$\|\boldsymbol{v}_h\|^2_{1,\omega} \;\geq\; \frac{1}{\|\tilde{h}\|_{\,L^\infty(\omega)}}\, \|\tilde{\boldsymbol{v}}_h\|^2_{\,1,\tilde{\Omega}}\;.$$

We can then write that :

$$\begin{aligned}
b_n(\,\tilde{\boldsymbol{v}}_h\,,\,\chi_h\,) \;&=\; -\int_\omega \chi_h\;\mathrm{div}_2\,(\int_b^{\tilde{\eta}} \tilde{\boldsymbol{v}}_h\;dz)\;d\omega \\
&=\; -\int_\omega \chi_h\;\mathrm{div}_2\,(\tilde{h}\;\boldsymbol{v}_h)\;d\omega \\
&=\; \tilde{b}^{2d}(\,\boldsymbol{v}_h\,,\,\chi_h\,) \\
&\text{using (6.40)} \\
&\geq\; \tilde{\beta}_h^{2d}\;\|\boldsymbol{v}_h\|_{\,1,\omega}\;\|\chi_h\|_{\,0,\omega} \\
&\geq\; \frac{\tilde{\beta}_h^{2d}}{(\|\tilde{h}\|_{\,L^\infty(\omega)}\,)^{1/2}}\;\|\tilde{\boldsymbol{v}}_h\|_{\,1,\tilde{\Omega}}\;\|\chi_h\|_{\,0,\omega} \\
&\geq\; \tilde{\beta}_h\;\|\tilde{\boldsymbol{v}}_h\|_{\,1,\tilde{\Omega}}\;\|\chi_h\|_{\,0,\omega}
\end{aligned}$$

where

$$\tilde{\beta}_h \;=\; \frac{\tilde{\beta}_h^{2d}}{(\|\tilde{h}\|_{\,L^\infty(\omega)}\,)^{1/2}}$$

$\diamondsuit$

Let us now choose the finite element spaces $\mathcal{Y}_h$, $\tilde{\mathcal{X}}_h$ and $\mathcal{M}_h$.

For this purpose we introduce a function $\varphi_b : T_R \longrightarrow \mathbb{R}$, defined on the reference triangle T_R and such that :

$$\varphi_b \in H_0^1(T_R)\,, \qquad 0 \;\leq\; \varphi_b \;\leq 1 \qquad \text{and} \quad \varphi_b(C_R) \;=\; 1\,,$$

where C_R is the barycenter of the reference triangle T_R. We denote by $B_R \;=\; Vect\,\{\,\varphi_b\,\}$ the vectorial space generated by the functions φ_b defined on the reference triangle corresponding to each element T in $\mathcal{T}_h^{2d}$.

A well-known example of a function φ_b is the *bubble* function, defined as the product of the barycentric coordinates in the triangle. The notion of bubble function was introduced by Crouzeix and Raviart in [16]. Arnold, Brezzi and Fortin proved in [6] the compatibility for the Stokes operator of the discrete spaces associated to the $\mathbb{P}_1$-bubble/$\mathbb{P}_1$ element, where the $\mathbb{P}_1$-bubble element consists of the set of piecewise linear functions enriched with a bubble function on each triangle. As a result, it has four degrees of freedom on each triangle, situated at the three vertices of the triangle and at the barycenter.

A variant of this function is the *quasi-bubble* function, first introduced in [94]. The reference triangle T_R is divided in 3 sub-triangles by linking the three vertices to the triangle's barycenter C_R, as illustrated by figure 6.3. The function φ_b is then defined as the piecewise linear function on each one

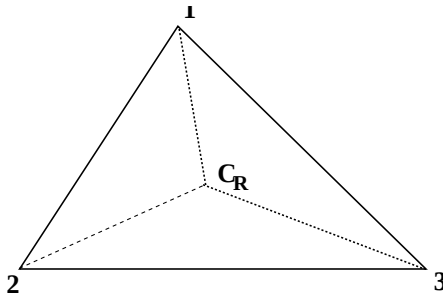


FIG. 6.3 – The $\mathbb{P}_1$-quasi-bubble element

of the sub-triangles which has the value 1 at C_R and 0 at the vertices of T_R. The advantage of this function with respect to the *bubble* function is its lower polynomial degree. The compatibility analysis made in [6] for the $\mathbb{P}_1$-bubble/$\mathbb{P}_1$ element can be applied to the $\mathbb{P}_1$-quasi-bubble/$\mathbb{P}_1$ element. This element is used for instance in the Telemac-2d system to approach the two-dimensional z-averaged horizontal velocity and the water height when solving a particular part of the Shallow-Water equations – see [65].

Let us introduce the following finite element spaces on $\mathcal{T}_h^{2d}$:

$$\mathcal{F}_{1,1}(\mathcal{T}_h^{2d}) = \Big\{ v_h : \omega \to \mathbb{R} \mid \quad v_h \in \mathcal{C}^0(\bar{\omega}), \quad v_h\Big|_T \circ \mathcal{M}_1^T \in \mathbb{P}_1(T_R), \quad \forall T \in \mathcal{T}_h^{2d} \Big\}, \tag{6.42}$$

$$\mathcal{B}_{1,1}(\mathcal{T}_h^{2d}) = \Big\{ v_h : \omega \to \mathbb{R} \mid \quad v_h \in \mathcal{C}^0(\bar{\omega}), \; v_h\Big|_T \circ \mathcal{M}_1^T \in B_R(T_R), \quad \forall T \in \mathcal{T}_h^{2d} \Big\}. \tag{6.43}$$

The dimension of $\mathcal{B}_{1,1}(\mathcal{T}_h^{2d})$ – equal to the number of triangles in the two-dimensional mesh –, will be denoted by N_T. We define the following set of basis functions :

$$\{ \varphi_{b,k} \}_{k=1,\ldots,N_T}, \quad \text{with} \quad \varphi_{b,k} = \varphi_b \circ (\mathcal{M}_1^{T_k})^{-1} \qquad \text{in} \quad \omega, \tag{6.44}$$

where T_k denotes the triangle in $\mathcal{T}_h^{2d}$ referenced by k.

The discrete functional space $\mathcal{Y}_h$ is then defined on the two-dimensional domain as follows :

$$\mathcal{Y}_h = \Big(\mathcal{F}_{1,1}(\mathcal{T}_h^{2d}) \oplus \mathcal{B}_{1,1}(\mathcal{T}_h^{2d}) \Big)^2, \tag{6.45}$$

which corresponds to the $\mathbb{P}_1$-bubble element.

Denoting by $\mathbb{P}_{1,xy}$ – respectively $\mathbb{P}_{1,z}$ – the set of piecewise linear functions in x and y (respectively z), we introduce the three-dimensional space $\mathcal{P}_1 = \mathbb{P}_{1,xy} \times \mathbb{P}_{1,z}$, *i.e.*,

$$\mathcal{P}_1 = \{ \, p(x,y,z) = q(x,y)\, r(z), \; q \in \mathbb{P}_{1,xy}, \; r \in \mathbb{P}_{1,z} \, \}. \tag{6.46}$$

We can now introduce the first order prismatic finite element space on $\mathcal{T}_h$, illustrated in figure 6.4 and defined by :

$$\mathcal{P}_{1,1}(\mathcal{T}_h) = \Big\{ \psi_h : \tilde{\Omega} \to \mathbb{R} \mid \quad \psi_h \in \mathcal{C}^0(\bar{\omega}), \quad \psi_h\Big|_P \circ \mathcal{M}^P \in \mathcal{P}_1(P_R), \quad \forall P \in \mathcal{T}_h \, \}. \tag{6.47}$$

We will denote by N_{3d} the dimension of $\mathcal{P}_{1,1}$, equal to the number of nodes in the three-dimensional mesh. Moreover, $\{ \psi_i \}_{i=1,\ldots,N_{3d}}$ will indicate a particular set of basis functions.

The discrete functional space $\tilde{\mathcal{X}}_h$ is defined on the three-dimensional configuration and is thus generated by basis functions defined on each prism P_0 in $\mathcal{T}_h$. We choose some of these basis functions to be independent of z. In

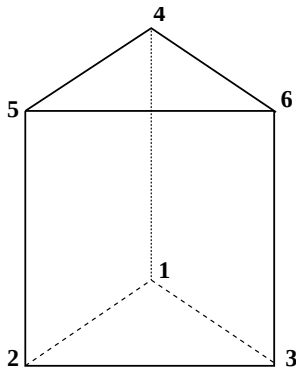


FIG. 6.4 – The first order prismatic element

fact, the finite element space chosen to approximate the horizontal velocity coincides with $\mathcal{P}_{1,1}(\mathcal{T}_h)$, enriched with a bubble function on each triangle T in $\mathcal{T}_h^{2d}$. Thus, $\tilde{\mathcal{X}}_h$ can be written as follows :

$$\tilde{\mathcal{X}}_h \;=\; \Big(\mathcal{P}_{1,1}(\mathcal{T}_h) \;\oplus\; \mathcal{B}_{1,1}(\mathcal{T}_h^{2d}) \Big)^2. \tag{6.48}$$

As a result, it has $(N_{3d} + N_T)$ degrees of freedom, situated at the vertices of the prisms in $\mathcal{T}_h$ and at the barycenters of each triangle in $\mathcal{T}_h^{2d}$, as illustrated by figure 6.5.

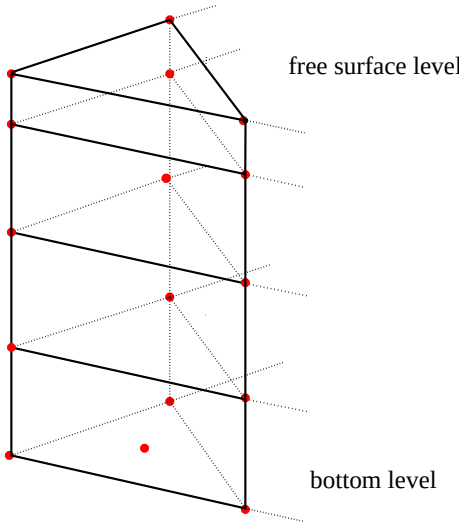


FIG. 6.5 – Degrees of freedom in a column of prisms in $\mathcal{T}_h$

Clearly, the dimension of $\tilde{\mathcal{X}}_h$ is $(N_{3d} + N_T)$, and any function v_h in $\tilde{\mathcal{X}}_h$ can be decomposed as follows :

$$v_h(x,y,z) \;=\; \sum_{i=1}^{N_{3d}} v_i \;\psi_i(x,y,z) \;+\; \sum_{k=1}^{N_T} v_k \;\varphi_{b,k}(x,y), \tag{6.49}$$

where v_i and v_k denote the coefficient of v_h associated, respectively, to the node i and the triangle k. Note that v_i is equal to the value of v_h at node i.

Finally, we define

$$\mathcal{M}_h \;=\; \mathcal{F}_{1,1}(\mathcal{T}_h^{2d}). \tag{6.50}$$

For the finite element spaces just defined, we have the following results.

We recall that a set of meshes $(\mathcal{T}_h^{2d})_h$ is regular if there is a constant $\sigma > 0$, independent of h, such that

$$\forall\; h, \qquad \max_{T \in \mathcal{T}_h^{2d}} \frac{h_T}{\rho_T} \;\leq\; \sigma.$$

Proposition 6.12 *If the affine set of meshes* $(\mathcal{T}_h^{2d})_h$ *is regular, the discrete* inf-sup *condition (6.40) holds uniformly with respect to* h*, which means that the constant* $\tilde{\beta}_h^{2d}$ *is independent of* h*.*

The proof of this proposition is a variant of the proof of compatibility for the Stokes operator of spaces $\mathcal{Y}_{h,0}$ and $\mathcal{M}_h$ made in [6]. We will now prove the same property for the bilinear form $\tilde{b}^{2d}(\,\cdot\,,\,\cdot\,)$. In this proof we will make use of the following lemma, owed to Fortin [49].

Lemma 6.3 *Consider two Hilbert spaces* $\mathcal{X}$ *and* $\mathcal{M}$*, a bilinear form* $b(\cdot,\cdot)$ *on* $\mathcal{X} \times \mathcal{M}$*, and assume the existence of a constant* $\beta > 0$ *such that*

$$inf_{q \in \mathcal{M}} \; sup_{\boldsymbol{v} \in \mathcal{X}} \; \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_X \; \|q\|_M} \;\geq\; \beta. \tag{6.51}$$

Then, two finite element spaces $\mathcal{X}_h \subset \mathcal{X}$ *and* $\mathcal{M}_h \subset \mathcal{M}$ *are compatible for* $b(\cdot,\cdot)$ *if and only if*

$$\begin{gathered} \exists\; \gamma_h > 0 \quad \textit{such that} \quad \forall \boldsymbol{v} \in \mathcal{X} \quad \exists\; \boldsymbol{\Pi}_h(\boldsymbol{v}) \in \mathcal{X}_h \quad \textit{such that} \\ \forall\; q_h \in \mathcal{M}_h, \quad b(\boldsymbol{v}, q_h) \;=\; b(\boldsymbol{\Pi}_h(\boldsymbol{v}), q_h) \quad \textit{and} \quad \|\boldsymbol{\Pi}_h(\boldsymbol{v})\|_X \;\leq\; \gamma_h \;\|\boldsymbol{v}\|_X. \end{gathered} \tag{6.52}$$

Proof. (Proposition 6.12)

Following the same arguments as for Lemma 6.2, we can prove that the continuous *inf-sup* condition (6.51) holds for $\tilde{\mathcal{X}} = H_0^1(\omega)^2$, $\mathcal{M} = L^2(\omega)$ and $b(\cdot,\cdot) = \tilde{b}^{2d}(\cdot,\cdot)$.

Consider a function $\boldsymbol{v}$ in $\mathcal{Y}_0$. We wish to build $\boldsymbol{\Pi}_h(\boldsymbol{v}) \in \mathcal{Y}_{h,0}$ such that

$$\forall \eta_h \in \mathcal{M}_h, \qquad \int_\omega \eta_h \, \mathrm{div}_2 \left(\tilde{h}\, \boldsymbol{\Pi}_h(\boldsymbol{v})\right) \, d\omega \;=\; \int_\omega \eta_h \, \mathrm{div}_2 \left(\tilde{h}\, \boldsymbol{v}\right) \, d\omega.$$

Since $\mathcal{M}_h \subset H^1(\omega)$ this identity is equivalent to

$$\forall \eta_h \in \mathcal{M}_h, \qquad \sum_{T \in \mathcal{T}_h^{2d}} \int_T \tilde{h}\, \boldsymbol{\Pi}_h(\boldsymbol{v}) \cdot \boldsymbol{\nabla} \eta_h \; dT \;=\; \sum_{T \in \mathcal{T}_h^{2d}} \int_T \tilde{h}\, \boldsymbol{v} \cdot \boldsymbol{\nabla} \eta_h \; dT.$$

Moreover $\boldsymbol{\nabla} \eta_h \in \mathbb{P}_0(T)^2$, this relation thus holds if

$$\forall T \in \mathcal{T}_h^{2d}, \qquad \int_T \tilde{h}\, \boldsymbol{\Pi}_h(\boldsymbol{v}) \; dT \;=\; \int_T \tilde{h}\, \boldsymbol{v} \; dT. \tag{6.53}$$

Let us now introduce the *Clement* operator (cf. [30]), that we will denote by $\boldsymbol{\mathcal{C}}_h(\cdot)$. According to the Clement Lemma for the continuous space $\mathcal{Y}_0$ and its approximation $\mathcal{Y}_{h,0}$, if the mesh is regular, there is a constant $c_{cl} > 0$ such that :

$$\forall \boldsymbol{v} \in \mathcal{Y}_0, \quad \exists\, \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) \in \mathcal{Y}_{h,0} \quad \text{such that}$$

$$\forall\, T \in \mathcal{T}_h^{2d} \qquad \left\| \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) - \boldsymbol{v} \right\|_{0,T} \;\leq\; c_{cl}\; h_T \; \|\boldsymbol{v}\|_{1,D_T} \tag{6.54}$$

and

$$\left\| \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) \right\|_{0,T} \;\leq\; c_{cl} \; \|\boldsymbol{v}\|_{1,D_T} \;, \tag{6.55}$$

where $D_T = \cup_{\overline{T'} \cap \overline{T} \neq \emptyset}$.

We define :

$$\boldsymbol{\Pi}_h(\boldsymbol{v}) \;=\; \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) \;+\; \sum_{k=1}^{N_T} \left(\gamma_k^1 \; \boldsymbol{e}^1 + \gamma_k^2 \; \boldsymbol{e}^2 \right) \; \varphi_{b,k} \tag{6.56}$$

where $(\boldsymbol{e}^1, \boldsymbol{e}^2)$ is the canonic basis of $\mathbb{R}^2$. Clearly, $\boldsymbol{\Pi}_h(\boldsymbol{v}) \in \mathcal{Y}_{h,0}$.

Relation (6.53) holds if, for $k = 0, \ldots, N_T$ and $l = \{1, 2\}$,

$$\gamma_k^l \;=\; \frac{\displaystyle\int_{T_k} \tilde{h}\,\left(\;\boldsymbol{v}^l \;-\; \boldsymbol{\mathcal{C}}_h(\boldsymbol{v})^l\;\right)\,dT}{\displaystyle\int_{T_k} \tilde{h}\,\varphi_{b,k}\;dT}\,,$$

where $\boldsymbol{v}^l$ and $\boldsymbol{\mathcal{C}}_h(\boldsymbol{v})^l$ denote the cartesian components of $\boldsymbol{v}$ and $\boldsymbol{\mathcal{C}}_h(\boldsymbol{v})$.

We now wish to bound $|\boldsymbol{\Pi}_h(\boldsymbol{v})|_{1,\omega}$. On one hand, the mesh being affine, we have that

$$|\,det\,(J_{\mathcal{M}_1^T})\,| \;=\; \frac{meas(T)}{meas(T_R)}.$$

Therefore :

$$|\int_{T_k} \tilde{h}\,\varphi_{b,k}\;dT\;| \quad\geq\quad \inf_{\omega}(\tilde{h})\;|\int_{T_k} \varphi_{b,k}\;dT\;| \quad=\quad c_R\;\inf_{\omega}(\tilde{h})\;meas(T_k)\,,$$

where

$$c_R \;=\; \frac{\displaystyle\int_{T_R} \varphi_b\;dT_R}{meas(T_R)}\;.$$

On the other hand, $\mathcal{T}_h^{2d}$ being regular, there is a constant $c > 0$ such that for any triangle T_k :

$$|\,\varphi_b\,|_{1,\omega} \;\leq\; c\;h_T^{-2}\;meas(T).$$

Let us now bound $|\,\gamma_k^l\,|$:

$$\begin{aligned} |\,\gamma_k^l\,| \;&\leq\; \frac{\displaystyle\int_T |\,\tilde{h}\,|\;|\,\boldsymbol{v}^l \;-\; \boldsymbol{\mathcal{C}}_h(\boldsymbol{v})^l\,|\;dT}{\displaystyle|\int_T \tilde{h}\;\varphi_{b,k}\;dT\;|} \\ &\leq\; \frac{\|\tilde{h}\|_{L^\infty(\omega)}\;\|\boldsymbol{v} \;-\; \boldsymbol{\mathcal{C}}_h(\boldsymbol{v})\|_{0,T}}{c_R\;\inf_\omega(|\,\tilde{h}\,|)\;meas(T_k)^{1/2}} \end{aligned}$$

Using (6.54) :

$$|\,\gamma_k^l\,| \;\leq\; \frac{c_{cl}\;h_T\;\|\tilde{h}\|_{L^\infty(\omega)}}{c_R\;\inf_\omega(|\,\tilde{h}\,|)\;meas(T_k)^{1/2}} \quad \|\boldsymbol{v}\|_{1,D_T}\,.$$

We then have that

$$
\begin{aligned}
|\sum_{k=1}^{N_T} \left(\gamma_k^1 \, \boldsymbol{e}^1 \, + \, \gamma_k^2 \, \boldsymbol{e}^2 \right) \, \varphi_{b,k} \, |_{1,\omega}^2 \quad &\leq \quad \sum_{k=1}^{N_T} \left(\, |\, \gamma_k^1 \, |^2 \, + \, |\, \gamma_k^2 \, |^2 \, \right) \, |\varphi_{b,k}|_{1,\omega}^2 \\
&\leq \quad 2 \; c \; \frac{(c_{cl})^2}{(c_R)^2} \quad \frac{\|\tilde{h}\|_{L^\infty(\omega)}^2}{\inf_\omega(\tilde{h})^2} \quad \sum_{k=1}^{N_T} \|\boldsymbol{v}\|_{1,D_T}^2 \\
&\leq \quad C_1 \quad \frac{\|\tilde{h}\|_{L^\infty(\omega)}^2}{\inf_\omega(\tilde{h})^2} \quad \|\boldsymbol{v}\|_{1,\omega}^2
\end{aligned}
$$

Using relation (6.55) we can bound $|\, \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) \,|_{1,\omega}^2$:

$$
\begin{aligned}
|\, \boldsymbol{\mathcal{C}}_h(\boldsymbol{v}) \,|_{1,\omega}^2 \; &\leq \; c_P \; \|\boldsymbol{\mathcal{C}}_h(\boldsymbol{v})\|_{0,\omega}^2 \; \leq \; c_P \sum_{k=1}^{N_T} \|\boldsymbol{\mathcal{C}}_h(\boldsymbol{v})\|_{0,T_t}^2 \\
&\leq \; c_P \; c_{cl} \sum_{k=1}^{N_T} \|\boldsymbol{v}\|_{1,D_T}^2 \; \leq \; C_2 \; \|\boldsymbol{v}\|_{1,\omega}^2 .
\end{aligned}
$$

Finally we obtain that :

$$
|\, \boldsymbol{\Pi}_h(\boldsymbol{v}) \,|_{1,\omega} \quad \leq \quad \left(C_2 \; \frac{\|\tilde{h}\|_{L^\infty(\omega)}}{\inf_\omega(\tilde{h})} \; + \; C_1 \right) \; |\boldsymbol{v}|_{1,\omega} \; ,
$$

where C_1 and C_2 are two strictly positive constants.

We just have shown that, setting

$$
\tilde{\gamma}_h \; = \; \left(C_1 \; \frac{\|\tilde{h}\|_{L^\infty(\omega)}}{\inf_\omega(|\, \tilde{h} \,|)} \; + \; C_2 \right) \; > \; 0, \tag{6.57}
$$

for any $\boldsymbol{v} \in \mathcal{Y}_0$ there is a discrete function $\boldsymbol{\Pi}_h(\boldsymbol{v}) \in \mathcal{Y}_{h,0}$ such that, for all q_h in $\mathcal{M}_h$,

$$
\tilde{b}^{2d}(\, \boldsymbol{v} \, , \, q_h \,) \; = \; \tilde{b}^{2d}(\, \boldsymbol{\Pi}_h(\boldsymbol{v}) \, , \, q_h \,) \quad \text{and} \quad |\, \boldsymbol{\Pi}_h(\boldsymbol{v}) \,|_{1,\omega} \; \leq \; \tilde{\gamma}_h \; |\, \boldsymbol{v} \,|_{1,\omega}^2 .
$$

This allows us to apply lemma 6.3 and conclude that the finite element spaces $\mathcal{Y}_{h,0}$ and $\mathcal{M}_h$ are compatible for the bilinear form $\tilde{b}^{2d}(\, \cdot \, , \, \cdot \,)$, *i.e.* the discrete *inf-sup* condition (6.40) holds. $\diamondsuit$

Let us now consider the three-dimensional function space $\tilde{\mathcal{Y}}_h$ defined from $\mathcal{Y}_h$ by

$$\tilde{\mathcal{Y}}_h = \{ \tilde{\boldsymbol{v}} : (x,y,z) \longrightarrow \mathbb{R} \quad | \quad \tilde{\boldsymbol{v}}_h(x,y,z) \ = \ \boldsymbol{v}_h(x,y), \ \boldsymbol{v}_h \in \mathcal{Y}_h \ \}.$$

We are in position to state the following proposition.

Proposition 6.13 *Consider the finite element spaces $\mathcal{Y}_h$, $\mathcal{M}_h$ and $\tilde{\mathcal{X}}_h$ defined respectively by (6.45), (6.50) and (6.48). If the two- and three-dimensional meshes $\mathcal{T}_h^{2d}$ and $\mathcal{T}_h$ are such that $\tilde{\mathcal{Y}}_h \subset \tilde{\mathcal{X}}_h$ then, the finite element spaces $\tilde{\mathcal{X}}_h$ and $\mathcal{M}_h$ are compatible for the bilinear form $\tilde{b}(\cdot,\cdot)$,* i.e., *the discrete* inf-sup *condition (6.37) holds.*

6.4 Approximation of the $\boldsymbol{u} - \eta$ problem with weak treatment of the imperviousness condition at the lateral boundaries

6.4.1 Treatment of the imperviousness condition

The strong treatment of the imperviousness condition (5.11) at the solid lateral boundaries – *i.e.* $\boldsymbol{u} \cdot \boldsymbol{n}_{xy} = 0$ on $\Gamma_{sol,t}$ – is difficult to implement.

The natural idea is to define the discrete space $\tilde{\mathcal{X}}_{h,0}$ approximating the horizontal velocity as follows :

$$\mathcal{X}_{h,0} \ = \ \{ \ \boldsymbol{v}_h \in \tilde{\mathcal{X}}_h \quad | \quad \boldsymbol{v}_h \ = \ 0 \text{ on } \Gamma_{h,in} \quad \text{and} \quad \boldsymbol{v}_h \cdot \boldsymbol{n}_{h,xy} \ = \ 0 \text{ on } \Gamma_{h,sol} \ \},$$

where we denote by $\Gamma_{h,in}$ the inflow boundary and by $\Gamma_{h,sol}$ the solid lateral boundaries of the discretized domain Ω_h. Now the construction of a finite element space of functions with zero normal component is delicate, especially if the domain is polyhedral and the degrees of freedom on the boundary are located at the vertices.

Indeed, it requires the construction of the outgoing normal vectors $\boldsymbol{n}_h$ at the vertices of the solid boundaries, and these are not defined on a polyhedral domain, as illustrated in figure 6.6. Note that most of the discretized domains are polyhedral. Several methods exist to overcome this problem. When required, the undefined normal vector can be constructed such that it preserves some particular properties satisfied at the continuous level. For instance, $\boldsymbol{n}_h$ can be defined at each node of the boundary $\partial\Omega_h$ of a discrete domain Ω_h such that, for any discrete function $\psi_h : \Omega_h \longrightarrow \mathbb{R}$,

$$\int_{\Omega_h} \operatorname{div} \psi_h \ d\boldsymbol{x} \ = \ \int_{\partial\Omega_h} \psi_h \cdot \boldsymbol{n}_h \ d\Gamma.$$

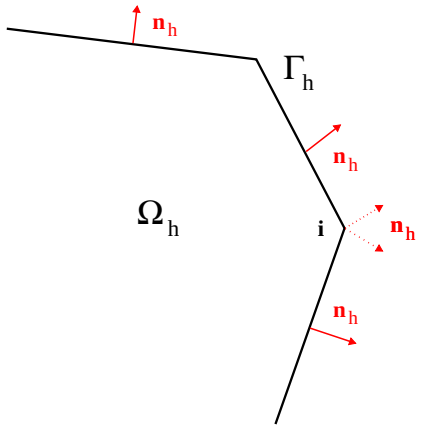


FIG. 6.6 – Polyhedral domain whose normal vector is undefined at the boundary nodes.

See [39] for a description of this method. One can also use the normal vector to the regular domain. This particular approximation of the imperviousness condition has been studied by Verfürth [137] in the framework of the Stokes problem. He shows that the resulting scheme is still convergent – provided compatible finite element spaces are used –, but the error is in $O(h^{1/2})$ independently of the finite element spaces chosen. Note that such a scheme is costly because the imperviousness condition couples the components of the fluid velocity.

One way to overcome these difficulties is to impose the condition weakly. This idea was first introduced in the framework of compressible fluids – see for instance [121] –, and has been applied in several other contexts. In particular, it has been widely studied by Pares [109] in the case of the Stokes problem. Our purpose is to apply this method to the $\boldsymbol{u} - \eta$ problem studied here.

6.4.2 Weak treatment of the imperviousness condition

In order to prescribe the imperviousness condition (5.11) at the lateral boundaries as a natural condition, we could derive an alternative variational formulation of the continuous $\boldsymbol{u} - \eta$ problem using the bilinear form

$$\tilde{b}_w(\boldsymbol{\psi}, \chi) = \int_{\tilde{\Omega}} \boldsymbol{\nabla}_2 \chi \cdot \boldsymbol{\psi} \, d\boldsymbol{x} \tag{6.58}$$

instead of

$$\tilde{b}(\boldsymbol{\psi}, \chi) = -\int_{\omega} \chi \, \mathrm{div}_2 \left(\int_b^{\tilde{\eta}} \boldsymbol{\psi} \, dz \right) d\omega \, .$$

Assuming that the free surface function η as well as the related test function χ belong to $H^1(\omega)$, the momentum equation can be written as

$$\delta\,(\,\boldsymbol{u}\,,\,\boldsymbol{\psi}\,)_{\tilde{\Omega}} \;+\; \alpha\;\tilde{d}(\,\boldsymbol{u}\,,\,\boldsymbol{\psi}\,) \;+\; \tilde{b}_w(\,\eta\,,\,\boldsymbol{\psi}\,) \;=<\boldsymbol{G},\boldsymbol{\psi}>_{\tilde{\mathcal{X}}'\times\tilde{\mathcal{X}}}\,,$$

and the free surface equation as

$$\delta\,(\,\eta\,,\,\chi\,)_{\omega} \;-\; \tilde{b}_w(\,\chi\,,\,\boldsymbol{u}\,) \;+\; \int_{\partial\omega}\chi\left(\int_b^{\eta}\boldsymbol{u}\;dz\right)\cdot\boldsymbol{n}_{\omega}\;d\gamma \;=<H,\chi>_{\mathcal{M}'\times\mathcal{M}}\;.$$

Using the natural imperviousness condition at the lateral boundaries, and imposing strongly the Dirichlet condition on the free surface at the outflow boundary, the boundary term $\displaystyle\int_{\partial\omega}\chi\left(\int_b^{\eta}\boldsymbol{u}\;dz\right)\cdot\boldsymbol{n}_{\omega}\;d\gamma$ vanishes. This leads to a weak formulation for which the functional spaces defining the variables $\boldsymbol{u}$ and η are, respectively,

$$\tilde{\mathcal{X}}_w \;=\; \{\;\boldsymbol{\psi}\in H^1(\tilde{\Omega})^2 \quad/\quad \boldsymbol{\psi} \;=\; 0 \quad\text{on}\quad \Gamma_{in}\;\} \tag{6.59}$$

and

$$\mathcal{N}_0 \;=\; \{\;\chi\in H^1(\omega) \quad/\quad \chi \;=\; 0 \quad\text{on}\quad \partial\omega_{out}\;\}\,. \tag{6.60}$$

Clearly, $\tilde{\mathcal{X}}_w$ is much easier to construct than $\tilde{\mathcal{X}}_0$ – defined by (6.5). However, this alternative weak formulation of the continous problem cannot be analysed as done in section 6.2. Indeed, on one hand, the corresponding bilinear form

$$\begin{aligned} L_w\,(\,(\,\boldsymbol{\psi}_1\,,\,\chi_1\,)\,,\,(\,\boldsymbol{\psi}_2\,,\,\chi_2\,)\,) \;&=\; \delta\,(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,)_{\tilde{\Omega}} \;+\; \alpha\;\tilde{d}(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,) \\ &+\; \tilde{b}_w(\,\boldsymbol{\psi}_2\,,\,\chi_1\,) \;-\; \tilde{b}_w(\,\boldsymbol{\psi}_1\,,\,\chi_2\,) \;+\; \delta\,(\,\chi_1\,,\,\chi_2\,)_{\omega} \end{aligned} \tag{6.61}$$

is symmetric and continuous, but it is not coercive on $(\tilde{\mathcal{X}}_0\times\mathcal{N}_0)^2$. The Lax-Milgram theorem can therefore not be applied to prove the existence and uniqueness of a solution. On the other hand, the analysis in the mixed problem framework seems difficult with this alternative formulation, because it requires the proof that the bilinear form $\tilde{b}_w(\,\cdot\,,\,\cdot\,)$ satisfies the *inf-sup* condition on the functional spaces $\tilde{\mathcal{X}}_w\times\mathcal{N}_0$, where $\mathcal{N}_0\subset H^1(\omega)$.

Therefore, we only use this idea at the discrete level. For this purpose, we consider two Lagrangian finite element spaces $\tilde{\mathcal{X}}_h$ and $\mathcal{N}_h$ approximating

respectively $\tilde{\mathcal{X}} = H^1(\tilde{\Omega})^2$ and $\mathcal{M} = L^2(\omega)$, and such that

$$\tilde{\mathcal{X}}_h \ \subset \ H^1(\tilde{\Omega})^2 \quad \text{and} \quad \mathcal{N}_h \ \subset \ H^1(\omega). \tag{6.62}$$

In addition we define :

$$\tilde{\mathcal{X}}_{h,w} \ = \ \{ \ \boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_h \ \ / \ \ \boldsymbol{\psi}_h \ = \ 0 \ \ \text{on} \ \ \Gamma_{in} \ \} \tag{6.63}$$

and

$$\mathcal{N}_{h,0} \ = \ \{ \ \chi_h \in \mathcal{N}_h \ \ / \ \ \chi_h \ = \ 0 \ \ \text{on} \ \ \partial\omega_{out} \ \}. \tag{6.64}$$

The following alternative approximation of the $\boldsymbol{u} - \eta$ problem is considered. Given $\tilde{\boldsymbol{u}}_h$ and $\tilde{\eta}_h$,

$$\text{find} \quad \boldsymbol{u}_h \ \in \ \tilde{\mathcal{X}}_{h,w} \quad \text{and} \quad \eta_h \ \in \ \mathcal{N}_{h,0} \quad \text{such that,}$$

$$\forall \, \boldsymbol{\psi}_h \ \in \ \tilde{\mathcal{X}}_{h,w} \quad \text{and} \quad \forall \, \chi_h \ \in \ \mathcal{N}_{h,0},$$

$$\left\{ \begin{array}{r} \delta \ (\ \boldsymbol{u}_h \ , \ \boldsymbol{\psi}_h \ )_{\tilde{\Omega}} \ + \ \alpha \ \tilde{d}(\ \boldsymbol{u}_h \ , \ \boldsymbol{\psi}_h \ ) \ + \ \tilde{b}_w(\ \eta_h \ , \ \boldsymbol{\psi}_h \ ) \\ \\ = \ < \boldsymbol{G}_h, \boldsymbol{\psi}_h >_{\tilde{\mathcal{X}}_h' \times \tilde{\mathcal{X}}_h} \ , \\ \\ \delta \ (\ \eta_h \ , \ \chi_h \ )_{\omega} \ - \ \tilde{b}_w(\ \chi_h \ , \ \boldsymbol{u}_h \ ) \ = < H_h, \chi >_{\mathcal{N}' \times \mathcal{N}} \ . \end{array} \right. \tag{6.65}$$

It is worthwhile to note that this is not an internal approximation of the continuous weak $\boldsymbol{u} - \eta$ problem (6.9). Indeed, the bilinear form $\tilde{b}_w(\cdot, \cdot)$ is different from $\tilde{b}(\cdot, \cdot)$. Therefore, we cannot use the standard techniques – as for instance the Céa Lemma – in order to estimate the error between the solution of the discrete problem and the solution of the continuous problem.

In fact, a more complex analysis would be needed to estimate this error. Note that in [109], Parès analyses an approximation of the Stokes problem with weak treatment of the imperviousness condition. Using non standard techniques, he establishes an error estimate of the solution of the discrete problem for the case where the velocity and the pressure are approximated in the $\mathbb{P}_1$-bubble/$\mathbb{P}_1$ space. This error estimate is in $O(h^{1/2})$. He also shows that this estimate is independent of the finite element spaces chosen, provided the *inf-sup* condition is verified.

We will not make such an analysis. We will only prove the existence and uniqueness of a solution to the discrete problem (6.65) introduced here, both in the Lax-Milgram and in the mixed problem frameworks.

6.4.3 Analysis in the Lax-Milgram framework

Under the regularity assumptions (6.3) and (6.4) on the domain $\tilde{\Omega}$, the bilinear form $L_w(\cdot,\cdot)$ defined in (6.61) is symmetric, continuous and coercive on the discrete space $(\tilde{\mathcal{X}}_{h,w} \times \mathcal{N}_{h,0})^2$ for the norms $\|\cdot\|_{1,\tilde{\Omega}}$ and $\|\cdot\|_{0,\omega}$. Therefore, the Lax-Milgram theorem can be applied, leading to the following proposition.

Proposition 6.14 *Under the regularity assumptions (6.3) and (6.4) on the domain $\tilde{\Omega}$, for every $\tilde{\boldsymbol{u}}_h \in (\tilde{\mathcal{X}}_h)'$, problem (6.65) has a solution $(\boldsymbol{u}_h, \eta_h)$, which is unique in $\tilde{\mathcal{X}}_{h,w} \times \mathcal{N}_{h,0}$ and satisfies :*

$$\|\boldsymbol{u}_h\|^2_{1,\tilde{\Omega}} \; + \; \|\eta_h\|^2_{0,\omega} \;\leq\; \left(\frac{f_h(\delta)}{\alpha_{L_w}(\alpha,\delta)} \right)^2 , \tag{6.66}$$

where $c_{Lw}(\alpha,\delta)$ is the coercivity coefficient of $L_w(\cdot,\cdot)$ and $f_h(\delta)$ the continuity coefficient of the right hand side.

We indicate that :

$$\begin{aligned} \alpha_{L_w}(\alpha,\delta) \;&=\; \mathrm{Min}\,(\alpha\,,\,\delta) \\ \text{and} \quad f_h(\delta) \;&=\; \sqrt{2}\;\delta\;\max\left(\|\tilde{\boldsymbol{u}}_h\|_{0,\Omega} \;+\; \|\tilde{\eta}_h\|_{0,\omega}\right). \end{aligned}$$

6.4.4 Analysis in the mixed problem framework

Since $\tilde{\mathcal{X}}_{h,w} \subset H^1(\tilde{\Omega})^2$, the continuity and coercivity of the bilinear form $\tilde{a}(\cdot,\cdot)$ is verified on $(\tilde{\mathcal{X}}_{h,w})^2$ for the norm $\|\cdot\|_{1,\tilde{\Omega}}$. On the other hand, since $\mathcal{N}_{h,0} \subset L^2(\omega)$, the bilinear form $m(\cdot,\cdot)$ is continuous and coercive on $(\mathcal{N}_{h,0})^2$ for the norm $\|\cdot\|_{0,\omega}$. Therefore, Theorem 1.2 in [20] can be applied for the discrete spaces $\tilde{\mathcal{X}}_{h,w}$ and $\mathcal{N}_{h,0}$, and for the norms $\|\cdot\|_{1,\tilde{\Omega}}$ and $\|\cdot\|_{0,\omega}$. It leads to the following proposition concerning the existence and uniqueness of a solution to the discrete problem.

Proposition 6.15 *Under the regularity assumptions (6.3) and (6.4) on the domain $\tilde{\Omega}$, if the discrete spaces $\tilde{\mathcal{X}}_h$ and $\mathcal{N}_h$ are chosen such that the following discrete* inf-sup *condition holds :*

$$\begin{aligned} &\exists\ \tilde{\beta}_{h,w} > 0 \quad \textit{such that} \quad \forall\ \chi_h \in \mathcal{N}_{h,0} \quad \exists\ \boldsymbol{\psi}_h \in \tilde{\mathcal{X}}_{h,w} \quad \textit{such that} \\ &\qquad \tilde{b}_w(\,\boldsymbol{\psi}_h\,,\,\chi_h\,) \;\geq\; \tilde{\beta}_{h,w}\;\|\boldsymbol{\psi}_h\|_{1,\tilde{\Omega}}\;\|\chi_h\|_{0,\omega}\,, \end{aligned} \tag{6.67}$$

then, for every $\tilde{\boldsymbol{u}}_h \in (\tilde{\mathcal{X}}_h)'$, the discrete problem (6.65) has a solution $(\boldsymbol{u}_h, \eta_h)$, which is unique in $\tilde{\mathcal{X}}_{h,w} \times \mathcal{N}_{h,0}$ and satisfies

$$\|\boldsymbol{u}_h\|_{1,\tilde{\Omega}} + \|\eta_h\|_{0,\omega} \leq K_1 \|\boldsymbol{G}_h\|_{1,\tilde{\Omega}} + K_2 \|H_h\|_{0,\omega}, \tag{6.68}$$

where

$$K_1 = \frac{1}{c_A(\alpha)}\left(4+\frac{1}{c_A(\alpha)}\right)\left(1+2\left(\frac{a^n(\alpha,\delta)}{\tilde{\beta}_{h,w}}\right)^2\right) + \frac{2}{(\tilde{\beta}_{h,w})^2}$$

and $K_2 = \left(\frac{1}{c_A(\alpha)} + 2\left(\frac{a^n(\alpha,\delta)}{c_A(\alpha)\,\tilde{\beta}_{h,w}}\right)^2\right)\left(1+2\left(\frac{a(\alpha,\delta)}{\tilde{\beta}_{h,w}}\right)^2\right).$

We recall that $a(\alpha)=\alpha$ and $c_A(\alpha,\delta)=\delta+\alpha$ are respectively the continuity and coercivity coefficients of the bilinear form $\tilde{a}(\,\cdot\,,\,\cdot\,)$. In addition,

$$\|\boldsymbol{G}_h\|_{\mathcal{X}_h'} = \delta\,\|\tilde{\boldsymbol{u}}\|_{0,\tilde{\Omega}} \quad \text{and} \quad \|H_h\|_{\mathcal{N}_h'} = \delta\,\|\eta\|_{0,\omega}.$$

Note also that the finite element spaces introduced in section 6.3.3 are compatible for the bilinear form $\tilde{b}_w(\,\cdot\,,\,\cdot\,)$ as well, with the norms in $H^1(\tilde{\Omega})$ and $L^2(\omega)$. Indeed, choosing

$$\tilde{\mathcal{X}}_h = \left(\mathcal{P}_{1,1}(\mathcal{T}_h) \oplus \mathcal{B}_{1,1}(\mathcal{T}_h^{2d})\right)^2$$

and

$$h^n = \mathcal{F}_{1,1}(\mathcal{T}_h^{2d}),$$

where $\mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$ is defined by (6.42), $\mathcal{B}_{1,1}(\mathcal{T}_h^{2d})$ by (6.43) and $\mathcal{P}_{1,1}(\mathcal{T}_h)$ by (6.47), the discrete inf-sup condition (6.67) is verified.

This result can be proven following almost the same arguments as in the proof of Proposition 6.13. Indeed, defining the finite element space

$$\mathcal{Y}_h = \left(\mathcal{F}_{1,1}(\mathcal{T}_h^{2d}) \oplus \mathcal{B}_{1,1}(\mathcal{T}_h^{2d})\right)^2,$$

we can show that Proposition 6.11 is still valid for the bilinear form $\tilde{b}_w(\cdot,\cdot)$ and the discrete spaces $\tilde{\mathcal{X}}_{h,w}$ and $\mathcal{N}_{h,0}$, observing that the restriction to $\mathcal{Y}_{h,0}\times\mathcal{N}_{h,0}$ of bilinear form $\tilde{b}_w(\,\cdot\,,\,\cdot\,)$ is equal to $\tilde{b}(\,\cdot\,,\,\cdot\,)$.

Chapitre 7

Application to the Telemac-3D system

7.1 Introduction

The Telemac-3D system solves the three-dimensional hydrostatic and non-hydrostatic free surface problems using the finite element method.

To deal with the moving domain, it uses the sigma-transformation approach, which is a particular ALE method described in Chapter 4. The two-dimensional domain ω is discretized by a triangular mesh $\mathcal{T}_h^{2d}$, so that the three-dimensional mesh $\mathcal{T}_h^n$ at each time t^n is prismatic, as illustrated by Figure 6.2. The three-dimensional velocity $\boldsymbol{U}$ and the dynamic pressure p_d are discretized by means of functions in the first order prismatic finite element space $\mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$ defined in (6.47), while the free surface η is discretized by means of two-dimensional piecewise linear functions – belonging to the finite element space $\mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$ defined in (6.42). For a precise description we refer the reader to [66].

The equations are splitted using the first order in time fractional step method. The solution of the hydrostatic, as well as the non-hydrostatic equations, involves the solution of a sub-problem which is similar to the $\boldsymbol{u} - \eta$ problem introduced in the previous chapter, though it doesn't have the same mathematical properties. We have modified the structure of the Telemac-3D system so that exactly the $\boldsymbol{u} - \eta$ problem analysed in the previous chapter is solved. We have then implemented the approximation of the $\boldsymbol{u} - \eta$ problem into the system.

But the hydrodynamic problem also involves the resolution of the linear advection equations of the velocity components and the tracer. Several methods – all of them explicit – exist in the Telemac-3D system to solve these

equations : a characteristic method, the SUPG method and two MURD schemes. However, the latter were implemented for the particular use of the classical sigma transformation defined in section 4.8. When we generalized the Telemac-3D system to the use of a larger range of sigma transformations – *i.e.* of discrete ALE mappings – as described in section 4.8, we observed that the conservation property of these MURD schemes was deteriorated. Therefore, we updated these schemes using the ALES-MURD approach described in Chapter 3, so that they be compatible with any mapping.

In this chapter, we will describe the modifications opered on the algorithm solving the hydrostatic equations in the Telemac-3D system. Note that these modifications also affect the solution of the non-hydrostatic equations, since they involve a hydrostatic intermediate step. Especially, we will describe the implementation of the approximation of the $\boldsymbol{u}-\eta$ problem into the system. Some numerical results obtained with this *new algorithm* will then be shown and compared to the results obtained on the same test cases with the *standard algorithm*. Then, we will describe how the ALE-MURD schemes introduced in Chapter 3 can be applied to solve the advection equation of the horizontal velocity and the tracer in the framework of the three-dimensional hydrostatic free surface problem. Note that we used this description when adapting the advection module in Telemac-3D to the generalized sigma coordinate system.

But let us first describe the general structure of the Telemac-3D system, more precisely the different fractional steps solved. In particular, we describe the standard algorithm solving the hydrostatic equations.

7.2 The Standard Telemac-3D system

7.2.1 The standard hydrostatic algorithm.

After time discretization, the hydrostatic problem (5.38)-(5.40) is splitted almost as described in section 5.13, except that the diffusion term in the momentum equation on the horizontal velocity $\boldsymbol{u}$ is also splitted. The system is thus divided into 4 sub-systems, solved independently :

– (h_1) the advection of the horizontal velocity $\boldsymbol{u}$,

$$\frac{1}{\Delta t}\ (\tilde{\boldsymbol{u}}-\boldsymbol{u}^n)\ +\ \boldsymbol{u}^{a_n}\cdot\boldsymbol{\nabla}_2\boldsymbol{u}^n\ +\ (w^*)^{a_n}\ \frac{\partial \boldsymbol{u}^n}{\partial z}\quad =\quad 0\quad \text{in}\quad \Omega^n,$$

where w^* is the "pseudo-vertical velocity" defined in (4.22), accounting for the domain movement, and t^{a_n} is a particular time in $[t^{n-1},t^n]$,

– (h_2) the diffusion of $\boldsymbol{u}$,

$$\frac{1}{\Delta t}\ (\boldsymbol{u}^d-\tilde{\boldsymbol{u}})\ +\ \text{div}\left(\nu\boldsymbol{\nabla}\boldsymbol{u}^d\right)\quad =\quad \boldsymbol{f}^{hor}\quad \text{in}\quad \Omega^n,$$

– (h_3) the problem coupling $\boldsymbol{u}$ and η,

$$\left\{ \begin{array}{rcl} \dfrac{1}{\Delta t}\ (\boldsymbol{u}^{n+1} - \boldsymbol{u}^{d}) \ +\ g\ \boldsymbol{\nabla}_2 \eta^{\theta_n} & = & 0 \quad \text{in} \quad \Omega^n \\ \dfrac{g}{\Delta t}\ (\eta^{n+1} - \eta^{n}) \ +\ g\ \mathrm{div}_2\ (\displaystyle\int_b^{\eta^n} \boldsymbol{u}^{\theta_n}\ dz) & = & 0 \quad \text{in} \quad \Omega^n. \end{array} \right.$$

– (h_4) the continuity equation that allows to recover the vertical velocity w,

$$\mathrm{div}\, \boldsymbol{U}^{c_n} \ = \ 0 \quad \text{in} \quad \Omega^n,$$

where $\boldsymbol{U}^{c_n} = (\boldsymbol{u}^{c_n}, w^{c_n})$ and t^{c_n} is a particular time in $[t^n, t^{n+1}]$.

As in section 5.13, the velocity divergence and the free surface gradient terms in the $\boldsymbol{u} - \eta$ problem are semi-implicited using the θ-method. Time t^{θ_n} is determined by an implicitness parameter $0 \leq \theta \leq 1$, defining

$$t^{\theta_n} = \theta\ t^{n+1} \ +\ (1-\theta)\ t^n$$

for each $n = 0, \ldots, N_t - 1$. The corresponding variables are then

$$\eta^{\theta_n} \ = \ \theta\ \eta^{n+1} \ +\ (1-\theta)\ \eta^n \qquad \text{and} \qquad \boldsymbol{u}^{\theta_n} \ = \ \theta\ \boldsymbol{u}^{n+1} \ +\ (1-\theta)\ \boldsymbol{u}^n.$$

7.2.2 The standard non-hydrostatic algorithm.

The non-hydrostatic problem (5.46)-(5.49) is solved by splitting the time-discrete equations into a hydrostatic and a hydrodynamic part. The hydrostatic part corresponds almost to the hydrostatic problem (5.38)-(5.40) : it is composed of the momentum equation (5.38) on the horizontal velocity $\boldsymbol{u}$ and the free surface equation (5.39). But it also includes the advection and diffusion of the vertical velocity w. The hydrodynamic part consists in the correction by the hydrodynamic pressure term p_d of the total velocity $\boldsymbol{U} = (u, v, w)$ so that it becomes divergence free.

Each part is then splitted independently, dividing the non-hydrostatic system in the following 6 sub-systems :

– (nh_1) the advection of the horizontal velocity $\boldsymbol{u}$,

$$\frac{1}{\Delta t}\ (\tilde{\boldsymbol{u}} - \boldsymbol{u}^n) \ +\ \boldsymbol{u}^{a_n} \cdot \boldsymbol{\nabla}_2 \boldsymbol{u}^n \ +\ (w^*)^{a_n}\ \frac{\partial \boldsymbol{u}^n}{\partial z} \quad = \quad 0 \quad \text{in} \quad \Omega^n,$$

– (nh_2) the advection of the vertical velocity w,

$$\frac{1}{\Delta t}\ (\tilde{w} - w^n) \ +\ w^{a_n} \cdot \boldsymbol{\nabla}_2 \boldsymbol{u}^n \ +\ (w^*)^{a_n}\ \frac{\partial w^n}{\partial z} \quad = \quad 0 \quad \text{in} \quad \Omega^n,$$

– (nh_3) the diffusion of $\boldsymbol{u}$,

$$\frac{1}{\Delta t}\ (\boldsymbol{u}^d - \tilde{\boldsymbol{u}})\ +\ \text{div}\left(\nu \boldsymbol{\nabla} \boldsymbol{u}^d\right)\ =\ \boldsymbol{f}^{hor}\ \text{ in }\ \Omega^n,$$

– (nh_4) the diffusion of w,

$$\frac{1}{\Delta t}\ (w^d - \tilde{w})\ +\ \text{div}\left(\nu \boldsymbol{\nabla} w^d\right)\ =\ f_z\ \text{ in }\ \Omega^n,$$

– (nh_5) the problem coupling $\boldsymbol{u}$ and η,

$$\left\{\begin{array}{rcl} \dfrac{1}{\Delta t}\ (\boldsymbol{u}^h - \boldsymbol{u}^d)\ +\ g\ \boldsymbol{\nabla}_2 \eta^{\theta_n} & = & 0\ \text{ in }\ \Omega^n \\ \dfrac{g}{\Delta t}\ (\eta^{n+1} - \eta^n)\ +\ g\ \text{div}_2\ (\displaystyle\int_b^{\eta^n} \boldsymbol{u}^{\theta_n}\ dz) & = & 0\ \text{ in }\ \Omega^n\,, \end{array}\right.$$

providing the free surface function.

– (nh_6) the so-called continuity step,

$$\left\{\begin{array}{rcl} \dfrac{1}{\Delta t}\ (\boldsymbol{U}^{n+1} - \boldsymbol{U}^h)\ +\ \dfrac{1}{\rho}\ \boldsymbol{\nabla} p_d^{n+1} & = & 0\ \text{ in }\ \Omega^n \\ \text{div}\, \boldsymbol{U}^{n+1} & = & 0\ \text{ in }\ \Omega^n\,, \end{array}\right.$$

in which the total velocity is updated ensuring the continuity equation.

Note that the sub-systems (nh_1), (nh_3) and (nh_5) correspond exactly to (h_1), (h_2) and (h_3) in the hydrostatic algorithm presented in the previous section, except for the step ensuring the continuity equation. Therefore, in the Telemac-3D system, the solution of the non-hydrostatic problem involves the solution of the hydrostatic equations .

We indicate that sub-system (nh_6) is manipulated to obtain a Poisson equation on the hydrodynamic correction p_d^{n+1}. The total velocity $\boldsymbol{U}^{n+1}$ is then updated according to :

$$\boldsymbol{U}^{n+1}\ =\ \boldsymbol{U}^h\ -\ \frac{\Delta t}{\rho}\ \boldsymbol{\nabla} p_d^{n+1}\ \text{ in }\ \Omega^n\,.$$

We will now focus on the resolution of the problem coupling the horizontal velocity and the free surface function.

7.2.3 Standard algorithm solving the problem coupling $\boldsymbol{u}$ and η.

We describe now the practical procedure used up to now in the Telemac-3D system to solve sub-system (h_3) in the hydrostatic algorithm presented previously – or sub-system (nh_5) in the non-hydrostatic algorithm. We point out that this problem is equivalent to the $\boldsymbol{u}-\eta$ problem analysed in Chapter 6 without diffusion and without source term. However, it is solved in a very particular way in the Telemac-3D system : the equations are *depth-integrated* at the discrete level. Indeed, at each time t^{n+1}, the variables sought are on one hand the depth-averaged horizontal velocity

$$\bar{\boldsymbol{u}}^{n+1} \;=\; \frac{1}{h^n}\int_b^{\eta^n} \boldsymbol{u}^{n+1}\, dz\,,$$

and on the other hand the water height h^{n+1}. Sub-system 3. is thus solved in two steps :

– $(h_{3,a})$

$$\left\{\begin{array}{rcl} \dfrac{1}{\Delta t}\;(\bar{\boldsymbol{u}}^{n+1}-\bar{\boldsymbol{u}}^{d}) \;+\; g\;\boldsymbol{\nabla}_2 h^{\theta_n} & = & -\,g\boldsymbol{\nabla}_2 b \quad \text{in} \quad \omega\,, \\ \dfrac{g}{\Delta t}\;(h^{n+1}-h^n) \;+\; g\;\mathrm{div}_2\,(h^n\;\bar{\boldsymbol{u}}^{\theta_n}) & = & 0 \quad \text{in} \quad \omega. \end{array}\right.$$

This system is defined on the *two-dimensional* domain ω, and we observe that it is very similar to the *Saint-Venant* system without advection and without friction term. In fact, it is solved in the Telemac-2D system, which treats the two-dimensional shallow water equations.

– $(h_{3,b})$

$$\boldsymbol{u}^{n+1}(x,y,z) \;=\; \boldsymbol{u}^{d}(x,y,z) \;-\; \bar{\boldsymbol{u}}^{d}(x,y) \;+\; \bar{\boldsymbol{u}}^{n+1}(x,y)\,.$$

The vertical profile of $\boldsymbol{u}$ is retrieved afterwards using the averaged value $\bar{\boldsymbol{u}}$ and the value $\boldsymbol{u}^d$ obtained at the diffusion step.

Let us now focus on system $(h_{3,a})$. The imperviousness condition on the solid lateral boundaries is imposed weakly. Note that an analysis was made by Goutal in [56] of the time-discretized Saint-Venant system without advection and with essential treatment of the imperviousness condition at the solid lateral boundaries. Now this analysis cannot be applied to problem $(h_{3,a})$, since the diffusion is treated explicitly and the imperviousness condition is treated weakly. We will not analyse this problem either. However, we can make some relevant remarks.

Since there is no diffusion term on $\bar{\boldsymbol{u}}$, the operator on $\bar{\boldsymbol{u}}$ is not coercive on $H^{div}(\omega)^2 \times H^{div}(\omega)^2$: the existence of a solution to the space-continuous

problem can therefore not be proven in the Lax-Milgram framework. The analysis can nevertheless be made in the mixed problem framework. But this suggests the necessity, at the discrete level, of satisfying a particular *inf-sup condition.*

The algebraic form of the space-discrete problem is, for $n = 0, \ldots, N_t$,

$$\begin{pmatrix} \dfrac{1}{\Delta t}\, M_U^{2d} & C^T \\ -\, E^n & \dfrac{g}{\Delta t}\, M_H^{2d} \end{pmatrix} \begin{pmatrix} \bar{U}^{n+1} \\ H^{n+1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q}_{n+1}^{2d} \\ R_{n+1}^{2d} \end{pmatrix}, \qquad (7.1)$$

where $\bar{U}^{n+1}$ and H^{n+1} are the unknown vectors containing, respectively, the discrete values of the depth-averaged horizontal velocity $\bar{\boldsymbol{u}}_h$ and the water height h_h at $t = t^{n+1}$. M_U^{2d} and M_H^{2d} are the *two-dimensional* mass matrices in ω_h for $\bar{\boldsymbol{u}}$ and h respectively. The matrices C and E^n are defined by

$$C_{ij} \;=\; g\,\theta \int_{\omega_h} \boldsymbol{\nabla}_2\, \chi_i(x,y) \cdot \boldsymbol{\varphi}_j(x,y) \;\; dx\, dy,$$

and

$$(E^n)_{ij} \;=\; g\,\theta \int_{\omega_h} h^n\; \boldsymbol{\nabla}_2\, \chi_i(x,y) \cdot \boldsymbol{\varphi}_j(x,y) \;\; dx\, dy,$$

where χ_i denotes the basis function corresponding to the i-th degree of freedom for the water height h, and $\boldsymbol{\varphi}_j$ is the basis function corresponding to the j-th degree of freedom for the variable $\bar{\boldsymbol{u}}$. The terms $\boldsymbol{Q}_{n+1}^{2d}$ and R_{n+1}^{2d} account for boundary conditions and for the explicit terms.

We observe that the linear system to solve is not symmetric. In this form, we cannot determine easily whether or not it is invertible. However, spurious oscillations have been noticed in some numerical test cases. They seem to be due to the fact that the finite element spaces used to approximate the depth-averaged velocity $\bar{\boldsymbol{u}}$ and the water height h do not satisfy a particular inf-sup condition required. In [65], it is shown that these oscillations can be suppressed by either using a higher accuracy or by adopting a compatible discretization. For this reason, the so-called *quasi-bubble* element was implemented to enrich the linear finite element space $\mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$ for the approximation of the depth-averaged velocity $\bar{\boldsymbol{u}}$. This element is described in section 6.3.3.

Another algorithm, referenced to as the *pseudo wave equation*, has been implemented in the Telemac-2D system to avoid these spurious oscillation.

It consists in solving the dual system of problem $h_{3,a}$, that is

$$\left(\frac{g}{\Delta t}\ M_H^{2d}\ +\ G\right)\ H^{n+1}\ =\ R_{n+1}^{2d}\ +\ E^n\ \left(M_U^{2d}\right)^{-1}\boldsymbol{Q}_{n+1}^{2d}$$
$$\bar{U}^{n+1}\ =\ -\Delta t\ \left(M_U^{2d}\right)^{-1}C^T\ H^{n+1}\ +\ \left(M_U^{2d}\right)^{-1}\boldsymbol{Q}_{n+1}^{2d}\ ,$$

where

$$G\ =\ \Delta t\ E^n\ \left(M_U^{2d}\right)^{-1}C^T.$$

To simplify the problem, the matrix G is approximated by the two-dimensional diffusion matrix D_H^{2d}, with a particular coefficient $(c_n)^2$ depending on the time step Δt, the implicitness parameter θ, the variable h^n and on g :

$$G\ \approx\ (c_n)^2\ D_H^{2d}\ .$$

This choice is justified by the term in $\mathrm{div}_2\left(\boldsymbol{\nabla}_2 h\right)$ appearing when the same manipulation is made on system $h_{3,a}$ at the continuous level – *i.e.* eliminating $\bar{\boldsymbol{u}}$ in the continuity equation. The problem finally reduces to an equation of the form

$$\left(\frac{g}{\Delta t}\ M^{2d}\ +\ (c_n)^2\ D^{2d}\right)\ H\ =\ F,$$

which does not require the verification of the inf-sup condition to be well posed. Note that in dimension 1, by dividing it by Δt, this equation can be interpreted as a wave propagation of celerity c. For a more detailed description of this algorithm see [66] and [65]. We indicate that this solver is faster than the primal problem solver. However, spurious oscillations have been detected in some cases, and they appear to be lower when solving the primal problem with the quasi-bubble discretization.

In any case, the depth-integration of the equations has the obvious advantage of simplifying their resolution, since it leads to a considerably smaller linear system. Moreover, this system did not require an important implementation effort since the two-dimensional module was available. However, this approach also has some drawbacks.

1. It requires the resolution of an additional step, the diffusion step, which is three-dimensional and quite costly.
2. By splitting the diffusion term in the momentum equation, it modifies an important property of the operator on $\boldsymbol{u}$: its coercivity, which ensures the existence and uniqueness of a solution to the discrete problem when the viscosity coefficient is not negligible, without having to satisfy a particular *discrete inf-sup condition*. In addition, by depth-integrating the equations, the symmetry of the problem's mixed formulation is lost.

3. It forces the system to use a vertically-structured mesh in order to allow the exact depth-integration of the discrete functions.

We propose in the next section a new approach to solve the 3D hydrostatic problem – and through it also the non-hydrostatic problem.

7.3 New hydrostatic algorithm for the Telemac-3D system

7.3.1 The new hydrostatic algorithm.

We have modified the hydrostatic part of the Telemac-3D system in order to circumvent some of the drawbacks described previously. The new hydrostatic algorithm comprises the same sub-sytems (h_1) and (h_4) as the standard hydrostatic algorithm – see section 7.2.1 – , but it puts together sub-systems (h_2) and (h_3) into one unique problem coupling the horizontal velocity and the free surface : that is the $\boldsymbol{u}-\eta$ problem analysed in Chapter 6. The *new hydrostatic algorithm* is the following :

– (h_1') the advection of the horizontal velocity $\boldsymbol{u}$,

$$\frac{1}{\Delta t}\ (\tilde{\boldsymbol{u}}-\boldsymbol{u}^n)\ +\ \boldsymbol{u}^{a_n}\cdot\boldsymbol{\nabla}_2\boldsymbol{u}^n\ +\ (w^*)^{a_n}\ \frac{\partial \boldsymbol{u}^n}{\partial z}\quad =\quad 0\quad \text{in}\quad \Omega^n,$$

where w^* is the "pseudo-vertical velocity" defined in (4.22), accounting for the domain movement, and t^{a_n} is a particular time in $[t^{n-1},t^n]$,

– (h_2') the $\boldsymbol{u}-\eta$ problem,

$$\left\{\begin{array}{l} \dfrac{1}{\Delta t}\ (\boldsymbol{u}^{n+1}-\tilde{\boldsymbol{u}})\ -\ \text{div}\left(\nu\boldsymbol{\nabla}\boldsymbol{u}^{\theta_n}\right)\ +\ g\ \boldsymbol{\nabla}_2\eta^{\theta_n}\quad =\quad 0\quad \text{in}\quad \Omega^n \\ \\ \qquad \dfrac{g}{\Delta t}\ (\eta^{n+1}-\eta^n)\ +\ g\ \text{div}_2\ (\displaystyle\int_b^{\eta^n}\boldsymbol{u}^{\theta_n}\ dz)\quad =\quad 0\quad \text{in}\quad \Omega^n. \end{array}\right.$$

– (h_3') the continuity equation that allows to recover the vertical velocity w,

$$\text{div}\,\boldsymbol{U}^{c_n}\ =\ 0\quad \text{in}\quad \Omega^n,$$

where $\boldsymbol{U}^{c_n}=(\boldsymbol{u}^{c_n},w^{c_n})$ and t^{c_n} is a particular time in $[t^n,t^{n+1}]$.

Sub-system (h_2') is approximated by the discrete $\boldsymbol{u}-\eta$ problem (6.65) presented in section 6.4, involving the weak treatment of the imperviousness condition.

We would like to highlight the fact that this new algorithm implies no splitting of the diffusion term and no depth-integration of the equations. There are several reasons for testing such a solver.

On one hand, we have analysed the $\boldsymbol{u} - \eta$ problem in chapter 6. This analysis has shown that, when the viscosity coefficient is not negligible, the continuous problem as well as its approximation have a unique solution regardless of the satisfaction of a particular *inf-sup condition*. Now if the diffusion term is dropped, we don't know if this property holds true. In this case, the existence and uniqueness of a solution can be proven in the mixed problem framework provided a particular *inf-sup condition* is satisfied. It may thus be necessary to use compatible finite element spaces to approximate the problem in absence of the diffusion term. Therefore, it seems more convenient to keep the diffusion term inside the $\boldsymbol{u}$-η system for more stability.

On the other hand, solving the complete three-dimensional $\boldsymbol{u}-\eta$ problem may be less costly than solving successively the three-dimensional diffusion step (h_2), the depth-integrated problem on $\boldsymbol{u}$ and η ($h_{3,a}$), and deducing the three-dimensional horizontal velocity using equation ($h_{3,b}$). In addition, a reduction of the number of equation splittings improves the treatment of the boundary conditions.

Finally, the possibility for the Telemac-3D system to be extended to vertically-unstructured meshes seems interesting. Now the depth-integration of the equations in the system ($h_{3,a}$) coupling $\boldsymbol{u}$ and η does not allow such a discretization. On the contrary, solving the complete three-dimensional $\boldsymbol{u} - \eta$ problem (h_2'), especially with a weak treatment of the imperviousness condition at the solid lateral boundaries, allows any type of vertical discretization. Note however that the adaptation of the whole system to a vertically-unstructured mesh would in all cases require a considerable implementation effort, and has not yet been done.

7.3.2 Solving the discrete $\boldsymbol{u} - \eta$ problem.

As already mentioned, we approximate the $\boldsymbol{u} - \eta$ problem (h_2') by the discrete weak problem (6.65).

Its algebraic form is, for $n = 0, \ldots, N_t$:

$$\begin{pmatrix} A^n & (B^n)^T \\ -B^n & M_H^{2d} \end{pmatrix} \begin{pmatrix} U^{n+1} \\ S^{n+1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q}^{n+1} \\ R^{n+1}, \end{pmatrix} \tag{7.2}$$

where U^{n+1} and S^{n+1} are the unknown vectors containing the discrete values of, respectively, the *three-dimensional* horizontal velocity $\boldsymbol{u}^{n+1}$ and the *two-dimensional* free surface η^{n+1} at $t = t^{n+1}$. The time-dependent matrix A^n is defined as follows :

$$A^n \;=\; \frac{1}{\Delta t}\;\; M_n^{3d} \;+\; \theta\,\nu\; D^n,$$

where M_n^{3d} is the three-dimensional mass matrix in Ω_h^n for the variable $\boldsymbol{u}$ and D^n is the three-dimensional diffusion matrix in Ω_h^n.

The matrix B^n is the algebraic counterpart of the operator $\mathcal{B}$ defined in (6.8), that is :

$$(B^n)_{ij} \;=\; g\,\theta \int_{\Omega_h^n} \boldsymbol{\nabla}_2\, \chi_i(x,y) \cdot \boldsymbol{\psi}_j(x,y,z)\; d\boldsymbol{x}\,,$$

where χ_i denotes the *two-dimensional* basis function corresponding to the i-th degree of freedom for the water height h, and $\boldsymbol{\psi}_i$ the *three-dimensional* basis function corresponding to the i-th degree of freedom for the unknown $\boldsymbol{u}$.

Here again, the terms $\boldsymbol{Q}^{n+1}$ and R^{n+1} account for boundary conditions and for the explicit terms.

We point out that the linear system is *definite positive* whenever the viscosity coefficient is not zero – see the Lax-Milgram analysis of the discrete problem in section 6.4.3. In addition, the system has a symmetry property, since it gets *symmetric* by multiplying the second equation with -1. This makes it easier to analyse, and simplifies the implementation of its solution.

The structure of the matrices as well as the degrees of freedom of the unknown quantities depend on the particular choice of the space discretization. We have implemented the resolution of this linear system taking profit of the finite element basis already implemented in the Telemac-3D system. Therefore, we have chosen to approximate :

- the horizontal velocity $\boldsymbol{u}$ in the first order prismatic finite element space $\mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$ defined in (6.47),
- the free surface η in the Lagrangian finite element space $\mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$ defined in (6.42).

That means that the degrees of freedom of $\boldsymbol{u}$ and η are respectively the three-dimensional nodes in $\hat{\mathcal{T}}_h$, and the two-dimensional nodes in $\mathcal{T}_h^{2d}$. For the sake of clarity, we will denote by N_{3d} the amount of nodes in $\hat{\mathcal{T}}_h$ and by N_{2d} the amount of nodes in $\mathcal{T}_h^{2d}$. The dimensions of the vectors U^{n+1} and H^{n+1} are then, respectively, N_{3d} and N_{2d}.

It is not known if the inf-sup condition (6.67) is verified or not with these finite element spaces. We may therefore expect to obtain instabilities when the viscosity coefficient is negligible. In that case, the implementation of the resolution using the compatible spaces introduced in section 6.3.3 would be useful. Note that we have not yet implemented this alternative discretization.

Since the matrix B^n is quite particular, and its storage has been the most delicate part of the implementation, let us focus on its structure.

- As easily verified, B^n is a rectangular matrix of dimension $N_{2d} \times N_{3d}$.

No structure of this type exists in the finite element basis of the Telemac-3D system. Therefore, we have adapted an existing structure to the storage of the matrix. We point out however that this storage is not optimal.

Remark 7.3.1 The Telemac-3D system uses essentially the element-by-element (EBE) storage [55]. This technique is based on the fact that any matrix M can be decomposed on the elements of the mesh as follows :

$$M \;=\; \sum_{K} Q_K \; M_K \; (Q_K)^T,$$

where K denotes an element of the mesh, M_K is the elementary matrix corresponding to K and Q_K is the matrix which makes the transition between the element K and the global mesh – Q_K is derived from the connectivity table. Any matrix-vector product can then be performed easily using the elementary and the transition matrices. Therefore, the matrices are not assembled, only the elementary matrices are stored. This technique is easy to implement, and it is very efficient for finite element systems using iterative resolution methods – as it is the case in the Telemac-3D system –, because the essential matrix manipulation is then the multiplication with a vector. However, the EBE storage prevents from using preconditioning that requires the assembled matrix, for instance the uncomplete Cholesky decomposition or the Cahouet-Chabard [10] preconditioning developed for the Stokes problem.

– Another particularity of the matrix B^n is that it is defined as an interaction between two-dimensional and three-dimensional basis functions. Thus, particular care must be taken for the matrix-vector product. Indeed, B^n is decomposed in the following unusual way :

$$B^n \;=\; \sum_{P \in \mathcal{T}_h^{\,n}} Q_{T_P} \; B_P^n \; (Q_P)^T,$$

where Q_{T_P} is the transition matrix between the basis triangle T_P of prism P and the two-dimensional mesh $\mathcal{T}_h^{\,2d}$, while Q_P is the transition matrix between P and the three-dimensional mesh $\mathcal{T}_h^{\,n}$.
We have therefore adapted the modules performing the matrix-vector product to the treatment of the particular B^n matrix.

– The elementary matrices B_P^n have 3 rows and 6 columns. They are defined as follows for each element P in $\hat{\mathcal{T}}_h$:

$$(B_P^n)_{ij} \;=\; g\,\theta \int_P \boldsymbol{\nabla}_2 \chi_i(x,y) \cdot \boldsymbol{\psi}_j(x,y,z) \; dP\,,$$

where $i = 1, 2, 3$ is the local node in the basis triangle T_P of prism P, and $j = 1, \ldots, 6$ is the local node in prism P.
Since χ_i is a piecewise linear function in the two-dimensional mesh $\mathcal{T}_h^{2d}$, and since the three-dimensional mesh considered here is vertically structured – *i.e.* the prisms are arranged in columns with the same triangular basis, as illustrated in figure 6.5 –, its gradient is constant on each prism. Therefore :

$$(B_P^n)_{ij} \;=\; (\boldsymbol{\nabla}_2 \chi_i)_{|\,T_P} \cdot \int_P \boldsymbol{\psi}_j(x, y, z)\; dP,$$

where T_P denotes the triangular basis of the prism P. In addition, since the lateral faces of the prisms are vertical, we have that :

$$\int_P \boldsymbol{\psi}_1\; dP \;=\; \int_P \boldsymbol{\psi}_4\; dP,$$

$$\int_P \boldsymbol{\psi}_2\; dP \;=\; \int_P \boldsymbol{\psi}_5\; dP,$$

and

$$\int_P \boldsymbol{\psi}_3\; dP \;=\; \int_P \boldsymbol{\psi}_6\; dP,$$

where local nodes 1 and 4 – respectively 2 and 5, 3 and 6 – are located on the same vertical, as illustrated in figure 6.4.
As a result, each elementary matrix B_P^n has only 9 distinct values, we could thus store them in 3×3 - dimensional arrays. Unfortunately this is not possible in the Telemac system, because any value in an elementary matrix can be modified before the linear system is solved by the imposition of the Dirichlet boundary conditions. The symmetry described can then be broken.

Remark 7.3.2 Note that the third point is only valid for our particular discretization of the mesh and the variables. We point out however that the elementary matrices B_P^n can be constructed independently of this decomposition, for any choice of finite element spaces and on any three-dimensional mesh.

Once the matrices are constructed, system (7.2) can be solved. For this purpose, the efficient GMRES iterative method is used. We indicate that although it was already implemented in the Telemac-3D system, it had to be adapted to the structure of our system. No convenient preconditioning being available, the simple diagonal preconditioner is adopted. Of course, this is far from being optimal for the system we want to solve.

Remark 7.3.3 For the reasons previously exposed, the resolution of the complete $\boldsymbol{u} - \eta$ problem we have implemented is not optimal. We would like to point out that our aim up to now was to study the feasibility of the algorithm. The resolution could surely be optimized, on one hand, by using a structure adapted to the storage of the matrix B^n, improving the efficiency of the matrix-vector product. On the other hand, a better preconditioning could significantly improve the resolution of the linear system.

In the next section we present some results obtained with the new algorithm.

7.4 Evaluation of the new hydrostatic algorithm

In order to evaluate the *new hydrostatic algorithm*, it has been applied to a very large range of three-dimensional test cases. The results have been compared to those obtained on the same test cases with the *standard version* of the Telemac-3D system. In one particular case – the standing wave in a closed basin – the results have also been compared to the analytical solution of the linear equations for an inviscid fluid provided by the small amplitude wave theory.

We have chosen to present four of these numerical tests in this section. We will describe them and show the results obtained with both the standard and the new algorithms. The respective CPU times will not be specified, since the new implementation has not been optimized yet and the data would therefore not be significant. After analysing these results, we will conclude on the validity of the resolution and the interest of the new formulation.

7.4.1 Oscillating wave in a closed basin

We first present the results obtained on the test case of the oscillating wave in a closed basin, introduced in section 3.8.3 and already used twice in this work – see also test case 1 in section 4.8.4. A closed square basin with side L and with a constant bottom at $z = -H$ is considered, where H is the equilibrium water depth. In this test no tracer is contained in the fluid. At the initial time the fluid is at rest and the free surface profile imposed is the same as in section 3.8.3, *i.e.*

$$\begin{cases} \eta(x,y) = \eta_0 \cos(kx), \\ u = v = w = 0, \\ p_d(x,y,z) = -\rho_0 g \eta + \rho_0 g \eta_0 \dfrac{\cosh(k(z+H))}{\cosh(kH)} \cos(kx), \end{cases} \tag{7.3}$$

where p_d denotes the hydrodynamic pressure correction and where

$$\eta_0 = 0.1\,\mathrm{m},\ k = \frac{\Pi}{L},\ L = 10\,\mathrm{m} \text{ and } H = 10\,\mathrm{m}.$$

We recall that such an initial configuration generates a wave motion with continuous exchange of potential and kinetic energy. In addition, the fluid is assumed inviscid : there is no energy loss and the oscillation continues forever. Note also that since the initial free surface slope is parallel to the x-axis, no motion takes place in the y-axis direction.

In order to compare the results with the analytical formulas given by the small amplitude wave theory, the nonlinear terms are neglected. Therefore, and because there is no source term, the hydrostatic part of the discrete equations reduces to the $\boldsymbol{u} - \eta$ problem with no diffusion terms, that is subsystem (h_3) formulated in section 7.2.1 :

$$\left\{ \begin{array}{rcl} \dfrac{1}{\Delta t}\ (\boldsymbol{u}^{n+1} - \boldsymbol{u}^n)\ +\ g\ \boldsymbol{\nabla}_2 \eta^{\theta_n} & = & 0 \\ & & \text{in} \quad \Omega^n \\ \dfrac{g}{\Delta t}\ (\eta^{n+1} - \eta^n)\ +\ g\ \mathrm{div}_2\ (\displaystyle\int_b^{\eta^n} \boldsymbol{u}^{\theta_n}\ dz) & = & 0. \end{array} \right.$$

In this case, the new resolution of the hydrostatic problem is equivalent to its standard resolution. We therefore expect to obtain the same results with the two methods. This test case can in fact be used as a *validation test of the newly implemented scheme.*

The simulation is performed with a time step length of 0.1 seconds for a time interval of 30 seconds. The resolution is of approximately 0.2 m in the horizontal direction and 1 m in the vertical direction. The initial three-dimensional mesh is the same as in section 3.8.3. It counts 11 levels, distributed along the vertical in the discrete domain using the classical sigma-transformation. Figure 7.2 shows its vertical cross-section at $y = 5$, and Figure 7.1 the two-dimensional grid used to build it. Note that a semi-implicit scheme has been used for all simulations, using the implicitness parameter $\theta = 0.55$.

The simulation is first performed using the hydrostatic model. The computed water height at the basin boundary $x = 0$ m is shown as a function of time in figure 7.3. As expected, the free surface profiles obtained with the standard and the new hydrostatic algorithms are exactly the same. The wave

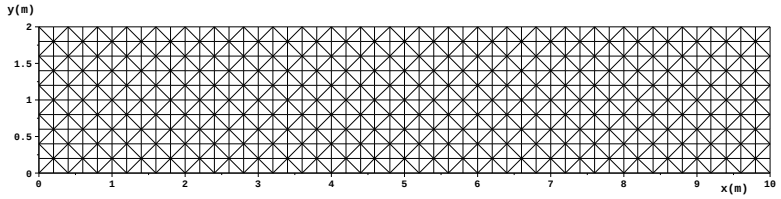


FIG. 7.1 – The horizontal 2D mesh.

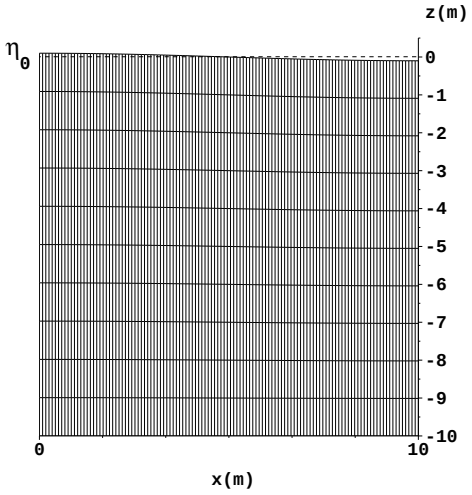


FIG. 7.2 – Vertical cross-section of the 3D mesh at $y = 5$ m.

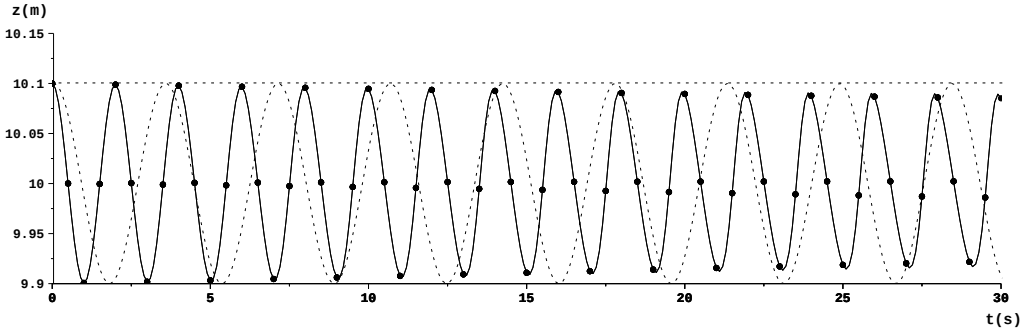


FIG. 7.3 – Temporal profile of the exact water height (dotted line) vs. the computed water height obtained using the hydrostatic model in Telemac-3D with the standard (solid line) and the new (solid line with plane circles) hydrostatic algorithms.

amplitude remains constant for a considerably long time, and is then progressively damped, due to numerical diffusion introduced by the fractional step method. This gives a measure of the energy loss. We also observe that the period of oscillation does not correspond to the analytical one. In fact, the hydrostatic model reproduces the unphysical long wave period.

We have therefore performed the same simulation using the non-hydrostatic model. The time series for water height obtained with this model are reproduced in Figure 7.4. Here again, the free surface profiles obtained with the two algorithms are identical. But this time the period of oscillation is reproduced exactly at the beginning of the simulation, although it progressively gets slightly longer. However, we observe that the wave damping is stronger with the non-hydrostatic model : the energy conservation is deteriorated. In fact, the numerical diffusion is increased because of the additional fractional step.

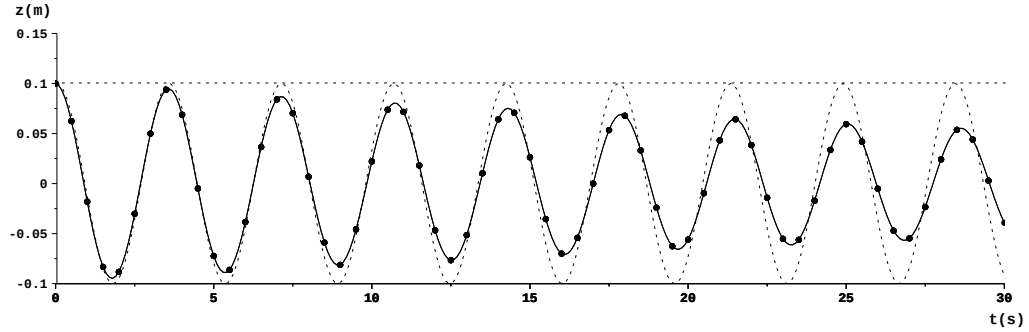


FIG. 7.4 – Temporal profile of the exact water height (dotted line) vs. the computed water height obtained using the non-hydrostatic model in Telemac-3D with the standard (solid line) and the new (solid line with plane circles) hydrostatic algorithms.

A comparison between the analytical velocity and the velocities obtained using the non-hydrostatic model in Telemac-3D with the standard and the new hydrostatic algorithms is given in vertical cross-sections at $y = 5\ m$. Figures 7.5, 7.6 and 7.7 show these results at respectively $t = 0.1$ s, $t = 1$ s and $t = 10$ s. We can say that the numerical results are in relative good agreement with the analytical solution. The velocities provided by the two different schemes are almost identical. Moreover, the conservation of the water mass is in both cases very satisfactory : the total mass-loss is of -0.5710^{-13} with the standard algorithm and -0.1210^{-13}, what can be considered as negligible.

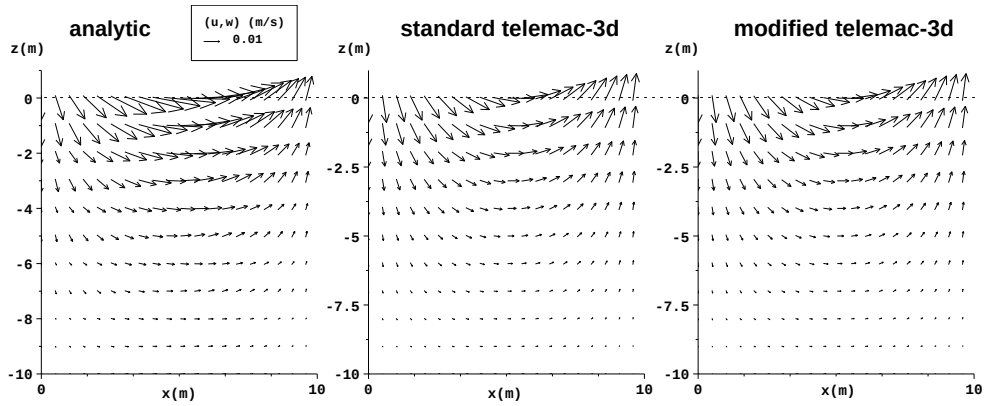


FIG. 7.5 – Analytical velocity at $t = 0.1$ s vs. the simulated velocity obtained using the non-hydrostatic model in Telemac-3D with the standard and the new hydrostatic algorithms.

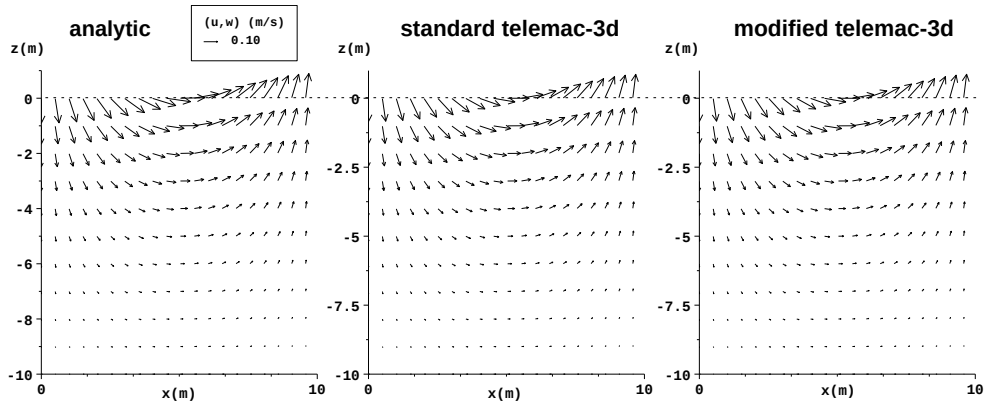


FIG. 7.6 – Analytical velocity at $t = 1$ s vs. the simulated velocity obtained using the non-hydrostatic model in Telemac-3D with the standard and the new hydrostatic algorithms.

7.4.2 Flow in an estuary

We now present the simulation of a flow in the transition zone between an estuary and the coastal waters, whose geometrical characteristics are very close to those of the Sado estuary in Portugal – see for instance [88]. The interest of this test case lies in the three-dimensional features of the flow caused by the strong bathymetry variations in the domain. Indeed, a recir-

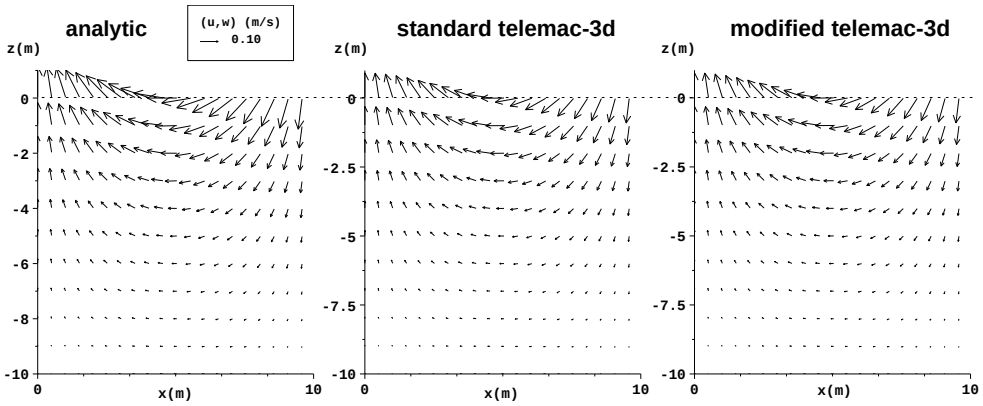


FIG. 7.7 – Analytical velocity at $t = 10$ s vs. the simulated velocity obtained using the non-hydrostatic model in Telemac-3D with the standard and new hydrostatic algorithms.

culation flow in the vertical plane is observed during the ebb period in the region of steep bathymetry. Note that this feature is considered important for the sediment transport in that region. Our purpose is to see if the numerical solutions provided by the two variants of the Telemac-3D system reproduce this phenomenon and if they are different.

We consider a chanel of length $L = 625$ m and width $l = 100$ m. Its ends are open, and the bottom elevation decreases from $z = 0$ to $z = -8$ in a non-uniform way. The initial three-dimensional mesh is shown in Figure 7.8, and its vertical cross-section at $y = 50$ m in Figure 7.9. It counts 10 levels along the vertical axis. The inflow boundary is situated at the left end of the chanel, the outflow boundary at its right end. Note that in order to improve the visualisation, a distorsion coefficient of 30 has been used for the vertical direction with respect to the horizontal one. The two-dimensional grid used to generate the 3D mesh is shown in Figure 7.10.

At the initial time, the free surface is constant at $z = 0$ m, and a uniform velocity field is prescribed in the x direction only. During the whole simulation, a constant discharge of $Q = 50$ m^2/s is prescribed at the inflow boundary, and a constant free surface $\eta_{out} = 0$ m is imposed at the outflow.

The water depth is very small with respect to the length of the chanel : the flow is therefore shallow water. However, the bottom gradients being relatively important, it is not clear whether or not the hydrostatic approxi-

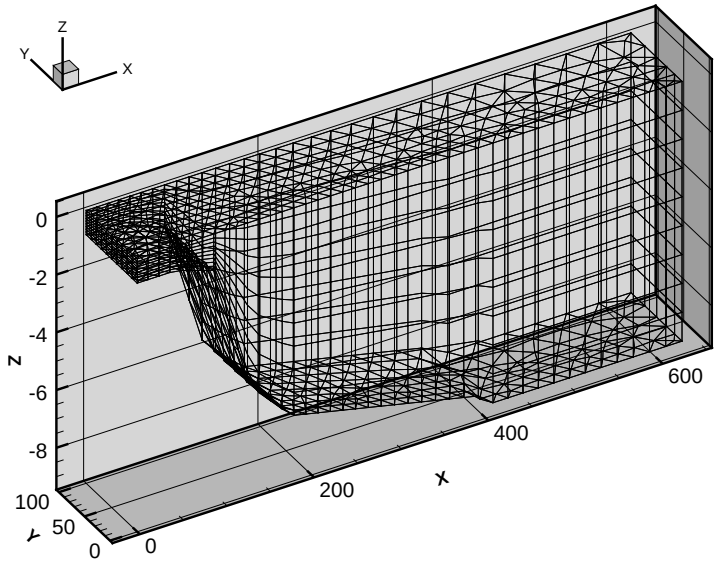


FIG. 7.8 – The initial 3D mesh.

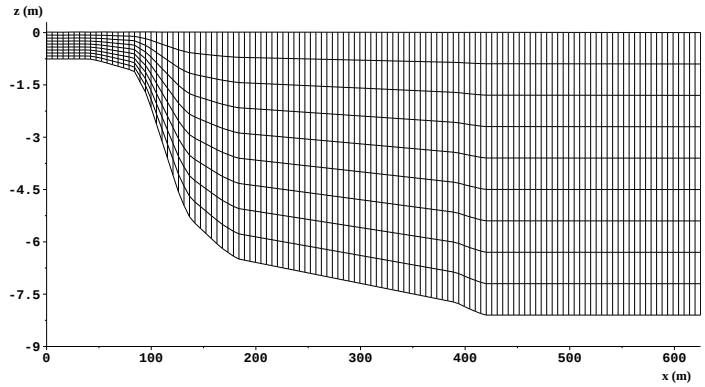


FIG. 7.9 – Vertical cross-section of the initial 3D mesh at $y = 50$m.

mation can be applied. Therefore the simulations are performed using both the hydrostatic and the non-hydrostatic models.

For the simulations we have used a time step of length $\Delta t = 2$ seconds, for a time interval of 2000 seconds. The first simulations have been performed with a horizontal viscosity of $\nu_h = 0.01$ m^2/s and a vertical viscosity coefficient of $\nu_v = 10^{-6}$ m^2/s. The friction is prescribed through a Strickler coefficient of 50. Note that the velocity advection step – subsystem (h_1) in

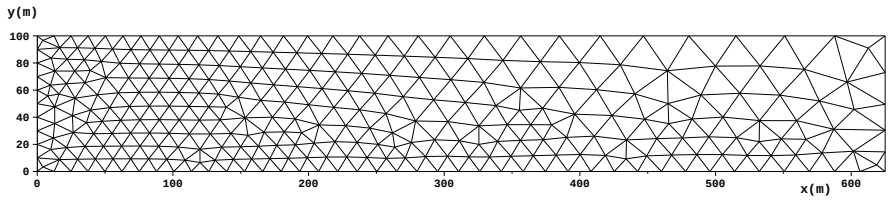


FIG. 7.10 – The horizontal 2D mesh.

section 7.2.1 for the hydrostatic model, subsystems (nh_1) and (nh_2) for the non-hydrostatic model – is solved using the Characteristic method.

The velocity profiles obtained after 2000 seconds using Telemac-3D with the new hydrostatic algorithm are shown in Figure 7.11, while those obtained using the standard algorithm are presented in Figure 7.12. Note that these figures show both the solution provided by the hydrostatic model (top figures) and the one provided by the non-hydrostatic model (bottom figures). We observe that the velocity profiles are very close : there is no relevant difference between the solutions provided by the standard and the new resolution, neither with the hydrostatic nor with the non-hydrostatic model.

The relative mass-loss at the end of the simulation is of 10^{-4} when using the new algorithm, and 2.10^{-4} when using the standard one.

As explained in [88], the recirculation appears when the high momentum of the top layers is not efficiently transported to the lower layers. Since the transport is mainly diffusive, that happens when the vertical diffusion of momentum is low, and this concerns as well the physical diffusion as the numerical one. In order to illustrate this statement, we have performed a second simulation with a higher vertical viscosity : $\nu_v = 0.02$ m^2/s. Figure 7.13 shows the results obtained with the hydrostatic model using the *standard* (top figure) and the *new* (bottom figure) algorithms.

We observe that for such a high vertical viscosity coefficient there is no recirculation. Again, there is no relevant difference between the results obtained with the two different algorithms.

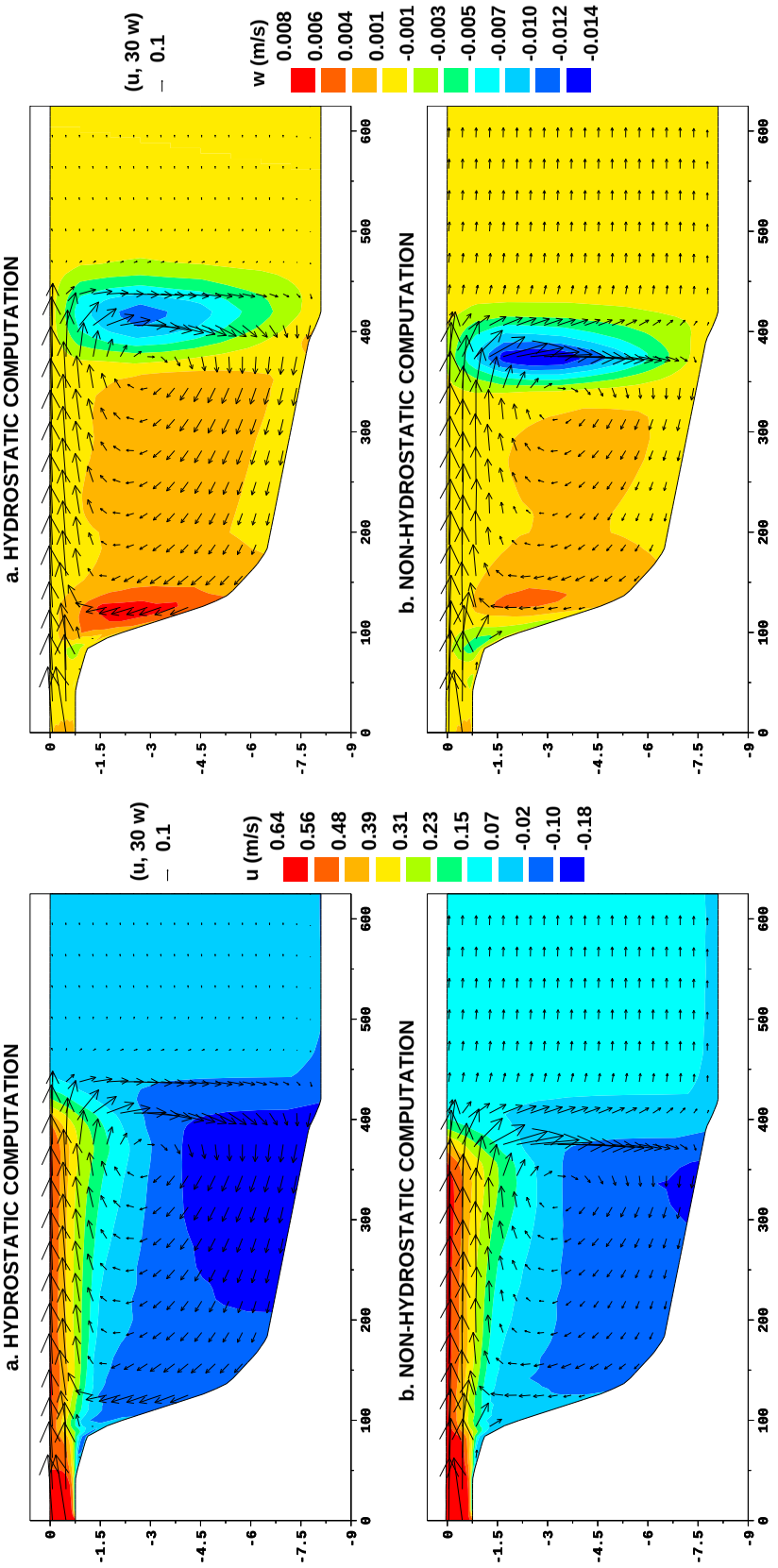


FIG. 7.11 – Horizontal and vertical velocities at $t = 2000$ s for $\nu_v = 10^{-6}$ m^2/s, using Telemac-3D with the NEW hydrostatic algorithm. Vertical cross-section at $y = 50$ m.

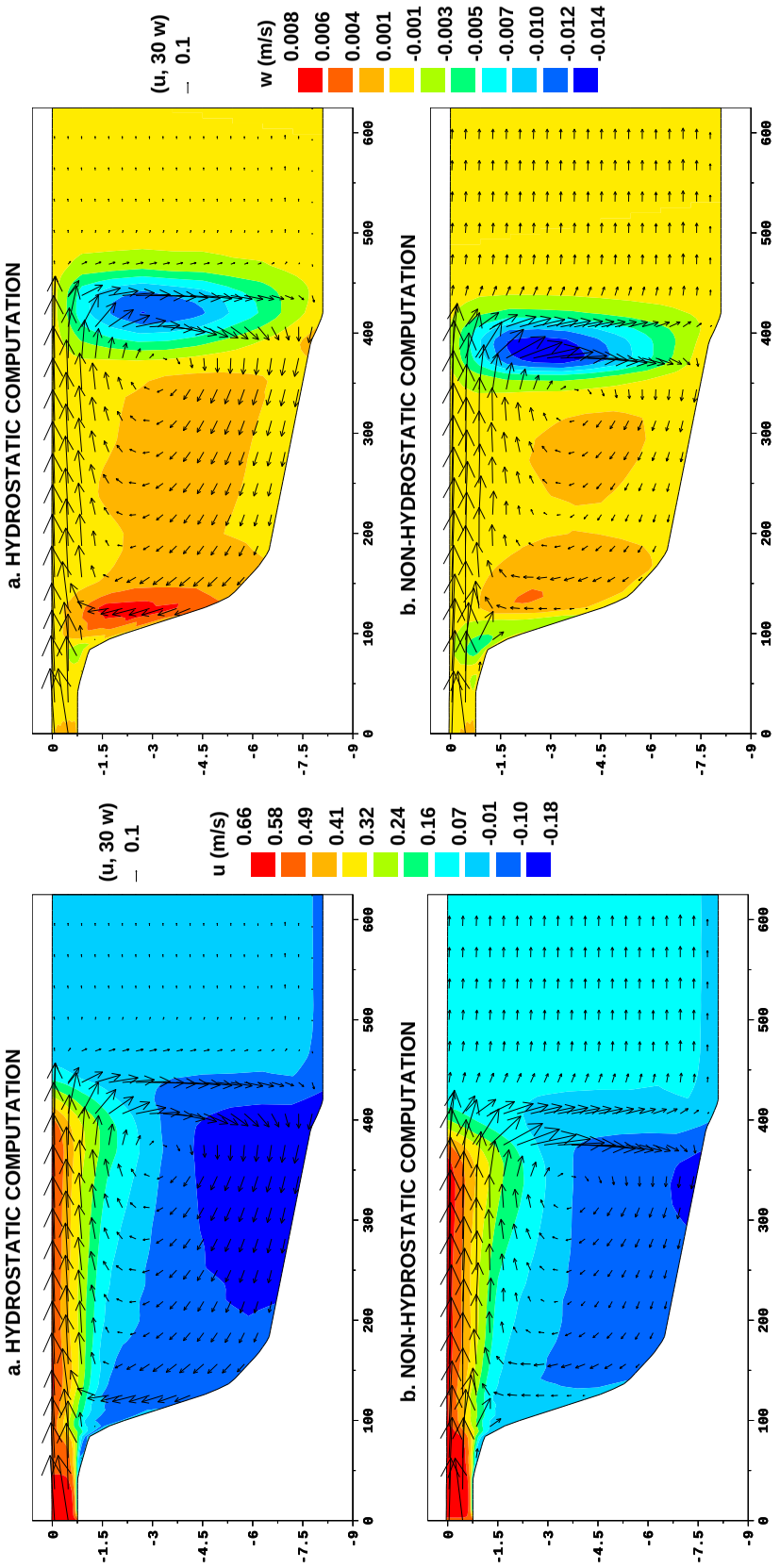


FIG. 7.12 – Horizontal and vertical velocities at $t = 2000$ s for $\nu_v = 10^{-6}$ m^2/s, using Telemac-3D with the STANDARD hydrostatic algorithm. Vertical cross-section at $y = 50$ m.

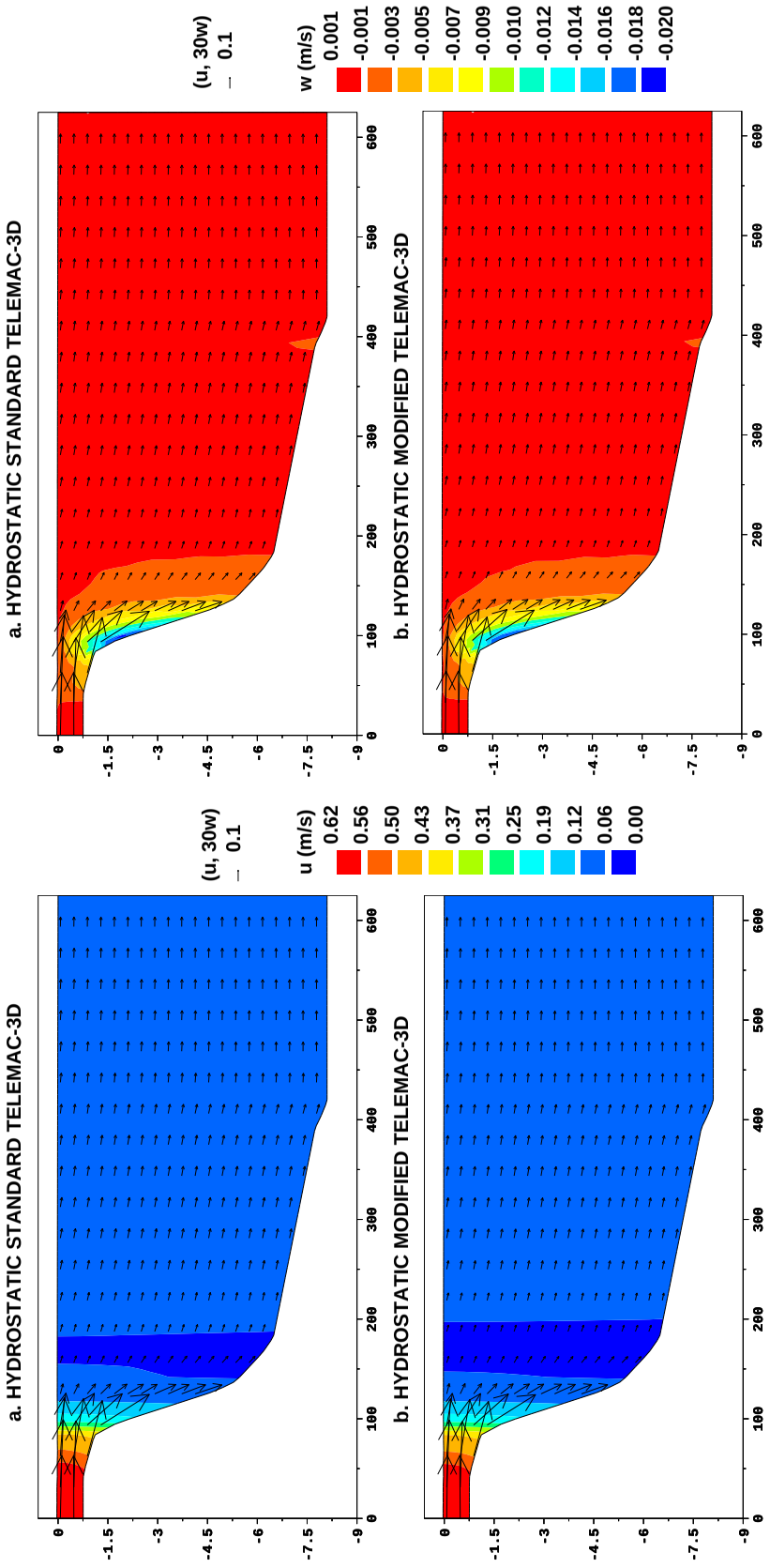


FIG. 7.13 – Horizontal and vertical velocities at $t = 2000$ s for $\nu_v = 0.02$ m^2/s, using the hydrostatic model of Telemac-3D with the (a) STANDARD and the (b) NEW hydrostatic algorithms.

7.4.3 Wind-driven circulation

An interesting test case is the simulation of a wind-driven flow in a long closed basin with shallow water. We consider a rectangular domain whose length, width and depth are respectively $L = 500$m, $l = 100$m and $b = -10.5$m. The initial three-dimensional mesh is shown in Figure 7.14. It has been obtained using the classical sigma transformation, distributing homogeneously 15 levels along the vertical. Note that a distorsion coefficient of 15 has been used for the vertical direction with respect to the horizontal one in order to improve the visualisation. The horizontal two-dimensional grid used is shown in Figure 7.15.

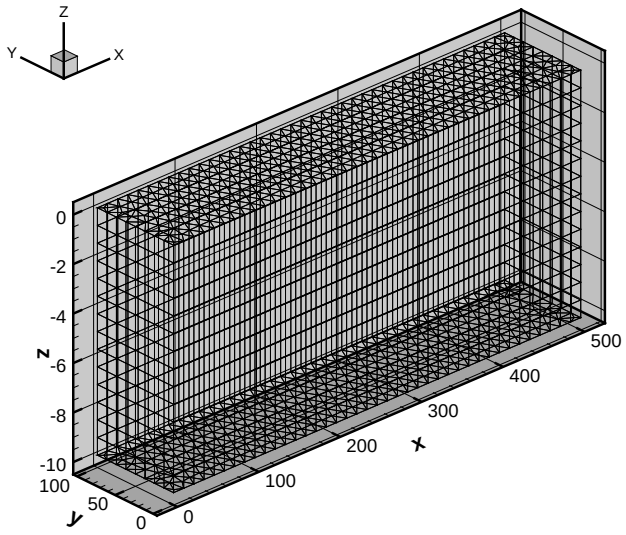


FIG. 7.14 – The initial 3D mesh.

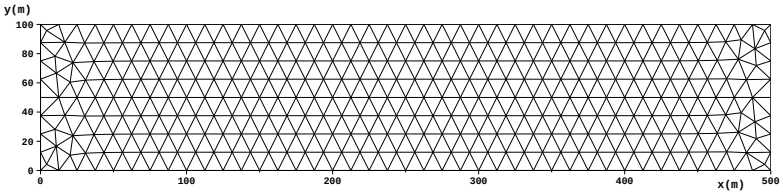


FIG. 7.15 – The horizontal 2D mesh.

At the initial time, the water height is constant and equal to 10.5 m, and

no motion is taken. A wind shear force is applied on the basin through the dynamic boudary condition at the free surface :

$$\boldsymbol{\nu}\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \;=\; \frac{\rho_{air}}{\rho}\, a_{wind}\, \boldsymbol{w}\, \|\boldsymbol{w}\| \quad \text{on} \quad \Gamma_{s,t}\,, \tag{7.4}$$

where ρ_{air} is the air density, taken equal to $1.29\, kg/m^3$, $\boldsymbol{u}$ is the horizontal velocity of the fluid and $\boldsymbol{w}$ is the wind velocity at 10 m above the water level. The coefficient a_{wind} is dimensionless and given by Flather [41]. (for a more detailed description of the wind forcing in the Telemac-3D system see [66]). In this simulation we consider a wind blowing only in the x-axis direction, whose velocity is 10 m/s.

The simulation is performed with a time step length of $\Delta t = 1$ s, for a time interval of 20000 seconds. The velocity advection step is solved using the PSI MURD scheme. The horizontal viscosity coefficient is taken constant at $\nu_h = 0.1\ m^2/s$, while the value of the vertical viscosity is succesively taken equal to $\nu_v = 0.1$, $\nu_v = 1.e-4$ and $\nu_v = 1.e-6\ m^2/s$. Note that although the flow is shallow water, we will use the non-hydrostatic model because of the strong accelerations of the vertical velocity generated by the wind.

We give an indication of the efficiency of the new hydrostatic algorithm compared to the standard one by indicating the number of GMRES iterations necessary to solve each step of these algorithms – see sections 7.2 and 7.3 to identify these steps.

Standard algorithm :

	$\nu_h = \nu_v = 0.1$	$\nu_h = 0.1,\ \nu_v = 10^{-4}$	$\nu_h = 0.1,\ \nu_v = 10^{-6}$
$(h_2)u$	1	2	2
$(h_2)v$	1	2	3
$(h_{3,a})$	1	1	1

New algorithm :

	$\nu_h = \nu_v = 0.1$	$\nu_h = 0.1,\ \nu_v = 10^{-4}$	$\nu_h = 0.1,\ \nu_v = 10^{-6}$
(h_2')	1	2	3

Note that $(h_2)_u$ and $(h_2)_v$ denote the systems corresponding to the diffusion of each component of the horizontal velocity, respectively u and v.

Test 1. : vertical viscosity $\nu_v = 0.1\ m^2/s$.

Figures 7.16 and 7.17 present, respectively, the computed water height and the velocities at $t = 20000$ s in the laminar case with a vertical viscosity of $\nu_v = 0.1\ m^2/s$. The results obtained using Telemac-3D with the standard and the new hydrostatic algorithms are almost exactly the same : the free surface leans towards the left end boundary of the domain, and a clear and widespread circulation is obtained over the entire area of the basin. Note that the fluid decelerates smoothly when reaching the wall.

In fact, the difference between the two resolutions only gets visible when a lower vertical viscosity is used. This can be explained by the fact that in that case the velocities are less homogenous along the vertical direction. We can thus expect that a three-dimensional resolution of the complete problem coupling the horizontal velocity and the free surface function provides different results than its standard resolution, which involves the splitting of the diffusion term and the depth-integration of the momentum equation.

Test 2. : vertical viscosity $\nu_v = 10^{-4}\ m^2/s$.

The results obtained at time $t = 20000$ seconds with a lower vertical viscosity coefficient of $\nu_v = 10^{-4}\ m^2/s$ are shown in Figures 7.18 and 7.19.

We observe clear spurious oscillations on the free surface, although they are slight since their maximal amplitude is of 3.10^{-3}m. They strongly look like oscillations due to a lack of *inf-sup conditions*. We could therefore think that the finite elements discretizing the horizontal velocity and the free surface do not verify the *inf-sup conditions* for the sub-problem $(h_{3,a})$ solved in the standard hydrostatic algorithm – see section 7.2.3.

The oscillations also occur when using the new hydrostatic algorithm, although they are lower. Since the vertical viscosity coefficient is very small – $\nu_v = 10^{-4} m^2/s$ – it is indeed possible that these oscillations arise from a lack of *inf-sup conditions*. Indeed, in Chapter 6 we have analysed the discrete $\boldsymbol{u} - \eta$ problem solved in the new hydrostatic algorithm – *i.e.* sub-problem (h_2'). We have proven for the case where the parameter α – corresponding to the viscosity coefficient – is not zero that the problem has a unique solution. We have also observed that the error estimation given by this analysis deteriorates when this parameter gets too small. For these cases – $\alpha = 0$ or α very small – we have proven the existence and uniqueness of a solution provided a particular *inf-sup condition* is verified. The oscillations observed in the actual test case could therefore come from the lack of this condition.

These oscillations have been detected very recently. Thus, we haven't really studied the problem yet. But we have some ideas on the matter that will be exposed in the conclusion of this section.

Let us now look at the velocity profiles shown in Figure 7.19. The profiles obtained using the two different hydrostatic algorithms are clearly different. We cannot say for sure which one is better, since the analytical solution is not known, but it looks like the velocity profile obtained with the standard hydrostatic algorithm is not correct. Indeed, we notice in figure b) that the velocity in the second top layer is inconsistent : this layer should be dragged by the first top layer on which the horizontal velocity is very strong. Now on the very left part of the second top layer the velocities change direction. This does not happen in the results shown in figure a) obtained with the new algorithm.

Test 3 : vertical viscosity $\nu_v = 10^{-6}\ m^2/s$.

Finally, the results obtained at time $t = 20000$ seconds with an almost negligible vertical viscosity coefficient of $\nu_v = 10^{-6}\ m^2/s$ are shown in Figures 7.20 and 7.21.

The oscillations detected in the previous test case are also visible on the results obtained with the standard hydrostatic algorithm in this test, and their amplitude has grown – the maximal amplitude is now 6.10^{-3} m. On the contrary, the oscillations are significantly lower using the new hydrostatic algorithm.

The velocity profiles obtained with the two algorithms are again very different. But this time, the results obtained with the new algorithm are clearly better. Indeed, figure b) presenting the profile obtained with the standard algorithm reveals a strong oscillation of the velocity on the four top layers. In addition, the vertical velocities on the left wall of the domain are inconsistent.

Therefore, it seems that the new hydrostatic algorithm is more stable than the standard algorithm in presence of low vertical viscosities.

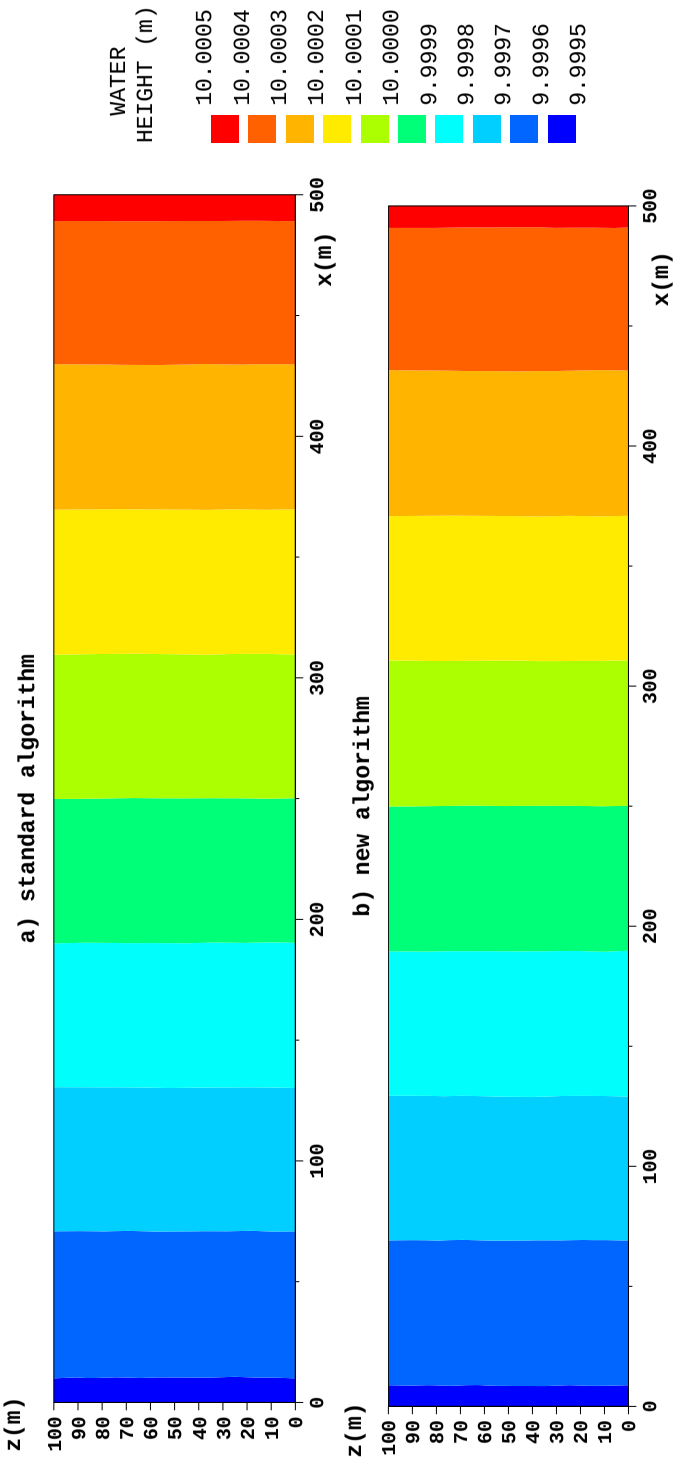


FIG. 7.16 – Free surface at $t = 2000$ s obtained for $\nu_h = \nu_v = 0.1\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the (a) standard and the (b) new hydrostatic algorithms.

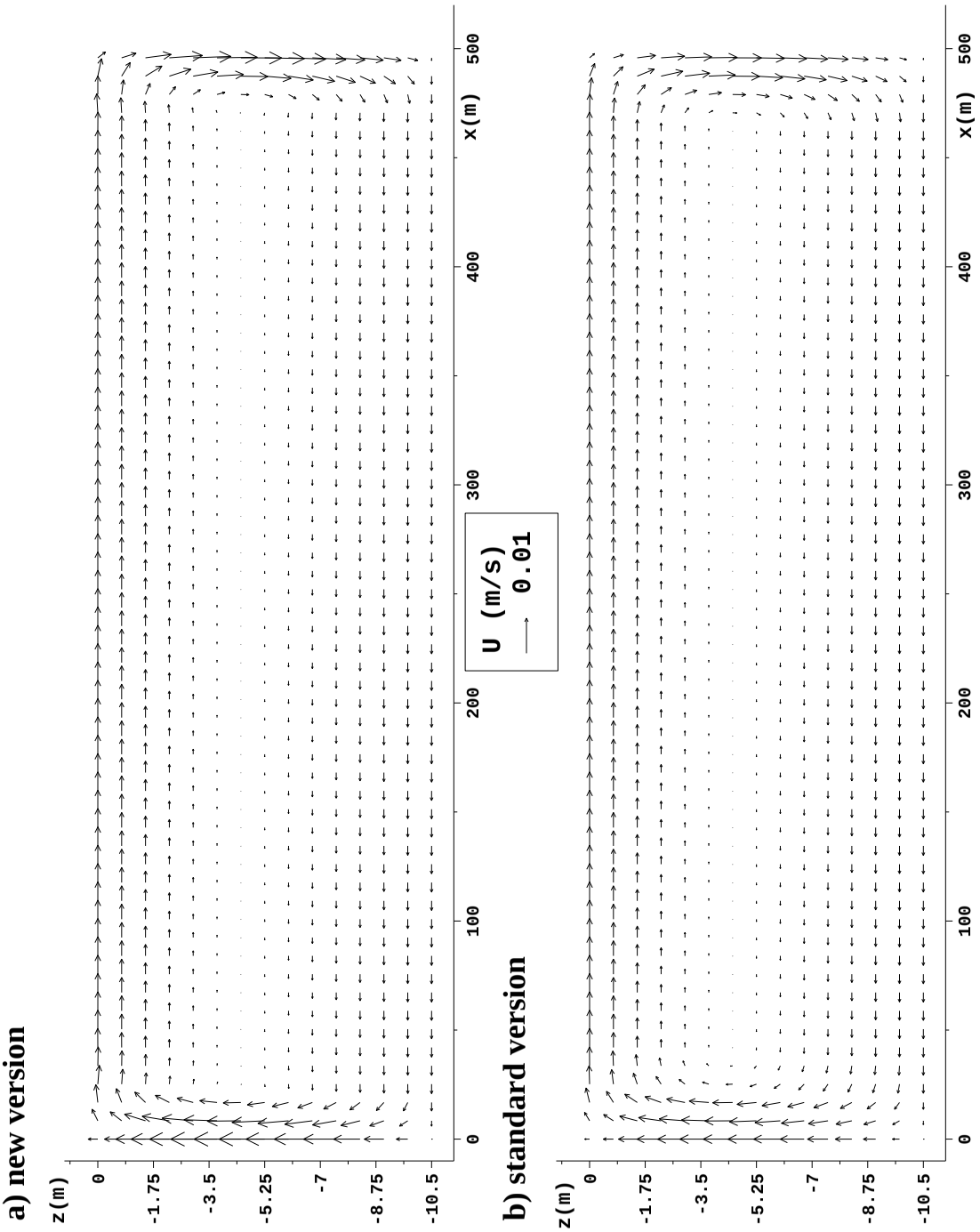


FIG. 7.17 – Velocity at $t = 2000$ s obtained for $\nu_h = \nu_v = 0.1\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the new (top figure) and standard (bottom figure) algorithms. Vertical cross-section at $y = 50$ m.

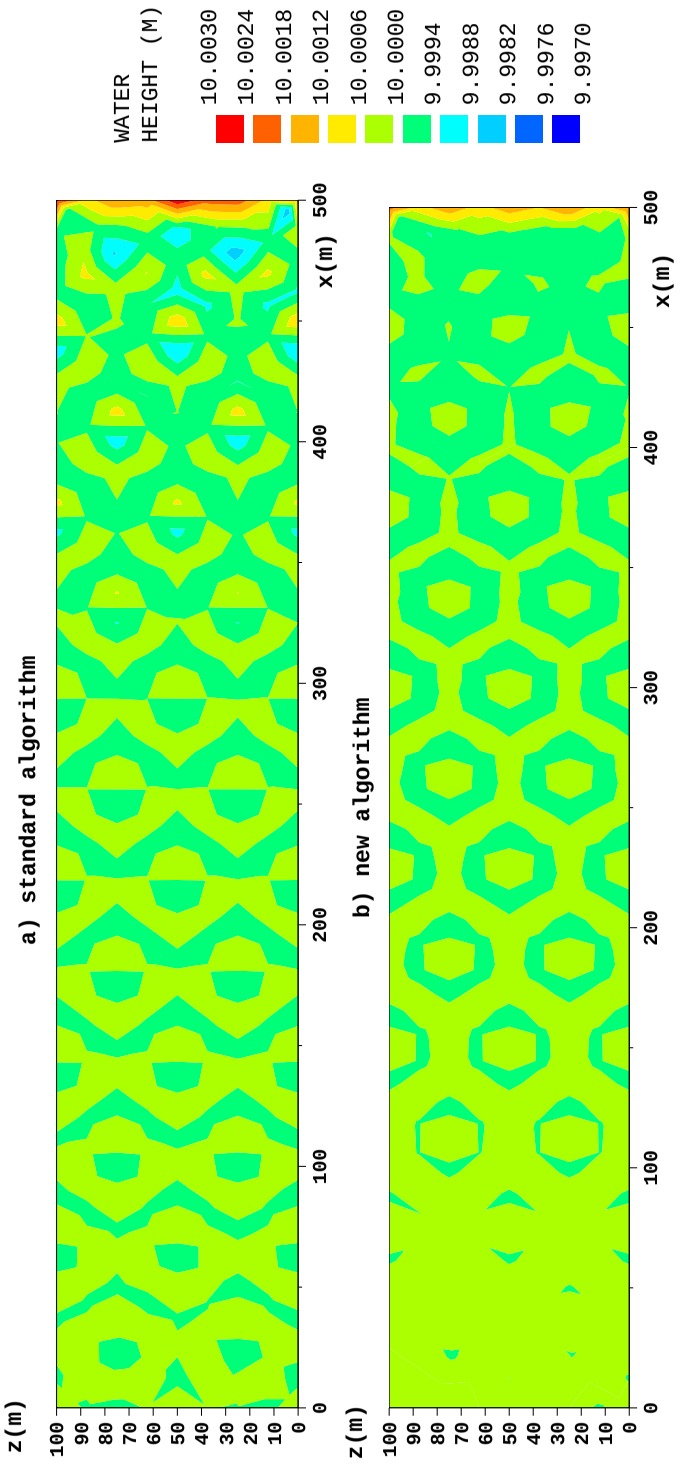


FIG. 7.18 – Free surface at $t = 2000$ s obtained for $\nu_h = 0.1\ m^2/s$ and $\nu_v = 10^{-4}\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the (a) standard and the (b) new hydrostatic algorithms.

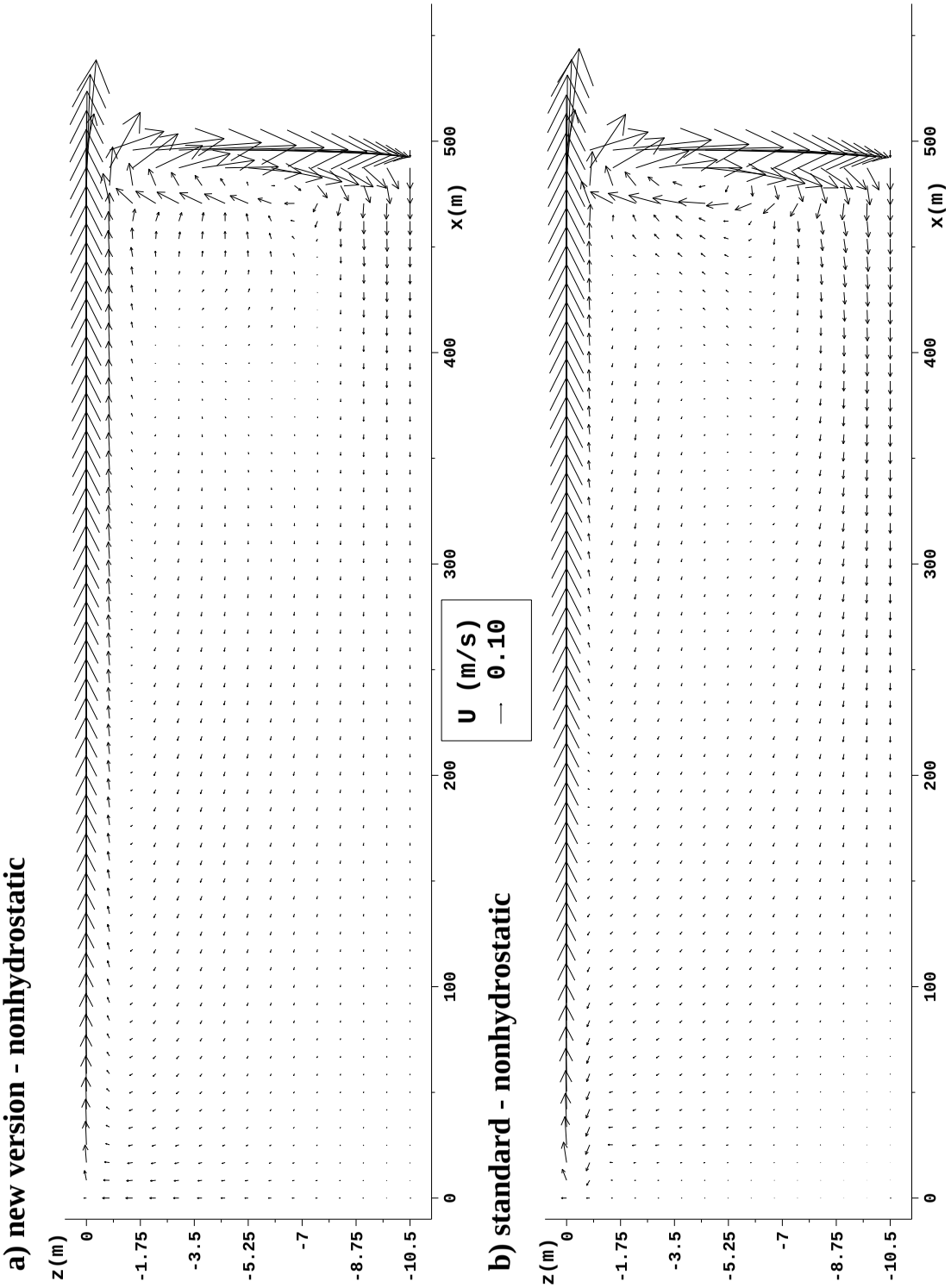


FIG. 7.19 – Velocity at $t = 2000$ s obtained for $\nu_h = 0.1\ m^2/s$ and $\nu_v = 10^{-4}\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the new (top figure) and standard (bottom figure) algorithms. Vertical cross-section at $y = 50$ m.

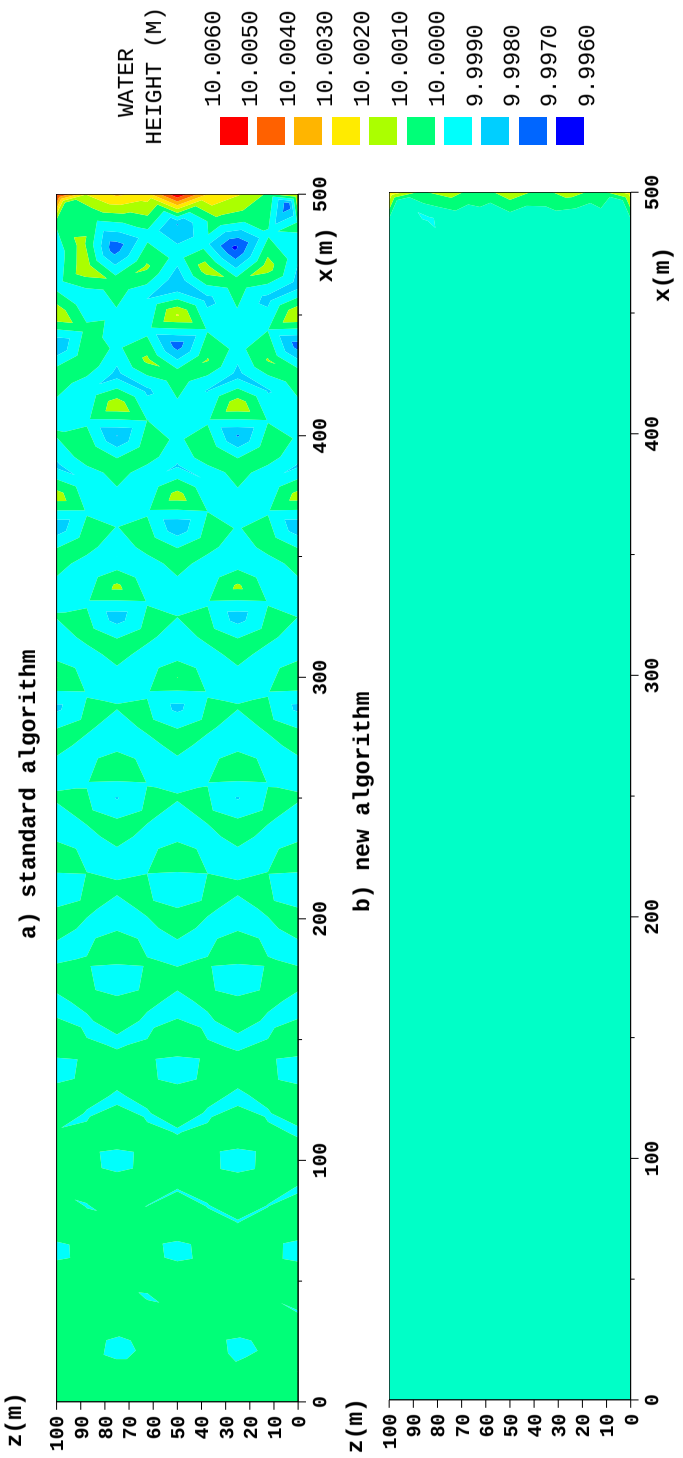


FIG. 7.20 – Free surface at $t = 2000$ s obtained for $\nu_h = 0.1\ m^2/s$ and $\nu_v = 10^{-6}\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the (a) standard and the (b) new hydrostatic algorithms.

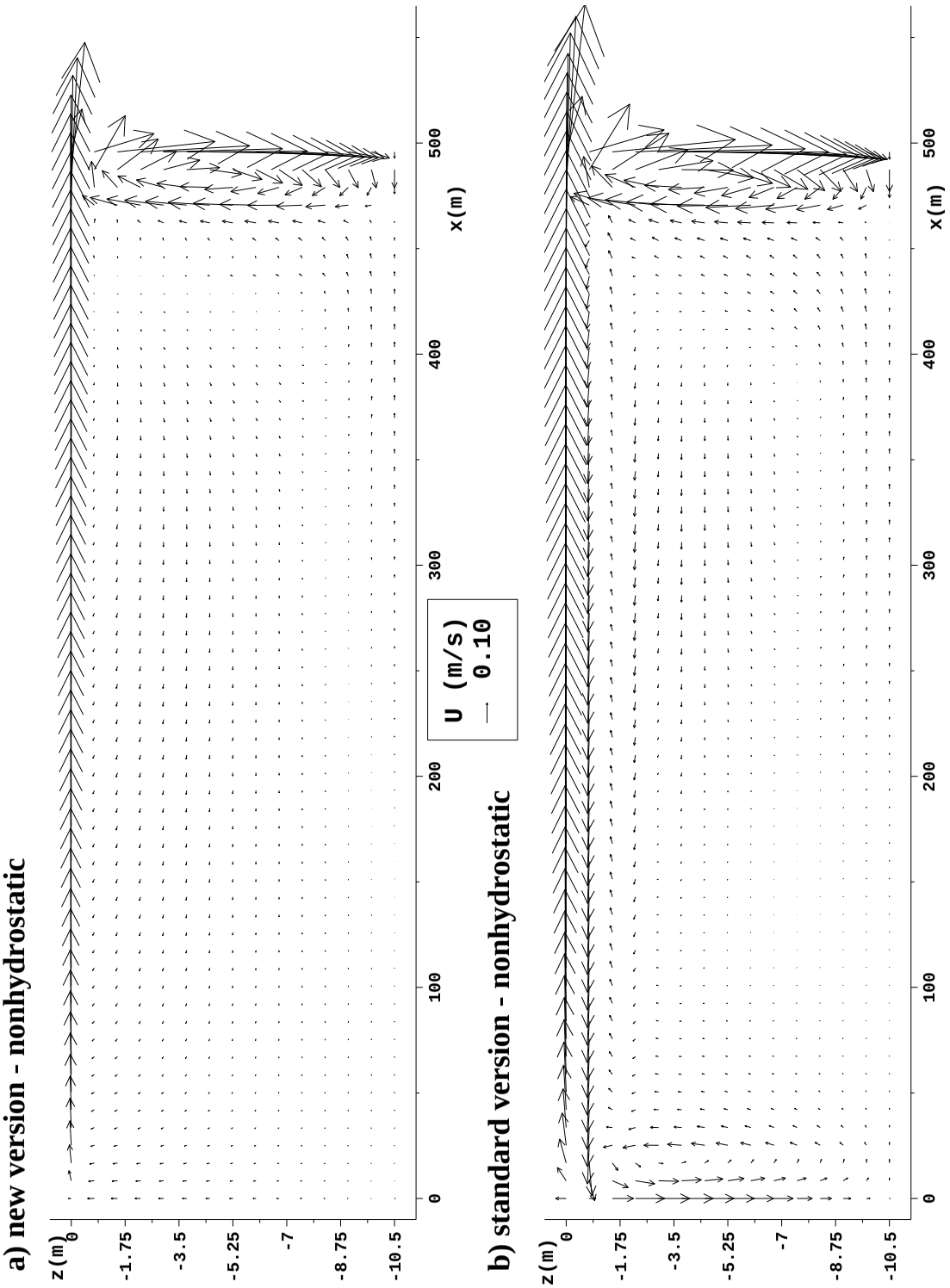


FIG. 7.21 – Velocity at $t = 2000$ s obtained for $\nu_h = 0.1\ m^2/s$ and $\nu_v = 10^{-6}\ m^2/s$, using the non-hydrostatic model of Telemac-3D with the new (top figure) and standard (bottom figure) algorithms. Vertical cross-section at $y = 50$ m.

7.4.4 Transcritical flow over a bump.

We consider now the classical test of the stationary flow over a parabolic bump, with a hydraulic jump downstream. It takes place in a chanel which is 2 m wide and about 21 m long. Its bottom is flat, at $z = -0.2$ m, except between the abscissae $x = 7.12$ m and $x = 12.88$ m, where there is a parabolic bump culminating at $x = 10$ m at elevation 0, according to the formula :

$$b(x, y) = \max(-0.2, -0.0246875(x - 10)^2)\,. \tag{7.5}$$

A vertical cross-section at $y = 1$ m of the initial three-dimensional mesh is shown in Figure 7.22. It counts 6 levels along the vertical axis, and the two-dimensional grid is shown in Figure 7.23.

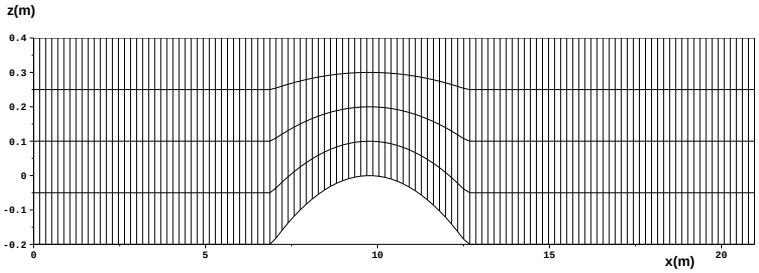


FIG. 7.22 – Vertical cross-section of the 3D mesh at $y = 5$ m.

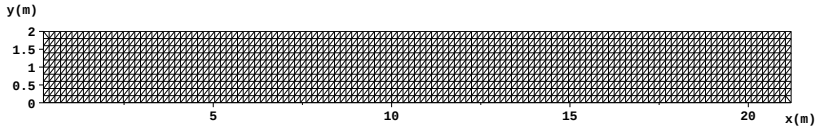


FIG. 7.23 – 2D mesh.

At the initial time the free surface is constant at $z = 0.4\,m$ and no motion is taken. But after the first time step a constant discharge of $Q = 2\ m^3/s$ is prescribed at the chanel's inflow boundary, as well as a depth of 60 cm at the outflow. The simulation is continued until a steady state is achieved. The flow is assumed to be viscous, with a constant horizontal viscosity coefficient $\nu_h = 10^{-4}$ m^2/s and a vertical viscosity coefficient given by a *mixing length turbulence model*. The friction is given trough a Strickler coefficient of 50.

For the simulations we have used a time step of length $\Delta t = 0.01$ seconds. Although the flow is shallow water, the flow is non-hydrostatic because of the strong bottom gradient and the strong surface gradient of the solution.

We therefore perform the simulations with both the hydrostatic and the non-hydrostatic models of Telemac-3D, and for each model we show the results obtained using the standard and the new hydrostatic algorithms. Note that the velocity advection step is solved using the Characteristic method.

Let us first indicate the number of GMRES iterations that have been necessary to solve the hydrostatic equations using the standard and the new algorithms :

standard algorithm :

(h_2) u	4
(h_2) v	6
$(h_{3,a})$	1

new algorithm :

(h_2')	1

The results obtained using the *hydrostatic model* of Telemac-3D are shown first. Figure 7.24 presents the three-dimensional profiles of the solutions obtained at $t = 100$ s. Note that in this figure, the inflow boundary is situated at the right end of the chanel, the outflow boundary at its left end. We observe strong oscillations both along the x- and the y-axis. We think that they arise from a serious problem in the treatment of the boundary condition on the vertical velocity at the inflow. This is a very common problem in hydrostatic models, since there is no advection nor diffusion on the vertical velocity w, and therefore no Dirichlet condition is to be imposed on w. This problem has been detected in several other test cases using the Telemac-3D system, but we have not yet tried to solve it. Anyway, it seems that the solution b) obtained with the new hydrostatic algorithm is more stable.

Figure 7.25 compares the free surface profiles obtained solving the hydrostatic model with the standard (solid line) and the new (dotted line) algorithms. The profiles are given at two different times $t = 50$ s and $t = 100$ s. We have verified that the schemes have converged at $t = 100$ s. Figure 7.26 compares the horizontal and vertical velocities obtained solving the hydrostatic model with the standard (a) and the new (b) algorithms. We observe slight differences in the free surface and the velocity profiles.

These differences disappear when using the non-hydrostatic Telemac-3D system. The surface profiles obtained with the standard and the new hydrostatic algorithms are shown in Figure 7.27 : they are identical. The velocity profiles are also identical. Therefore we only show one profile in Figure 7.28. In addition, there are no oscillations on the free surface.

Note that the dynamic pressure correction, shown at the bottom of Figure 7.28, reveals that the flow is hydrostatic everywhere except at the location of the hydraulic jump. Note also that the hydraulic jump is less marked and is

followed by an undulation, which is more in conformity with the observation.

The stability of this solution confirms our idea that the oscillations obtained with the hydrostatic model come from the bad prescription of the boundary condition on the vertical velocity w. Indeed, this problem does not exist in the non-hydrostatic model since the vertical velocity is advected and diffused, and therefore the Dirichlet boundary conditions on w can be prescribed correctly.

7.4.5 Conclusion and perspectives

We have implemented a new hydrostatic algorithm into the Telemac-3D system, involving the resolution of the $\boldsymbol{u} - \eta$ problem analysed in Chapter 6.

The main motivation for this alternative algorithm was the symmetry property of the $\boldsymbol{u} - \eta$ problem and its well-posedness whenever the viscosity coefficient ν is not negligible. Indeed, the space-continuous $\boldsymbol{u} - \eta$ problem as well as its approximation are definite positive for a non-zero value of ν.

But the evaluation of the feasability of this new algorithm was also motivated by the fact that it reduces the number of equation splittings and allows a complete three-dimensional treatment of the free-surface problem, opening a way to the extension of the Telemac-3D system to vertically-unstructured meshes.

The new algorithm has been validated through several test cases, in particular the case of an oscillating wave in a closed basin – see section 7.4.1 : the results of the simulation performed with the nonhydrostatic model in Telemac-3D using the new algorithm are in good agreement with the analytical results for the linear equations. In addition, they are identical to the results obtained using the standard algorithm.

Unfortunately, we have not found a test case showing a satisfactory improvement in the stability of the solution through the new algorithm. In fact, we were expecting to find a case in which the solution provided by the Telemac-3D system using the standard algorithm would have shown typical *inf-sup* oscillations for a non-negligible value of the horizontal viscosity coefficient. Because of the well-posedness of the approximation of the $\boldsymbol{u} - \eta$ problem when the diffusion coefficient is non-zero, these oscillations would have disappeared when using the new hydrostatic algorithm. But this "ideal" test case has not been found.

Nevertheless, in the test cases presented in sections 7.4.3 and 7.4.4, the results obtained using the new hydrostatic algorithm are less unstable than those obtained using the standard one.

The first one of these tests, corresponding to the simulation of a wind-driven circulation, has been performed recently. It is more interesting since spurious oscillations are observed for very small values of the vertical viscosity coefficent, which strongly look like oscillations due to a lack of *inf-sup conditions*. We could have verified this matter by analysing the kernel of the matrix inverted in this particular test case, but it seemed difficult because the matrices are not assembled in the Telemac-3D system. Note that the idea that the oscillations could come from a lack of inf-sup conditions is confirmed by the analysis of the $\boldsymbol{u} - \eta$ problem made in Chapter 6. Indeed, we have shown that the analysis in the Lax-Milgram framework is not appropriate when the viscosity coefficent is zero – or negligible . The existence and unicity of a solution to the $\boldsymbol{u} - \eta$ problem can then be proven in the mixed problem framework provided that a particular *inf-sup condition* is verified. This condition may not be satisfied at the discrete level in this particular test case, and that would explain the spurious oscillations observed when ν is very small.

We point out that a pair of compatible finite element spaces approximating the horizontal velocity and the free surface function has been proposed in section 6.3.3. The particular *inf-sup condition* possibly required by the $\boldsymbol{u} - \eta$ problem is satisfied on these spaces. It would therefore be worthwhile to implement this discretization and see if it makes the spurious oscillations disappear.

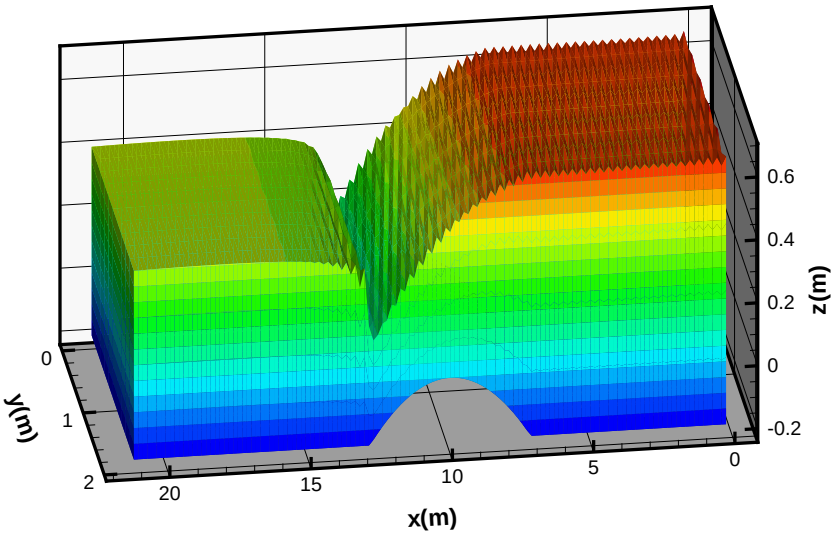
However, it can be shown that the vertical diffusion term is not necessary to the well-posedness of the $\boldsymbol{u} - \eta$ problem. Indeed, the analysis in the Lax-Milgram framework of the space-continuous and discrete problems – see sections 6.2.1 and 6.3.1 – can also be made in the case where the diffusion is only horizontal, *i.e.* when the bilinear form $\tilde{d}(\,\cdot\,,\,\cdot\,)$ defined in (6.6) is replaced by

$$\tilde{d}(\,\boldsymbol{\psi}_1\,,\,\boldsymbol{\psi}_2\,) \;=\; \int_{\tilde{\Omega}} \boldsymbol{\nabla}_2 \boldsymbol{\psi}_1 \cdot \boldsymbol{\nabla}_2 \boldsymbol{\psi}_2 \; d\boldsymbol{x}\,,$$

where $\boldsymbol{\nabla}_2$ is the horizontal gradient operator. An explicit treatment of the vertical diffusion in the $\boldsymbol{u} - \eta$ problem may therefore be sufficient to stabilize the problem when the vertical viscosity coefficient is very small.

We intent to test both ideas in a very near future. We also indicate that the new hydrostatic algorithm we have implemented may be optimized and integrated to the Telemac-3D system.

(a) standard hydrostatic algorithm.



(b) new hydrostatic algorithm.

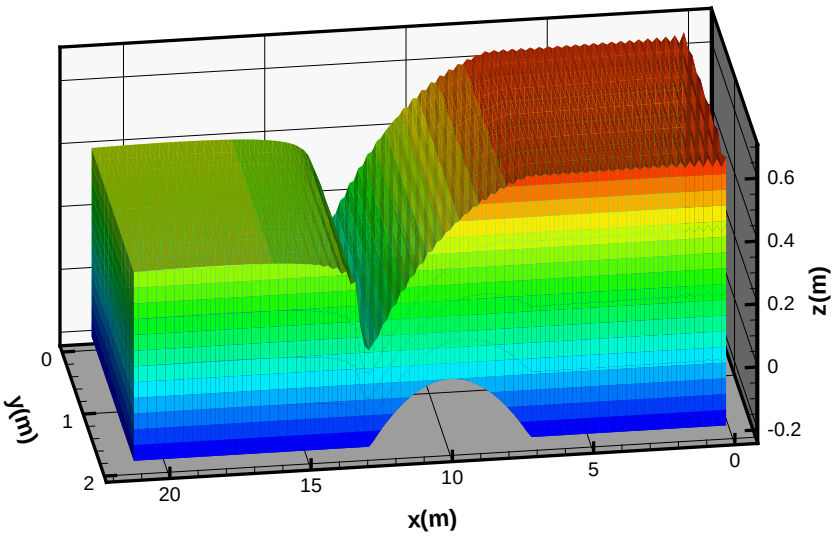
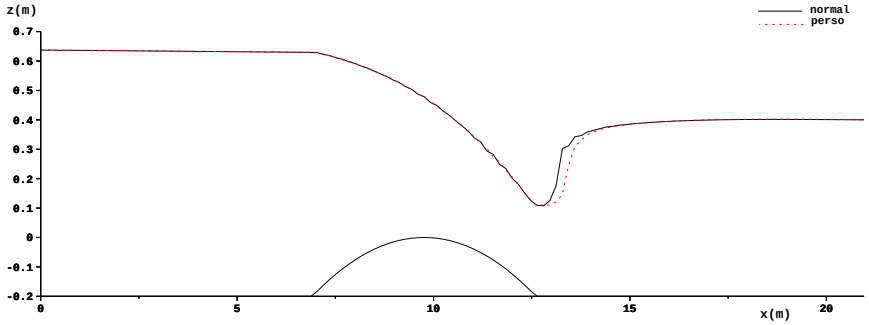


FIG. 7.24 – Three-dimensional results obtained at $t = 100$ s using the *hydrostatic model* of Telemac-3D with the standard (a) and the modified (b) hydrostatic algorithms.

(a) $t = 50$ s



(b) $t = 100$ s

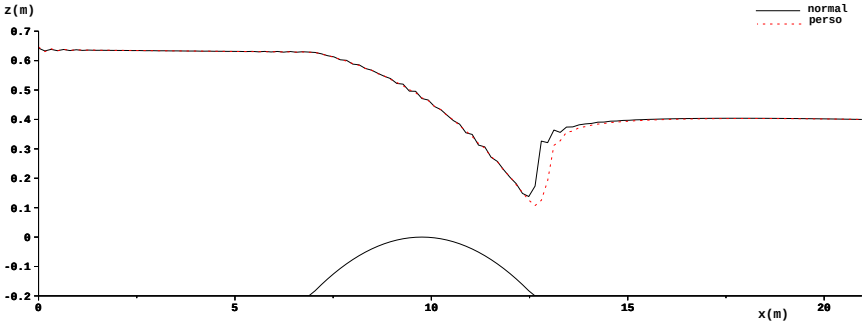


FIG. 7.25 – Free surface profiles obtained at (a) $t = 50$ seconds and (b) $t = 100$ seconds using the hydrostatic model of Telemac-3D with the standard (black solid line) and the new (red dotted line) algorithms. Vertical cross-section at $y = 1\ m$.

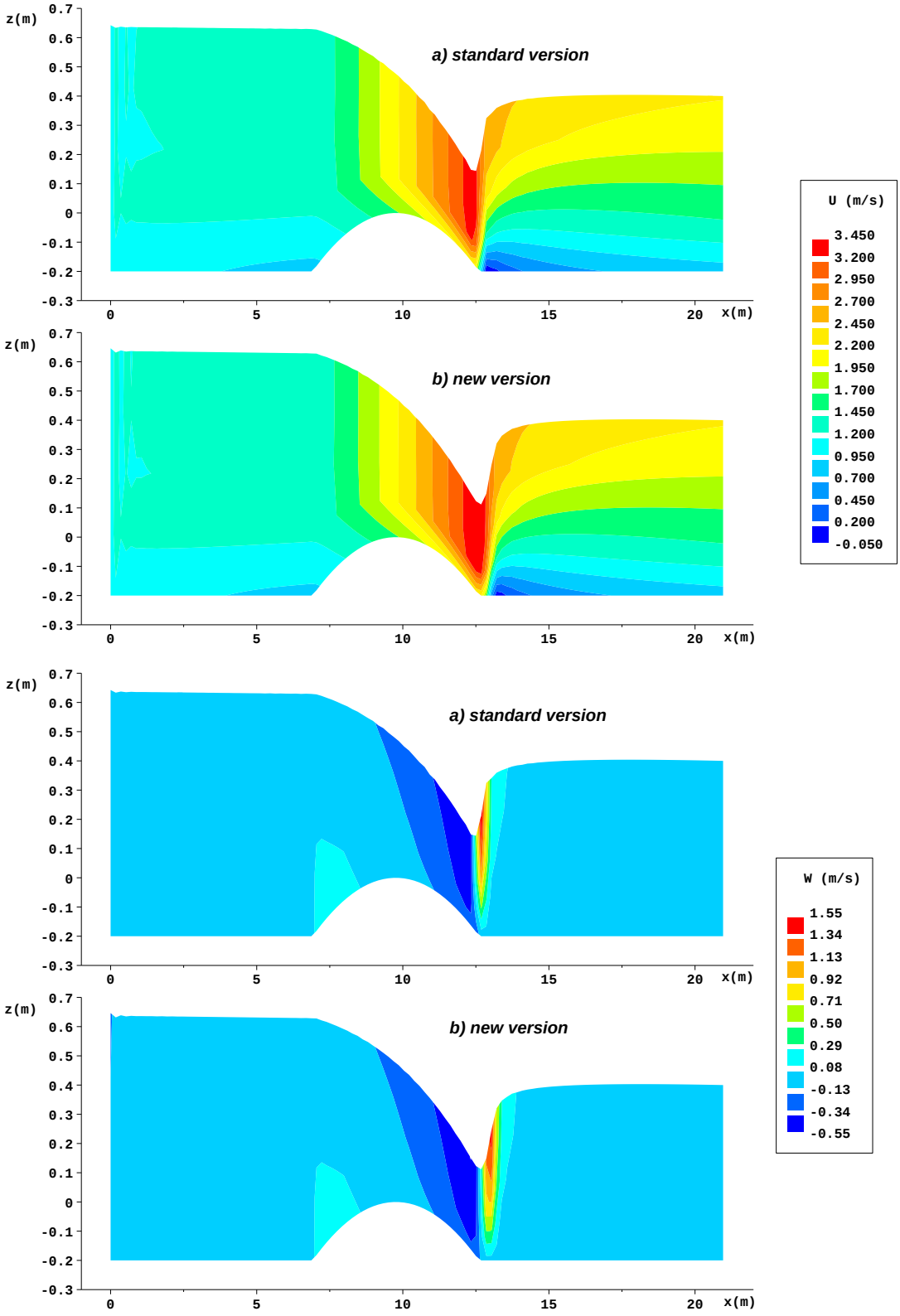
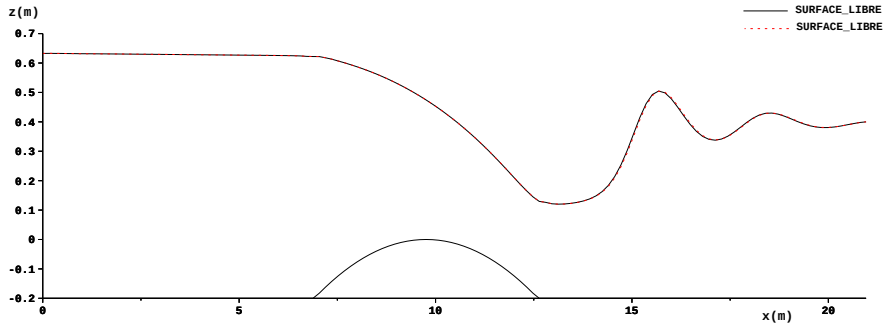


FIG. 7.26 – Horizontal (top figures) and vertical (bottom figures) velcoity profiles obtained at $t = 100$ s, with the *hydrostatic model*, using the standard (figures (a) and (c)) and the modified (figures (b) and (d)) algorithm.

(a) $t = 50$ s



(b) $t = 100$ s

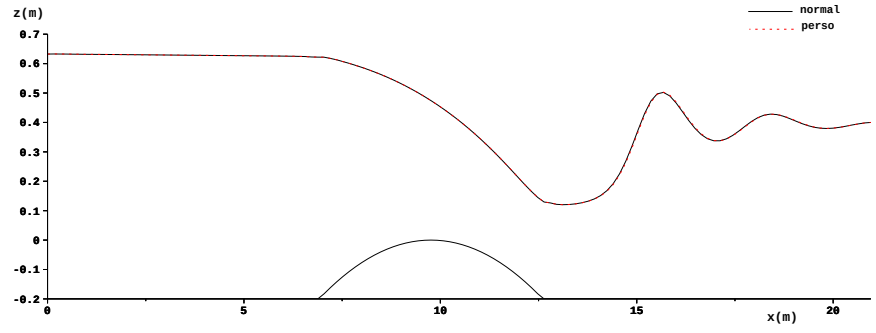


FIG. 7.27 – Free surface profiles obtained at (a) $t = 50$ seconds and (b) $t = 100$ seconds using the non-hydrostatic model of Telemac-3D with the standard (black solid line) and the new (red dotted line) algorithms. Vertical cross-section at $y = 1$ m.

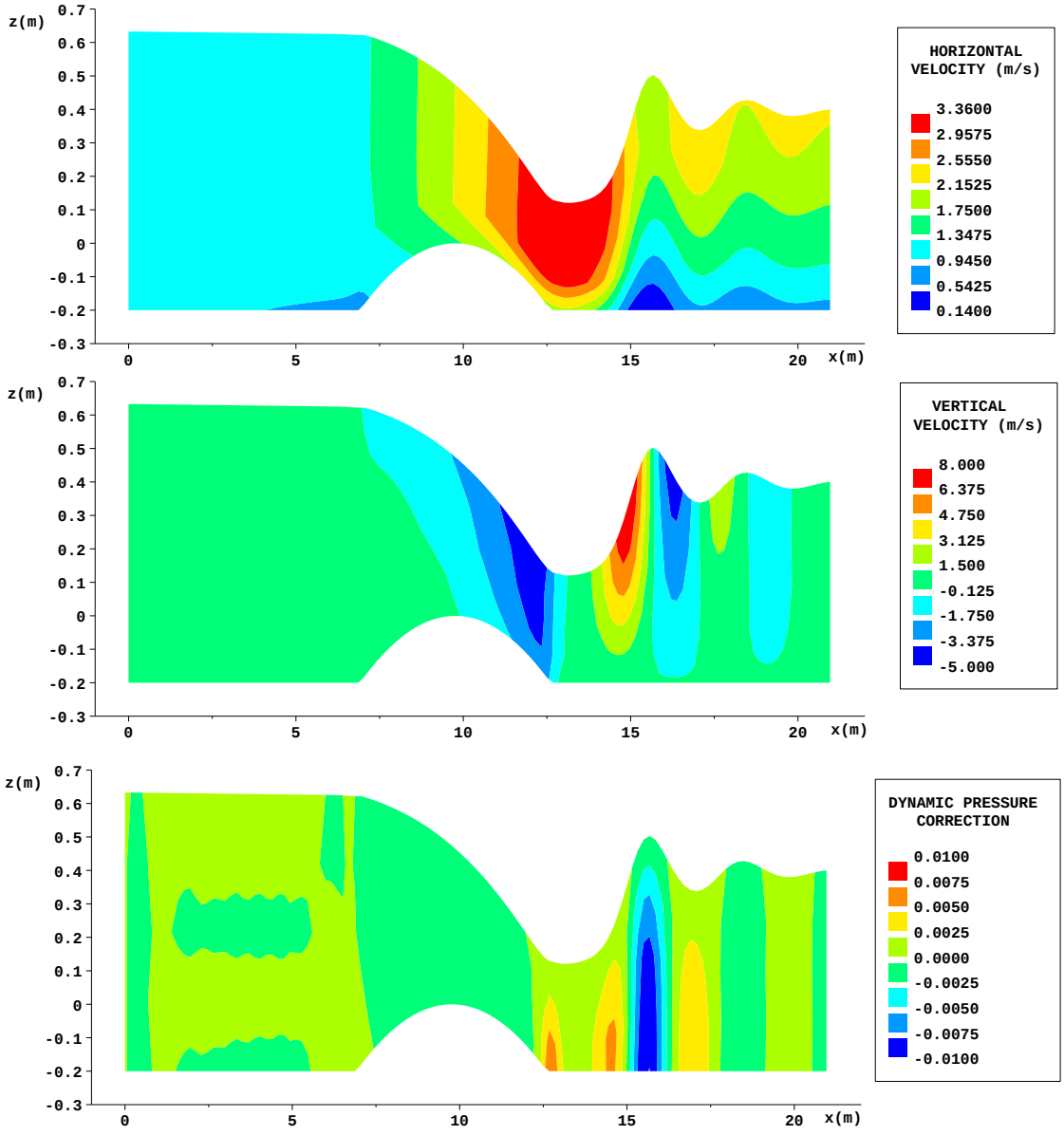


FIG. 7.28 – Horizontal (top figures) and vertical (bottom figures) velocity profiles obtained at $t = 100$ s, with the *nonhydrostatic model*, using the standard (figures (a) and (c)) and the modified (figures (b) and (d)) algorithm.

7.5 Application of an ALE-MURD scheme to the linear advection equations in the hydrostatic model

We are now interested in the resolution of the advection equations present in the hydrostatic model, using the ALE-MURD schemes introduced in Chapter 3. As explained in section 5.13, the fractional step method allows to isolate the ALE advective terms in the momentum equation (5.38) on the horizontal velocity $\boldsymbol{u}$, as well as in the tracer equation (5.20). An ALE-MURD scheme can thus be applied to the resolution of the linear advection equations on the horizontal velocity and on the tracer.

We recall that, in order to ensure the correct conservation of the advected quantities on a moving domain, the scheme must satisfy a particular constraint. In the framework analysed here, although the domain is three-dimensional and only moves in the vertical direction, we do not proceed as in section 3.8 to satisfy this constraint. Indeed, the vertical velocity w is not sought separately in the hydrostatic model presented in this work – see section 5.9. In fact, the vertical ALE velocity $\bar{w} = w - c$ is sought. It is recovered a posteriori from the horizontal fluid velocity $\boldsymbol{u}$ by solving the semi-discrete continuity equation (5.75). Therefore, the total velocity $\boldsymbol{U}$ can not be assumed divergence free in the sense of (3.30) and Proposition 3.10 – stating the conservation property under this assumption of the ALE-MURD schemes described in Chapter 3 – cannot be applied.

On the other hand, the conservation constraint of these schemes in the framework of the hydrostatic model coincides with a particular discretization of the weak formulation (5.40) of the continuity equation written in the ALE frame. This was already observed by Janin in [74], where he proposed an extension of the *N* and *PSI* distributive schemes to the advection of an active tracer driven by a three-dimensional hydrostatic free-surface flow, in the particular configuration of the Telemac-3D system. But his results were limited to the use of the classical sigma transformation, which is a particular ALE mapping described in Chapter 4. We have extended this observation to the general form of ALE mappings for three-dimensional domains moving in the vertical direction only. This has led us to deduce how to ensure the conservation property of the ALE-MURD schemes in the framework of the hydrostatic free surface flow problem.

Following this result, we have extended the MURD schemes implemented in the Telemac-3D system to the ALE-MURD aproach. They are now adapted to the use of any ALE mapping for three-dimensional domains moving in the vertical direction only. In particular, they are conservative for the use of any of these mappings, including the generalized sigma coordinate system introduced in section 4.8 and implemented lately into the Telemac-3D

system.

We will first recall the semi-discrete weak ALE formulation of the advection equations to be solved in the hydrostatic model considered. The conservation constraint of the ALE-MURD schemes in this particular framework will then be expressed. We will then show how to solve the three-dimensional continuity equation so that the advection scheme is conservative. Finally, we will present some results obtained with an ALE-MURD scheme in the Telemac-3D system. In particular, we will show the correct conservation of the advected quantities when using the generalized sigma coordinate system.

7.5.1 Formulation of the advection equation on the horizontal velocity

Let us first define $\hat{\mathcal{Y}} = H^1(\hat{\Omega})$ and its time-discrete counterpart at each time t^n :

$$\mathcal{Y}^n \;=\; \{\; \boldsymbol{\psi} : \Omega^n \longrightarrow \mathbb{R}, \quad \boldsymbol{\psi} \;=\; \hat{\boldsymbol{\psi}} \circ \mathcal{A}_n^{-1}, \quad \hat{\boldsymbol{\psi}} \in \hat{\mathcal{Y}} \;\}\,. \tag{7.6}$$

For the sake of simplicity, we will assume that the Dirichlet condition (5.16) on the horizontal velocity at the inflow is homogeneous, *i.e.* that

$$\boldsymbol{u}(x,y,z,t) \;=\; 0 \qquad \text{on} \quad \Gamma_{in,t}. \tag{7.7}$$

The horizontal velocity $\boldsymbol{u}^n$ at $t = t^n$ will therefore be sought in

$$\mathcal{Y}_0^n \;=\; \{\; \boldsymbol{\psi} \in \mathcal{Y}^n \quad / \quad \boldsymbol{\psi} = 0 \quad \text{on} \quad \Gamma_{in}^n \;\}\,. \tag{7.8}$$

For any time independent function $\hat{\boldsymbol{\psi}} \in \hat{\mathcal{Y}}$, we will use the notation :

$$\boldsymbol{\psi}^n \;=\; \hat{\boldsymbol{\psi}} \circ \hat{\mathcal{A}}_n \quad \text{in} \quad \Omega^{n_2}, \quad \text{for } n = 0, \ldots, N_t.$$

After splitting the advection terms in the momentum equation in (5.70), we obtain the *time-discrete non-conservative weak ALE form (5.73) of the linear advection equation on the horizontal velocity* $\boldsymbol{u}$, that is :

$$\text{given the initial condition } \boldsymbol{u}_0, \text{ for } n = 0, \ldots, N_t - 1,$$

$$\text{find} \quad \tilde{\boldsymbol{u}} = (\tilde{u}, \tilde{v}) \;\in\; (\mathcal{Y}_0^n)^2 \quad \text{such that} \quad \forall\, \hat{\boldsymbol{\psi}} = (\hat{\psi}_u, \hat{\psi}_v) \;\in\; (\hat{\mathcal{Y}})^2,$$

$$\int_{\Omega^n} (\tilde{u} \;-\; u^n)\; \psi_u^n \; d\boldsymbol{x} \;+\; \int_{\Omega^{a_n}} \left(\bar{\boldsymbol{U}}^{a_n} \cdot \boldsymbol{\nabla} u_{a_n}^n\right)\; \psi_u^{a_n} \; d\boldsymbol{x} \;=\; 0\,, \tag{7.9}$$

$$\int_{\Omega^n} (\tilde{v} \;-\; v^n)\; \psi_v^n \; d\boldsymbol{x} \;+\; \int_{\Omega^{a_n}} \left(\bar{\boldsymbol{U}}^{a_n} \cdot \boldsymbol{\nabla} v_{a_n}^n\right)\; \psi_v^{a_n} \; d\boldsymbol{x} \;=\; 0\,, \tag{7.10}$$

where, for each $n = 0, \ldots, N_t - 1$, t^{a_n} is a particular time in $[t^{n-1}, t^n]$. We recall that $\bar{\boldsymbol{U}} = \boldsymbol{U} - \boldsymbol{c}$ is the ALE advective velocity, with $\boldsymbol{U}$ the total velocity of the fluid and $\boldsymbol{c} = (0, 0, c)^T$ the domain velocity.

7.5.2 Formulation of the advection equation on the tracer

Again, for the sake of simplicity, we assume that the Dirichlet condition (5.21) on the tracer at the inflow is homogeneous, *i.e.* that

$$T(x, y, z, t) \;=\; 0 \qquad \text{on} \quad \Gamma_{in,t}. \tag{7.11}$$

The following weak ALE formulation of the tracer equation (5.20) is considered :

$$\begin{gathered}\text{find} \quad T \in \mathcal{Y}_0(\Omega_t, t) \quad \text{such that} \quad \forall \psi \in \mathcal{Y}(\Omega_t), \quad \forall t \in I, \\ \int_{\Omega_t} \Big(\frac{\partial T}{\partial t}\Big|_{\hat{\boldsymbol{x}}} \;+\; (\boldsymbol{U} - \boldsymbol{c}) \cdot \boldsymbol{\nabla} T\Big)\; \psi \; d\boldsymbol{x} \;+\; \int_{\Omega_t} \nu_T \boldsymbol{\nabla} T \cdot \boldsymbol{\nabla} \psi \; d\boldsymbol{x} \\ = \; \int_{\Omega_t} f_T \; \psi \; d\boldsymbol{x} \;+\; \int_{\Gamma_{s,t}} (aT + b) \; \psi \; d\Gamma \;+\; \int_{\Gamma_{out,t}} T \; \psi \; d\Gamma \,.\end{gathered} \tag{7.12}$$

Note that this weak form has been obtained imposing weakly boundary condition (5.24).
The equation is then discretised in time as follows :

$$\begin{aligned}\int_{\Omega^{n+1}} (T^{n+1} - T^n_{n+1}) \; \psi^{n+1} \; d\boldsymbol{x} \;+\; \Delta t \int_{\Omega^{q_n}} \big(\bar{\boldsymbol{U}}^{q_n} \cdot \boldsymbol{\nabla} T^n_{q_n}\big) \; \psi^{q_n} \; d\boldsymbol{x} \\ + \int_{\Omega^{n+1}} \nu_T \boldsymbol{\nabla} T^{n+1} \cdot \boldsymbol{\nabla} \psi^{n+1} \; d\boldsymbol{x} \\ = \; \int_{\Omega^{n+1}} f^{n+1}_T \; \psi^{n+1} \; d\boldsymbol{x} \\ + \int_{\Gamma^{n+1}_s} (aT^n_{n+1} + b) \; \psi^{n+1} \; d\Gamma \;+\; \int_{\Gamma^{n+1}_{out}} T^n_{n+1} \; \psi^{n+1} \; d\Gamma \,,\end{aligned} \tag{7.13}$$

where, for each $n = 0, \ldots, N_t - 1$, t^{q_n} is a particular time in $[t^n, t^{n+1}]$.

A splitting finally leads to the following *semi-discrete non-conservative weak ALE form of the linear advection equation on the tracer T.*

$$\begin{gathered}\text{Given the initial condition } T_0, \text{ for } n = 0, \ldots, N_t - 1, \\ \text{find} \quad \tilde{T} \in \mathcal{Y}_0^{n+1} \quad \text{such that} \quad \forall \hat{\psi} \in \hat{\mathcal{Y}}, \\ \int_{\Omega^{n+1}} (\tilde{T} - T^n_{n+1}) \; \psi^{n+1} \; d\boldsymbol{x} \;+\; \Delta t \int_{\Omega^{q_n}} \big(\bar{\boldsymbol{U}}^{q_n} \cdot \boldsymbol{\nabla} T^n_{q_n}\big) \; \psi^{q_n} \; d\boldsymbol{x} \;=\; 0 \,.\end{gathered} \tag{7.14}$$

Note that in (7.14), the tracer is advected in the current domain between times t^n and t^{n+1}, *i.e.*, the movement of the domain from Ω^n to Ω^{n+1} is taken into account. On the contrary, in (7.9) and (7.10), the horizontal velocity is advected in the domain between times t^{n-1} and t^n, *i.e.* : there is a time delay of one time step, so that it is the movement of the domain from Ω^{n-1} to Ω^n which is taken into account during this advection.

7.5.3 Space discretization

Let us now consider a space discretization of the semi-discrete hydrostatic flow problem (5.73)-(5.74)-(5.75) through the Galerkin method. We denote by $\mathcal{T}_h^{2d}$ the triangulation of the two-dimensional domain ω, and by $\hat{\mathcal{T}}_h$ the triangulation of the three-dimensional reference configuration $\hat{\Omega}$. The mapping and the different functional spaces are assumed to be discretized following the description in Chapter 2, sections 2.1.6 and 2.1.7.

We indicate that the continuous spaces $\hat{\mathcal{X}} = H^1(\hat{\Omega})^2$, $\hat{\mathcal{V}} = L^2(\hat{\Omega})$ and $\mathcal{M} = L^2(\omega)$ are approximated by some finite element spaces denoted by $\hat{\mathcal{X}}_h$, $\hat{\mathcal{V}}_h$ and $\mathcal{M}_h$. At each time, the horizontal velocity $\boldsymbol{u}(t^n)$ and the vertical ALE velocity $\bar{w}(t^n)$ are approximated respectively in $\mathcal{X}_h^n$ and $\mathcal{V}_h^n$ – defined as in (3.3). The free surface $\eta(t^n)$ is approximated in $\mathcal{M}_h$.

We choose $\hat{\mathcal{Y}} = H^1(\hat{\Omega})$ to be approximated by the Lagrangian finite element space $\hat{\mathcal{Y}}_h = (\mathcal{F}_{n',k}(\hat{\mathcal{T}}_h))^2$ for some particular degree $n' \geq k$. At each time , the corresponding space $\mathcal{Y}^n$ on the current domain is then approximated by $\mathcal{Y}_h^n = \mathcal{F}_{n',k}(\mathcal{T}_h^n)$.

Moreover, in order to apply the ALE-MURD scheme, we choose the test functions for the advection equations (7.9), (7.10) and (7.14) to be approximated by constant functions on the whole three-dimensional mesh $\hat{\mathcal{T}}_h$ – the equation will then be solved in $L^1(\Omega_h)$.

This way, the *fully-discretized formulation of the linear advection equations* to be solved corresponds to (3.4) ; that is for each $n = 0, \ldots, N_t - 1$,

$$\left\{ \begin{array}{l} \text{find} \quad (\tilde{u}_h, \tilde{v}_h) \in (\mathcal{Y}_{h,0}^n)^2 \quad \text{such that} \\ \\ \displaystyle\int_{\Omega^n} \tilde{u}_h - u_h^n \, d\boldsymbol{x} + \Delta t \int_{\Omega_h^{a_n}} \bar{\boldsymbol{U}}_h^{a_n} \cdot \boldsymbol{\nabla} u_{h,\,a_n}^n \, d\boldsymbol{x} = 0 \,, \\ \\ \displaystyle\int_{\Omega^n} \tilde{v}_h - v_h^n \, d\boldsymbol{x} + \Delta t \int_{\Omega_h^{a_n}} \bar{\boldsymbol{U}}_h^{a_n} \cdot \boldsymbol{\nabla} v_{h,\,a_n}^n \, d\boldsymbol{x} = 0 \,, \end{array} \right. \tag{7.15}$$

and

$$\left\{ \begin{array}{c} \text{find} \quad \tilde{T}_h \ \in \ \mathcal{Y}_{h,0}^{n+1} \quad \text{such that} \\ \displaystyle\int_{\Omega^{n+1}} \tilde{T}_h - T_{h,n+1}^n \, d\boldsymbol{x} \ + \ \Delta t \int_{\Omega_h^{q_n}} \bar{\boldsymbol{U}}_h^{q_n} \cdot \boldsymbol{\nabla} T_{h,\,q_n}^n \, d\boldsymbol{x} \ = \ 0 \,. \end{array} \right. \tag{7.16}$$

The ALE-MURD scheme detailed in section 3.3, can now be constructed. As explained in this description, if the distribution chosen for the scheme is locally positive, it is monotonic. Moreover, it satisfies the required conservation property provided a particular constraint is satisfied.

7.5.4 Conservation constraint

Let us first express the discrete conservation property to be satified by the advection schemes just presented. Following Definition 3.6.1 and according to the time and space discretization (7.15) and (7.16) of each problem, the *discrete conservation property* required at each time step for the horizontal velocity is expressed by :

$$\int_{\Omega_h^n} \tilde{\boldsymbol{u}}_h \, d\boldsymbol{x} - \int_{\Omega_h^{n-1}} \boldsymbol{u}_h^n \, d\boldsymbol{x} \ = \ - \, \Delta t \int_{\Gamma_{h,liq}^{a_n}} \boldsymbol{u}_{h,\,a_n}^n \, (\boldsymbol{u}_h^{a_n} \cdot \boldsymbol{n}_{h,xy}) \, d\Gamma \,, \tag{7.17}$$

and for the tracer :

$$\int_{\Omega_h^{n+1}} \tilde{T}_h \, d\boldsymbol{x} - \int_{\Omega_h^n} T_h^n \, d\boldsymbol{x} \ = \ - \, \Delta t \int_{\Gamma_{h,liq}^{q_n}} T_{h,\,q_n}^n \, (\boldsymbol{u}_h^{q_n} \cdot \boldsymbol{n}_{h,xy}) \, d\Gamma \,. \tag{7.18}$$

As already anounced in section 3.6, in the particular case of the hydrostatic flow problem we apply Proposition 3.8 to determine the condition making the ALE-MURD scheme conservative.

In the following, $N_{\mathcal{Y}}$ will denote the dimension of $\hat{\mathcal{Y}}_h$ and $\{\hat{\psi}_i\}_{i=1,\ldots,N_h}$ a set of basis functions for $\hat{\mathcal{Y}}_h$.We recall that, at each time t^n, a corresponding set of basis functions $\{\psi_i^n\}_{i=1,\ldots,N_h}$ of $\mathcal{Y}_h^n$ can be defined by $\psi_i^n = \hat{\psi}_i \circ \mathcal{A}_{h,n}^{-1}$ for each node i.

For the advection of the velocities u and v, the conservation constraint is :

$$\left\{ \begin{array}{c} \forall\, n = 0, \ldots, N_t - 1, \quad \forall\, i = 1, \ldots, N_h, \\ \displaystyle\int_{\Omega_h^{n+1}} \psi_i^{n+1}\, d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n\, d\boldsymbol{x} \;-\; \Delta t \int_{\Omega_h^{a_n}} \bar{\boldsymbol{U}}_h^{a_n} \cdot \boldsymbol{\nabla} \psi_i^{a_n}\, d\boldsymbol{x} \\ \displaystyle +\quad \Delta t \int_{\Gamma_{h,liq}^{a_n}} \psi_i^{a_n}\, \boldsymbol{u}_h^{a_n} \cdot \boldsymbol{n}_{h,xy}\, d\Gamma \;=\; 0. \end{array} \right. \tag{7.19}$$

For the tracer advection, the expression of the conservation constraint is the following :

$$\left\{ \begin{array}{c} \forall\, n = 0, \ldots, N_t - 1, \quad \forall\, i = 1, \ldots, N_h, \\ \displaystyle\int_{\Omega_h^{n+1}} \psi_i^{n+1}\, d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n\, d\boldsymbol{x} \;-\; \Delta t \int_{\Omega_h^{q_n}} \bar{\boldsymbol{U}}_h^{q_n} \cdot \boldsymbol{\nabla} \psi_i^{q_n}\, d\boldsymbol{x} \\ \displaystyle +\quad \Delta t \int_{\Gamma_{h,liq}^{q_n}} \psi_i^{q_n}\, \boldsymbol{u}_h^{q_n} \cdot \boldsymbol{n}_{h,xy}\, d\Gamma \;=\; 0\,. \end{array} \right. \tag{7.20}$$

We would like to highlight that the results we present in the following sections are based on the work by Janin [74].

7.5.5 Conservativity and continuity

Let us now consider the following space-discretization of the semi-discrete continuity equation (5.75), providing an approximation of the vertical ALE advection velocity $\bar{w} = w - c$. For $n = 0, \ldots, N_t$:

$$\text{find} \quad \bar{w}_h^{c_n} \in \mathcal{V}_{h,0}^{c_n} \quad \text{such that}, \quad \forall\, \hat{\phi}_h \in \hat{\mathcal{W}}_{h,0}\,,$$

$$\int_{\Omega_h^{c_n}} \bar{w}_h^{c_n}\, \frac{\partial \phi_h^{c_n}}{\partial z}\, d\boldsymbol{x} \;+\; \int_{\Omega_h^{c_n}} \boldsymbol{u}_h^{c_n} \cdot \boldsymbol{\nabla}_2 \phi_h^{c_n}\, d\boldsymbol{x} \;-\; \int_{\Gamma_{h,liq}^{c_n}} \phi_h^{c_n}\, \boldsymbol{u}_h^{c_n} \cdot \boldsymbol{n}_{h,xy}\, d\Gamma$$
$$-\quad \frac{1}{\Delta t}\; \Big(\int_{\Omega_h^{n+1}} \phi_h^{n+1}\, d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \phi_h^n\, d\boldsymbol{x}\, \Big) \;=\; 0\,, \tag{7.21}$$

where t^{c_n} is a particular time between t^n and t^{n+1}, and where $\hat{\mathcal{W}}_h$ is the discrete space approximating $\hat{\mathcal{W}} = H^1(\hat{\Omega})$. In addition,

$$\mathcal{V}_{h,0}^n \;=\; \{\; \phi_h \in \mathcal{V}_h^n \quad | \quad \phi_h\; n_{h,z} \;=\; \boldsymbol{u}_h^{n+1} \cdot \boldsymbol{n}_{h,xy} \quad \text{on} \quad \Gamma_{h,b}\; \}$$

and

$$\mathcal{W}_{h,0}^n \;=\; \{\; \phi_h \in \mathcal{W}_h^n \quad | \quad \phi_h \; n_{h,z} \;=\; \boldsymbol{u}_h^{n+1} \cdot \boldsymbol{n}_{h,xy} \quad \text{on} \quad \Gamma_{h,b} \;\}.$$

Assume the continous space $\hat{\mathcal{W}}$ is approximated by the same Lagrangian finite element space $\mathcal{F}_{n',k}(\hat{\mathcal{T}}_h)$ as $\hat{\mathcal{Y}}$, so that at each time t^n :

$$\mathcal{W}_h^n \;=\; \mathcal{F}_{n',k}(\mathcal{T}_h^n) \;=\; \mathcal{Y}_h^n \,. \tag{7.22}$$

Note that the dimension of $\hat{\mathcal{W}}_{h,0}$ is then equal to N_h minus the number of nodes situated on the bottom $\Gamma_{h,b}$ of the domain. Indeed, a set of basis functions for $\hat{\mathcal{W}}_{h,0}$ is

$$\{\; \hat{\psi}_i \;\}_{\; i \,\in\, \hat{\mathcal{T}}_h,\; i \,\notin\, \Gamma_{h,b}} \,,$$

where, with an obvious abuse of notation, $i \in \hat{\mathcal{T}}_h$ means that the node i belongs to $\hat{\mathcal{T}}_h$ and $i \notin \Gamma_{h,b}$ means that it does not belong to the bottom boundary.

The discrete vertical ALE advection velocity $\bar{w}_h$ is then sought such that $\bar{w}_h^{c_n} \in \mathcal{V}_{h,0}^{c_n}$ at each time t^{c_n} and

$$\left\{ \begin{array}{c} \forall\, n = 0, \ldots, N_t - 1, \quad \forall\, i \,\in\, [1, N_h] \text{ such that } i \,\notin\, \Gamma_{h,b} \,, \\ \displaystyle\int_{\Omega_h^{c_n}} \bar{w}_h^{c_n} \, \frac{\partial \psi_i^{c_n}}{\partial z} \; d\boldsymbol{x} \;+\; \int_{\Omega_h^{c_n}} \boldsymbol{u}_h^{c_n} \cdot \boldsymbol{\nabla}_2 \psi_i^{c_n} \; d\boldsymbol{x} \;-\; \int_{\Gamma_{h,liq}^{c_n}} \psi_i^{c_n} \; \boldsymbol{u}_h^{c_n} \cdot \boldsymbol{n}_h \; d\Gamma \\ \displaystyle -\quad \frac{1}{\Delta t} \;\Big(\int_{\Omega_h^{n+1}} \psi_i^{n+1} \; d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n \; d\boldsymbol{x} \Big) \;=\; 0 \,. \end{array} \right. \tag{7.23}$$

We observe that, taking $t^{q_n} = t^{c_n}$ for each $n = 0, \ldots, N_t - 1$, the relation associated to the node i in (7.20) is identical to component of the continuity equation (7.23) associated to the same node, except for those nodes i situated on the bottom $\Gamma_{h,b}$ of the domain. In that case, the conservation constraint for the tracer advection reduces to :

$$\left\{ \begin{array}{c} \forall\, n = 0, \ldots, N_t - 1 \,, \quad \forall\, i \in \Gamma_{h,b} \,, \\ \displaystyle\int_{\Omega_h^{n+1}} \psi_i^{n+1} \; d\boldsymbol{x} \;-\; \int_{\Omega_h^n} \psi_i^n \; d\boldsymbol{x} \;-\; \Delta t \int_{\Omega_h^{c_n}} \bar{\boldsymbol{U}}_h^{c_n} \cdot \boldsymbol{\nabla} \psi_i^{c_n} \; d\boldsymbol{x} \\ \displaystyle +\quad \Delta t \int_{\Gamma_{h,liq}^{c_n}} \psi_i^{c_n} \; \bar{\boldsymbol{U}}_h^{c_n} \cdot \boldsymbol{n}_h \; d\Gamma \;=\; 0 \,, \end{array} \right. \tag{7.24}$$

where t^{c_n} is the particular time between t^n and t^{n+1} used in the continuity equation (7.21) for the ALE velocity $\bar{\boldsymbol{U}}_h$. We can state the following intermediate result.

Lemma 7.1 *If the discrete ALE velocity of the fluid $\bar{\boldsymbol{U}}_h$ satisfies the discrete continuity equation (7.21) and if, for $n = 0, \ldots, N_t - 1$, $t^{q_n} = t^{c_n}$, then the advective scheme (7.16) for the tracer T is conservative in the sense of (7.18) if relation (7.24) holds.*

Let us now assume that $t^{a_n} = t^{c_{n-1}}$ for each $n = 0, \ldots, N_t - 1$. The component of the continuity equation (7.23) associated to each node $i \notin \Gamma_{h,b}$ is identical to the relation associated to the same node i in (7.19) In this case, the conservation constraint for the horizontal velocity also reduces to (7.24).

We can thus state a second intermediate result.

Lemma 7.2 *If the discrete ALE velocity of the fluid $\bar{\boldsymbol{U}}_h$ satisfies the discrete continuity equation (7.23) and if, for $n = 1, \ldots, N_t$, $t^{a_n} = t^{c_{n-1}}$, then the advective scheme (7.15) for the horizontal velocity $\boldsymbol{u}$ is conservative in the sense of (7.17) if relation (7.24) holds.*

This leads us to the following proposition.

Proposition 7.1 *In the framework of the semi-discrete hydrodynamic problem (5.73)-(5.74)-(5.75), the advection scheme (7.15) for the horizontal velocity $\boldsymbol{u}$ is conservative in the sense of (7.17) and the advection scheme (7.16) for the tracer T is conservative in the sense of (7.18) if*

i) *the tracer T and the two components of the horizontal velocity $\boldsymbol{u}$ are approximated by the same Lagrangian finite element space as the test functions of the continuity equation (5.75),*

ii) *the time orders for the ALE advection velocity $\bar{\boldsymbol{U}}_h$ satisfy $t^{a_n} = t^{c_{n-1}}$ and $t^{q_n} = t^{c_n}$ for each $n = 0, \ldots, N_t - 1$,*

iii) *relation (7.24) holds.*

In the next section, we give the example of a finite element space approximating $\hat{\mathcal{Y}}$ for which the conservation constraint (7.24) can easily be satisfied.

7.5.6 Conservation and free surface equation

Consider a triangulation $\mathcal{T}_h^{2d}$ of the two-dimensional domain ω composed of triangles. Assume that $\hat{\mathcal{T}}_h$ is a prismatic mesh obtained by piling up $\mathcal{T}_h^{2d}$ along the vertical in $\hat{\Omega}$, as illustrated by Figure 6.2 in section 6.3.3. We will denote by N_l the number of horizontal levels in $\hat{\mathcal{T}}_h$ and by $\hat{z}_k$, for $k = 1, \ldots, (N_l - 1)$, the height of level k. The layer delimited by levels k and $k+1$ will be denoted by $l_{k+1/2}$ and by $\Delta\hat{z}_{k+1/2}$ its width.

We choose to approximate the continuous space $\hat{\mathcal{Y}} = H^1(\hat{\Omega})$ by the finite element space $\mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$ defined in (6.47) :

$$\hat{\mathcal{Y}}_h \;=\; \mathcal{P}_{1,1}(\hat{\mathcal{T}}_h). \tag{7.25}$$

In the following we will drop the subscript $_h$ on the discrete variables and functions for the sake of clarity.

Remark 7.5.1 Any function $\hat{\psi}$ in $\mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$ can be written as the product of a function ψ^h defined on the two-dimensional discrete domain ω_h and a function $\hat{\psi}^v$ defined on the vertical of the three-dimensional domain $\hat{\Omega}$:

$$\hat{\psi}(x, y, \hat{z}) \;=\; \psi^h(x, y)\; \hat{\psi}^v(\hat{z}), \quad \forall\, \hat{\boldsymbol{x}} = (x, y, \hat{z}) \in \hat{\Omega}.$$

Moreover, each basis function $\hat{\psi}_i$ of $\mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$ associated to node i can be decomposed as follows on each prism $\hat{P}$ belonging to $\hat{\mathcal{T}}_h$ and containing i :

$$\hat{\psi}_i\,(x, y, z) \;=\; \psi^h_{i_h}\,(x, y)\; \hat{\psi}^v_{i_v}(z) \quad \text{in} \quad \hat{P}, \tag{7.26}$$

where i_h denotes the associated node in $\mathcal{T}_h^{2d}$ and i_v the level on which node i is situated, as illustrated by Figure 7.29.

Function $\hat{\psi}^v_{i_v}$ is chosen such that its support is composed of layers $l_{iv+1/2}$ and $l_{iv-1/2}$ – when they exist. Its derivative with respect to $\hat{z}$ is then constant and defined by :

$$\frac{\partial \hat{\psi}^v_{i_v}}{\partial \hat{z}} \;=\; \begin{cases} \dfrac{-1}{\Delta\hat{z}_{iv+1/2}} & \text{in} \quad [\hat{z}_k, \hat{z}_{k+1}] \\[2ex] \dfrac{1}{\Delta\hat{z}_{iv-1/2}} & \text{in} \quad [\hat{z}_{k-1}, \hat{z}_k] \\[2ex] 0 & \text{in} \quad [0, \hat{z}_{k-1}] \cup [\hat{z}_{k+1}, 1]. \end{cases} \tag{7.27}$$

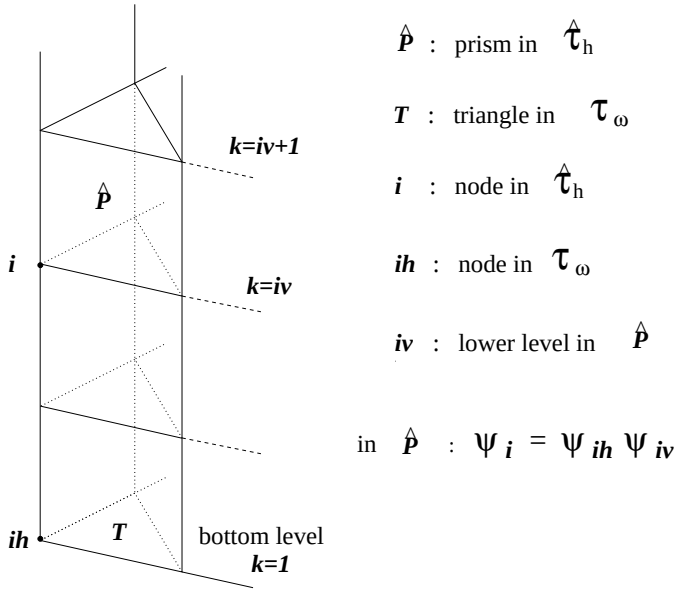


FIG. 7.29 – A prism column in the reference mesh $\hat{\mathcal{T}}_h$

We recall that the ALE mapping used in this problem is of the form (5.25). This is precisely the general form of the mappings used in the ALE-Sigma (ALES) approach described in Chapter 4, section 4.3, for which the formulae (4.17)–(4.21) have been derived. We can therefore apply them here to any function $\psi : \Omega_h^n \longrightarrow \mathbb{R}$, at any time t^n.

In particular, we can reformulate as follows the conservation constraint for the advection of $\boldsymbol{u}$ and T – *i.e.* system (7.19) for $a_n = c_n$ or equivalently system (7.20) for $q_n = c_n$:

$$\left\{ \begin{array}{l} \forall\, n = 0, \ldots, N_t - 1, \quad \forall\, i \in [1, N_h] \text{ such that } i \notin \Gamma_{h,b}\,, \\ \displaystyle\int_{\hat{\Omega}} (\hat{J}^{n+1} - \hat{J}^n)\ \hat{\psi}_i\ d\hat{\boldsymbol{x}} \ - \ \Delta t \int_{\hat{\Omega}} \hat{J}^{c_n}\ \hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{\nabla}_2 \hat{\psi}_i\ d\hat{\boldsymbol{x}} \\ \displaystyle - \quad \Delta t \int_{\hat{\Omega}} (\,\hat{J}\, w^*\,)^{c_n}\ \frac{\partial \hat{\psi}_i}{\partial \hat{z}}\ d\hat{\boldsymbol{x}} \ + \ \Delta t \int_{\hat{\Gamma}_{liq}} \hat{J}^{c_n}\ \hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{n}\ \hat{\psi}_i\ d\Gamma \ = \ 0\,, \end{array} \right. \tag{7.28}$$

where $\hat{J}\, w^*$ is a "pseudo-vertical velocity" defined in $\hat{\Omega}$ by

$$\hat{J}\, w^* \ = \ \hat{w} \ - \ \hat{c} \ - \ \hat{u}\, \frac{\partial z}{\partial x}\Big|_{\hat{\boldsymbol{x}}} \ - \ \hat{v}\, \frac{\partial z}{\partial y}\Big|_{\hat{\boldsymbol{x}}}\,. \tag{7.29}$$

We recall that for any function $\psi : \Omega^n \longrightarrow \mathbb{R}$, $\hat{\psi} = \psi \circ \hat{\mathcal{A}}_n$ is the corresponding function on the ALE frame.

Remark 7.5.2 Note that the continuous "pseudo-vertical velocity" $\hat{J}_t\ w^*$ was already introduced in section 4.4 to simplify the *ALE-Sigma* (ALES) formulation of the three-dimensional free surface Navier-Stokes equations.

System (7.28) can now be decomposed using (7.26), leading to :

$$\left\{ \begin{array}{l} \forall\, n = 0, \ldots, N_t - 1, \quad \forall\, ih = 1, \ldots, N_\omega \quad \text{and} \quad \forall\, iv = 1, \ldots, N_l, \\ \displaystyle\int_{\hat{\Omega}} (\,\hat{J}^{n+1} - \hat{J}^n\,)\ \psi^h_{i_h}\ \hat{\psi}^v_{i_v}\ d\hat{\boldsymbol{x}} \ - \ \Delta t \int_{\hat{\Omega}} \hat{J}^{c_n}\ (\,\hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{\nabla}_2\, \psi^h_{i_h}\,)\ \hat{\psi}^v_{i_v}\ d\hat{\boldsymbol{x}} \\ \qquad - \quad \Delta t \displaystyle\int_{\hat{\Omega}} (\,\hat{J}\, w^*_h\,)^{c_n}\ \psi^h_{i_h}\ \frac{\partial \hat{\psi}^v_{i_v}}{\partial \hat{z}}\ d\hat{\boldsymbol{x}} \\ \qquad\qquad + \quad \Delta t \displaystyle\int_{\hat{\Gamma}_{liq}} \hat{J}^{c_n}\ \psi^h_{i_h}\ \hat{\psi}^v_{i_v}\ \hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{n}\ d\Gamma \quad = \quad 0. \end{array} \right. \tag{7.30}$$

Let us now consider the following equation :

$$\int_{\omega} (\eta^{n+1} - \eta^n)\ \psi^h_{i_h}\ d\omega \ - \ \Delta t \int_{\omega_h} \psi^h_{i_h}\ \mathrm{div}_2\ (\int_b^{\eta^{c_n}} \boldsymbol{u}^{c_n}\ dz)\ d\omega \ = \ 0\,, \tag{7.31}$$

where η^n denotes the approximation in $\mathcal{M}_h$ of the free surface $\eta(t^n)$ at time t^n, for $n \in [0, N_t - 1]$.

We can state the following result.

Lemma 7.3 *By summing up the components of system (7.30) associated to the nodes situated on the same vertical axis in $\hat{\mathcal{T}}_h$* – i.e.*, the components corresponding to the same node ih in $\mathcal{T}_h^{2d}$ – we obtain equation (7.31).*

Proof. The sum over the levels iv of functions $\hat{\psi}^v_{i_v}(\hat{z})$ is equal to 1 at any $\hat{z}$ in $[0, 1]$. In addition, the sum of $\partial_{\hat{z}}\, \hat{\psi}^v_{i_v}$ is zero. Therefore, summing up the components of system (7.30) over $iv = 1, \ldots, N_l$ we obtain :

$$\begin{array}{l} \displaystyle\int_{\hat{\Omega}} (\,\hat{J}^{n+1} - \hat{J}^n\,)\ \psi^h_{i_h}\ d\hat{\boldsymbol{x}} \quad - \quad \Delta t \displaystyle\int_{\hat{\Omega}} \hat{J}^{c_n}\ \hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{\nabla}_2\, \psi^h_{i_h}\ d\hat{\boldsymbol{x}} \\ \qquad\qquad + \quad \Delta t \displaystyle\int_{\hat{\Gamma}_{liq}} \hat{J}^{c_n} \psi^h_{i_h}\ \hat{\boldsymbol{u}}^{c_n} \cdot \boldsymbol{n}\ d\Gamma \quad = \quad 0\,. \end{array}$$

Since this relation only contains partial derivatives with respect to the horizontal coordinates x and y, it can be reformulated in the current domain as follows :

$$\int_{\Omega^{n+1}} \psi_{i_h}^h \, d\boldsymbol{x} \; - \; \int_{\Omega^n} \psi_{i_h}^h \, d\boldsymbol{x} \quad - \quad \Delta t \int_{\Omega^{c_n}} \boldsymbol{u}^{c_n} \cdot \boldsymbol{\nabla}_2 \, \psi_{i_h}^h \, d\boldsymbol{x}$$
$$+ \quad \Delta t \int_{\Gamma_{liq}^{c_n}} \psi_{i_h}^h \, \boldsymbol{u}^{c_n} \cdot \boldsymbol{n} \, d\Gamma \quad = \quad 0 \, ,$$

$$\Leftrightarrow \qquad \int_{\omega} (h^{n+1} - h^n) \, \psi_{i_h}^h \, d\omega \quad - \quad \Delta t \int_{\omega} \boldsymbol{\nabla}_2 \, \psi_{i_h}^h \cdot \left(\int_b^{\eta^{c_n}} \boldsymbol{u}^{c_n} \, dz \right) \, d\omega$$
$$+ \quad \Delta t \int_{\partial\omega} \psi_{i_h}^h \, \left(\int_b^{\eta^{c_n}} \boldsymbol{u}^{c_n} \, dz \right) \cdot \boldsymbol{n}_\omega \, d\gamma \quad = \quad 0 \, .$$

Assuming that the Green Formula can be applied to the horizontal flux term, we obtain equation (7.31). Note that the Green Formula can be applied provided the local flux terms $\Phi_{K_0}^{n+1}$ defined by (3.7) are computed exactly in the ALE-MURD advection scheme.
$\diamondsuit$

We notice that equation (7.31) is a particular space-discretization of the semi-discrete free surface equation in (5.74). Indeed, taking $t^{\theta_n} = t^{c_n}$ for each $n = 0, \ldots, N_t - 1$, and choosing $\mathcal{M}_h = \mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$, the semi-discrete weak ALE formulation of the free surface equation in (5.74) can be approximated by equation (7.31).

This result shows that, if the relations composing system (7.30) are satisfied for $ih = 1, \ldots, N_\omega$ and $iv = 2, \ldots, N_l$, and if the discrete form of the free surface equation solved is (7.31), then the relations composing system (7.30) are also satisfied for $iv = 1$, which denotes the bottom level. In this case, the components of the conservation constraint associated to the nodes situated on the bottom $\Gamma_{h,b}$ of the domain do not have to be ensured.

The following proposition can finally be stated.

Proposition 7.2 *In the framework of the semi-discrete hydrodynamic problem (5.73)-(5.74)-(5.75), if*

i) *the two-dimensional domain ω is discretized using a triangular mesh $\mathcal{T}_h^{2d}$, and the three-dimensional reference configuration $\hat{\Omega}$ is triangularized by a prismatic mesh $\hat{\mathcal{T}}_h$ obtained piling up $\mathcal{T}_h^{2d}$ along the vertical,*

ii) *the continuous space $\mathcal{M}$ defining the free surface function η and the test functions for the free surface equation in (5.74) is approximated by the finite element space $\mathcal{M}_h = \mathcal{F}_{1,1}(\mathcal{T}_h^{2d})$,*

iii) *the continuous space $\hat{\mathcal{W}}$ defining the test functions for the continuity equation (5.75) is approximated by the finite element space $\hat{\mathcal{W}}_h = \mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$,*

iv) *the continuous space $\hat{\mathcal{Y}}$ defining the tracer T and the components of the horizontal velocity u and v is approximated by the finite element space $\hat{\mathcal{Y}}_h = \hat{\mathcal{W}}_h = \mathcal{P}_{1,1}(\hat{\mathcal{T}}_h)$,*

v) *for each $n = 0, \ldots, N_t$, $t^{a_n} = t^{\theta_{n-1}}$ and $t^{q_n} = t^{c_n} = t^{\theta_n}$,*

then the advection scheme (7.15) for the horizontal velocity $\boldsymbol{u}$ is conservative in the sense of (7.17) and the advection scheme (7.16) for the tracer T is conservative in the sense of (7.18).

7.5.7 A particular time and space discretization of the hydrostatic flow problem leading to a conservative advective ALE-MURD scheme

Following the successive results obtained in this section, we are now able to propose a particular time and space discretization of the weak hydrostatic flow problem (5.57) and the tracer equation (7.12), allowing the use of a conservative ALE-MURD scheme to solve the advection equations.

For this purpose we first set

$$\hat{\mathcal{Y}}_h = \mathcal{P}_{1,1}(\hat{\mathcal{T}}_h), \quad \hat{\mathcal{W}}_h = \mathcal{P}_{1,1}(\hat{\mathcal{T}}_h) \quad \text{and} \quad \mathcal{M}_h = \mathcal{F}_{1,1}(\mathcal{T}_h^{2d}).$$

Here again we drop the subscript $_h$ on the discrete variables for the sake of clarity.

The fully-discretized hydrostatic flow problem mentioned is the following.

Given the initial conditions $\boldsymbol{u}_0$ and η_0, for each $n = 0, \ldots, N_t - 1$,

find $\tilde{u} \in \mathcal{Y}^n_{h,0}$ such that

$$\int_{\Omega^n_h} (\tilde{u} - u^n)\, d\boldsymbol{x} + \int_{\Omega^{\theta_{n-1}}_h} \bar{\boldsymbol{U}}^{\theta_{n-1}} \cdot \boldsymbol{\nabla} u^n_{\theta_{n-1}}\, d\boldsymbol{x} = 0, \tag{7.32}$$

find $\boldsymbol{u}^{n+1} \in \mathcal{X}^{n+1}_{h,0}$ and $\eta^{n+1} \in \mathcal{M}_h$ such that,

$$\forall\, \boldsymbol{\psi} \in \mathcal{X}^n_{h,0} \quad \text{and} \quad \forall\, \chi \in \mathcal{M}_h,$$

$$\frac{1}{\Delta t} (\boldsymbol{u}^{n+1}_n - \boldsymbol{u}^n, \boldsymbol{\psi})_{\Omega^n_h} + \nu (\boldsymbol{\nabla} \boldsymbol{u}^{n+1}_n, \boldsymbol{\nabla} \boldsymbol{\psi})_{\Omega^n_h}$$

$$- g \int_{\omega_h} \eta^{\theta_n} \operatorname{div}_2 \left(\int_b^{\eta^n} \boldsymbol{\psi}\, dz \right) d\omega$$

$$= < \boldsymbol{F}^{n+1}_h, \boldsymbol{\psi} >_{(\mathcal{X}^n_h)' \times \mathcal{X}^n_h} \tag{7.33}$$

$$\frac{g}{\Delta t} (\eta^{n+1} - \eta^n, \chi)_\omega - g\, b_n(\chi, \boldsymbol{u}^{\theta_n}) = 0, \tag{7.34}$$

find $\bar{w}^{\theta_n} \in \mathcal{V}^{\theta_n}_{h,0}$ such that $\forall\, \hat{\psi} \in \hat{\mathcal{W}}_{h,0}$

$$\int_{\Omega^{\theta_n}} \bar{w}^{\theta_n} \frac{\partial \psi^{t^{\theta_n}}}{\partial z}\, d\boldsymbol{x} + \int_{\Omega^{\theta_n}} \boldsymbol{u}^{\theta_n} \cdot \boldsymbol{\nabla}_2 \psi^{t^{\theta_n}}\, d\boldsymbol{x}$$

$$- \int_{\Gamma^{\theta_n}_{liq}} \psi^{t^{\theta_n}} \boldsymbol{u}^{\theta_n} \cdot \boldsymbol{n}\, d\Gamma \tag{7.35}$$

$$+ \frac{1}{\Delta t} \left(\int_{\Omega^{n+1}} \psi^{t^{n+1}}\, d\boldsymbol{x} - \int_{\Omega^n} \psi^{t^n}\, d\boldsymbol{x} \right) = 0,$$

find $\tilde{T} \in \mathcal{Y}^{n+1}_{h,0}$ such that

$$\int_{\Omega^{n+1}} (\tilde{T} - T^n_{n+1})\, d\boldsymbol{x} + \Delta t \int_{\Omega^{q_n}} \bar{\boldsymbol{U}}^{\theta_n} \cdot \boldsymbol{\nabla} T^n_{q_n}\, d\boldsymbol{x} = 0, \tag{7.36}$$

where $t^{\theta_n} = \theta\, t^{n+1} + (1-\theta)\, t^n$ is a particular time in $[t^n, t^{n+1}]$, determined

by the implicitness parameter $0 \leq \theta \leq 1$. Moreover,

$$\eta^{\theta_n} = \theta\, \eta^{n+1} + (1-\theta)\, \eta^n\,, \qquad \boldsymbol{u}^{\theta_n} = \theta\, \boldsymbol{u}^{n+1} + (1-\theta)\, \boldsymbol{u}^n\,,$$
$$\bar{w}^{\theta_n} = \theta\, \bar{w}^{n+1} + (1-\theta)\, \bar{w}^n\,, \qquad \bar{\boldsymbol{U}}^{\theta_n} = (\boldsymbol{u}^{\theta_n}, \bar{w}^{\theta_n})\,.$$

Proposition (7.2) can be applied to this system, proving the conservation property of the advective schemes (7.32) and (7.36).

7.5.8 Conservation of the ALE-MURD schemes in the Telemac-3D system

The time and space discretization proposed previously is exactly the one used in the Telemac-3D system to solve the hydrostatic part of the equations – as well for the hydrostatic as for the non-hydrostatic problem. This scheme was chosen for the same reasons we have given in our work, that is in order to ensure the conservation property of the two existing MURD schemes solving the linear advection equations of the hydrodynamic problem.

However, the demonstration – see the work by Janin [74] – was only made for the particular use of the classical sigma transformation – which is equivalent to using a particular ALE mapping, as explained in Chapter 4. Therefore it was not clear whether or not the conservation property would hold for any mapping.

In addition, the implementation of a generalized sigma coordinate system into the Telemac-3D system revealed a serious loss of the conservation property of the MURD schemes when a transformation other than the classical sigma transformation was used.

For these reasons we have generalized the demonstration in the ALE formulation, taking profit of the description of the ALE-MURD schemes presented in Chapter 3. This has lead to the results presented in sections 7.5.5 to 7.5.7.

These results have allowed us to find the reasons for the loss of advected quantity that occured systematically when using a different mapping than the classical sigma transformation.

On one hand, an approximation was made in the computation of the terms in the free surface equation (7.34) and in the continuity equation (7.35), so that the property stated by Lemma 7.3 did not hold true.

On the other hand, in some cases the mass term in the free surface equation is mass-lumped, and this was not taken into account up to now, leading to the same problem.

In fact, these points were corrected so that Lemma 7.3 could be applicable, but only for the particular case of the classical sigma transformation.

Thus, important conservation errors appeared when using a different mapping.

We have corrected these errors and made the ALE-MURD schemes conservative up to a very satisfactory level. Let us illustrate it trough a numerical test simulating the effect of a heat source in the middle of a closed basin. We are interested in evaluating the conservation property of the PSI ALE-MURD scheme implemented in Telemac-3D.

We first consider a square closed basin of side $L = 4000$ m, with a constant bottom at $z = -10$ m. At the initial time the water height is flat ($h = 10$ m) and no motion is taken. Figure 7.30 shows the vertical cross-section of the initial three-dimensional mesh at $y = 2000$ m and Figure 7.31 the horizontal two-dimensional grid used to build it.

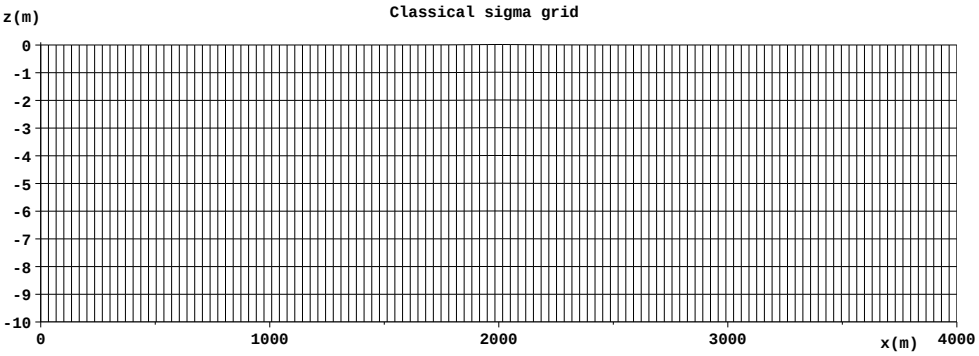


FIG. 7.30 – Vertical cross-section at $y = 2000$ m of the 3D mesh.

After one time step, water is discharged with an excess temperature of $T = 333^o$C from a point source located in the middle of the water column at the centre of the model. The coordinates of this source point are $x = 2000$ m, $y = 2000$ m and $z = -5$ m. The water discharge of the source is $Q = 20m^3/s$. A constant horizontal viscosity coefficient of $\nu_h = 10^{-4}$ and no vertical diffusion are chosen for the fluid velocities. No diffusion at all is considered on the tracer, so that only the advection scheme is evaluated. We indicate that all the simulations are performed with a time step length of 5 seconds for a time interval of 30 minutes.

We first simulate the advection of a passive tracer : the temperature is assumed to have no influence on the fluid. The advection is performed using successively the characteristic method and the PSI ALE-MURD scheme. Figure 7.32 gives the evolution of the relative loss of tracer quantity during the

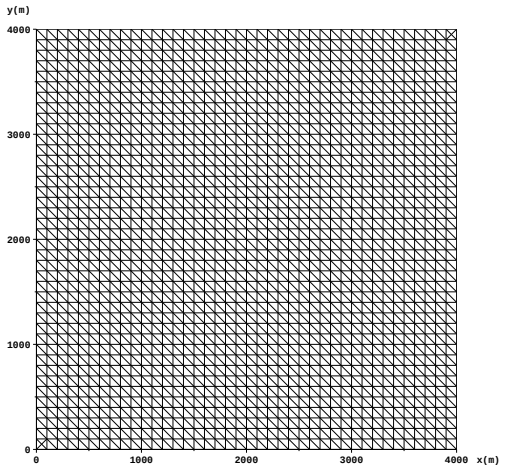


FIG. 7.31 – The horizontal 2D mesh.

simulation. It reveals that the characteristic method has a serious problem with conservation. On the contrary, the conservation property of the MURD scheme is excellent : the relative loss of tracer quantity during the simulation is of about 10^{-12}.

The temperature profiles obtained with both schemes at $t = 1800$ seconds are shown in vertical cross-sections at $y = 2000$ m in Figure 7.33. Since the tracer is assumed to be passive, the heat spreads out in all directions.

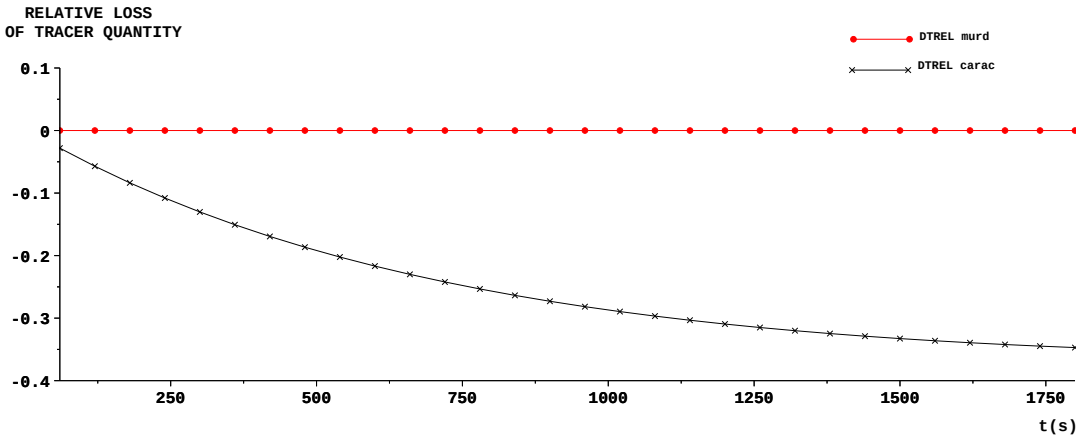


FIG. 7.32 – Passive tracer : relative loss of advected quantity using the characteristic method (solid red line with plane circles) and the ALE-MURD scheme (solid black line with crosses).

A second simulation is performed in the same configuration but considering the temperature as an active tracer : a state equation relates the fluid

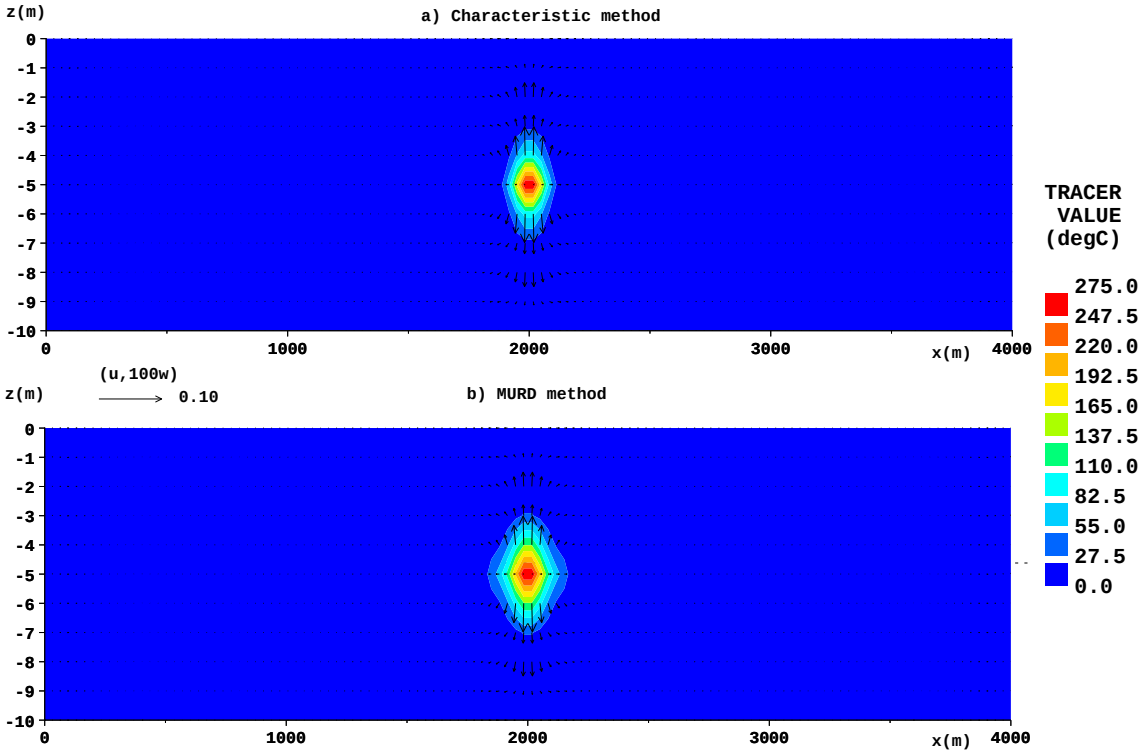


FIG. 7.33 – Passive tracer : temperature profiles obtained at $t = 1800$ s using (a) the characteristic method and (b) the PSI ALE-MURD scheme.

density to the temperature (see section 5.3.1). Figure 7.34 gives the evolution of the relative loss of tracer quantity during the simulation. Again, the conservation of the advected quantity is not satisfactory with the characteristic method, whereas it is excellent with the MURD scheme.

The temperature profiles after 30 minutes are shown in Figure 7.35. The effect of the heat on the flow has been taken into account : the velocities are mainly vertical and their amplitude is ten times larger than in the previous results where the tracer was considered passive. In addition, these velocities are oriented in the direction of the free surface : the heat goes upwards. We observe that the heat does not reach the free surface when advected with the characteristic method, what reveals a further problem of this method. However the MURD scheme seems to be more diffusive.

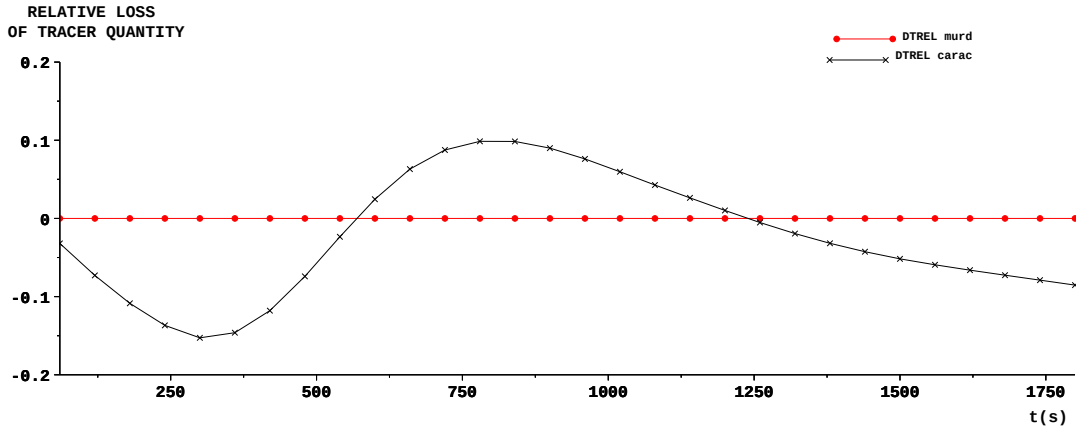


FIG. 7.34 – Active tracer : relative loss of advected quantity using the characteristic method (solid red line with plane circles) and the ALE-MURD scheme (solid black line with crosses).

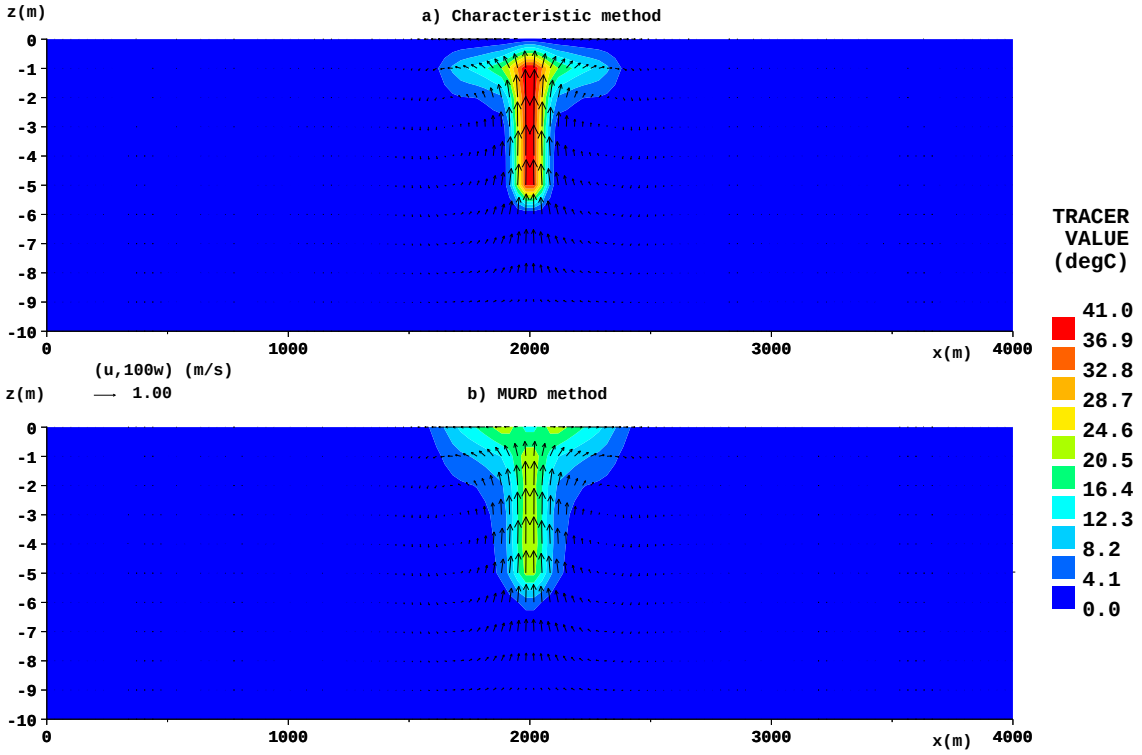


FIG. 7.35 – Active tracer : temperature profiles obtained at $t = 1800$ s using (a) the characteristic method and (b) the PSI ALE-MURD scheme.

The simulations presented next have been performed on a different configuration. A cone-shaped obstacle is placed on the bottom at the middle of the domain. The influence of the temperature on the fluid is taken into consideration again, but this time we only use the ALE-MURD scheme to convect the tracer. In fact, we compare the results obtained with this scheme on two different meshes, whose vertical cross-sections are shown in Figure 7.36 : mesh (a) has been obtained using the classical sigma transformation whereas mesh (b) has been obtained using another mapping, included in the recently implemented generalized sigma coordinate system. Note that mesh (b) has a fixed horizontal plane at $z = -4$, so that the layers containing the heat source do not move during the entire simulation.

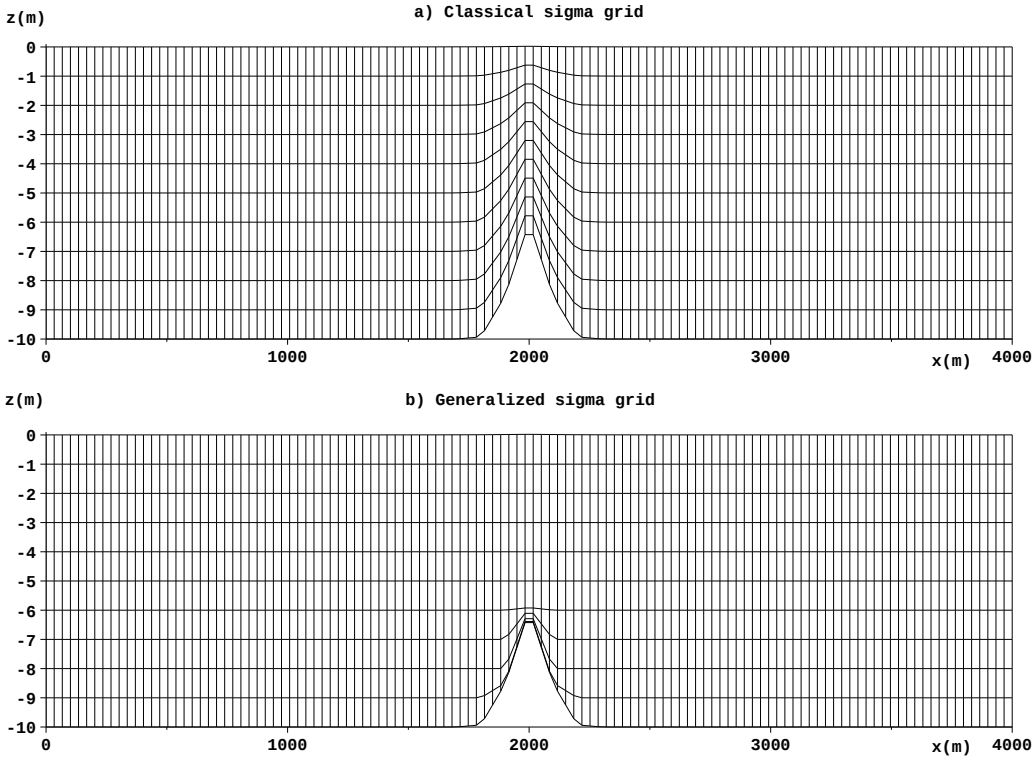


FIG. 7.36 – Mesh (a) has been obtained using the classical sigma transformation. Mesh (b) has been obtained using a more general sigma coordinate system.

Figure 7.37 gives the evolution of the relative loss of tracer quantity during the simulation. On both meshes, the ALE-MURD scheme is perfectly conservative.

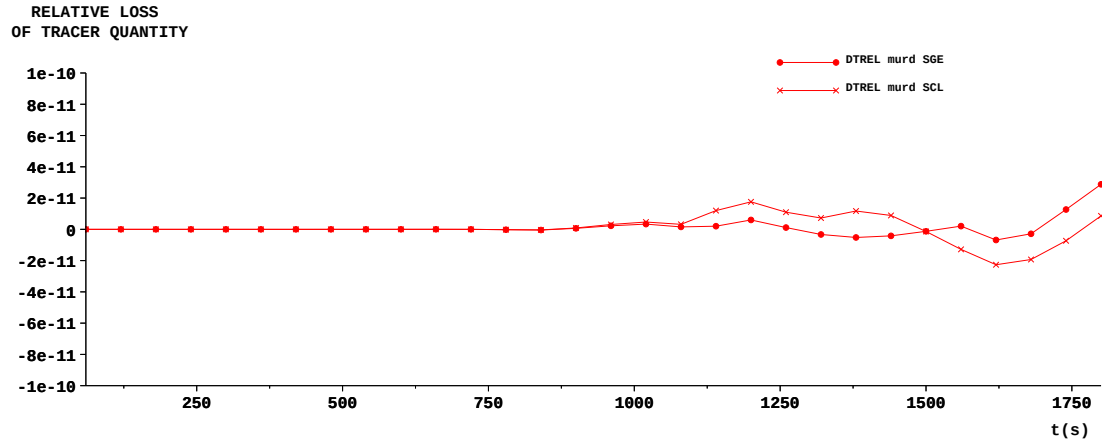


FIG. 7.37 – Relative loss of advected quantity using the classical sigma mesh a) (solid line with crosses) and the sigma generalized mesh b) (solid line with plane circles).

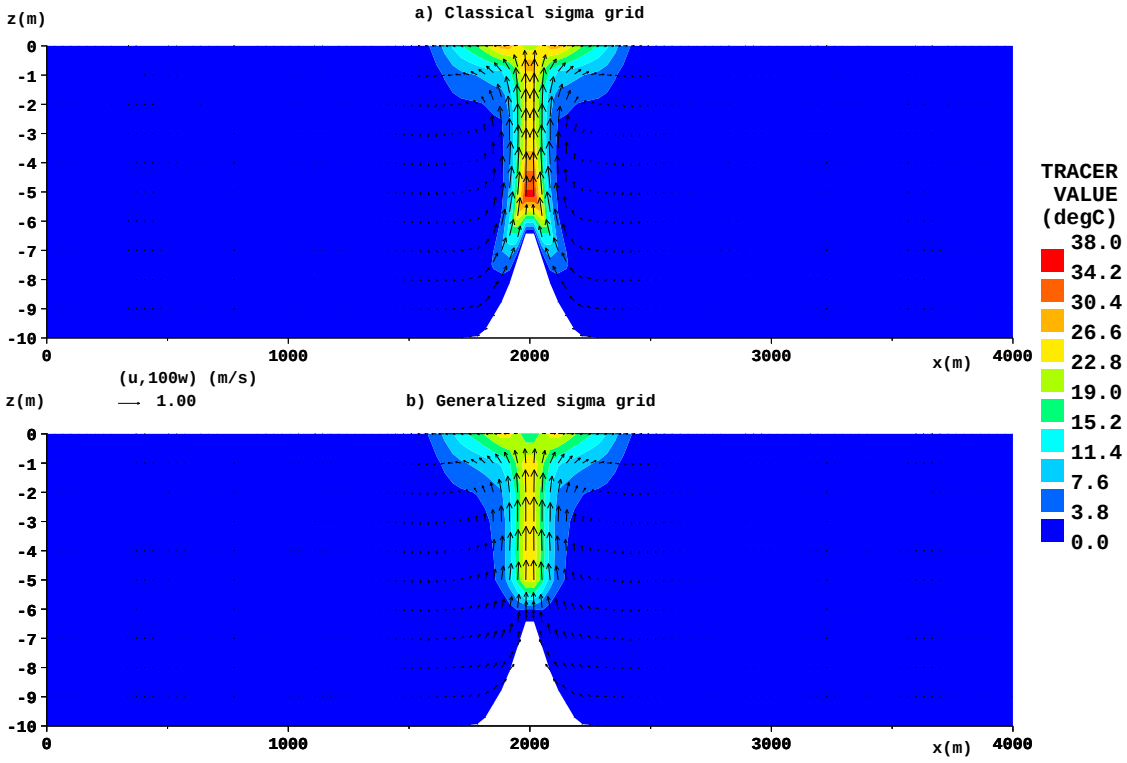


FIG. 7.38 – Temperature profiles obtained at $t = 1800$ s using (a) the classical sigma mesh and (b) the sigma generalized mesh.

We recall that the advected quantity was not conserved on a mesh obtained with a different mapping than the classical sigma transformation before updating the MURD schemes in Telemac-3D by extending them to the ALE-MURD approach.

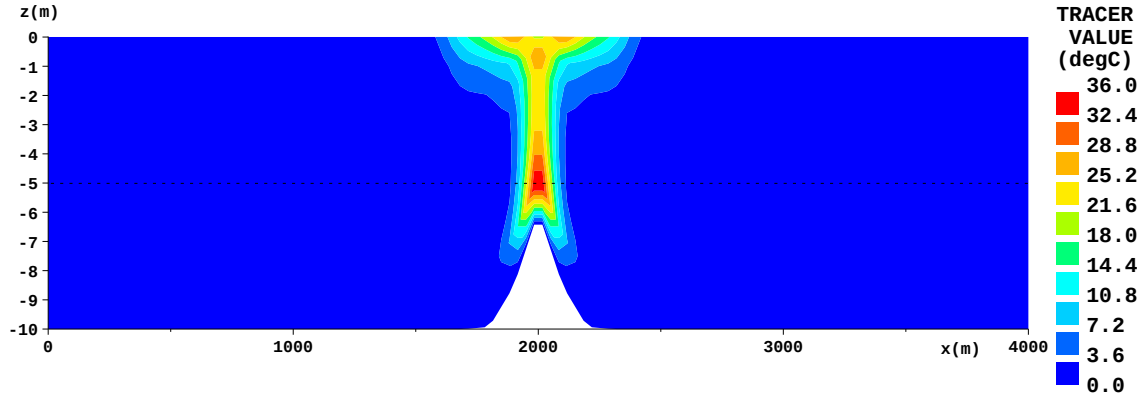
Figure 7.38 (and Figure 7.39 for more clarity) shows the temperature profiles obtained after 360 time steps, using both meshes, as vertical-cross sections at $y = 2$ m. The result obtained using the generalized sigma mesh (b) is more satisfactory, since the heat only goes upwards. This does not hold when using the classical sigma mesh (a) : the temperature also spreads out in the bottom direction. We point out that this result confirms the advantage of the general sigma coordinate introduced in Chapter 4 and implemented in Telemac-3D : the mesh can be adapted to the particular needs of the applications.

7.5.9 Conclusion

We have updated the MURD schemes implemented in the Telemac-3D system taking profit of the ALE-MURD approach introduced in Chapter 3. These advection schemes are now compatible with any ALE mapping for a three-dimensional domain moving in the vertical direction only. In particular, they are conservative up to a very satisfactory level with any vertical discretization of the domain.

Note that the conservation property of the ALE-MURD schemes is ensured in the same way for the hydrostatic and the nonhydrostatic models in the Telemac-3D system. We think that this is not optimal for the nonhydrostatic model : it would be more convenient to apply the proceedure described in section 3.8, valid for linear advection problems with divergence free-velocity, posed on three-dimensional domains moving in the vertical direction only. Indeed, in the non-hydrostatic model the total velocity $\boldsymbol{U} = (\boldsymbol{u}, w)$ can easily be computed such that it is divergence-free at least in the sense of (3.30). If this relation is satisfied, and if the mesh velocity is considered constant within each time step, Proposition 3.10 states that the ALE-MURD schemes are conservative.

(a) Classical sigma grid



(b) Generalized sigma grid

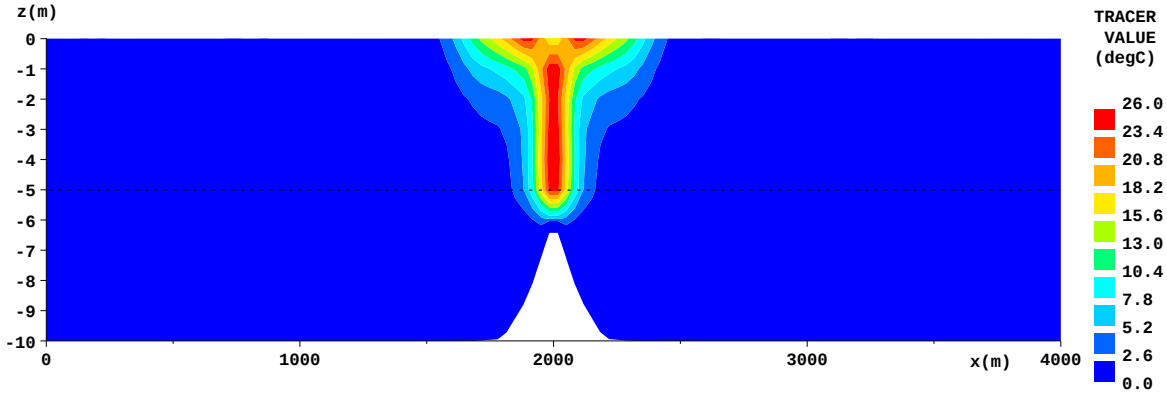


FIG. 7.39 – Temperature profiles obtained at $t = 1800$ s using (a) the classical sigma mesh and (b) the sigma generalized mesh.

Annexe A

Parameter identification for a one-dimensional blood flow model

This work has been carried out in collaboration with Vincent Martin, François Clément and Jean-Frédéric Gerbeau. It was published in ESAIM Proceedings [87]*, September 2005.*

Introduction

We focus in this study on the parameter estimation of a 1-d blood flow model, [116, 45] :

$$\begin{aligned} \frac{\partial A}{\partial t} + \frac{\partial Q}{\partial z} &= 0, \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial z}\left(\frac{\alpha Q^2}{A}\right) + \frac{A}{\rho}\frac{\partial P}{\partial z} + K_r\left(\frac{Q}{A}\right) &= 0, \end{aligned}$$

where the pressure P, the area A and the flux Q are the unknowns of the problem. We denoted by z the abscissa, t the time, ρ the density of the blood. Two parameters, the Coriolis coefficient α and the friction parameter K_r, are introduced in this model. This system is closed with a wall displacement law of the form

$$P(t,z) - P_{ext} = \tilde{\beta}\left(A^{1/2} - A_0^{1/2}\right),$$

that links the pressure and the area. We introduced the external pressure P_{ext}, the area at rest A_0 and a coefficient $\tilde{\beta}$ that takes into account the mechanical characteristics of the arterial wall. The parameters $(\alpha, \tilde{\beta}, K_r, A_0)$ used in this model are related to physiological data or to the velocity profile. Thus our aim is to identify some of these parameters.

Motivations Two main objectives can be thought of to motivate the parameter estimation in 1-d models. The first goal may have interesting clinical applications. Knowing some non-invasive clinical data measured on a patient, one would like to retrieve the actual physiological or mechanical constants of this patient. For instance, it is possible to measure unintrusively the mean fluxes and areas as a function of time at two or three different sections of an artery. From these data, one would like to identify the mechanical properties of the arterial wall : we could thus obtain the compliance of the wall and the pressure in a totally non-invasive manner.

The second objective is consists in making a *coarsening of models*. One is now able to solve the full 3D fluid structure interaction problem on real geometries, coming from real patients. However the resolution of this problem is quite expensive and this complex model cannot probably be afforded for an intensive numerical study requiring lots of resolutions. This can occur when one wishes to modify the configuration of the flux or the boundary conditions for instance. In this case, a single resolution of the accurate but expensive 3D model could provide *data*, such as the flux and the area for all sections of the mesh. Then one can estimate the 1-d parameters *from the 3D data* ; this would allow to make use of the cheap 1-d model for intensive computations, but using the parameters that take into account data coming from a real geometry and a physically more detailed model. This multiscale approach has already been used in a medical application : some 3D Navier-Stokes (without compliant walls) were performed to obtain a simple numerical/experimental law that was used in ulterior 0-d simulations, [96, 95]. The difference is that in this study, we base our parameter identification on sound mathematical tools, and we focus on 1-d models that provide a more accurate description of wave propagation in the large arteries. One of the conclusions of this work is that the coefficient estimated by solving the inverse problem is quite different from the coefficient one would have chosen *a priori* (an *a priori* expression for $\tilde{\beta}$ is provided in Section A.1). Thus this approach could provide 1-d models that are suitable for a 3D–1-d coupling in multiscale computations, see [44]. This illustrates the relevance of our approach.

To conclude on the motivations, one can either aim at estimating physical parameters from experimental unintrusive measurements, or at having a cheap 1-d model to be as close as possible to an expensive 3D model, in order to make realistic configuration studies using this cheap model. We note here that the methodology remains identical for both objectives, the only difference being the expression of the measurement operator.

Methodology We followed a standard nonlinear least squares approach to parameter estimation, [13, 14, 15]. It is based on the optimization of an appropriate cost function. The resolution of such a minimization problem generally requires the efficient and accurate computation of the gradient of

the cost function with respect to the parameters, [13]. We discretized the 1-d model with a second order Taylor-Galerkin scheme, [4]. We computed analytically the gradient of the discrete cost function, using an adjoint approach. The adjoint problem obtained is, as expected, a linear 1-d hyperbolic system, but has nonstandard discretization and boundary conditions, that are due to the differentiation of the Taylor-Galerkin scheme.

A previous attempt was made in [78] to estimate the elasticity of the arteries. However, instead of differentiating the discrete equations as we do here, the author differentiated first the continuous equations and then discretized the continuous adjoint problem thus obtained. We believe this is a possible reason for the very slow convergence he reached in the minimization process (about 1500 iterations for three parameters).

Finally, this approach allows to make a sensitivity analysis, [11, 79], that can provide information on the relevant parameters to estimate, and on the type of measurements to perform. But this goes beyond the scope of this study.

Results The analytical discrete gradient was implemented and validated by comparison with finite differences approximations. The adjoint problem was not computed by automatic code differentiation, but directly implemented. Thus we could easily control the memory required by the gradient code. We used a constrained optimization code based on a quasi-Newton method with active constraints.

We present some preliminary numerical results. In these numerical simulations, we mainly focused on determining the parameter $\tilde{\beta}$ that is linked to the mechanical properties, i.e. the compliance, of the arterial walls. The synthetic data we used to estimate the parameter were obtained from a numerical computation performed with a 3D fluid structure interaction model. We first used as data the values of the areas and fluxes at only two or three points of the domain (boundaries plus maybe the middle point). Although the data at two points do not seem to be enough to find a stable value, it seems that with three points, one can obtain a $\tilde{\beta}$ relatively stable (i.e. it is little changed when the estimation is made with all available spatial data). In the second numerical tests, we used all data available from the 3D computation. These first numerical results seem promising and should be followed by further developments.

In Section A.1, we present briefly the continuous 1-d model, that is derived in Section A.2 with the Taylor-Galerkin scheme. In Section A.3, the gradient of the least squares cost function is computed with the adjoint approach. Numerical results are presented in Section A.4, and some conclusions and perspectives in Section A.5.

A.1 Direct model : 1-d blood flow model

We present in this section a 1-d blood flow model based on the works in [46, 124]. See also [116]. It is a 1-d vectorial hyperbolic problem, with a 2×2 flux matrix that admits two real eigenvalues with opposite signs under physiological conditions.

We leave the problem of the parameterization and of the measurements for the next sections, see Section A.3.1.

A.1.1 Continuous blood flow model

Let $\Omega = (0, L)$ be a 1-d domain of length $L > 0$. Let $I = (0, T_f)$, with $T_f > 0$, the time interval of simulation. The continuous system of equations reads, for the abscissa $z \in \Omega$ and the time $t \in I$

$$\begin{array}{rcl} \dfrac{\partial A}{\partial t} + \dfrac{\partial Q}{\partial z} & = & 0, \quad z \in \Omega, \quad t \in I, \\ \dfrac{\partial Q}{\partial t} + \dfrac{\partial}{\partial z}\left(\dfrac{\alpha Q^2}{A}\right) + \dfrac{A}{\rho}\dfrac{\partial P}{\partial z} + K_r\left(\dfrac{Q}{A}\right) & = & 0, \quad z \in \Omega, \quad t \in I, \end{array} \tag{A.1}$$

where $\mathbf{U} = [A, Q]^\top$ is the vectorial unknown of the problem, made of the area A and of the flux Q. The pressure P is an intermediary unknown. We denoted by ρ the blood density that we assume perfeclty known in this study. Two parameters, the Coriolis coefficient α and the friction parameter K_r, are introduced when deriving this model. This system of partial differential equations is completed with an initial condition

$$\mathbf{U}(0, z) \quad = \quad \mathbf{U_0}(z), \quad z \in \Omega, \tag{A.2}$$

and some adapted boundary conditions

$$\begin{array}{rcl} \phi_0(\mathbf{U}(t, 0), \mathbf{p}) & = & q_0(t), \quad t \in I, \\ \phi_L(\mathbf{U}(t, L), \mathbf{p}) & = & q_L(t), \quad t \in I\ . \end{array} \tag{A.3}$$

In equations (A.2) and (A.3), $\mathbf{U_0}$, q_0 and q_L are given initial and boundary data. The definition of the real-valued boundary functions ϕ_0, ϕ_L is discussed in Section A.2.2.

The system (A.1) is closed with a wall displacement law of the form, see [126] for instance and the reference therein,

$$P(t, z) - P_{ext} = \psi(A; A_0, \boldsymbol{\beta}) = \beta_0 \left[\left(\frac{A}{A_0}\right)^{\beta_1} - 1\right], \tag{A.4}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1)$ is a pair of positive real parameters. The power coefficient β_1 is often taken equal to $1/2$, which means that the pressure difference is proportional to the wall displacement, and, in this case, a linear elastic law

can provide an expression for β_0, if the mechanical properties of the arterial wall are known *a priori*,

$$\beta_1 = 1/2, \quad \beta_0 = \frac{\sqrt{\pi} h_{0,w} E}{\sqrt{A_0}(1-\nu^2)}, \tag{A.5}$$

where $h_{0,w}$ is the wall thickness, E is the wall Young modulus, and ν is the Poisson coefficient. We can reformulate the wall displacement law in the following way :

$$P(t,z) - P_{ext} = \psi(A; A_0, \beta) = 2\rho\beta \left[A^{1/2} - A_0^{1/2} \right], \tag{A.6}$$

$$\text{with} \quad \beta = \frac{\beta_0}{2\rho A_0} = \frac{\sqrt{\pi} h_{0,w} E}{2\rho A_0 (1-\nu^2)}.$$

We introduce the following quantity

$$c = c(A; A_0, \boldsymbol{\beta}) = \sqrt{\frac{A}{\rho}\frac{\partial \psi}{\partial A}} = \sqrt{\frac{\beta_0 \beta_1}{\rho}\left(\frac{A}{A_0}\right)^{\beta_1}}, \tag{A.7}$$

which has the dimension of a velocity and is related to the speed of propagation of simple waves along the tube. We also introduce the integral of the square of the celerity c with respect to the area

$$C(A; A_0, \boldsymbol{\beta}) = \int_{A_0}^{A} c^2(\tau; A_0, \boldsymbol{\beta})\, d\tau = \frac{\beta_0 \beta_1 A_0}{\rho(\beta_1 + 1)} \left[\left(\frac{A}{A_0} \right)^{\beta_1 + 1} - 1 \right] . \tag{A.8}$$

Defining the flux function

$$\mathbf{F}(\mathbf{U}, \mathbf{p}) = \begin{bmatrix} Q \\ \alpha \dfrac{Q^2}{A} + C \end{bmatrix} \tag{A.9}$$

and the source term

$$\mathcal{B}(\mathbf{U}, \mathbf{p}) = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}, \qquad B_2 = K_r \frac{Q}{A} + \frac{A}{\rho}\frac{\partial \psi}{\partial A_0}\frac{dA_0}{dz} + \frac{A}{\rho}\frac{\partial \psi}{\partial \boldsymbol{\beta}}\frac{d\boldsymbol{\beta}}{dz} - \frac{\partial C}{\partial A_0}\frac{dA_0}{dz} - \frac{\partial C}{\partial \boldsymbol{\beta}}\frac{d\boldsymbol{\beta}}{dz}, \tag{A.10}$$

we can write the complete problem in a conservative form

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial}{\partial z}[\mathbf{F}(\mathbf{U}, \mathbf{p})] + \mathcal{B}(\mathbf{U}, \mathbf{p}) \quad = \quad \mathbf{0}, \quad z \in \Omega \quad t \in I \ . \tag{A.11}$$

We introduce the jacobian matrix

$$\mathbf{H}(\mathbf{U}, \mathbf{p}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} 0 & 1 \\ c^2 - \alpha \left(\dfrac{Q}{A} \right)^2 & 2\alpha \dfrac{Q}{A} \end{bmatrix}, \tag{A.12}$$

whose eigenvalues are real, distinct, see [116]. It is also noted that for common values of the blood flow in human arteries, these eigenvalues have *opposite signs*. We assume in the rest of the article that this hypothesis holds.

Remark 1 We can make two remarks concerning the initial and boundary terms. First, when simulating a phenomenon that is periodic in time such as the arterial flow, the initial data $\mathbf{U_0}$ should not interfere on the results. Therefore we did not consider it as one of the parameters to be estimated.

Second, as already noticed, the boundary conditions must be properly chosen. The number of boundary conditions to impose at each end of the vessel equals the number of characteristics entering the domain through that boundary. As the eigenvalues are distinct with opposite signs, i.e. the flow is sub-critical everywhere, one must impose exactly *one scalar boundary condition at* $z = 0$ *and* $z = L$, see [116].

Thus the boundary conditions in equation (A.3) are correctly set, provided that the scalar functions ϕ_0, ϕ_L are properly chosen.

A.2 Discrete blood flow model

In this section, we give the discretization of the continuous problem (A.11).

In the rest of this article, the upper-scripts will be devoted to time steps numbers, whereas the lower indices will in general denote the space indices, or the component index of a vector. For instance, for a vector v, the α^{th} component of v at time t^n and at abscissa z_i will be denoted $(v_\alpha)_i^n$.

A.2.1 Taylor-Galerkin scheme

We discretize our system by a second order Taylor-Galerkin scheme [4], which might be seen as the finite element counterpart of the Lax-Wendroff scheme. It has been chosen for its excellent dispersion error characteristic and its relative simplicity of implementation. See [45] for details concerning the derivation of the scheme for the blood flow model.

Let the interval $\Omega = (0, L)$ be subdivided into $N+1$ elements $e_{i+\frac{1}{2}} = [z_i, z_{i+1}]$, for $i = 0, \ldots, N$ and $z_{i+1} = z_i + h_{i+\frac{1}{2}}$, with $z_0 = 0$, and $\sum_{i=0}^{N} h_{i+\frac{1}{2}} = L$, where $h_{i+\frac{1}{2}} > 0$ is the local element size. Let $h > 0$ be the smallest diameter of all elements $e_{i+\frac{1}{2}}, i = 0, \ldots, N$. We discretize the time interval $I = (0, T_f)$ in the same way : let $(t^n), n = 0, \ldots, N_t$, be $N_t + 1$ instants such that $t_0 = 0 < \ldots < t^n < t^{n+1} < \ldots < t^{N_t} = T_f$. We call $\Delta t^{n+\frac{1}{2}} = t^{n+1} - t^n > 0$ the time step between t^n and t^{n+1}, $n = 0, \ldots, N_t - 1$.

The space discretization is carried out using the finite element method [115]. Let Ψ_h be the space of continuous piecewise linear finite element functions, also denoted P_1, and $\mathbf{\Psi}_h = [\Psi_h]^2$, while $\mathbf{\Psi}_{h,0} = [\Psi_{h,0}]^2 = \{\boldsymbol{\psi}_h \in \mathbf{\Psi}_h \,|\boldsymbol{\psi}_h = \mathbf{0} \text{ at } z = 0 \text{ and } z = L\}$. Let ψ_i be the P_1 linear finite element nodal function associated to the node at $z = z_i$, $i = i = 0, \ldots, N+1$. Thus one can write $\Psi_h = \text{span}\{\psi_i,\ i = 0, \ldots, N+1\}$ while $\Psi_{h,0} = \text{span}\{\psi_i,\ i = 1, \ldots, N\}$. We will denote a generic vector valued test function by $\boldsymbol{\psi}_h \in \mathbf{\Psi}_h$. The discrete

continuity and momentum equations are recovered by taking test functions of the form $\boldsymbol{\psi}_h = [\psi_h, 0]^T$ and $\boldsymbol{\psi}_h = [0, \psi_h]^T$, respectively, with $\psi_h \in \Psi_{h,0}$.

At each time step we seek the solution $\mathbf{U}_h \in \boldsymbol{\Psi}_h$ that we may write $\mathbf{U}_h^n(z,t) = \sum_{i=0}^{N+1} \mathbf{U}_i^n \psi_i(z,t)$, with $\mathbf{U}_i^n = [A_i^n, Q_i^n]$ the approximation of A and Q at mesh node z_i.

Further, we note the $\mathbf{L}^2(0,L) = \mathbf{L}^2(\Omega)$ scalar product $(\mathbf{u},\mathbf{v})_\Omega = \int_0^L \mathbf{u}\cdot\mathbf{v}dz$.

Introducing the notations for $\mathbf{U} \in \boldsymbol{\Psi}_h$, $\mathbf{B}_\mathbf{U} = \dfrac{\partial \mathcal{B}}{\partial \mathbf{U}}$,

$$\mathbf{F}_{LW}^n(\mathbf{U}) = \mathbf{F}(\mathbf{U}) - \frac{\Delta t^{n+\frac{1}{2}}}{2}\mathbf{H}(\mathbf{U})\mathcal{B}(\mathbf{U})$$

$$\text{and} \qquad \mathcal{B}_{LW}^n(\mathbf{U}) = \mathcal{B}(\mathbf{U}) - \frac{\Delta t^{n+\frac{1}{2}}}{2}\mathbf{B}_\mathbf{U}(\mathbf{U})\mathcal{B}(\mathbf{U}) \ ,$$

we write in equation (A.13) the Taylor-Galerkin discretization of the problem (A.11).

Given $\mathbf{U}_h^0$, for $n = 0, 1, \ldots, N_t - 1$, find $\mathbf{U}_h^{n+1} \in \boldsymbol{\Psi}_h$ such that for all $\boldsymbol{\psi}_h \in \boldsymbol{\Psi}_{h,0}$

$$\begin{aligned}
(\mathbf{U}_h^{n+1}, \boldsymbol{\psi}_h)_\Omega \;=\; & (\mathbf{U}_h^n, \boldsymbol{\psi}_h)_\Omega + \Delta t^{n+\frac{1}{2}} \left\{ \left(\mathbf{F}_{LW}^n(\mathbf{U}_h^n), \frac{d\boldsymbol{\psi}_h}{dz} \right)_\Omega \right. \\
& \left. - (\mathcal{B}_{LW}^n(\mathbf{U}_h^n), \boldsymbol{\psi}_h)_\Omega \right\} \\
& + \frac{(\Delta t^{n+\frac{1}{2}})^2}{2} \left\{ - \left(\mathbf{H}(\mathbf{U}_h^n) \frac{\partial \mathbf{F}}{\partial z}(\mathbf{U}_h^n), \frac{d\boldsymbol{\psi}_h}{dz} \right)_\Omega \right. \\
& \left. + \left(\mathbf{B}_\mathbf{U}(\mathbf{U}_h^n) \frac{\partial \mathbf{F}(\mathbf{U}_h^n)}{\partial z}, \boldsymbol{\psi}_h \right)_\Omega \right\} \ ,
\end{aligned} \tag{A.13}$$

$$+\text{B.C. } z = 0 \qquad \text{equation on } \mathbf{U}_h^{n+1}(0) \ ,$$

$$+\text{B.C. } z = L \qquad \text{equation on } \mathbf{U}_h^{n+1}(L) \ .$$

In the system (A.13), by taking internal test functions $\boldsymbol{\psi}_h = [\psi_i, 0]^T$ and $\boldsymbol{\psi}_h = [0, \psi_i]^T$, for $i = 1, \ldots N$, we obtain N discrete equations for continuity and momentum, respectively, for a total of $2(N+2)$ unknowns (A_i and Q_i for $i = 0, \ldots, N+1$). Thus boundary and compatibility conditions have to provide four additional scalar relations, see next section.

To fully discretize the equation in (A.13), we need to determine how the non-linear terms are computed and which numerical integration is performed. We choose to approximate the vectors $\mathbf{F}$ and $\mathcal{B}$ depending on $\mathbf{U}_h^n$

in (A.13) in the space P_1, and the matrices $\mathbf{H}$ and $\mathbf{B_U}$ in the space of element-wise constant functions $P_{0,h} := \left\{ v_h \in L^2(\Omega) \mid v_h|_{e_{i+\frac{1}{2}}} = \text{cst},\ i = 0, \ldots, N \right\}$. With this hypothesis, the integrations in (A.13) can easily be made exactly.

Thus for a vectorial function $v : \mathbf{\Psi}_h \to \mathbb{R}^2, \mathbf{U} \mapsto v(\mathbf{U}) = [v_1, v_2]^\top$, we define

$$v_h(\mathbf{U}) = \left[\sum_{i=0}^{N+1} v_1(\mathbf{U}_i)\psi_i \ , \sum_{i=0}^{N+1} v_2(\mathbf{U}_i)\psi_i \right]^\top \in \mathbf{\Psi}_h,$$

and for a matrix $M : \mathbf{\Psi}_h \to \mathbb{R}^{2\times 2}, \mathbf{U} \mapsto M(\mathbf{U}) = (M_{\alpha\beta})_{\alpha,\beta=1,2}$, we define

$$M_h(\mathbf{U}) = \left(M_{h,\alpha\beta}(\mathbf{U})\right)_{\alpha,\beta=1,2}, \quad M_{h,\alpha\beta}(\mathbf{U}) = \sum_{i=0}^{N} \tilde{M}_{h,\alpha\beta,\, i+\frac{1}{2}} 1_{i+\frac{1}{2}} \in P_{0,h} \tag{A.14}$$

with the mean value $\tilde{M}_{h,\alpha\beta,\, i+\frac{1}{2}} = \frac{1}{2}\left(M_{\alpha\beta}(\mathbf{U}_i) + M_{\alpha\beta}(\mathbf{U}_{i+1})\right)$ over the element $e_{i+\frac{1}{2}} = [z_i, z_{i+1}]$, and the characteristic function $1_{i+\frac{1}{2}}$ of the interval $[z_i, z_{i+1}]$, for $i = 0, \ldots, N$.

With these notations, we can write the fully discretized Taylor-Galerkin scheme (A.13). Using the notation

$$\mathbf{F}^n_{h,LW}(\mathbf{U}) = \mathbf{F}_h(\mathbf{U}) - \frac{\Delta t^{n+\frac{1}{2}}}{2}\mathbf{H}_h(\mathbf{U})\mathcal{B}_h(\mathbf{U})$$

$$\text{and} \qquad \mathcal{B}_{h,LW}(\mathbf{U}) = \mathcal{B}_h(\mathbf{U}) - \frac{\Delta t^{n+\frac{1}{2}}}{2}(\mathbf{B_U})_h(\mathbf{U})\mathcal{B}_h(\mathbf{U}),$$

and introducing the operator for $\mathbf{U}, \mathbf{W} \in \mathbf{\Psi}_h$

$$\begin{aligned} a^n_h(\mathbf{U}, \mathbf{W}; \mathbf{p}) \ = \ & \Delta t^{n+\frac{1}{2}} \left\{ \left(\mathbf{F}^n_{h,LW}(\mathbf{U}), \frac{d\mathbf{W}}{dz} \right)_\Omega - \left(\mathcal{B}^n_{h,LW}(\mathbf{U}), \mathbf{W} \right)_\Omega \right\} \\ & + \frac{(\Delta t^{n+\frac{1}{2}})^2}{2} \left\{ - \left(\mathbf{H}_h(\mathbf{U}) \frac{\partial \mathbf{F}_h}{\partial z}(\mathbf{U}), \frac{d\mathbf{W}}{dz} \right)_\Omega \right. \\ & \left. + \left((\mathbf{B_U})_h(\mathbf{U}) \frac{\partial \mathbf{F}_h}{\partial z}(\mathbf{U}), \mathbf{W} \right)_\Omega \right\}, \end{aligned} \tag{A.15}$$

the fully discretized scheme (A.13) yields :
Given $\mathbf{U}^0_h \in \mathbf{\Psi}_h$, find $\mathbf{U}_h := (\mathbf{U}^n_h)_{n=1,\ldots,N_t} \in (\mathbf{\Psi}_h)^{N_t}$ such that for $n = 0, 1, \ldots, N_t - 1$

$$\begin{aligned} (\mathbf{U}^{n+1}_h, \boldsymbol{\psi}_h)_\Omega \ = \ & (\mathbf{U}^n_h, \boldsymbol{\psi}_h)_\Omega + a^n_h(\mathbf{U}^n_h, \boldsymbol{\psi}_h; \mathbf{p}) \quad \forall \boldsymbol{\psi}_h \in \mathbf{\Psi}_{h,0}\ , \\ & \text{B.C. on } \mathbf{U}^{n+1}_{h,0}, \\ & \text{B.C. on } \mathbf{U}^{n+1}_{h,N+1}. \end{aligned} \tag{A.16}$$

In the discrete approximate problem (A.16), all the integrals can be computed exactly as they involve only the products of two functions belonging either to P_1 or to P_0.

We need to add boundary conditions and compatibility conditions at $z = 0$ and $z = L$ to close the system.

A.2.2 Boundary conditions

In this section and in the next sections A.2.3 and A.2.4, we will omit the subscript h in the variable names, for the sake of simplicity.

As noticed in Remark 1, the hyperbolic problem (A.11) is well posed under sub-critical flow hypothesis when one imposes as boundary conditions one scalar equation at each side of the tube. These conditions can be the prescribed incoming characteristics, the prescribed pressure or the prescribed flux, for instance. It is not possible to impose exactly at the same boundary both the flux and the pressure, see [45].

We decide to give a flux at the inlet and a pressure at the outlet :

$$\begin{array}{rcl} Q_0^{n+1} & = & \phi_0(q_0(t^{n+1}), \mathbf{p}) = q_0(t^{n+1}) \\ A_{N+1}^{n+1} & = & \phi_L(p_L(t^{n+1}), \mathbf{p}), \end{array} \tag{A.17}$$

where q_0 is the flux and p_L is the pressure to be imposed at the boundaries, whereas $\phi_i, i = 0, L$ are given functions. The function ϕ_0 is in this case the identity, and ϕ_L transforms the given pressure p_L at the outlet into the conservative unknown A. With the pressure law (A.4), it becomes $A = \phi_L(p_L, \mathbf{p}) = A_0 \left(1 + \frac{p_L}{\beta_0}\right)^{1/\beta_1}$.

A.2.3 Compatibility conditions

To close the discrete system we need two more conditions on the boundaries. These conditions are a numerical artefact that is linked to the type of scheme we adopted. They are called compatibility conditions and are chosen to be non-reflecting conditions at each side of the tube. We assume that these conditions are treated explicitly and linearly, imposing some conditions on the pseudo-characteristic variables, [116]. Thus they can take the following form :

$$\begin{array}{rcll} \mathbf{l}_2^\top(\mathbf{U}^n)\mathbf{U}_0^{n+1} - T_2(\mathbf{U}^n) & = & 0 & \text{at } z = 0, \\ \mathbf{l}_1^\top(\mathbf{U}^n)\mathbf{U}_{N+1}^{n+1} - T_1(\mathbf{U}^n) & = & 0 & \text{at } z = L, \end{array} \tag{A.18}$$

where $\mathbf{l}_i(\mathbf{U}^n) \in \mathbb{R}^2, i = 1, 2$ are the two left eigenvectors depending on $\mathbf{U}^n$ associated to the matrix $\mathbf{H}$ defined in equation (A.12), and T_1 and T_2 are some scalar functions depending on $\mathbf{U}^n$.

A.2.4 Fully discretized problem

Gathering the boundary and compatibility conditions (A.17) and (A.18) and reintroducing the parameters $\mathbf{p}$, we obtain

$$\begin{array}{rcl} \Theta_0(\mathbf{U}^n, \mathbf{p})\mathbf{U}_0^{n+1} - \mathbf{T}_0(\mathbf{U}^n, \mathbf{p}) & = & 0 \\ \Theta_L(\mathbf{U}^n, \mathbf{p})\mathbf{U}_{N+1}^{n+1} - \mathbf{T}_L(\mathbf{U}^n, \mathbf{p}) & = & 0 \ , \end{array} \tag{A.19}$$

with at $z = z_0 = 0$

$$\Theta_0(\mathbf{U}^n, \mathbf{p}) = \begin{bmatrix} \Theta_{0,1}^\top \\ \Theta_{0,2}^\top(\mathbf{U}^n, \mathbf{p}) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \mathbf{l}_2^\top(\mathbf{U}^n, \mathbf{p}) \end{bmatrix}, \quad \mathbf{T}_0(\mathbf{U}^n, \mathbf{p}) = \begin{bmatrix} q_0(t^{n+1}) \\ T_2(\mathbf{U}^n, \mathbf{p}) \end{bmatrix},$$

and at $z = z_{N+1} = L$

$$\Theta_L(\mathbf{U}^n, \mathbf{p}) = \begin{bmatrix} \Theta_{L,1}^\top(\mathbf{U}^n, \mathbf{p}) \\ \Theta_{L,2}^\top \end{bmatrix} = \begin{bmatrix} \mathbf{l}_1^\top(\mathbf{U}^n, \mathbf{p}) \\ 1 \quad 0 \end{bmatrix}, \quad \mathbf{T}_L(\mathbf{U}^n, \mathbf{p}) = \begin{bmatrix} T_1(\mathbf{U}^n, \mathbf{p}) \\ \phi_L(p_L(t^{n+1}), \mathbf{p}) \end{bmatrix}.$$

Finally, reintroducing the subscript h, adding the boundary and compatibility conditions (A.19) to the completely discretized scheme (A.16), the problem to be solved reads :
Given $\mathbf{U}_h^0 \in \mathbf{\Psi}_h$, find $\mathbf{U}_h := (\mathbf{U}_h^n)_{n=1,\dots,N_t} \in (\mathbf{\Psi}_h)^{N_t}$ such that for $n = 0, 1, \dots, N_t - 1$

$$\begin{array}{lcl} (\mathbf{U}_h^{n+1}, \boldsymbol{\psi}_h)_\Omega & = & (\mathbf{U}_h^n, \boldsymbol{\psi}_h)_\Omega + a_h^n(\mathbf{U}_h^n, \boldsymbol{\psi}_h; \mathbf{p}) \quad \forall \boldsymbol{\psi}_h \in \mathbf{\Psi}_{h,0} \ , \\ \Theta_0(\mathbf{U}_h^n, \mathbf{p})\mathbf{U}_{h,0}^{n+1} & = & \mathbf{T}_0(\mathbf{U}_h^n, \mathbf{p}) \ , \\ \Theta_L(\mathbf{U}_h^n, \mathbf{p})\mathbf{U}_{h,N+1}^{n+1} & = & \mathbf{T}_L(\mathbf{U}_h^n, \mathbf{p}) \ . \end{array} \tag{A.20}$$

Remark 2 The problem (A.11) is a 1-d nonlinear hyperbolic problem, but the Taylor-Galerkin discretization (A.20) contains (second order) parabolic terms. Thus, one needs to impose not only the boundary conditions (A.17) that are natural for a hyperbolic problem, but also the compatibility conditions (A.18). Therefore at each time step, before computing the solution at internal nodes, one has to compute the boundary values $\mathbf{U}_{h,0}^{n+1}$ and $\mathbf{U}_{h,N+1}^{n+1}$, by solving the two (2×2) linear systems (A.19).

A.3 Gradient of the discrete problem

We present in this section the gradient of the discrete problem (A.20) using the adjoint approach. We recall that we used and implemented the rule (advocated in [15], [13], [14]) : disretize first, differentiate after. In Section A.3.1, some notations are defined in order to present briefly in Section A.3.2 the adjoint state approach, and to compute analytically the gradient in Section A.3.3.

A.3.1 Forward operators

State equation

The equation (A.20) defines a *state* equation that relates the parameters $\mathbf{p}$ to the state variables $\mathbf{U}_h$

$$E(\mathbf{p}, \mathbf{U}_h) = 0 \ , \tag{A.21}$$

where the parameters are

$$\mathbf{p} = (\alpha, \beta_0, \beta_1, K_r, A_0, q_0, q_L)^\top \in \mathbb{R}^{\mathcal{P}} \ , \tag{A.22}$$

with the dimension $\mathcal{P} > 0$ of the parameter space to be defined further, and the state variable are

$$\mathbf{U}_h = (\mathbf{U}_h^n)_{n=1,\dots,N_t} = \left([A_i^n, Q_i^n]^\top \right)_{i=0,\dots,N+1;\ n=1,\dots,N_t} \in (\mathbf{\Psi}_h)^{N_t} = \mathbb{R}^{\mathcal{N}}, \tag{A.23}$$

with $\mathcal{N} = 2N_t(N+2)$.
In equation (A.23), we have assimilated the finite functional space $(\mathbf{\Psi}_h)^{N_t}$ and the space $\mathbb{R}^{\mathcal{N}}$. It will also be done for $\mathbf{\Psi}_h$ and $\mathbb{R}^{2(N+2)}$. In the rest of the article, for simplicity but with an abuse of notation, we may use the same notation $\mathbf{W}_h^n$ for an element of $\mathbf{\Psi}_h$ or of $\mathbb{R}^{2(N+2)}$.

As there exists a unique solution $\mathbf{U}_{\mathbf{p},h}$ to the problem (A.21) for a given set of parameters $\mathbf{p}$, see [45], we can define the direct application

$$\begin{array}{rcl} \varphi : \mathbb{R}^{\mathcal{P}} & \longrightarrow & \mathbb{R}^{\mathcal{N}}, \\ \mathbf{p} & \longmapsto & \varphi(\mathbf{p}) = \mathbf{U}_{\mathbf{p},h} \ , \end{array} \tag{A.24}$$

The parameters $\mathbf{p}$ in (A.22) are supposed to be constant with time and are to be estimated. The issue of the parameterization will be addressed in the section A.3.1.

Measurement operator

Let M be a measurement (or observation) operator

$$\begin{array}{rcl} M : \mathbb{R}^{\mathcal{N}} & \longrightarrow & \mathbb{R}^{\mathcal{M}}, \\ \mathbf{U} & \longmapsto & M(\mathbf{U}) = \mathbf{V}, \end{array} \tag{A.25}$$

with $\mathcal{M} > 0$ the dimension of the measurement space. Different measurement operators can be devised, linear or nonlinear, according to the type of data that are provided.

For instance, an experiment could provide the flux and area of a vessel at two different points $\zeta_0 < \zeta_1$, $\zeta_1 - \zeta_0 = L$, at given $(\mathcal{V}_t + 1)$ instants $\theta_0 = 0 < \ldots < \theta^\nu < \theta^{\nu+1} < \ldots < \theta^{\mathcal{V}_t} = T_f$. It is then reasonable to perform the 1-d simulation between these two points, and therefore take $z_0 = \zeta_0$ and $z_{N+1} = \zeta_1$. One can also assume for simplicity that the temporal discretization is chosen such that the instants $(\theta^\nu)_{\nu=0,\dots,\mathcal{V}_t}$ constitutes a subsequence of $(t^n)_{n=0,\dots,N_t}$. Therefore, let $\phi_t : \{0, 1, \dots, \mathcal{V}_t\} \to \{0, 1, \dots, N_t\}$ be the increasing function such that $t^{\phi_t(\nu)} = \theta^\nu$, with $\phi_t(0) = 0$ and $\phi_t(\mathcal{V}_t) = N_t$. With these hypotheses, the measurement operator is simply a weighted sampling of the direct solution $\mathbf{U}_{\mathbf{p},h}$. It is linear and can be expressed

$$M(\mathbf{U}) = \left([w_{i,\nu}^A \ A_i^{\phi_t(\nu)}, w_{i,\nu}^Q \ Q_i^{\phi_t(\nu)}]^\top \right)_{i=0 \text{ and } N+1;\ \nu=0,\dots,\mathcal{V}_t}$$

$$\text{and} \qquad \mathcal{M} = 2 \times (\mathcal{V}_t + 1) \times 2 \ ,$$

where the coefficients $w^A_{i,\nu}$ and $w^Q_{i,\nu}$ are positive weights attributed to each measure, representing the "degree of confidence", i.e. the inverse of the incertainty, one has in this measure.

When the data come from a numerical 3D fluid structure computation, one can exploit the numerical results on $(\mathcal{V}_z+1)$ internal sections $(\zeta_j)_{j=0,\ldots,\mathcal{V}_z}$ of the 3D domain : $\mathcal{V}_z = 1, 2$ or more. In this case, apply the same technique for the spatial discretization as the one shown above for the temporal sampling, and introduce the increasing indexation function ϕ_z in the same manner as ϕ_t. The measurement operator becomes this time

$$M(\mathbf{U}) = \left([w^A_{j,\nu} \ A^{\phi_t(\nu)}_{\phi_z(j)}, w^Q_{j,\nu} \ Q^{\phi_t(\nu)}_{\phi_z(j)}]^\top \right)_{j=0,\ldots,\mathcal{V}_z; \ \nu=0,\ldots,\mathcal{V}_t} \tag{A.26}$$

and

$$\mathcal{M} = 2(\mathcal{V}_t + 1)(\mathcal{V}_z + 1) \, , \tag{A.27}$$

with the positive weights $w^A_{j,\nu}$ and $w^Q_{j,\nu}$, $j = 0, \ldots, \mathcal{V}_z$; $\nu = 0, \ldots, \mathcal{V}_t$.

Forward function

Let $F = M \circ \varphi$ be the forward function, from the parameter space to the measure space

$$\begin{array}{rcl} F : \mathbb{R}^{\mathcal{P}} & \longrightarrow & \mathbb{R}^{\mathcal{M}}, \\ \mathbf{p} & \longmapsto & F(\mathbf{p}) = \mathbf{V}_{\mathbf{p},h} = M(\mathbf{U}_{\mathbf{p},h}) = M(\varphi(\mathbf{p})). \end{array} \tag{A.28}$$

Cost function

The resolution of the inverse problem consists in minimizing a cost function that computes the least squares error between the measured data and the numerical results computed with the 1-d model. Thus the inverse problem amounts to an optimization problem. To solve it, it is necessary to define the cost function to minimize, that depends on the type of problem under consideration. Then one must choose an adequate technique to perform the minimization of this cost function.

We assume that the measurement operator is defined by (A.26). Consider a given data vector $Z \in \mathbb{R}^{\mathcal{M}}$ and we call $Z_w \in \mathbb{R}^{\mathcal{M}}$ this vector multiplied by the positive weights introduced in (A.26),

$$Z = ([Ad^\nu_j, \ Qd^\nu_j]^\top)_{j=0,\ldots,\mathcal{V}_z; \ \nu=0,\ldots,\mathcal{V}_t} \ \in \mathbb{R}^{\mathcal{M}} \ ,$$

$$Z_w = ([w^A_{j,\nu} \ Ad^\nu_j, \ w^Q_{j,\nu} \ Qd^\nu_j]^\top)_{j=0,\ldots,\mathcal{V}_z; \ \nu=0,\ldots,\mathcal{V}_t} \ \in \mathbb{R}^{\mathcal{M}} \ ,$$

where Ad^ν_j, and respectively Qd^ν_j, represent the data area, respectively the data flux, given at space index j and time index ν.

Let J be the cost function

$$\begin{aligned} J : \mathbb{R}^{\mathcal{P}} &\longrightarrow \mathbb{R}, \\ \mathbf{p} &\longmapsto J(\mathbf{p}) = \frac{1}{2}\|Z_w - M(\mathbf{U}_{\mathbf{p},h})\|^2_{\mathbb{R}^{\mathcal{M}}} \ , \end{aligned} \tag{A.29}$$

where $\|\cdot\|_{\mathbb{R}^n}$ is the discrete l^2 norm in $\mathbb{R}^n$, associated with the scalar product $< \cdot, \cdot >_{\mathbb{R}^n}$, $n > 0$.

The cost function can be expressed easily, with $\mathbf{U}_{\mathbf{p},h} = ([A_i^n, Q_i^n]^\top) \in \mathbb{R}^{\mathcal{N}}$,

$$J(\mathbf{p}) = \frac{1}{2}\sum_{\nu=0}^{\mathcal{V}_t}\sum_{j=0}^{\mathcal{V}_z}\left\{(w^A_{j,\nu})^2 \ \left(Ad^\nu_j - A^{\phi_t(\nu)}_{\phi_z(j)}\right)^2 + (w^Q_{j,\nu})^2 \ \left(Qd^\nu_j - Q^{\phi_t(\nu)}_{\phi_z(j)}\right)^2\right\} \ . \tag{A.30}$$

Parameterization

It is important to define the parameterization properly, see [31], and a sensitivity analysis can provide some information on this matter. In this section, we consider a general parameterization. The parameters are only supposed to satisfy some constraints determined by some physiological considerations, such that the parameters $\mathbf{p}$ belong to a subset C of $\mathbb{R}^{\mathcal{P}}$.

A.3.2 Optimization : adjoint state approach

In order to perform correctly the optimization of the cost function (A.29) with a descent method, one needs to compute the gradient of the cost function in a fast and accurate way. It is known that the finite difference method is not efficient, see [13, 11]. One could rely on a *direct* computation of the jacobian of the forward map F, with a cost that is proportional to the number $\mathcal{P}$ of parameters. We decided to compute the gradient using an adjoint state approach that we present briefly here. As in our numerical experiments, see section A.4, the number of parameters to identify is very low, we could have fruitfully taken the former approach. However, we preferred the adjoint method, that allows to identify, in a second step, more parameters, with a cost that remains independent of the dimension of the parameter space.

To compute the gradient of a function $G = G(\mathbf{p}, \mathbf{V})$ which is an *explicit function* of the parameter vector $\mathbf{p}$ and the output vector $\mathbf{V} = F(\mathbf{p})$, we introduce the Lagrangian

$$\mathcal{L}(\mathbf{p}, \mathbf{U}, \lambda) = G(\mathbf{p}, M(\mathbf{U})) + < E(\mathbf{p}, \mathbf{U}), \lambda >_{\mathbb{R}^{\mathcal{N}}} \ , \tag{A.31}$$

where the Lagrange multiplier $\lambda \in \mathbb{R}^{\mathcal{N}}$ is the adjoint variable of $\mathbf{U}$.

With these notations, we can state the following proposition, (see [15]) :

Proposition 1 *Let (A.21) be a state equation and (A.25) an observation operator. Let $G(\mathbf{p}, \mathbf{V})$ and $F(\mathbf{p})$, given by (A.28), two regular enough functions. Let $\mathcal{L}(\mathbf{p}, \mathbf{U}, \lambda)$ be a Lagrangian defined by (A.31) associated with the equation (A.21).*

Then $\mathbf{p} \in C \subset \mathbb{R}^{\mathcal{P}} \mapsto G(\mathbf{p}, F(\mathbf{p})) \in \mathbb{R}$ is differentiable, and its gradient ∇G is given by the gradient equation *:*

$$< \nabla G, \delta\mathbf{p} >_{\mathbb{R}^{\mathcal{P}}} = \frac{\partial \mathcal{L}}{\partial \mathbf{p}}(\mathbf{p}, \mathbf{U}_{\mathbf{p}}, \lambda_{\mathbf{p}})\delta\mathbf{p} \quad \forall \delta\mathbf{p} \in \mathbb{R}^{\mathcal{P}} \ , \tag{A.32}$$

where

* $\mathbf{U}_{\mathbf{p}} \in \mathbb{R}^{\mathcal{N}}$ *is the solution of the* direct equation

$$E(\mathbf{p}, \mathbf{U}) = 0 \ , \tag{A.33}$$

* $\lambda_{\mathbf{p}} \in \mathbb{R}^{\mathcal{N}}$ *is the solution of the* adjoint equation

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}}(\mathbf{p}, \mathbf{U}_{\mathbf{p}}, \lambda)\delta\mathbf{U} = 0 \quad \forall \delta\mathbf{U} \in \delta\mathbb{R}^{\mathcal{N}} \ . \tag{A.34}$$

In this context, $\mathbf{U}_{\mathbf{p}}$ is called the direct state*, and $\lambda_{\mathbf{p}}$ the* adjoint state*.*

From the formula (A.34), the adjoint equation yields :

$$\left\{ \frac{\partial E}{\partial \mathbf{U}}(\mathbf{p}, \mathbf{U}) \right\}^{\top} \lambda + M'(\mathbf{U})^{\top} \nabla_{\mathbf{V}} G(\mathbf{p}, \mathbf{V}) = 0. \tag{A.35}$$

This presentation has the advantage of synthetizing the different applications according to the choice on function $G(\mathbf{p}, \mathbf{V})$:

- if $G(\mathbf{p}, \mathbf{V}) =< \mathbf{V}, e_i >_{\mathbb{R}^{\mathcal{M}}}$ where e_i is a basis vector of the measure space, then :
$$\nabla G = F'(\mathbf{p})^{\top} e_i,$$
and the adjoint approach enables a line-wise computation of the jacobian of the function $F(\mathbf{p})$.
- if $G(\mathbf{p}, \mathbf{V}) =< \mathbf{V}, g_v >_{\mathbb{R}^{\mathcal{M}}}$ where g_v is a given vector in $\mathbb{R}^{\mathcal{M}}$, then :
$$\nabla G = F'(\mathbf{p})^{\top} g_v.$$
- if $G(\mathbf{p}, \mathbf{V}) = \frac{1}{2}\|Z_w - \mathbf{V}\|^2$, i.e. is equal to the cost function $J(\mathbf{p})$, then :
$$\nabla G = \nabla J.$$

A.3.3 Gradient of the discrete problem

We compute in this section the gradient of the discrete problem (A.20) using the adjoint approach. First, we define the Lagrangian in section A.3.3, then we differentiate the Lagrangian with respect to the state variable to obtain the adjoint problem in section A.3.3. Finally, the gradient is computed by differentiation with respect to the parameters in section A.3.3.

Lagrangian of the discrete problem

We define the Lagrangian of our particular discrete 1-d blood flow model (A.20). Two sets of Lagrange multipliers are introduced. The first ones are associated with the boundary conditions (A.19) at $z = 0$ and $z = L$:

$$\mu_{zi,h} = (\mu^n_{zi,h})_{n=1,\dots,N_t} \in (\mathbb{R}^2)^{N_t} \ , \quad i = 0, L, \tag{A.36}$$

and the second ones are associated with the equations on the internal nodes (A.16) and live in the space of the test functions :

$$\lambda_h = (\lambda^n_h)_{n=1,\dots,N_t} \in (\mathbf{\Psi}_{h,0})^{N_t} = (\mathbb{R}^2)^{N_t \times N} \ . \tag{A.37}$$

The Lagrangian associated with the forward discrete problem (A.20) and to the generic function $G = G(\mathbf{p}, \mathbf{V})$ is hence :

$$\begin{aligned}
\mathcal{L}(\mathbf{p}, \mathbf{U}, \lambda_h, \mu_{z0,h}, \mu_{zL,h}) \ = \ & G(\mathbf{p}, M(\mathbf{U}_h)) \\
& + \sum_{n=0}^{N_t-1} \left[(\mathbf{U}^{n+1}_h, \lambda^{n+1}_h)_\Omega - (\mathbf{U}^n_h, \lambda^{n+1}_h)_\Omega - a^n_h(\mathbf{U}^n_h, \lambda^{n+1}_h; \mathbf{p}) \right] \\
& + \sum_{n=0}^{N_t-1} \sum_{i=0,L} \left[\Theta_i(\mathbf{U}^n_h, \mathbf{p}) \mathbf{U}^{n+1}_{h,i} - \mathbf{T}_i(\mathbf{U}^n_h, \mathbf{p}) \right] \cdot \mu^{n+1}_{zi,h} \ .
\end{aligned} \tag{A.38}$$

where $\mathbf{U}_{h,L} = \mathbf{U}_{h,N+1} = \mathbf{U}_h(t, L)$.

Discrete adjoint problem

We differentiate the Lagrangian (A.38) with respect to the state $\mathbf{U}_h$, remarking that $\delta\mathbf{U}^0_h = 0$, and omitting for the sake of conciseness the dependency to the parameters $\mathbf{p}$, to obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \mathbf{U}_h}(\mathbf{p}_h, \mathbf{U}_h, \lambda_h, \mu_{z0,h}, \mu_{zL,h}) \delta\mathbf{U}_h \quad = \quad & \left(\frac{\partial M}{\partial \mathbf{U}_h}(\mathbf{U}_h) \right)^\top \left(\frac{\partial G}{\partial \mathbf{V}} \right)^\top \cdot \delta\mathbf{U}_h \\
& + \sum_{n=1}^{N_t-1} \left[(\delta\mathbf{U}^{n+1}_h, \lambda^{n+1}_h)_\Omega - (\delta\mathbf{U}^n_h, \lambda^{n+1}_h)_\Omega - \frac{\partial a^n_h}{\partial \mathbf{U}_h}(\mathbf{U}^n_h, \delta\mathbf{U}^n_h, \lambda^{n+1}_h) \right] \\
& + \sum_{n=1}^{N_t-1} \sum_{i=0,L} \left[\Theta_i(\mathbf{U}^n_h) \, \delta\mathbf{U}^{n+1}_{h,i} + \frac{\partial \Theta_i}{\partial \mathbf{U}_h}(\mathbf{U}^n_h) \, (\delta\mathbf{U}^n_h, \mathbf{U}^{n+1}_{h,i}) - \frac{\partial \mathbf{T}_i}{\partial \mathbf{U}_h}(\mathbf{U}^n_h) \, \delta\mathbf{U}^n_h \right] \cdot \mu^{n+1}_{zi,h} \\
& + (\delta\mathbf{U}^1_h, \lambda^1_h)_\Omega + \sum_{i=0,L} \left[\Theta_i(\mathbf{U}^0_h) \, \delta\mathbf{U}^1_{h,i} \right] \cdot \mu^1_{zi,h} \\
& \qquad = 0 \qquad \forall \delta\mathbf{U}_h \in (\mathbf{\Psi}_h)^{N_t} \ .
\end{aligned} \tag{A.39}$$

In equation (A.39), we formulated explicitely the dependency of the term $\frac{\partial a^n_h}{\partial \mathbf{U}_h}$ with respect to $\mathbf{U}_h$ and $\delta\mathbf{U}_h$, that derives from the nonlinearity of the

operator a_h^n. A discrete integration by parts in time to order the terms in function of $\delta\mathbf{U}_h^n$ yields

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\mathbf{U}_h}(\mathbf{p}_h,\mathbf{U}_h,\lambda_h,\mu_{z0,h},\mu_{zL,h})\delta\mathbf{U}_h &= [\nabla_{\mathbf{U}_h}M(\mathbf{U}_h)\,\nabla_{\mathbf{V}}G]\cdot\delta\mathbf{U}_h\\ &+(\delta\mathbf{U}_h^{N_t},\lambda_h^{N_t})_\Omega+\sum_{i=0,L}\left[\Theta_i(\mathbf{U}_h^{N_t-1})\,\delta\mathbf{U}_{h,i}^{N_t}\right]\cdot\mu_{zi,h}^{N_t}\\ &+\sum_{n=1}^{N_t-1}\left[(\delta\mathbf{U}_h^n,\lambda_h^n)_\Omega-(\delta\mathbf{U}_h^n,\lambda_h^{n+1})_\Omega-\frac{\partial a_h^n}{\partial\mathbf{U}_h}(\mathbf{U}_h,\delta\mathbf{U}_h^n,\lambda_h^{n+1})\right]\\ &+\sum_{n=1}^{N_t-1}\sum_{i=0,L}\left[\Theta_i(\mathbf{U}_h^{n-1})\,\delta\mathbf{U}_{h,i}^n\right]\cdot\mu_{zi,h}^n\\ &+\sum_{n=1}^{N_t-1}\sum_{i=0,L}\left[\frac{\partial\Theta_i}{\partial\mathbf{U}_h}(\mathbf{U}_h^n)\,(\delta\mathbf{U}_h^n,\mathbf{U}_{h,i}^{n+1})-\frac{\partial\mathbf{T}_i}{\partial\mathbf{U}_h}(\mathbf{U}_h^n)\,\delta\mathbf{U}_h^n\right]\cdot\mu_{zi,h}^{n+1}\\ &=0\qquad\forall\delta\mathbf{U}_h\in(\boldsymbol{\Psi}_h)^{N_t}\ .\end{aligned}\tag{A.40}$$

First, taking care of the final time t^{N_t}, i.e. $\delta\mathbf{U}_h=\delta\mathbf{U}_h^{N_t}$, we take successively internal test functions $\delta\mathbf{U}_h^{N_t}=\boldsymbol{\psi}_h\in\boldsymbol{\Psi}_{h,0}$ and boundary test functions $\delta\mathbf{U}_h^{N_t}=\boldsymbol{\psi}_{h,i}\in\boldsymbol{\Psi}_h,\ i=0,L$, where $\boldsymbol{\psi}_{h,0}(0)=\boldsymbol{\psi}_{h,L}(L)=1$. Thus we have to solve the two following problems at time t^{N_t} :

Find $\lambda_h^{N_t}\in\boldsymbol{\Psi}_{h,0}$ such that, for all $\boldsymbol{\psi}_h\in\boldsymbol{\Psi}_{h,0}$

$$(\lambda_h^{N_t},\boldsymbol{\psi}_h)_\Omega = -\left[\nabla_{\mathbf{U}_h}M(\mathbf{U}_h^{N_t})\,\nabla_{\mathbf{V}}G\right]\cdot\boldsymbol{\psi}_h\ ,\tag{A.41}$$

and then find $\mu_{zi,h}^{N_t}\in\mathbb{R}^2, i=0,L$ such that

$$\Theta_i^\top(\mathbf{U}_h^{N_t-1})\mu_{zi,h}^{N_t} = -\left[\nabla_{\mathbf{U}_h}M(\mathbf{U}_h^{N_t})\,\nabla_{\mathbf{V}}G\right]\cdot\boldsymbol{\psi}_{h,i}-(\lambda_h^{N_t},\boldsymbol{\psi}_{h,i})_\Omega.\tag{A.42}$$

Next, we take care successively of time steps $n, n=N_t-1,N_t-2,\ldots,1$. We use as test functions $\delta\mathbf{U}_h^n=\boldsymbol{\psi}_h\in\boldsymbol{\Psi}_{h,0}$ and $\delta\mathbf{U}_h^n=\boldsymbol{\psi}_{h,i},\ i=0,L$. Thus we have to solve the two following problems at time t^n :

Find $\lambda_h^n\in\boldsymbol{\Psi}_{h,0}$ such that, for all $\boldsymbol{\psi}_h\in\boldsymbol{\Psi}_{h,0}$

$$\begin{aligned}(\lambda_h^n,\boldsymbol{\psi}_h)_\Omega &= -[\nabla_{\mathbf{U}_h}M(\mathbf{U}_h^n)\,\nabla_{\mathbf{V}}G]\cdot\boldsymbol{\psi}_h+(\lambda_h^{n+1},\boldsymbol{\psi}_h)_\Omega+\frac{\partial a_h^n}{\partial\mathbf{U}_h}(\mathbf{U}_h^n,\boldsymbol{\psi}_h,\lambda_h^{n+1})\\ &-\sum_{i=0,L}\left[\frac{\partial\Theta_i}{\partial\mathbf{U}_h}(\mathbf{U}_h^n)\,(\boldsymbol{\psi}_h,\mathbf{U}_{h,i}^{n+1})-\frac{\partial\mathbf{T}_i}{\partial\mathbf{U}_h}(\mathbf{U}_h^n)\,\boldsymbol{\psi}_h\right]\cdot\mu_{zi,h}^{n+1}\ ,\end{aligned}\tag{A.43}$$

and then find $\mu^n_{zi,h} \in \mathbb{R}^2, i = 0, L$ such that

$$\begin{aligned}\Theta_i^\top(\mathbf{U}_h^{n-1})\mu^n_{zi,h} &= -[\nabla_{\mathbf{U}_h} M(\mathbf{U}_h^n)\ \nabla_{\mathbf{V}} G]\cdot \boldsymbol{\psi}_{h,i} \\ &\quad -(\lambda_h^n, \boldsymbol{\psi}_{h,i})_\Omega + (\lambda_h^{n+1}, \boldsymbol{\psi}_{h,i})_\Omega + \frac{\partial a_h^n}{\partial \mathbf{U}_h}(\mathbf{U}_h^n, \boldsymbol{\psi}_{h,i}, \lambda_h^{n+1}) \\ &\quad - \sum_{j=0,L}\left[\frac{\partial \Theta_j}{\partial \mathbf{U}_h}(\mathbf{U}_h^n)\,(\boldsymbol{\psi}_{h,i}, \mathbf{U}_{h,j}^{n+1}) - \frac{\partial \mathbf{T}_j}{\partial \mathbf{U}_h}(\mathbf{U}_h^n)\,\boldsymbol{\psi}_{h,i}\right]\cdot \mu^{n+1}_{zj,h}\ .\end{aligned} \tag{A.44}$$

The equations (A.41)-(A.42) and (A.43)-(A.44) consitute the *adjoint* problem associated with the direct problem (A.20). Some details concerning the derivatives that appeared in the equations (A.43)-(A.44), in particular a way to express analytically $\dfrac{\partial a_h^n}{\partial \mathbf{U}_h}$, will be given in a forecoming article.

Several remarks can be made.

Remark 3 First, one can note that equations at final time t^{N_t+1} (A.41)--(A.42) are deduced from the general equations (A.43)-(A.44), with the assumption that $\lambda^{N_t+1} \equiv 0$ and $\mu^{N_t+1}_{zi,h} \equiv 0,\ i = 0, L$.

Second, the adjoint problem is, as expected, linear in $(\lambda_h, \mu_{z0,h}, \mu_{zL,h})$ and backward in time. It requires the full knowledge of $\mathbf{U}_h$ at all instants and nodes. As our problem is 1-d, we can afford to store this information, and no special technique is required to solve the adjoint problem.

Third, the adjoint problem has the same form as the direct problem (A.20). It is essentially a 1-d linear hyperbolic problem, but as we used a Taylor-Galerkin discretization for the direct problem that contains (second order) parabolic terms, the formulation (A.43)-(A.44) contains also nontrivial second order terms. The generic function G (or the cost function J) introduces a source term in the adjoint equation.

Fourth, the equations on the boundaries can be solved at each time step n, after the computation of the internal Lagrange multiplier λ_h^n. They involve the inversion of two 2×2 matrices that are the transpose of the ones in the direct systems (A.19).

Discrete gradient equation

At this stage, to compute the gradient, one has simply to apply the proposition 1. The equation (A.32) yields for a $\delta\mathbf{p} \in \mathbb{R}^{\mathcal{P}}$

$$\begin{aligned}< \nabla G, \delta\mathbf{p} >_{\mathbb{R}^{\mathcal{P}}} &= \nabla_{\mathbf{p}} G(\mathbf{p}, M(\mathbf{U}_h))\cdot\delta\mathbf{p} + \sum_{n=0}^{N_t-1} -\frac{\partial a_h^n}{\partial \mathbf{p}}(\mathbf{U}_h^n, \lambda_h^{n+1}; \mathbf{p})\cdot\delta\mathbf{p} \\ &\quad + \sum_{n=0}^{N_t-1}\sum_{i=0,L}\left[\frac{\partial \Theta_i}{\partial \mathbf{p}}(\mathbf{U}_h^n, \mathbf{p})\,\mathbf{U}_{h,i}^{n+1} - \frac{\partial \mathbf{T}_i}{\partial \mathbf{p}}(\mathbf{U}_h^n, \mathbf{p})\right]\ \mu^{n+1}_{zi,h}\cdot\delta\mathbf{p}\ .\end{aligned} \tag{A.45}$$

A.4 Numerical results

In this section, some first numerical results obtained with the parameter estimation method presented in the previous sections are shown. In this study, only synthetic data were used for the optimization and a single constant parameter was estimated : in this section, the parameter vector is defined as $\mathbf{p} = \beta \in \mathbb{R}$, where β is the compliance given in the equation (A.6), such that $\boldsymbol{\beta} = (2\rho A_0 \beta, 1/2)$. In order to simplify the optimization for these preliminary tests, β was supposed constant in space.

After validating the computation of the discrete cost function gradient, the estimation method was tested using different data and different amounts of data, in order to evaluate the sensitivity of the method with respect to the information given.

A.4.1 Computing the $3d$ synthetic data

The data used for the parameter estimation was computed from two different 3D fluid structure interaction models.

The first simulation was made using a shell model for the structure [53, 12].

For the second simulation, the model used is implemented in the `lifev` code, [48]. A linear Venant Kirchoff model was used for the structure. The fluid was modeled by the incompressible Navier-Stokes equations for a Newtonian fluid. The interaction algorithm uses an exact Jacobian preconditioner, [40].

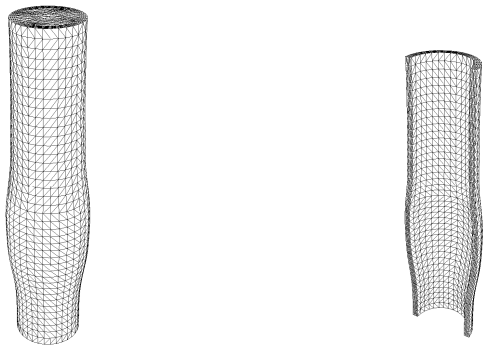


FIG. A.1 – Left : fluid mesh. Right : structure mesh.

The domain used for both simulations is a cylindrical tube of length $L_{3d} = 5$ along the z axis, with a circular basis of radius at rest $R_0 = 0.5$. A relatively coarse mesh was chosen for the simulation, as shown at Figure A.1 : the mesh is made of 30720 tetrahedra for the structural mesh and 68160 for the fluid mesh. The initial mesh is unstructured in the Oxy plane, but all

nodes are contained in one of the 41 sections at altitude $z = i \times L_{3d}/40,\ i = 0, 1, \ldots, 40$.

The wall density ρ_w was taken equal to 1.2 and the Poisson coefficient ν_w equal to 0.3. Two different values of the Young modulus E and the wall thickness at rest $h_{0,w}$ were used. For the first simulation, $E_1 = 3.E6$ and $h^1_{0,w} = 0.05$, whereas for the second one $E_2 = 4.E6$ and $h^2_{0,w} = 0.1$. The density of the fluid chosen was $\rho_f = 1.0$ and its viscosity $\nu_f = 0.03$.

The temporal discretization used was quite fine ($dt = 1E-4$) in order to obtain a sufficient amount of data for the optimization. The number of timesteps computed was 99 for the first simulation and 79 for the second one. The simulations were stopped once the pressure wave arrived at the end of the tube, in order to avoid unphysiological reflexion.

The data obtained from each simulation was post-processed in order to have, at each time step and at each section of the tube, the values of the area Ad, the flux Qd and the pressure Pd of the blood flow (see the notations in section A.3.1).

Figures A.2 and A.3 show two examples at the same instant of a 3D model solution for two different Young modulus (the second is computed using a much coarser mesh, for visualizations purposes only).

One notes the effect when one increases the Young modulus : the pressure

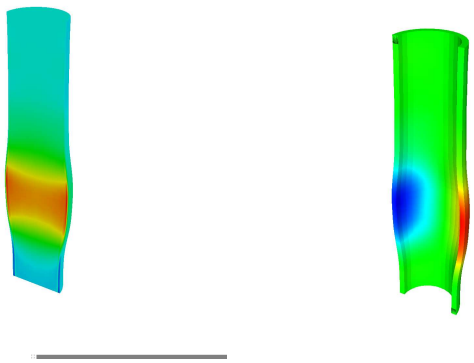


FIG. A.2 – Solution of the 3D model at time $t = 50E-4$ for a Young modulus $E_1 = 4E6$. Left : fluid velocity along the z axis. Right : displacement of the structure along y axis.

wave is extended in space and decreased in amplitude.

A.4.2 Solving the 1-d-model

In order to find the optimal value of a parameter, the algorithm does a series of simulations of the direct 1-d model.

The parameter estimated here was $\beta (= \frac{\beta_0}{2\rho A_0})$; all other parameters were kept constant : $\beta_1 = 1/2$, $\alpha = 1$, $K_r = 8\pi\nu \approx 0.75$, $\rho = 1$, $A_0 = \pi R_0^2 \approx 0.78$.

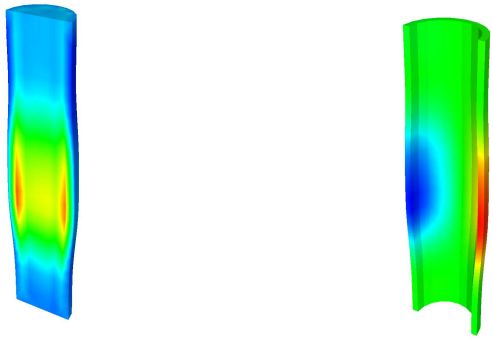


FIG. A.3 – Solution of the 3D model at time $t = 50E{-}4$ for a Young modulus $E_1 = 6E6$. Left : fluid velocity along the z axis. Right : displacement of the structure along y axis.

The 1-d domain, denoted by Ω, was chosen to be slightly shorter than the cylindrical tube used for the 3D simulations : $\Omega = [0.5, 4.5]$, and its length was therefore $L = 4$. This was done in order to avoid the problems arising from the difference in the type of the boundary conditions between the 3D and the 1-d models.

The 1-d computations were made with a spatial discretization of $N = 128$ elements of constant length $L/N = 4/128$, and a time step $dt = 1.e-5$, that is ten times smaller than the 3D time step. These values ensure the Courant condition to be always respected.

A.4.3 Measurement operator and cost function

For each test case, a series of estimations of β was done using different amounts of data. This means that the measurement operator used for the optimization was changed, and so was, consequently, the expression of the cost function J.

As explained in section A.3.1, the measurement operator is a weighted sampling of the direct solution expressed by (A.26). The weights used here were $w^A \equiv 1$ and $w^Q \equiv 0.1$ for all measures. They were chosen to compensate the orders of magnitude of the area ($\mathcal{O}(1)$) and the flux ($\mathcal{O}(10)$) in the vessel. Three different space samplings were chosen : the first one uses 33 measures taken at $\zeta_j = 0.5 + j \times (4/32)$, for $j = 0, \ldots, 32$. The second one uses 3 measures, at $\zeta_1 = 0.5$, $\zeta_2 = 2.5$ and $\zeta_3 = 4.5$, and the last one only two, at $\zeta_1 = 0.5$ and $\zeta_2 = 4.5$. The time sampling consists in using the data at either all the instants available or only at every ten of them. For the first simulation, that means the measures are taken at $\theta_\nu = \nu \times dt$ for $\nu = 0, \ldots, 99$, or at $\theta_\nu = 10\nu \times dt$ for $\nu = 0, \ldots, 9$. For the second simulation, they are taken at $\theta_\nu = \nu \times dt$ for $\nu = 0, \ldots, 79$ or at $\theta_\nu = 10\nu \times dt$ for $\nu = 0, \ldots, 7$.

Let us note that the cost function minimized during each estimation is normalized. Its expression is the following (cf. (A.30)) :

$$J(\beta) = \frac{\sum_{\nu=0}^{\mathcal{V}_t} \sum_{j=0}^{\mathcal{V}_z} \left\{ \left(Ad_j^\nu - A_{\phi_z(j)}^{\phi_t(\nu)} \right)^2 + (0.1)^2 \left(Qd_j^\nu - Q_{\phi_z(j)}^{\phi_t(\nu)} \right)^2 \right\}}{\sum_{\nu=0}^{\mathcal{V}_t} \sum_{j=0}^{\mathcal{V}_z} \left\{ (Ad_j^\nu)^2 + (0.1)^2 (Qd_j^\nu)^2 \right\}} .$$

A.4.4 Validating the computation of the cost function gradient

In order to validate the computation of the cost function gradient using the method presented here, its value was compared with the value obtained with the finite difference method. This method consists in computing an approximation of the derivative of the cost function $J(\beta)$ with respect to β, that is :

$$\frac{\partial J}{\partial \beta}(\beta) = \lim_{h \to 0} \frac{J(\beta + h) - J(\beta)}{h} \tag{A.46}$$

The cost function gradient was computed in some test cases with both methods and the results, made in double precision, showed 6 to 7 identical numbers.

A.4.5 Expected results of the estimation

Let us recall the role of the parameter β. As exposed in section 2, to derive the 1-d blood model from the 3D one, a wall displacement law (cf. (A.4)) is used, that introduces two positive real parameters β_0 and β_1. In these tests, the power coefficient β_1 is taken equal to $1/2$ and the law is reformulated into the expression (A.6), that only involves one positive parameter β.

A linear elastic law provides the following expression for β :

$$\beta_{3d} = \frac{\sqrt{\pi} h_{0,w} E}{2\rho A_0 (1 - \nu_w^2)}, \tag{A.47}$$

where E, $h_{0,w}$ and ν_w are the wall Young modulus, the wall thickness at rest and the Poisson coefficient *used to compute the 3D data*.

The optimal value of β, minimizing the difference between the results of the 1-d and 3D models, is therefore expected to be β_{3d}. It will be shown that this is not the case : the results obtained computing the 1-d model with the estimated optimal value, β_{estim}, fit better to the 3d data than the results obtained with β_{3d}. The accuracy of the estimation algorithm is therefore not to be evaluated from the accuracy on β, but from the accuracy on the error between the 1-d results and the data.

A.4.6 Test case 1 : shell structure, $E = 3.E6$ and $h_{0,w} = 0.05$

The first test was made using the 3D data obtained with a shell / Navier-Stokes coupling. The structure has a Young modulus equal to $3.E6$ and a wall thickness equal to 0.05. The number of iterations needed for the optimization varied between 6 and 10 for all the different estimations.

Results of the optimization

The results obtained are shown in tables A.1 and A.2 : for each estimation, the optimal value β_{estim} is given, as well as the corresponding value of the minimized cost function J and the value $1 - \sqrt{J}$, that gives the order of explained data.
In order to better compare the precision of the different estimations, we also give the value of a general cost function, called "fine cost function", and denoted by J_{fine}. This function J_{fine} computes the least squares error between the 3D data measured at each of the 99 timesteps and the 33 sections, and the 1-d data obtained with the estimated value β_{estim}.

Nb z_{meas}	33	3	2
β_{estim}	$1.3E5$	$1.32E5$	$1.41E5$
J	0.048	0.027	$1.6E-5$
$1-\sqrt{J}$	78.1%	83.4%	99.6%
J_{fine}	0.048	0.048	0.052
$1-\sqrt{J_{fine}}$	78.1%	78%	77.2%

TAB. A.1 – Results using 99 time measurements and 33, 3 or 2 spatial measurements for a 3D data such that : $E = 3E6$ and $h_{0,w} = 0.05$.

Nb z_{meas}	33	3	2
β_{estim}	$1.32E5$	$1.33E5$	$1.41E5$
J	0.038	0.02	$3E-6$
$1-\sqrt{J}$	80.6%	85.9%	99.8%
J_{fine}	0.048	0.048	0.052
$1-\sqrt{J_{fine}}$	78.0%	78.0%	77.2%

TAB. A.2 – Results using 9 time measurements and 33, 3 or 2 spatial measurements for a 3D data such that : $E = 3E6$ and $h_{0,w} = 0.05$.

We can see that, logically, the most accurate estimation is obtained in the case where the biggest amount of 3D data is used for the optimization. But the results do not degenerate too much when the number of time or space measurements is reduced. In fact, the value of β estimated from only 9 time and 2 space measurements, provides results that explain $77,2\%$ of the data.

The results are actually quite stable with respect to the amount of data used. Particularly, when the number of time measures varies from 99 to 9, the estimated value of β remains almost constant. On the other hand, we notice that the value of β_{estim} is more sensitive to the variation of the number of space measures. But this sensitivity remains relatively low, the maximal variation of β_{estim} being of 10%, when the amount of space measurements changes from 33 to 2.

Results of the 1-d model using the estimated value β_{estim}

For this test case, the expected value of β, according to relation (A.47), is $\beta_{3d} = 1.87E5$. In order to compare this value with the results of the different estimations, the 1-d model has also been computed using β_{3d}. The percentage of explained data obtained with this simulation, measured with the function $1-\sqrt{J_{fine}}$, is of 64%, which is far bellow the one obtained with the different estimations made (see tables A.1 and A.2).

We show next the results of the 1-d model computed using the estimated value $\beta_{estim} = 1.30E5$, obtained from 99 time and 33 space measurements of the 3D data. The results are compared with the 3D data and with the 1-d results computed using β_{3d}. Note that there is a difference of about 30% between the values β_{estim} and β_{3d}. Figures A.4 and A.5 represent the area of the vessel and the flux of the blood flow in the whole domain, after 50 and 90 time steps. Figure A.6 shows the time evolution of the area and the flux at the middle of the tube, that is at $z = L/2$.

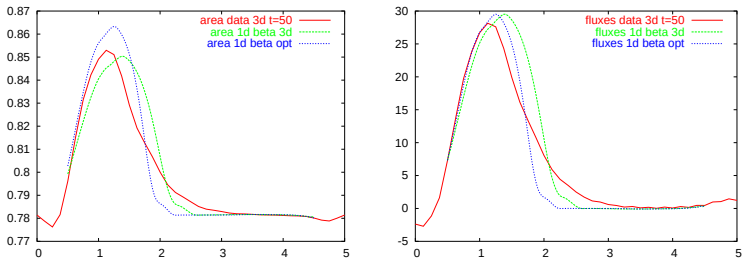


FIG. A.4 – Area and flux obtained from the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), after 50 timesteps.

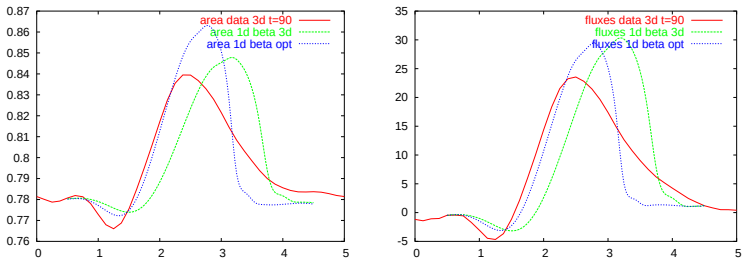


FIG. A.5 – Area and flux obtained from the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), after 90 timesteps.

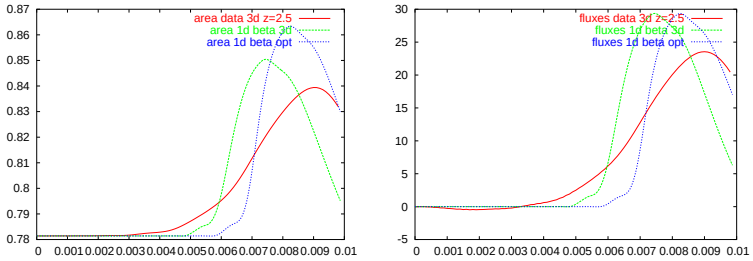


FIG. A.6 – Area and flux obtained from the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), at $z = L/2$.

As an additional information, the pressure of the blood flow was computed from the 1-d results according to the wall displacement law (A.6). It is displayed after 50 and 90 time steps in Figure A.7.

On these figures, we observe that the results of the 1-d model manage to capture quite well the phase of the waves of the 3D data, but not its shape and amplitude. Particularly, the waves of the 3D data are significantly larger. This can be explained by one of the assumptions made for the derivation of the 1-d model, that considers the vessel as a sequence of independent rings. On the contrary, the 3D model used to compute the data takes into account the propagation in the structure. On the other hand, the difference in the amplitude can be due to the more important diffusivity of the 3D models.

In addition, we can see that the results obtained with the estimated optimal value β_{estim} are closer to the 3D data than those computed using the expected value β_{3d}. This becomes particularly true after 90 timesteps and at the middle of the tube, where the waves obtained with β_{3d} are clearly in advance with respect to the data. On the contrary, the phase error of the waves obtained with β_{estim} remains relatively small.

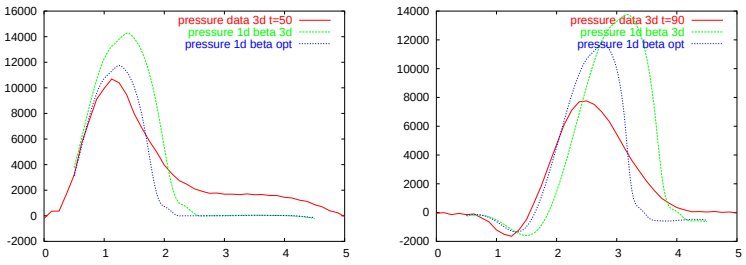


FIG. A.7 – Pressure obtained from the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line). Left : after 50 timesteps. Right : after 90 timesteps.

This result has to be emphasized, because it proves the existence of a more optimal value of β than the one given by (A.47), a relation resulting from physical considerations.

Sensitivity of the estimation with respect to the space measurements

Next, the optimal values β_{estim} obtained from different amounts of space measurements are compared, the number of time measurements being constant and equal to 99. The results of the 1-d model computed with these values are shown in figures (A.8) and (A.9). The comparison between the values estimated from 9 time measures is not made because, as deduced from tables A.1 and A.2, there is almost no difference.

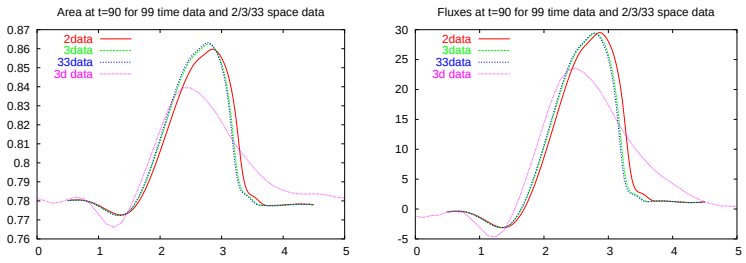


FIG. A.8 – Results of the 1-d model using the value β_{estim} obtained from 99 time measurements and 33 (blue dotted line), 3 (green dashed line) or 2 (red line) space measurements, compared to the 3D data (mauve small-dotted line), after 90 timesteps.

We can see that the optimal values of β estimated from 33 and 3 space measures provide almost the same 1-d results. But when changing to 2 space measurements, the phase error of the 1-d results grows significantly, so that

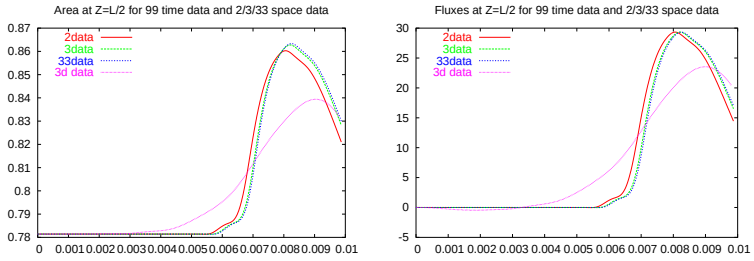


FIG. A.9 – Results of the 1-d model using the value β_{estim} obtained from 99 time measurements and 33 (blue dotted line), 3 (green dashed line) or 2 (red line) space measurements, compared to the 3D data (mauve small-dotted line), at $z = L/2$.

the advance of the area and flux waves with respect to the data increases largely. This means that with 2 space measurements, the estimation is not yet stable.

Sensitivity of the estimation with respect to the time measurements

The comparison is now made between the optimal values of β obtained from 99 and 9 time measurements, the number of space measurements being constant and equal to 2, which is the minimum. The results of the 1-d model computed with these values are shown in figures (A.10) and (A.11).

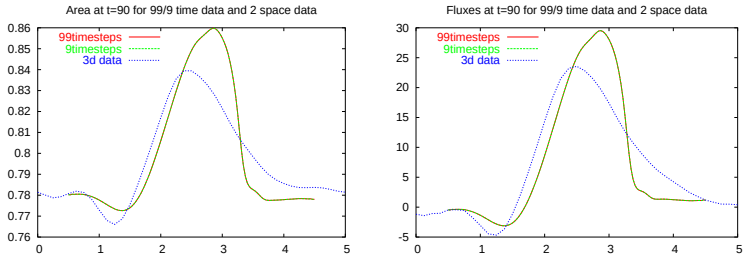


FIG. A.10 – Results of the 1-d model using the value β_{estim} obtained from 2 space measurements and 99 (blue dotted line) or 9 (red line) time measurements, compared to the 3D data (mauve small-dotted line), after 90 timesteps. The two curves with 9 and 99 are almost identical.

We observe that, in both cases, the curves are perfectly superposed. The quantity of explained data is thus already stable at 9 time measures.

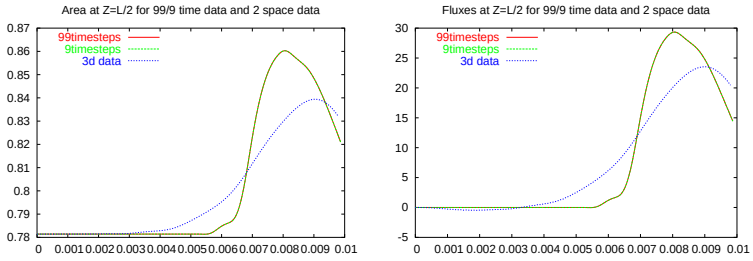


FIG. A.11 – Results of the 1-d model using the value β_{estim} obtained from 2 space measurements and 99 (blue dotted line) or 9 (red line) time measurements, compared to the 3D data (mauve small-dotted line), at $z = L/2$. The two curves with 9 and 99 are almost identical.

A.4.7 Test case 2 : 3D structure, $E = 4.E6$ and $h_{0,w} = 0.1$

The second test was made using the 3D data obtained with a 3D structure model with a Young modulus of $4.E6$ and a wall thickness of 0.1. Here again, the number of iterations necessary for the optimization was bounded by 6 and 10.

Results of the optimization

The results obtained are shown in tables A.3 and A.4. The fine cost function computes here the least squares error between the 3D data measured at each of the 79 timesteps and the 33 sections, and the 1-d data obtained with the estimated value β_{estim}.

Nb z_{meas}	33	3	2
β_{estim}	$3.13E5$	$3.20E5$	$3.67E5$
J	0.013	0.0049	$1.4E-5$
$1-\sqrt{J}$	88.8%	93%	99.6%
J_{fine}	*idem*	0.013	0.025
$1-\sqrt{J_{fine}}$	*idem*	88.6%	84.1%

TAB. A.3 – Results using 79 time measurements and 33, 3 or 2 spatial measurements for a 3D data such that : $E = 4E6$ and $h_{0,w} = 0.1$.

As expected, the most accurate estimation is obtained with the biggest amount of 3D data, that is 79 time and 33 space measurements. The percentage of explained data, measured with the function $1-\sqrt{J_{fine}}$, is in this case of 88.8%. But in the case where only 7 time and 2 space measurements are used, this percentage is of 87.4%, which is not much lower. Here again, the

Nb z_{meas}	33	3	2
$\beta_{0,estim}$	$5.64E5$	$5.74E5$	$5.98E5$
β_{estim}	$3.19E5$	$3.24E5$	$3.38E5$
J	0.0087	0.0017	$1.1E-5$
$1-\sqrt{J}$	90.7%	95.9%	99.7%
J_{fine}	0.013	0.013	0.016
$1-\sqrt{J_{fine}}$	88.7%	88.4%	87.4%

TAB. A.4 – Results using 7 time measurements and 33, 3 or 2 spatial measurements for a 3D data such that : $E = 4E6$ and $h_{0,w} = 0.1$.

accuracy of the estimation does not decrease significantly with the amount of data.

But the results of the estimation are less stable than in the first case. Particularly, the variation of the estimated value when the time sampling changes is quite important. Indeed, when only 2 space measurements are considered, the relative variation of β_{estim} reaches 17% when the amount of time measurements is decreased from 79 to 7.

Results of the 1-d model using the estimated value β_{estim}

For this test case, the expected value of β, according to the relation (A.47), is $\beta_{3d} = 4.96E5$. This value provides 1-d results that explain 71% of the data, which is again far bellow the percentage obtained with the different estimations made (see tables A.3 and A.4). Note that the maximal difference between the estimated value β_{estim} and β_{3d} is of about 38%.

Next, we show the results of the 1-d model, computed using the estimation $\beta_{estim} = 3.13E5$, which provides the highest value of $1-\sqrt{J}$. They are compared with the 3D data and with the 1-d results computed using β_{3d}. Figures A.12 and A.13 represent the area of the vessel and the flux the blood flow in the whole domain, after 50 and 70 time steps. Figure A.14 shows the time evolution of the area and the flux at the middle of the tube, that is at $z = L/2$.

Firstly, we observe that in this case the results of the 1-d model fit much better to the 3D data than in the test case one. In fact, not only the phase of the waves is captured, but their shape and amplitude is also well approached. Let us note that the 3D model used here to compute the data is linear for the structure. This is probably the reason why, in this case, the linear 1-d model is closer to the 3D one.

We can also see that the results obtained with the value β_{estim} are much better than those obtained with β_{3d}. Indeed, the waves obtained with this

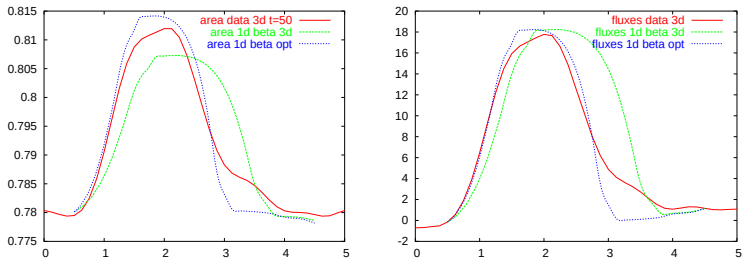


FIG. A.12 – Results of the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), after 50 timesteps.

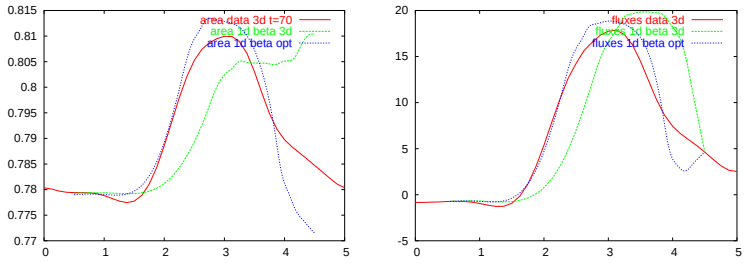


FIG. A.13 – Results of the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), after 70 timesteps.

last value are much too advanced with respect to the data, whilst the phase error of the waves obtained with β_{estim} remains relatively small, even after 70 timesteps.

Sensitivity of the estimation with respect to the space measurements

The optimal values β_{estim} obtained from different amounts of space measurements are now compared, the number of time measurements being fixed at 79. The results of the 1-d model computed with these values are shown in figures (A.15) and (A.16).

These figures show that, here again, a significant variation in the 1-d results only occures when dicreasing the amount of space measurements from 3 to 2. In that case, the phase of the waves is lost, so that the area and flux appear to be much more in advance with respect to the data. Thus, we can say that the estimation starts to be stable from 3 space measurements.

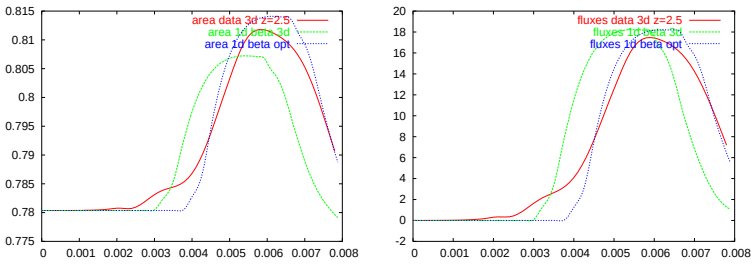


FIG. A.14 – Results of the 1-d model using β_{estim} (blue dotted line) and β_{3d} (green dashed line), compared to the 3D data (red line), at $z = L/2$.

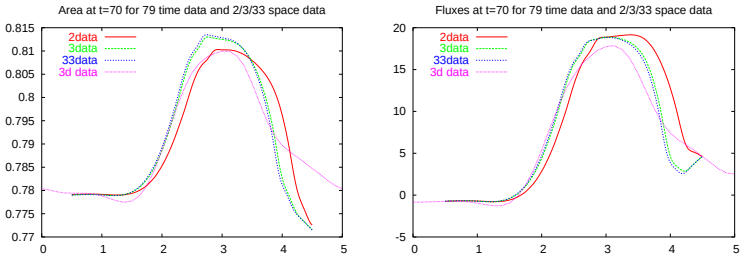


FIG. A.15 – Results of the 1-d model using the value β_{estim} obtained from 79 time measurements and 33 (blue dotted line), 3 (green dashed line) or 2 (red line) space measurements, compared to the 3D data (mauve small-dotted line), after 70 timesteps.

Sensitivity of the estimation with respect to the time measurements

Here, the comparison is made between the estimated values of β obtained using 79 and 7 time measurements. The amount of space measurements is first fixed at 33 and later at 2. The results of the first comparison are shown in figures A.17 and A.18. For the second one, the results are shown in figures A.19 and A.20.

From the observation of these figures, we can say that, when 33 space measurements are used, the estimation is stable with respect to the amount of time measurements. On the contrary, when the data is only considered at 2 points in space, the estimation is unstable and varies significantly with the amount of time measurements provided.

These numerical tests show that it is possible to identify the parameter β in a stable way from 3 measures in space and 7 measures in time (in the first test case already from 2 measures in space and 9 in time). Moreover, the estimations obtained are more optimal than the expected value β_{3d}, as they provide 1-d results that explain a bigger percentage of data.

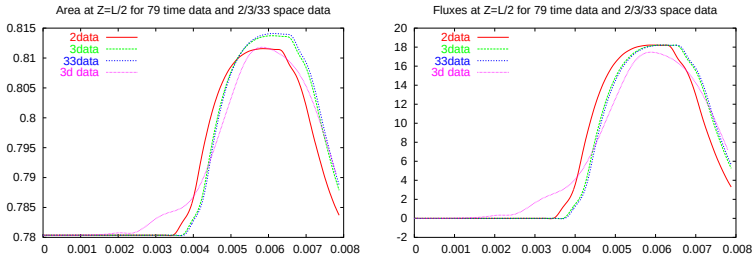


FIG. A.16 – Results of the 1-d model using the value β_{estim} obtained from 79 time measurements and 33 (blue dotted line), 3 (green dashed line) or 2 (red line) space measurements, compared to the 3D data (mauve small-dotted line), at $z = L/2$.

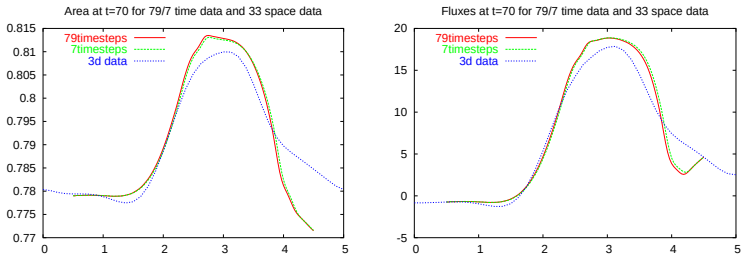


FIG. A.17 – Results of the 1-d model using the value β_{estim} obtained from 33 space measurements and 79 (blue dotted line) or 7 (red line) time measurements, compared to the 3D data (mauve small-dotted line), after 70 timesteps.

A.5 Conclusions

In this work, the parameter identification of a 1-d blood flow model using the adjoint approach was studied. Starting from the discretization of the model with a second order Taylor-Galerkin scheme, the discrete adjoint problem was derived. We obtained a linear 1-d hyperbolic system, with non-standard discretization and boundary conditions. The resolution of the adjoint problem allows to compute the gradient of a cost function, which evaluates the least squares error between the measured data and the 1-d results. The analytical discrete gradient is necessary to minimize the cost function in order to find the optimal values of the parameters. Its computation was therefore implemented and validated using the finite difference method. Some preliminary numerical tests were performed. A single parameter was estimated – the arterial compliance β – which is the most important parameter from a biomedical point of view. In order to simplify the estimation, β was assumed to be constant in space. The optimization was made using data provided by two different 3D models, and for two different values of the Young modulus

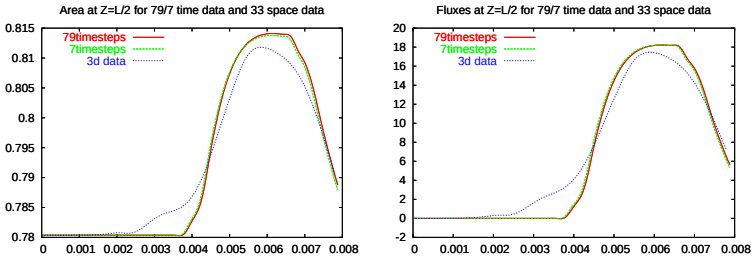


FIG. A.18 – Results of the 1-d model using the value β_{estim} obtained from 33 space measurements and 79 (blue dotted line) or 7 (red line) time measurements, compared to the 3D data (mauve small-dotted line), at $z = L/2$.

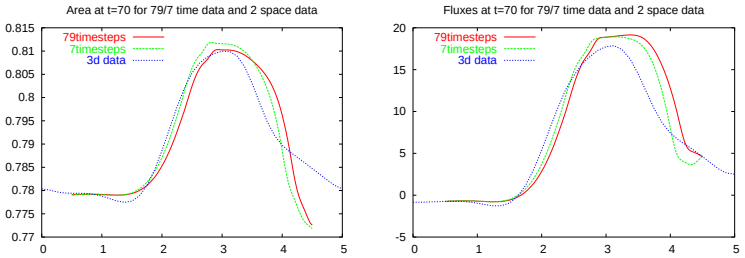


FIG. A.19 – Results of the 1-d model using the value β_{estim} obtained from 2 space measurements and 79 (blue dotted line) or 7 (red line) time measurements, compared to the 3D data (mauve small-dotted line), after 70 timesteps.

and the wall thickness at rest. In all cases, the estimated values of β obtained were very different from the expected value β_{3d}[1], the difference going from 30 to 38%. However, the 1-d results computed with the estimated values appeared to be significantly closer to the data than those obtained with β_{3d}. Actually, the percentage of explained data obtained with these estimations was of about 75% to 89%, depending on the test case and the amount of data used for the optimization (this is not surprising as in the first case the data were computed using a shell model with a non-linear constitutive law). Thus, it is possible to retrieve the pressure of the blood flow in an artery in a non-intrusive way.

The interest of this work first lies in the application of a parameter estimation method to a 1-d blood flow model, using a rigorous mathematical approach. Second, the method used for the optimization is efficient. Indeed, once the adjoint state has been derived from the discrete 1-d model, the estimation can quite simply be extended to more parameters and especially

[1] given by an analytical formula and dependent on the values of the wall Young modulus E, the wall thickness at rest $h_{0,w}$ and the Poisson coefficient ν_w.

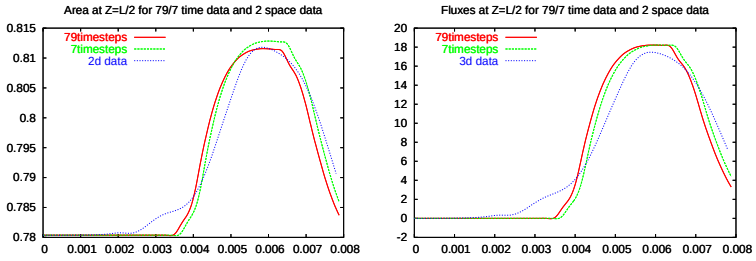


FIG. A.20 – Results of the 1-d model using the value β_{estim} obtained from 2 space measurements and 79 (blue dotted line) or 7 (red line) time measurements, compared to the 3D data (mauve small-dotted line), at $z = L/2$.

to parameters depending on space.

This work is a first step in the parameter identification for the 1-d blood flow model. The tests were for the moment limited to the estimation of a single parameter, assumed to be constant in space and time. The optimization should now be extended to other parameters, possibly depending on space. After a sensitivity analysis, it would be interesting to estimate the area at rest A_0 at each discretization point of the 1-d domain, because this could allow to localize a possible tapering in the vessel. Another relevant application of the method would be the parameter estimation involved in a bifurcation.

Bibliographie

[1] Abgrall R., Status of MURD schemes and Applications in Aeronautics. *AIAA paper* 2000-2328, June 2000.

[2] Abgrall R., Toward the Ultimate Conservative Scheme : Following the Quest. *J. Comput. Phys.* 167(2) : 277-315, 2001.

[3] Abgrall R., Barth T., Residual Distribution Schemes for the Conservation laws via Adaptative Quadrature. *SIAM Journal on Scientific Computing* 24(3) : 732-769, 2002.

[4] Ambrosi, D and Quartapelle L. A Taylor-Galerkin method for simulating nonlinear dispersive water waves. *J. Comput. Phys.* 146(2) : 546-569, 1998.

[5] Aris R., *Vectors, Tensors, and the Basic Equations of fluid Mechanics.* Dover Publications Inc., 1989.

[6] Arnold D.N., Brezzi F., Fortin M., A stable finite element for the Stokes equations. *Calcolo* **21** : 337–344, 1984.

[7] Audusse E., *Modélisation hyperbolique et analyse numérique pour les écoulements en eaux peu profondes.* Thèse de l'Université P. et M. Curie (Paris 6), 2004.

[8] Audusse E., A multilayer Saint-Venant model. *Discrete Cont. Dyn. Syst. Ser. B* 5 No 2 : 189-214, 2005.

[9] Audusse E., Bristeau M.O., Decoene A., 3D free surface flows simulations using a multilayer Saint-Venant model. Comparisons with Navier-Stokes solutions. *Proceedings of the Enumath Conf.*, Santiago, Spain, July 2005, (submitted).

[10] Cahouet J., Chabard J.-P., Some fast 3D finite elements for the generalized Stokes problem. *Int. Jou. Num. Meth. Fluids* **8** : 869-895, 1988.

[11] Cacuci D.G., Ionescu-Bujor M., A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Systems - I : Deterministic Methods. *Nuclear Science and Engineering* 147(3) : 189–203, 2004.

[12] Chapelle D. and Bathe K. J., *The finite element analysis of shell - Fundamentals.* Springer-Verlag, 2004.

[13] Chavent G., *Identification of distributed parameter systems : about the output least square method, its implementation and identifiability*, In R. Isermann, editor, Identification and system parameter identification (Proc. Fifth IFAC Symposium on Identification, Darmstadt), p.85–97, Pergamon Press, Amsterdam, 1979.

[14] Chavent G., On the theory of practice of nonlinear least-squares. Parameter identification in ground water flow, transport, and related processes, Part I. *Adv. in Water Res.* 14(2) : 55-63, 1991.

[15] Chavent G., *Non linear Least Squares for Inverse problems*, 2005 (in preparation).

[16] Crouzeix M. et Raviart P.-A., Conforming and nonconforming finite element methods for solving the stationnary Stokes equations. *R.A.I.R.O., Anal. Numér.* **3** : 33–75, 1973.

[17] Beckman A. and Haidvogel D., Numerical simulation of flow around a tall isolated seamount. *J. Phys. Oceanogr.* 23 : 1736-1753, 1993.

[18] Bell M.J., *Vortex stretching and bottom torque in the Bryan-Cox ocean circulation model.* Applications T.N. 17, The Met. Office, UK.

[19] Blumberg A.F. and Mellor G.L., *A description of a three-dimensional coastal ocean circulation model.* In N. Heaps (ed.), Three-dimensional Coastal Ocean Models, Vol.4, American Geophysical Union, Washington D.C., 1-16, 1987.

[20] Brezzi F. and Fortin M., *Mixed and hybrid finite element methods.* Springer Series in Computational Mathematics Vol. 15, Springer Verlag, New York, 1991.

[21] Bryan K., A numerical model for the study of the circulation of the world oceans. *J. Comput. Phys.* 4 : 347-376, 1969.

[22] Burchard H. and Petersen O., Hybridization between σ- and z-coordinates for improving the internal pressure gradient calculation in marine models with steep bottom slopes . *Int. J. Numer. Methods Fluids* 25 : 1003-1023, 1997.

[23] Cairncross R.A., Shunk P.R., Baer T.A., Rao R.R., Sackinger P.A., A finite element method for free surface flows of incompressible fluids in three dimensions, Part I : Boundary fitted mesh motion. *Int. J. Numer. Methods Fluids* 33 : 375-403, 2000.

[24] Casulli V., Cheng R., Gartner J., Tidal, Residual, Intertidal Mudflat (TRIM) model and its applications to the San Francisco Bay, California. *Estuarine, Coastal, and Shelf Science* 36 : 235-280, 1993.

[25] Casulli V., Cattani E., Stability, accuracy and efficiency of a semi-implicit method for three-dimensional shallow water flow. *Computer Math. Applic.* 27(4) : 99-112. 1994.

[26] Casulli V., Stelling G., Simulation of three-dimensional, non-hydrostatic free surface flows for estuary and coastal seas, *Estuarine and Coastal Modeling, Proceedings of the* 4-*th International Conference*, San Diego. M.Spaulding and R. Cheng, Eds., ASCE, New York, 1-12, 1995.

[27] Casulli V., Walters R.A., An unstructured, three-dimensional model based on the shallow water equations. *Int. J. Numer. Methods Fluids* 32 : 331-348, 2000.

[28] Casulli V., Zanolli P., Semi-Implicit Numerical Modelling of Non-Hydrostatic Free-Surface Flows for Environmental Problems. *Mathematical and Computer Modelling* 36 : 1131-1141, 2002.

[29] Céa J., Approximation variationnelle des problèmes aux limites. *Ann. Inst. Fourier (Grenoble)* **9** : 77–84, 1975.

[30] Clément P., Approximation by finite element functions using local regularization. *R.A.I.R.O., Anal. Numer.* **14** : 345–444, 1964.

[31] Clément F., Khvoenkova N., Cartalade A. and Montarnal Ph. Analyse de sensibilité et estimation de paramètres de transport pour une équation de diffusion, approche par état adjoint. Report 5132, Inria, Rocquencourt, France, 2004.

[32] Decoene A., Equations de Navier-Stokes à surface libre en dimension 3 avec hypothèse hydrostatique. Rapport EDF R&D-LNHE, HP-75/04/020/A. 2004.

[33] Deardoff J.A., A numerical study of three-dimensional turbulent channel flow at large Reynolds number. *Journal of Fluid Mechanics* 41(part 2) : 453-480, 1970.

[34] Deconinck H., Paillère H., Struijs R., and Roe P.L., Multidimensional Upwind schemes based on fluctuation splitting for systems of conservation laws. *Computational Mechanics* 11 : 323-340, 1993.

[35] Deconinck H. and Paillère H., Multidimensional Upwind Residual Distribution Schemes for the 2D Euler Equations. In Problemes non lineaires appliques. Schemas distribues en mecanique des fluides et applications. Support de cours Ecoles CEA-EDF-INRIA, 25-28 avril 2000.

[36] Deleersnijder E. and Beckers J.M., On the use of the σ coordinate system in regions of large bathymetric variations. *J. Marine Syst.* 3 : 381-390, 1992.

[37] Deleersnijder E. and Ruddick K.G., A generalized vertical coordinate for 3D marine models. *Bull. Soc. R. Sci. Liège* 61(6) : 498-502, 1992.

[38] Delft Hydraulics, Delft-3D-Flow manual. Delft Hydraulics, Release 3.05, 1999.

[39] Engelman M.S., Sani R.L., Gresho P.M., The implementation of normal and/or tangential velocity boundary conditions in finite element codes

for incompressible fluid flow. *Int. J. Numer. Methods Fluids* 2 (3) : 225-238, 1982.

[40] Fernández M. and Moubachir M., A Newton method using exact jacobians for solving fluid-structure coupling. *Computers and Structures* 83(2-3) : 127-142, 2005.

[41] Flather R.A., Results from surge prediction model of the North-West European continental shelf for April, November and December 1973. Institute of Oceanographic (UK), report number 24, 1976.

[42] Fontana F., Miglio E., Quarteroni A. and Saleri F., A Finite Element Method for 3D Hydrostatic Water Flow. *Computing and Visualization in Science* 2(2-3) : 85–93, 1999.

[43] Formaggia L., Nobile F., A Stability Analysis for the Arbitrary Lagrangian Eulerian Formulation with Finite Elements. *East-West Journal of Numerical Mathematics* 7 :105-132, 1999.

[44] Formaggia, L., Gerbeau J.-F., Nobile F. and Quarteroni A., On the coupling of 3D and 1D Navier-Stokes equations for flow problems in compliant vessels. *Comput. Methods Appl. Mech. Engrg.* 191(6-7) : 561-582, 2001.

[45] Formaggia L. and Veneziani A., Reduced and multiscale models for the human cardiovascular system. Mox, Politecnico di Milano, Report 21, 2003.

[46] Formaggia L., Lamponi D. and Quarteroni A., One-dimensional models for blood flow in arteries. Mathematical modelling of the cardiovascular system. *J. Engrg. Math.* 47(3-4) : 251-276, 2003.

[47] Formaggia L. and Nobile F., Stability Analysis of second order time accurate the Arbitrary Lagrangian Eulerian Formulation with Finite Elements. *Comput. Methods Appl. Mech. Engrg* 193 : 4097-4116, 2004.

[48] Formaggia L., Gerbeau J.F. and Prud'homme C., LifeV Developer Manual. Technical Report 2004. Software and documentation available electronically, (`http ://www.lifev.org/documentation/lifev-dev.pdf`).

[49] Fortin M., An analysis of the convergence of mixed finite element methods. *R.A.I.R.O, Anal. Numér.* **11** : 341–354, 1977.

[50] Galland J.-C., Goutal N., Hervouet J.-M., TELEMAC : A new numerical model for solving shallow water equations. *Adv. Water Resources* 14(3) : 138-148, 1991.

[51] Gary J.M., Estimate of truncation error in transformed coordinate primitive equation atmospheric models. *J. Atmos. Sci.* 30 : 223-233, 1973.

[52] Gerbeau J.-F., Lelièvre T., Le Bris C., Simulations of MHD flows with moving boundaries. *Comput. Methods Appl. Mech. Eng.* 190(11-12) : 14671-1482, 2000.

[53] Gerbeau J.-F. and Vidrascu M., A quasi-Newton algorithm based on a reduced model for fluid-structure interaction problems in blood flows. *M2AN Math. Model. Numer. Anal.* 37(4) : 631-647, 2003.

[54] Gerdes R., A primitive equation ocean circulation model using a general vertical coordinate transformation. *Journal of Geophysical Research* 98, 14 : 683-701, 1993.

[55] van Gijzen M.B., Large-scale of finite element computation with GMRES-like methods on a CRAY Y-MP. *Appl. Num. Math.* 19 : 51-62, 1995.

[56] Goutal N., *Résolution des équations de Saint-Venant en régime transcritique par une méthode d'éléments finis. Application aux bancs découvrants.* Thèse de doctorat de l'Université Paris VI, 1987.

[57] Gueuzaine P., Grandmont C., Farhat C. Design and analysis of ALE schemes with provable second-order time accuracy for inviscid and viscous flow simulations. *Journal of Computational Physics* 191 : 206-227, 2003.

[58] Guillard H., Farhat C., On the significance of the geometric conservation law for flow computations on moving boundaries. *Comput. Methods Appl. Mech. Eng.* 190(11-12) : 14671-1482, 2000.

[59] Hall M.G., *Cell-vertex multigrid schemes for the solution of the Euler Equations.* Numerical Methods for Fluid Mechanics, II (K.W. Morton and M.J. Baines, eds.). Oxford University Press, 1986.

[60] Haney T.L., On the pressure gradient force over steep topography in sigma coordinate ocean models. *J. Phys. Oceanogr.* 21 : 610-619, 1991.

[61] Harlow F., Welch J., Numerical calculation of time-dependent viscus incompressible Navier-Stokes equations. *The Physics of Fluids* 8(12) : 2182-2189, 1965.

[62] Heaps N., *Three-dimensional coastal ocean models.* Vol 4 of Coastal and estuarine sciences, American Geophysical Union, Washington D.C., 1987.

[63] Hervouet J.-M. and van Haren L., Recent advances in numerical methods for fluid flows. Tech. Rep. HE-43/97/015/B, Electricité de France (EDF-DER), Département Laboratoire National d'Hydraulique, 1997.

[64] Hervouet J.-M., Atkinson J.H., Westerink J.J., Similarities Between the Wave Equation and the Quasi-Bubble Solutions to the Shallow Water Equation. *International Journal for Numerical Methods in Fluids* 45 : 689-714, 2004.

[65] Hervouet J-M., On spurious oscillations in primitive shallow water equations. XIII International Conference on Computational Methods in Water Resources. Calgary, Canada. 25-29 Juin 2000

[66] Hervouet J.-M., *Hydrodynamique des écoulements à surface libre. Modélisation numérique avec la méthode des éléments finis.* Presses de l'école nationale des Ponts et Chaussées, 2003.

[67] Hirt C.W., Amsden A.A., Cook J.L., An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *J. Comput. Phys.* 14(3) : 227-253, 1974.

[68] Hirt C.W., Nichols B., Volume of Fluid (VOF) Method for the dynamics of free boundaries. *J. Comput. Phys.* 39 : 201-225, 1981.

[69] Hodges B.R., *Numerical Simulation of nonlinear free-surface waves on a turbulent open-chanel flow.* PhD Diss., Dept. Civil Engineering. Stanford University, 235 p, 1997.

[70] Huerta A., Liu W.K., Viscous flow with large surface motion. *Comput. Methods Appl. Mech. Engrg.* 69 : 277-324, 1988.

[71] Hughes T.J.R. and Brooks A., A multi-dimensional upwind scheme with no cross-wind diffusion. *Finite Element Methods for Convection Dominated Flows, ASME* 34, 1979.

[72] Hughes T.J.R. and Mallet M., A New Finite Element Formulation for Computational Fluid Dynamics : III. The Generalized Streamline Operator For Advective-Diffusive Systems. *Computer Methods in Applied Mechanics and Engineering* 58 : 305-328, 1986.

[73] Ibler B., *Résolution des équations de Navier-Stokes par une méthode d'éléments finis.* Thèse de doctorat de l'Université Paris-Sud, janvier 1981.

[74] Janin J.M., Conservativité et positivité dans un module de transport de scalaire écrit en éléments finis. Application à Telemac-3D. Rapport EDF R&D-LNHE, 1996.

[75] Jankowski J.A., *A non-hydrostatic model for free surface flows.* PhD thesis, Universität Hannover, Institut für Strömungsmechanik und ERiB, 1998.

[76] Janjic Z.I., Pressure gradient force and advection scheme used for forecasting with steep and small scale topography. *Contrib. Atmos. Phys.* 50 : 186-199, 1977.

[77] Kocyigit M., Falconer R., Lin B., Three-dimensional numerical modeling of free surface flows with non-hydrostatic pressure. *Int. J. Numer. Methods Fluids* 40 : 1145-1162, 2002.

[78] Lagrée P.-Y., An inverse technique to deduce the elasticity of a large artery. *Eur. Phys. J./AP* 9 : 153-163, 2000.

[79] Laporte E. and Le Tallec P., *Numerical methods in sensitivity analysis and shape optimization.* Modeling and Simulation in Science, Engineering, & Technology, Birkhäuser Boston Inc., Boston, MA, 2003. With 1 CD-ROM (Unix, Macintosh).

[80] Lesoinne M., Farhat C., Geometric conservation laws for flow problems with moving boundaries and deformable meshes and their impact on aeroelastic computations. *Comput. Methods Appl. Mech. Eng.* 134 : 71-90, 1996.

[81] Luettich R.A., Westerink J.J., Scheffner N.W., ADCIRC, an advanced three-dimensional circulation model for shelves, coasts and estuaries. Dept. of the Army, U.S. Army Corps of Engineers, Washington D.C., August 1991.

[82] Mahadevan A., Oliger J., Street R., A non-hydrostatic meso-scale ocean model. *Journal of Physical Oceanography* 26 : 1881-1900, 1996.

[83] Malcherek A., *Mathematische modellierung von Strömungen und Stofftransportprozessen in Ästuaren.* PhD thesis, Universität Hannover, Institut für Strömungsmechanik und ERiB, 1995.

[84] Malcherek A., Markofsky M. , Zielke W., Peltier E., Le Norman C., Teisson C., Cornelisse J., Molinaro P., Corti S. and Grego G., Three-dimensional numerical modelling of cohesive sediment transport processes in estuarine environments. Final Report to the EC Contract MAS2-CT92-0013, Institut für Strömungsmechanik und ERiB, Universität Hannover, EDF, Delft Hydraulics, ENEL, 1996.

[85] Marcos F., Janin J.M., Nouveaux développements dans l'étape de convection-diffusion de Telemac-3D. Rapport EDF R&D-LNHE, HE-42/94/025/A. 1997.

[86] Maronnier V., *Simulation numérique des écoulements de fluides incompressibles avec surface libre.* Thèse de l'Ecole Polytechnique Fédérale de Lausanne, 2000.

[87] Martin V., Clément F., Decoene A., Gerbeau J.-F., Parameter identification for a one-dimensional blood flow model. *ESAIM : Proceedings*, September 2005 vol. 14, 174-200. Eric Cancès and Jean-Frédéric Gerbeau, Editors.

[88] Martins F., Leitao P., Silva A., Neves R., 3D modelling in the Sado estuary using a new generic vertical discretization approach. *Oceanologica Acta* 24, 1-2, 2001.

[89] Maury B., Characteristics ALE Method for the Unsteady 3D Navier-Stokes Equations with a Free Surface. *International Journal of Computational Fluid Dynamics* 6 : 175-188,1996.

[90] McCalpin J.D., A comparison f second-order and fourth-order pressure gradient algorithm in a sigma-co-ordinate model. *Int. J. Numer. Methods Fluids* 18 : 361-383, 1994.

[91] Mellor G.L., Ezer T. and Oey L.-Y., The pressure gradient conundrum of sigma coordinate ocean models. *J. Atmosph. Oceanic. Technol.* 11 : 1126-1134, 1994.

[92] Mellor G.L., Oey L.-Y. and Ezer T., Sigma Coordinate Pressure Gradient and the Seamount Problem. *J. Atmosph. Oceanic. Technol.* 15 : 1122-1131, 1997.

[93] Mellor G., Häkkinen S., Ezer T. and Patchen R. *A Generalization of a Sigma Coordinate Ocean Model and an Intercomparison of Model Vertical Grids.* Ocean Forecasting : Conceptual Basis ans Applications. N. Pinardi, J. Woods (Eds.), Springer Verlag, Berlin, 55-72, 2002.

[94] Mewis P., Holz K.-P., A quasi-bubble function approach for shallow water waves. *Advances in Hydro-Sciences and Engineering.* Vol. 1, 1993.

[95] Migliavacca F., Pennati G., Dubini G., Fumero R., Pietrabissa R., Hsia T.-Y. and de Leval M.R., Computational model of the fluid dynamics in systemic-to-pulmonary shunts. *J. Biomech.* 33 (5) : 549–557, 2000.

[96] Migliavacca F., Pennati G., Dubini G., Fumero R., Pietrabissa R., Urcelay G., Bove E., Hsia T.-Y. and de Leval M.R., Modeling of the Norwood circulation : effects of shunt size, vascular resistances, and heart rate. *Am. J. Physiol. Heart. Circ. Physiol.* 280 : H2076–H2086, 2001.

[97] Miglio E., Quarteroni A. and Saleri F., Finite Element Approximation of Quasi-3D Shallow Water Equations. *Comput. Methods Appl. Mech. Engrg* **174**(3-4) : 355–369, 1999.

[98] Miglio E., Saleri F., An existence and uniqueness result for the 3D Shallow Water Equations. Internal Report n. **421/*P***, Dipartimiento di Matematica Francesco Brioschi, Politecnico di Milano, 2000.

[99] Monaghan J.J., Simulating free surface flows with SPH. *Journal of Computational Physics* 110 : 399-406, 1994.

[100] Morton K.W, Crumpton P.I., Mackenzie J.A., Cell-vertex methods for inviscid and viscous flows. *Journal of Computers and Fluids* 22(2/3) : 91-102, 1993.

[101] Muzaferija S., Peric M., *Computation of free surface flows using interface-tracking and interface-capturing methods.* Mahrenholtz and Markiewicz eds. Nonlinear Water Wave Interaction, Chap. 2, Computational Mechanics Publications, Southampton, 1998.

[102] Nakayama T., Mori. M., An Eulerian finite element method for time-dependent free surface problems in hydrodynamics. *Int. J. Numer. Methods Fluids* 22 : 175-194, 1996.

[103] Namin M, Lin B., Falconer R., An implicit numerical algorithm for solving non-hydrostatic free-surface flow problems. *Int. J. Numer. Methods Fluids* 35 : 341-356, 2001.

[104] Navier C., *Mémoire sur les lois du mouvement des fluides.* Mémoires de l'Académie des Sciences, 6, 380-398, 1823.

[105] Ni R.-H., A multiple Grid Scheme for Solving the Euler Equations. *AIAA Journal* 20 : 1565-1571, 1981.

[106] Paillère H., *Multidimensional Upwind Residual Distribution Schemes for the Euler and Navier-Stokes Equations on Unstructured Grids.* Thèse de l'Université Libre de Bruxelles, 1995.

[107] Paillère H. and Deconinck H., Compact Vertex Convection Schemes on unstructured Meshes. In Problemes non lineaires appliques. Schemas distribues en mecanique des fluides et applications. Support de cours Ecoles CEA-EDF-INRIA, 25-28 avril 2000.

[108] Paillère H., Deconinck H., van der Weide, Upwind Residual Distributive methods for compressible flows : an alternative to FV and FE methods. VKI LS 1997-02, Computational Fluid Dynamics, Belgium, March 3-7, 1997.

[109] Parès C., Approximation d'un modèle de turbulence par une méthode de Lagrange-Galerkin. *Rev. mat. Apl* 15 : 63-124, 1995.

[110] Perthame B. and Gerbeau J.-F., Derivation of Viscous Saint-Venant System for Laminar Shallow Water ; Numerical Validation. *Discrete Cont. Dyn. Syst. Ser. B* 1 no 1, 89-102, 2001.

[111] Phillips N.A., A coordinate system having some special advantages for numerical forecasting. *J. Meteorol.* 14 : 184-185, 1957.

[112] Pironneau O., On the transport-diffusion algorithm and its applications to the Navier-Stokes equations. *Numer. Math.* 38 : 309-332, 1982.

[113] Pironneau O., *Méthodes des éléments finis pour les fluides.* Masson 1988.

[114] Powell K.G. and van Leer B., A genuinely multi-dimensional upwind cell-vertex scheme for the Euler equations. AIAA-89-0095, 1989.

[115] Quarteroni A. and Valli A., *Numerical Approximation of Partial Differential Equations*, Springer-Verlag, Berlin, 1994.

[116] Quarteroni A. and Formaggia L., *Mathematical modelling and numerical simulation of the cardiovascular system.* Handbook of numerical analysis. Vol. XII, 3–127, North-Holland, Amsterdam, 2004.

[117] Raviart P.-A., Thomas J.-M., A mixed finite element method for S2-nd oder elliptic problems. in Galligani I., Magenes E. eds., Mathematical Aspects of Finite Element Methods, Lecture Notes in Maths. 606, 292-315, Springer Berlin, 1977.

[118] Reynolds O., *On the dynamical theory of incompressible viscous fluids and the determination of the criterion.* Philosophical Transactions of the Royal Society of London, A186, 1895.

[119] Roe P.L., *Fluctuations and Signal - A Framework for Numerical Evolution Problems.* Numerical Methods for Fluid Dynamics. Academic Press, 1982.

[120] Roe P.L., Optimum upwind advection on a triangular mesh. Technical Report 90-75, ICASE, 1990.

[121] Rostand Ph., Kinetic boundary layers, Numerical Simulations. Rapport de Recherche INRIA, 728, 1987.

[122] Saint-Venant A.J.C., *Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit.* Compte-Rendu à l'Académie des Sciences de Paris, 73, 147-154, 1871.

[123] Shchepetkin A.F., McWilliams J.C., The Regional Ocean Model System (ROMS) : A split-explicit, free-surface, topography-following coordinate ocean model. *Ocean Modelling* 9 : 347-404, 2005.

[124] Sherwin S. J., Formaggia L., Peiró J. and Franke V., Computational modelling of 1D blood flow with variable mechanical properties and its application to the simulation of wave propagation in the human arterial system. ECCOMAS Computational Fluid Dynamics Conference, Part II (Swansea, 2001), *Internat. J. Numer. Methods Fluids* 43(6-7) : 673-700, 2003.

[125] Slordal L.H., The pressure gradient force in sigma-co-ordinate ocean models. *Int. J. Numer. Methods Fluids* 24 : 987-1017, 1997.

[126] Smith N.P., Pullan A.J. and Hunter P.J., An anatomically based model of transient coronary blood flow in the heart. *SIAM J. Appl. Math.* 62 : 990–1018, 2001/2002.

[127] Song Y. and Haidvogel D., A Semi-implicit Ocean Circulation Model Using a Generalized Topography-Following Coordinate System. *J. of Comput. Phys.* 115 : 228-244, 1994.

[128] Song Y.T., A General Pressure Gradient Formulation for Ocean Models. Part 1. Scheme Design and Diagnostic Analysis. *Mon. Weather Rev.* 126 : 3213-3230, 1998.

[129] Soulaïmani A., Saad Y., An arbitrary Lagrangian-Eulerian finite element method for solving three-dimensional free surface flows. *Comput. Methods Appl. Mech. Engrg.* 162 : 79-106, 1998.

[130] Spiegel E., Veronis G., On the Boussinesq approximation for a compressible fluid. *Astrophys. J.* 131 : 442-447, 1960.

[131] Spurk H.S., *Fluid Mechanics.* Springer Verlag, Berlin, Germany, 1997.

[132] Stelling G.S. and Van Kester J.A.TH.M., On the approximation of horizontal gradients in sigma-coordinates for bathymetry with steep bottom slopes. *Int. J. Numer. Methods Fluids* 18 : 915-935,1994.

[133] Stokes G., *On the theories of the internal friction of fluids motion and of the equilibrium and motion of elastic solids.* Trans. Cambridge Phil. Soc., 8, 287-305, 1845.

[134] Struijs R., Deconinck H., Roe P.L., Fluctuation Splitting Schemes for the 2D Euler Equations. VKI LS 1991-01, Computational Fluid Dynamics, 1991.

[135] Struijs R., *A Multi-Dimensional Upwind Discretization Method for the Euler Equations on Unstructured Grids.* PhD thesis, The University of Delft, The Netherlands, 1994.

[136] Temam R., *Navier-Stokes equations.* North Holland, 1977.

[137] Verfürth R., Finite Element Approximations on Steady Navier-Stokes Equations with Mixed Boundary Conditions. *R.A.I.R.O., Mod. Math. Anal. Numer.* 19 : 461-475, 1984.

[138] Versteeg H., Malalasekera W., *An introduction to computational fluid dynamics. The Finite Volume method.* Longman Scientific and Technical, 1995.

[139] Violeau D., Issa R., La méthode numérique SPH appliquée à l'hydraulique. Une méthode Lagrangienn aux applications variées. *Revue Européenne des Eléments finis* 12(2-3), 2003.

[140] Wilcox D.C., *Turbulence modeling for CFD.* DCW Industries Inc., La Canada, USA, 1993.

[141] Yuan H., Wu C.H., A two-dimensional vertical non-hydrostatic σ model with an implicit method for free-surface flows. *Int. J. Numer. Meth. Fluids* 44 : 811-835, 2004.

[142] Yost S., *Three-dimensional on-hydrostatic modeling of free-surface flows and transport of cohesive sediment.* PhD thesis, University of Michigan, 1995.

[143] Zhang H., Reggio M., Trèpanier J.Y., Camarero R., Geometric form of the GCL for moving meshes and its implementation in CFD schemes. *Comput. Methods Appl. Mech. Eng.* 134 : 71-90, 1996.

[144] Zhou J.G., Stansby P.K., An arbitrary Lagrangian-Eulerian Sigma (ALES) model with non-hydrostatic pressure for shallow water. *Comput. Methods Appl. Mech. Engrg.* 178 : 199-214, 1999.