Numerical Model Building Based on XFEM/Level Set Method to Simulate Ledge Freezing/Melting in Hall-Héroult Cell

Thèse

Min LI

Doctorat en génie civil
Philosophiæ doctor (Ph.D.)

Québec, Canada

Numerical Model Building Based on XFEM/Level Set Method to Simulate Ledge Freezing/Melting in Hall-Héroult Cell

Thèse

Min LI

Sous la direction de:

Mario FAFARD, directeur de recherche
Jean-Loup ROBERT, codirecteur de recherche

Résumé

Au cours de la production de l'aluminium via le procédé de Hall-Héroult, le bain gelé, obtenu par solidification du bain électrolytique, joue un rôle significatif dans le maintien de la stabilité de la cellule d'électrolyse. L'objectif de ce travail est le développement d'un modèle numérique bidimensionnel afin de prédire le profil du bain gelé dans le système biphasé bain liquide/bain gelé, et ce, en résolvant trois problèmes physiques couplés incluant le problème de changement de phase (problème de Stefan), la variation de la composition chimique du bain et le mouvement de ce dernier. Par souci de simplification, la composition chimique du bain est supposée comme étant un système binaire. La résolution de ces trois problèmes, caractérisés par le mouvement de l'interface entre les deux phases et les discontinuités qui ont lieu à l'interface, constitue un grand défi pour les méthodes de résolution conventionnelles, basées sur le principe de la continuité des variables. En conséquence, la méthode des éléments finis étendus (XFEM) est utilisée comme alternative afin de traiter les discontinuités locales inhérentes à chaque solution tandis que la méthode de la fonction de niveaux (level-set) est exploitée pour capturer, implicitement, l'évolution de l'interface entre les deux phases.

Au cours du développement de ce modèle, les problématiques suivantes : 1) l'écoulement monophasique à densité variable 2) le problème de Stefan couplé au transport d'espèces chimiques dans un système binaire sans considération du phénomène de la convection et 3) le problème de Stefan et le mouvement du fluide qui en résulte sont investigués par le biais du couplage entre deux problèmes parmi les problèmes mentionnées ci-dessus. La pertinence et la précision de ces sous-modèles sont testées à travers des comparaisons avec des solutions analytiques ou des résultats obtenus via des méthodes numériques conventionnelles. Finalement, le modèle tenant en compte les trois physiques est appliqué à la simulation de certains scénarios de solidification/fusion du système bain liquide-bain gelé. Dans cette dernière application, le mouvement du bain, induit par la différence de densité entre les deux phases ou par la force de flottabilité due aux gradients de température et/ou de concentration, est décrit par le problème de Stokes.

Ce modèle se caractérise par le couplage entre différentes physiques, notamment la variation de la densité du fluide et de la température de fusion en fonction de la concentration des espèces chimiques. En outre, la méthode XFEM démontre sa précision et sa flexibilité pour traiter différents types de discontinuité tout en considérant un maillage fixe.

Abstract

During the Hall-Héroult process for smelting aluminium, the ledge formed by freezing the molten bath plays a significant role in maintaining the internal working condition of the cell at stable state. The present work aims at building a vertically two-dimensional numerical model to predict the ledge profile in the bath-ledge two-phase system through solving three interactive physical problems including the phase change problem (Stefan problem), the variation of bath composition and the bath motion. For the sake of simplicity, the molten bath is regarded as a binary system in chemical composition. Solving the three involved problems characterized by the free moving internal boundary and the presence of discontinuities at the free boundary is always a challenge to the conventional continuum-based methods. Therefore, as an alternative method, the extended finite element method (XFEM) is used to handle the local discontinuities in each solution space while the interface between phases is captured implicitly by the level set method.

In the course of model building, the following subjects: 1) one-phase density driven flow 2) Stefan problem without convection mechanism in the binary system 3) Stefan problem with ensuing melt flow in pure material, are investigated by coupling each two of the problems mentioned above. The accuracy of the corresponding sub-models is verified by the analytical solutions or those obtained by the conventional methods. Finally, the model by coupling three physics is applied to simulate the freezing/melting of the bath-ledge system under certain scenarios. In the final application, the bath flow is described by Stokes equations and induced either by the density jump between different phases or by the buoyancy forces produced by the temperature or/and compositional gradients.

The present model is characterized by the coupling of multiple physics, especially the liquid density and the melting point are dependent on the species concentration. XFEM also exhibits its accuracy and flexibility in dealing with different types of discontinuity based on a fixed mesh.

Contents

List of Tables

List of Figures

Nomenclature

Acronyms

ALE	Arbitrary Lagrangian-Eulerian
CPP	Closest Point Projection
CR	Cryolite Ratio
DOF	Degree Of Freedom
FEM	Finite Element Method
FESh++	Finite Element Shell in C++
GLS	Galerkin/Least-Squares
LBB	Ladyženskaja-Babuška-Brezzi
MAC	Marker-And-Cell
MHD	MagnetoHydroDynamics
NS	Navier-Stokes
PSPG	Pressure-Stabilizing/Petrov-Galerkin
SUPG	Streamline-Upwind/Petrov-Galerkin
VOF	Volume Of Fluid
XFEM	eXtended Finite Element Method

Alphabetical symbols

A	Area, m^2
a	Additional degree of freedom
$\boldsymbol{c}$	Particle velocity relative to the mesh, m/s
c	Molar concentration, mol/m^3 or mol/l
C_p	Specific heat, $J/(kg \cdot K)$
$\underline{\underline{D}}$	Shear rate tensor, $1/s$
D	Diffusion coefficient, m^2/s
e	Internal energy per unit mass, J/kg
e^{sub}	Sub-element
E_{st}^{tot}	Total energy, J
$\vec{F}$	Force, N
$\vec{f}$	Body force per unit volume, N/m^3

F	Normal speed field, m/s
f	Any function
$\vec{g}$	Gravitational acceleration, m/s^2
δh	Enthalpy of unit volume, J/m^3
h	Heat transfer coefficient, $W/(m^2 \cdot K)$
	or enthalpy per unit mass J/kg
	or height, m
$\underline{\underline{I}}$	Identity tensor
$[J]$	Jacobian matrix
$\vec{J}$	Relative mass flux, $kg/(m^2 \cdot s)$
k	Thermal conductivity, $W/(m \cdot K)$
k_p	Partition coefficient
L	Specific latent heat, J/kg
	or distance, m
$\bar{M}$	Local enrichment function
M	Local enrichment function
$\bar{N}$	Shape function
$\mathcal{N}$	Set of total nodes
$\mathcal{N}_e$	Set of enriched nodes
$\mathcal{N}_{e,expanded}$	Set of nodes belonging to cut or blending elements
$\vec{N}$	Absolute mass flux, $kg/(m^2 \cdot s)$
$\vec{n}$	Unit outward normal
$\vec{n}_\Gamma$	Unit normal at the interface
N	Shape function
p	Pressure, N/m^2
$\vec{q}$	Heat flux, W/m^2
Q	Net heat, J
$q_{n,\omega}$	Mass flux in the normal direction, W/m^2
$q_{n,T}$	Heat flux in the normal direction, W/m^2
R	Ramping function
	or residual
r	Thermal resistivity, $K \cdot m/W$
s	Thermal source or sink, W/m^3
$\vec{t}$	Traction, N/m^2
T	Temperature, ℃
t	Time, s
T_a	Ambient temperature, ℃
T_m	Melting point, ℃
T_w	Wall temperature, ℃

$\vec{u}$	Velocity, m/s
u	Any unknown
	or component of velocity in x direction, m/s
$\hat{\boldsymbol{v}}$	Mesh node velocity, m/s
$\boldsymbol{v}$	Particle velocity, m/s
$\vec{v}$	External velocity to advect ϕ, m/s
v	Component of velocity in y direction, m/s
V_I	Interface normal speed, m/s
$\boldsymbol{w}$	Particle velocity seen from the mesh frame of reference, m/s
W	Net work
	or weight coefficient
	or weak formulation
w	Component of velocity in z direction, m/s
$\boldsymbol{x}$	Laboratory coordinates, m
$\vec{x}$	Point in the real element
$\vec{x}_m$	Master point in CPP method
$\vec{x}_s$	Slave point in CPP method

Greek Symbols

α	Thermal diffusivity, m^2/s
α_p	Thermal expansion coefficient, $1/K$
β	Penalty parameter
Γ	Boundary
δd	Parameter to indicate the distance range
δ	Variational operator
	or thickness, m
Δ	Increment
κ	Curvature
λ	The second coefficient of viscosity
μ	Dynamic viscosity, $kg/(m \cdot s)$
$\vec{\xi}$	Point in the reference element
ρ	Density, kg/m^3
$\tilde{\rho}$	Real density in Boussinesq approximation, kg/m^3
$\hat{\rho}$	Mass concentration of species in the mixture, kg/m^3
$\underline{\underline{\sigma}}$	Cauchy stress tensor, N/m^2
τ	Stabilization parameter
$\underline{\underline{\tau}}$	Viscous stress tensor, N/m^2
ϕ	Level set function, m
χ	Molar fraction

$\bar{\chi}$	Volume fraction
$\boldsymbol{\chi}$	Mesh coordinates, m
ψ	Global enrichment function
Ω	Domain
ω	Mass fraction

Superscripts

$+$	Liquid side of the interface
$-$	Solid side of the interface
$consv$	Conservative form related
e	Single element
i	Iteration step
n	Time step
T	Transpose

Subscripts

0	Initial value
b	Bath
c	Contacting surface
D	Dirichlet type
f	Frozen ledge
I	Interface
i	Component i
j	Component j
l	In liquid phase
N	Neumann type
ref	Reference value
s	In solid phase
sh	Shell
t	Related to tangential matrix
x	Derivative with respect to x or component in x axis
y	Derivative with respect to y or component in y axis

Other Symbols

[ ]%	Weight percent, $wt\%$
[ ]	Matrix
$\breve{}$	Enriched part
$'$	First derivative

$''$	Second derivative
$\langle\ \rangle$	Row vector
$[\![\ ]\!]$	Jump value
$\|\ \|$	Absolute value or determinant
$\nabla\cdot$	Divergence operator
∇	Gradient operator
$\|\|\ \|\|$	Norm of vector
$\{\ \}$	Column vector

Acknowledgements

The completion of this thesis relies on the support and help from many people as well as some organizations. I hereby would like to express my thanks to each of them who have contributed their efforts in any form to my project.

Firstly, my sincere gratitude goes to my supervisor Prof. Mario Fafard not only for his expert guidance throughout my PhD project but also for his patience, understanding, encouragement and help when I got problems and felt frustrated either in my scientific research or in the foreign campus life. His kindness and generosity have created a harmonious working atmosphere, so I can be completely absorbed in my research work. Since being his student, I've been constantly nourished by his deep insight and rich experience in numerical modelling and mathematics through the countless discussion with him. His guidance and support is the cornerstone for me to accomplish this thesis. His vigour and vitality will always affect me. Thanks again for his academic, financial and emotional support.

My sincere thanks also goes to my co-supervisor Prof. Jean-Loup Robert, who is very nice to work with. His guidance in the course of supervising my research work is really important. The regular meetings on discussing my progress are always more productive due to his presence. His insightful comments based on his expertise in the computational fluid dynamics are always instructive to me. I also would like to thank my industrial advisor Dr. Donald Ziegler, who works for the company Alcoa. He gave me a lot of valuable advice from the perspective of engineering as well as a lot of detailed explanations of what happens in actual production.

What's more, I am honoured to be a member of the research group Modelling of Aluminium Cell and Electrical Energy Efficiency (MACE^3), which has been a source of friendship and communication. I am grateful to all the previous contributors to develop the in-house code FESh++, that I mainly worked with in my project. I am particularly grateful to Dr. Patrice Goulet for introducing me to FESh++. I also particularly thank Dr. Hicham Chaouki, with whom I worked closely. His presence in the regular meetings always made the discussion more smooth. He also helped me a lot in doing the relevant mathematical analysis. He together with my supervisors revised my writing of this thesis. His kindness, hard-working attitude and erudition impressed me. I also would like to thank Prof. Houshang Alamdari and Prof. Louis Gosselin for their advice and comments on my research work at some certain points in

time. The experience of working in this research group, where every one is nice and friendly, will be my good memory.

My PhD project has been funded by several organizations. I am deeply thankful to the Chinese Scholarship Council (CSC) for the first four-years' living allowance. I also would like to acknowledge the financial support of the National Sciences and Engineering Research Council of Canada (NSERC), the company Alcoa as part of the industrial research chair NSERC-Alcoa MACE^3, and the Fonds de Recherche du Québec-Nature et Technologie (FRQNT) by the intermediary of the Aluminium Research Centre-REGAL.

I also would like to thank my friends for their company and encouragement. I won't forget those days we spent together for fun and relaxation.

Last but not the least, I would like to thank my family, who are always the source of my motivation to proceed. Thank my parents for giving all their best to me and supporting me silently in all my pursuits. Thank my sisters, Chao and Rui, who are always my pride. Thanks to my wife Cuihong, without whose unquestioning faith and support this thesis was never accomplished.

Foreword

The first two chapters are mainly dedicated to introduce the background of this work, including the motivation, research progress in the literature and the relevant mathematical models and numerical methods. From Chapter 3 through Chapter 6, each chapter represents one scientific article centering on one sub-problem after factorizing the main problem. Chapter 7 presents conclusions of the present study as well as some possibilities for the future work. All the articles contained in this thesis are listed below.

- Chapter 3
  M. Li, H. Chaouki, J. -L. Robert, D. Ziegler and M. Fafard. Finite element analysis of mass transport in a binary miscible solution of variable density, International Journal of Nonlinear Sciences and Numerical Simulation, under review

- Chapter 4
  M. Li, H. Chaouki, J. -L. Robert, D. Ziegler and M. Fafard. Numerical simulation of Stefan problem coupled with mass transport through XFEM/level set method, Journal of Scientific Computing, under review

- Chapter 5
  M. Li, H. Chaouki, J. -L. Robert, D. Ziegler and M. Fafard. Numerical simulation of Stefan problem with ensuing melt flow through XFEM/level set method, Finite Elements in Analysis and Design, under review

- Chapter 6
  M. Li, H. Chaouki, J. -L. Robert, D. Ziegler and M. Fafard. Numerical simulation of Stefan problem with ensuing melt flow and mass transport in the binary system through XFEM/level set method, manuscript ready to submit

All the work involved in the manuscripts listed above were supervised closely by Prof. Fafard, Prof. Robert and Dr. Ziegler. Under their guidance as well as the help of Dr. Chaouki, I was in charge of the concrete work including the design of each article and the model building in

the in-house code FESh++. Prof. Fafard, Prof. Robert and Dr. Ziegler contributed their expert knowledge and advice in building the numerical model and in interpreting the meaning of simulation results. Dr. Chaouki contributed his mathematical knowledge in helping me to do the concrete finite element analysis. What's more, in the article presented in chapter 5, Dr. Chaouki provided me with the Comsol result for the model validation. All the manuscripts are written by me and further revised according to the comments from Dr. Chaouki, Prof. Fafard, Prof. Robert and Dr. Ziegler.

Chapter 1

Introduction

1.1 Background and motivation

The Hall-Héroult process can trace back to one hundred years ago. It is the most widely used industrial process to produce aluminium at present. Fig. 1.1 shows the simplified cross sectional view of a typical modern aluminium electrolysis cell with prebaked anodes using Hall-Héroult process.

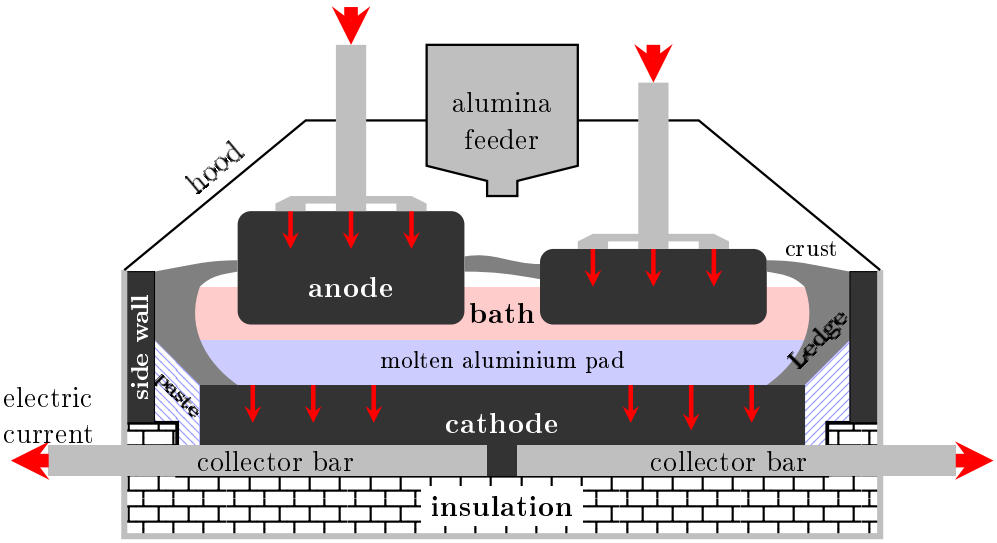


Figure 1.1 – Cross sectional view of the modern aluminium electrolysis cell with prebaked anodes using Hall-Héroult process.

The Hall-Héroult process can be briefly described as follows: alumina powder (Al_2O_3) is fed into the cell and dissolves in the molten electrolyte (also named bath) for ionization; as the electricity is supplied continuously, the overall oxidation-reduction reaction

$$2Al_2O_{3(s)} + 3C_{(s)} \longrightarrow 4Al_{(l)} + 3CO_{2(g)} \tag{1.1}$$

takes place; during this process, the carbon anode is consumed, the gas is generated on the anode surface, and the molten pure aluminium deposits on the cathode surface gradually. Refer to [24, 71] for more details.

The present work mainly focuses on the formation of the ledge that adheres to the side wall of the cell as shown in Fig.1.1. The ledge plays a crucial role in the Hall-Héroult process. Its importance can be demonstrated from the following several aspects.

- Being protective lining
  The ledge is crucial to the cell's lifespan. Inside the cell, the molten bath is highly corrosive. In practice, almost no lining materials can resist that corrosion, that means direct contact with the side walls would destroy the cell quickly. The ledge exists as a protective lining to prevent the side walls from chemical attack. It forms at the early stage of a newly-installed cell and remains during the whole left life of the cell.

- Interaction with thermal balance
  Energy efficiency is another concerned issue either from the perspective of economy or environmental protection. Aluminium production through Hall-Héroult process consumes a large amount of electrical energy, about $13 \sim 15 kWh/kg\ Al$ [8]. At present, the energy efficiency, i.e. the ratio that the unavoidable alumina reduction enthalpy required to propel the reaction 1.1 takes up in the entire consumed energy, of the best aluminium smelters is only in the order of 50% [26]. Another 50% energy is consumed by the Joule effect and lost in the form of heat given off to the surroundings. Part of the Joule heat is utilized to maintain the high temperature ($\sim$ 965℃) of the electrolyte in order to keep it being molten. In essence, the ledge formation results from the frozen of the electrolyte. This solidification phenomenon has two requirements. One is that, the bath temperature must be close to its liquidus line, typically $5 \sim 10$℃ super-heated. The other one requires sufficient heat dissipation through the side walls, then where the ledge resides on is the coldest region of the cell. In other words, the ledge formation is the result of the thermal behaviour inside the cell, and its profile also. From another perspective, the ledge serves as a buffer to balance the bath temperature. For example, if a sudden heat accumulation happens in the bath, the ledge would melt to absorb the extra heat as a response mechanism to prevent temperature increasing rapidly to some extent; and vice versa. Therefore, the ledge is important to maintain a stable temperature condition of the cell.

- Interaction with bath composition
  Table 1.1 lists a typical bath composition for modern aluminium electrolysis [8]. In general, the bath comprises Na_3AlF_6 (cryolite), additives and impurities. Impurities usually come into the bath with alumina powder. The most major component of bath is cryolite. Molten cryolite acts as solvent to dissolve alumina. The additives of bath are

Table 1.1 – A typical bath composition for modern aluminium electrolysis [8]

Bath Component	$wt\%$ (Approximately)
Na_3AlF_6	80
AlF_3	11
CaF_2	5.6
Al_2O_3	2.5
LiF	Trace

used to enhance some characters of cryolite, e.g., decreasing melting point, improving alumina's solubility, or raising current efficiency, etc. Among all the addtives, AlF_3 takes the largest portion and affects the bath composition considerably. Its existence can improve current efficiency, lower cryolite's melting point, but decrease alumina's solubility [72]. In practice, the content of AlF_3 in the bath is usually in the range $10 \sim 13wt\%$. Since the pure cryolite (Na_3AlF_6) can be regarded as 3 mol NaF plus 1 mol AlF_3, the additive portion of AlF_3 is often referred to as excess AlF_3. In industrial production, a practical parameter, named cryolite ratio and defined as $CR = \frac{\text{moles of } NaF}{\text{moles of } AlF_3}$, is usually used to indicate the bath composition. Due to the excess AlF_3, CR of the bath is usually around 2.2 in practice. Fig. 1.2 is the sketch of system $NaF - AlF_3$ phase diagram. It tells that at equilibrium state the first separated phase is cryolite [13, 37]. That means the ledge can be regarded as pure cryolite [14, 38], whose composition is 3 in terms of CR. Due to the different chemical composition between ledge and bath, the melting and freezing of ledge would change the bath composition. In addition, the phase diagram also tells that the bath composition would affect the melting point. It implies that the interaction between ledge formation and bath composition is two-way.

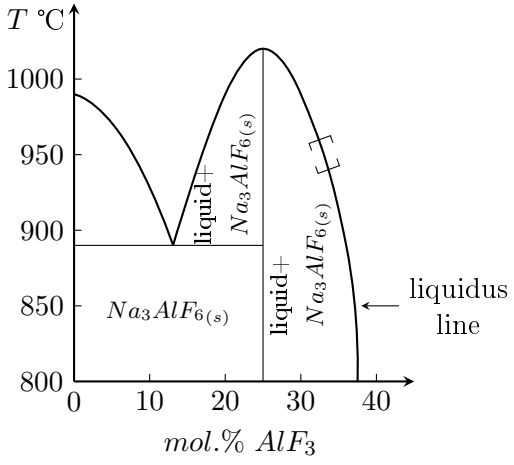


Figure 1.2 – Sketch of phase diagram of system $NaF - AlF_3$; the brackets indicate the normal operating range [13, 37].

- Interaction with bath motion

The influence of bath motion to the ledge formation or its profile is indirect. Bath motion exerts its influence by changing the thermal behaviour and the transport of species which would impact the chemical composition locally. Likewise, the ledge can influence bath motion through the interactions with thermal behaviour and bath composition, due to the fact some physical properties like density or viscosity are dependent on temperature or chemical composition. In addition, one direct impact is that the ledge profile (shape and thickness) determines the effective volume of the bath.

As mentioned above, the ledge is crucial to building a stable and optimised environment inside the cell, which is the foundation for the normal operation of an aluminium electrolysis cell. The process of ledge formation or profile variation is complicated due to the fact that it strongly interacts with other physical processes as mentioned above. To form and maintain a stable and optimised environment, it's necessary to figure out mechanisms behind and get deep insight and quantify the interrelationships of all related processes. However, the extremely harsh conditions (high temperature and strong corrosivity of the bath) make direct observation and experiments difficult. Then numerical simulations serve as a powerful alternative to study this subject. The established numerical model can help to design and operate the cell.

In essence, the ledge formation process belongs to phase change problem. The moving ledge front makes the numerical modelling fall into the category of free boundary problem, which is indeed a challenge to most numerical methods. Given this situation, the literature review in the following section is done from two perspectives. First, numerical techniques used to deal with free boundary problem is discussed. Then, the development of numerically modelling ledge profile is reviewed.

1.2 Literature review

1.2.1 Numerical techniques to deal with free boundary problems

The free boundary problems are commonly encountered in many engineering and scientific fields, such as phase change problems, free surface problems (like open-channel flow), multi-phase flow problems and fluid-structure interaction problems. The challenge of modelling the free boundary problems stem from the fact that the moving boundary is a priori unknown besides other conventional unknown variables. Extra work must be done to locate the free boundary at each time step. There are various techniques to locate the free boundary in the context of conventional numerical methods. In the literature, these methods are usually grouped into two types, namely front-tracking method and front-capturing method. By convention, the former is typically based on moving mesh while the latter on fixed grids. For details, readers are referred to some review articles [69, 70, 66, 32, 9]. Here, only three nowadays relatively popular methods, i.e. arbitrary Lagrangian-Eulerian (ALE) method, volume

of fluid(VOF) method and level set method are picked out to be briefly discussed.

The ALE mehtod is front-tracking method working with a moving mesh. It was first developed within the context of finite difference method [29], and subsequently introduced into finite element method context [33]. The ALE method can be regarded as generalization of the classical Lagrangian and Eulerian descriptions of motion. As shown in Fig. 1.3, neither like the Lagrangian description, in which the mesh nodes are completely attached to the material particles during motion, nor like the Eulerian description, in which the mesh nodes are fixed in space, ALE method allows the mesh to move with an arbitrary velocity. It combines the merits of both descriptions, i.e. it allows moving boundaries and the mesh nodes don't have to follow the material particles [12]. In ALE method, the fundamental conservation equations

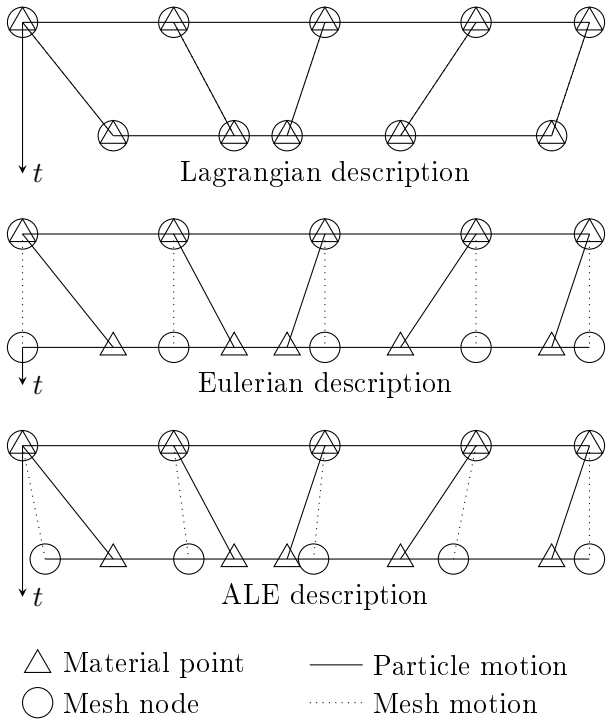


Figure 1.3 – One-dimensional example of Lagrangian, Eulerian and ALE mesh and particle motion [12].

have to be rewritten in the ALE form by replacing the particle velocity $\boldsymbol{v}$ in each convective term by the particle convective velocity $\boldsymbol{c}$, which is defined as [11]

$$\boldsymbol{c} := \boldsymbol{v} - \hat{\boldsymbol{v}} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\chi}} \cdot \boldsymbol{w} \tag{1.2}$$

where $\boldsymbol{w}$ is the particle velocity seen from the mesh frame of reference; $\boldsymbol{v}$, $\hat{\boldsymbol{v}}$ and $\boldsymbol{c}$ represent particle velocity, mesh node velocity and particle velocity relative to the mesh, respectively, and all seen from laboratory frame of reference; $\boldsymbol{x}$ denotes the laboratory coordinates; $\boldsymbol{\chi}$ denotes the mesh coordinates; $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\chi}}$ acts as mapping operator from the mesh frame of reference to the laboratory frame of reference. Usually, the moving boundary is described by the La-

grangian formulation ($\boldsymbol{c} = 0$), while the interior part far from the free boundary is described by Eulerian formulation (($\hat{\boldsymbol{v}} = 0$)). In between it's a transition zone. Due to the Lagrangian portion, ALE method has a clear delineation of free boundary. Compared with pure Lagrangian method, ALE method is able to cope with higher deformation without mesh entanglement. Nevertheless, mesh rebuilding might be required because mesh quality tends to deteriorate as mesh deformation gets large. What's more, it is hard to handle the case in which topological changes (splitting or merging) take place on the free boundary. During the past decades, ALE method has been widely used and further developed to improve its perfomance [2, 15, 44].

Both the VOF method and the level set method are front-capturing method based on a fixed mesh. Unlike the ALE method, these two methods don't manipulate the mesh but employ an additional scalar function to describe the free boundary motion implicitly. The VOF method can be regarded as evolving from the earlier marker-and-cell (MAC) method [27]. In MAC method, several marker particles are utilized to label one cell, which leads to large computational cost. By contrast, in VOF method, only one scalar function, defined as volume fraction occupied by the traced fluid, is required, that reduces the computational load significantly [30]. The level set method can trace back to 1988 and the scalar function of it is defined as a signed distance function away from the free boundary [51]. In both methods, the motion of the free boundary is accounted for by advecting each scalar function over the fixed domain. The governing equation takes the following form

$$\frac{\partial f}{\partial t} + \vec{u} \cdot \nabla f = 0 \tag{1.3}$$

where f holds the space for the scalar function. Compared with the ALE method, VOF and level set methods share some common advantages, such as the numerical implementation can be straightforwardly extended to three-dimensional case and they can deal with free moving boundary with topological changes. However, by comparison in between, each of them has its own strengths and drawbacks [50, 20]. The VOF method is superior to level set method at mass conservation but inferior at obtaining geometrical informations of the free boundary such as normal direction and curvature. Therefore, in the context of VOF method, research focuses on techniques of reconstruction of the free boundary [53, 55], while in the context of level set method, techniques to improve the performance in mass conservation are mainly discussed [50, 49, 34].

Unlike the ALE method, in which the free boundary always conforms with the element edge, the VOF and level set methods bring another challenge that the solution field might present discontinuity on itself or its derivative across the free boundary within one element since the free boundary usually cuts the elements. This behaviour violates the continuity assumption that most numerical methods are based on. One common practice is to smooth the physical properties over a strip of layer around the discontinuity [67]. It's equivalent to smear the sharp interface. Sometimes, to increase the accuracy, locally adaptive mesh refinement is

required [36]. The alternative solution to handle the discontinuity problem is to employ the extended finite element method (XFEM). XFEM has been widely used to deal with various discontinuity problems [19], since it's first developed in 1999 [3, 10]. In the context of XFEM, level set method is typically used to capture the interface due to the coupling in between is straight and simple.

1.2.2 Development of ledge simulation

Over the past several decades, numerical techniques have been widely utilized to analyse various physicochemical phenomena involved in the Hall-Héroult process. On the subject of thermal balance, various thermal models or thermo-electric models were built up to examine the thermal behaviour inside or through the cell [16, 54, 25, 26]. On the subject of bath composition's local variation, some works simulate the alumina distribution in the bath with a purpose of optimizing the feeding rate or feeders layout [17, 31, 73]. But actually the bath including the dissolved alumina is molten salt; it means that bath components are in ionic state. Then in some other works, the multi-component model was developed to analyse the ionic species transport in different scenarios [21, 62, 65]. On the subject of fluid dynamics, published research works mainly fall into two categories. One type investigates the instability of the interface between the bath and the molten metal pad or the internal flow field of the cell through magnetohydrodynamics (MHD) simulation [59, 22, 5]. The other one models the evolution of gas bubbles (or the bubble driven flow) near the anode surface [63, 75]. In these research works related to these three subjects, conventional continuum-based methods were commonly used. In addition, these three subjects mentioned here are interacted with each other in practice. The ledge formation process involves these three subjects. Instead of reviewing each subject respectively, the following part reviews the development of numerical simulation on the formation or profile variation of the ledge.

In the relatively early stage, the research mainly confined to the highly simplified one-dimensional model. In one-dimensional model, the system was usually reduced to a chain of serially connected thermal resistance. It also can be called multilayer system. In [28], the author derived a mathematical formula to calculate the ledge thickness purely according to the fact that the heat flow through each layer is conserved under equilibrium conditions, as shown in Fig. 1.4.

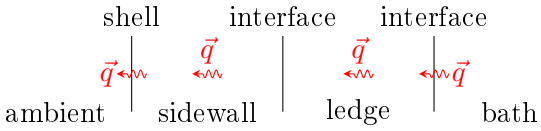


Figure 1.4 – Heat flow through each layer in one dimensional model.

The conservation law was expressed as

$$h_b A_b(T_b - T_m) = \frac{k_f A_f}{\delta_f}(T_m - T_w) = \frac{T_w - T_{sh}}{\sum_i \frac{\delta_i}{k_i A_i} + \sum_c \frac{r_c}{A_c}} = h_{sh} A_{sh}(T_{sh} - T_a) \tag{1.4}$$

where h, A, k, δ, r denote heat transfer coefficient, surface area, thermal conductivity, thickness and thermal resistivity, respectively; the subscripts b, f, i, sh and c represent bath, frozen ledge, sidewall component, shell and contacting surface, respectively; T_b, T_m, T_w, T_{sh}, T_a represent temperature of bath $1-2$ inches from ledge, liquidus temperature, wall temperature behind ledge, shell temperature and ambient temperature, respectively. In general, T_a, T_m and T_b would be provided, then the ledge thickness and temperature distribution could be calculated by solving Eq. 1.4. The models written in mathematical formula like Eq. 1.4 represent the earliest types of models for one-dimensional steady case [28, 6].

For the transient problem in one dimension, the dynamic model based on a similar multilayer cell as shown in Fig. 1.5 can be built up, see [74, 1] for example. In this dynamic model, the

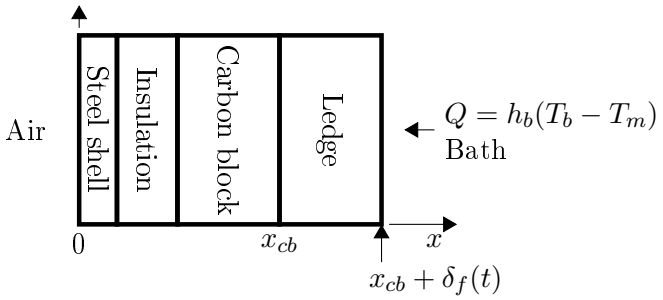


Figure 1.5 – The multilayer cell wall including side ledge [1].

following classical Stefan problem was solved by finite difference method

$$\rho(x,T)c_p(x,T)\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\Big(k(x,T)\frac{\partial T}{\partial x}\Big), \qquad x \in \Big(0, x_{cb} + \delta_f(t)\Big) \tag{1.5a}$$

$$T = T_m(t), \qquad x = x_{cb} + \delta_f(t) \tag{1.5b}$$

$$\rho_f L_f \frac{d\delta_f}{dt} = k_f \frac{\partial T}{\partial x} - h_b\Big(T_b(t) - T_m(t)\Big), \quad x = x_{cb} + \delta_f(t) \tag{1.5c}$$

where ρ is density; c_p is specific heat; L_f is the latent heat of frozen ledge fusion. Eq. 1.5b is one boundary condition as solving Eq. 1.5a. The calculation can start either by solving Eq. 1.5c or Eq. 1.5a. Eq. 1.5c and Eq. 1.5a are supposed to be solved alternately, because the former provides the latter with temperature distribution while the latter provides the former with ledge front position. It's equivalent to utilizing a moving mesh to handle the deforming border. In [74], the advection effect arising from grid movement was also taken into account. Through this dynamic model, the ledge response to some operational disturbances, such as AlF_3 addition (CR is changed), alumina feeding and anode effect, was analysed [74, 1]. For example, Fig. 1.6 is one result from [1] showing how the ledge thickness responses to a large dump of AlF_3.

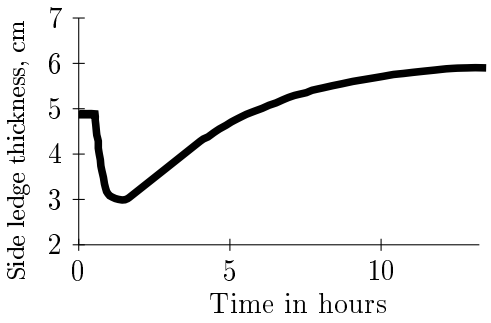


Figure 1.6 – Ledge response after AlF_3 ($200kg$) addition [1].

The one-dimensional models only predict the thickness of the ledge. To simulate the ledge profile varying in the vertical direction, the models require at least being two dimensions. In [35], the authors built a two-dimensional model by a commercial finite element code. The computational domain is shown in Fig. 1.7. It's purely a thermal model with a source term

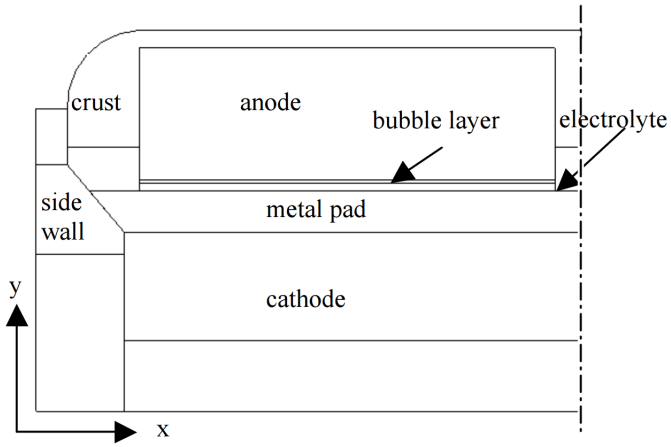


Figure 1.7 – The two-dimensional hypothetic cell used in [35].

to account for the heat generated by molten bath under the anode. The convective effect in the molten bath, which is treated as solid, is taken into account by the enhanced thermal conductivity. Enthalpy method is used to describe the ledge formation process. The ledge front can be extracted according to the temperature distribution. Fig. 1.8 presents the ledge profiles under different heat inputs.

The development of numerical simulations has shown a trend of coupling multi-physical fields to improve the prediction's accuracy. In [7], a coupled thermo-electrical model was built by solving the steady thermal conduction equation with Joule heating source term and charge conservation equation (for electric potential) through finite element method. The computational domain covers a half cross-section of the cell. The latent heat of fusion is not taken into account. The ledge front is extracted simply by comparing the local temperature with the melting point of the ledge. The enhanced thermal conductivity is used to account for convec-

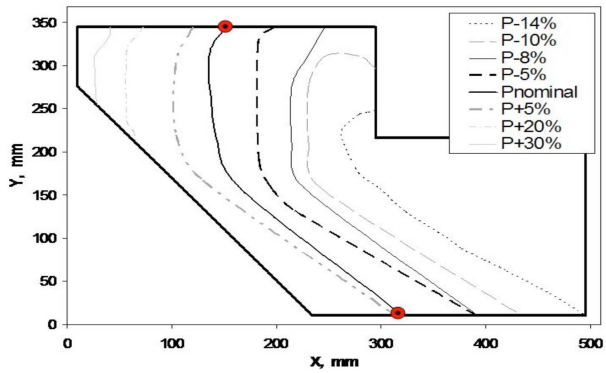


Figure 1.8 – Ledge profile varies with respect to different heat inputs [35].

tion in the liquid region. Fig. 1.9 presents the simulation result of isolines for the temperature and the corresponding ledge profile is shown in Fig. 1.10.

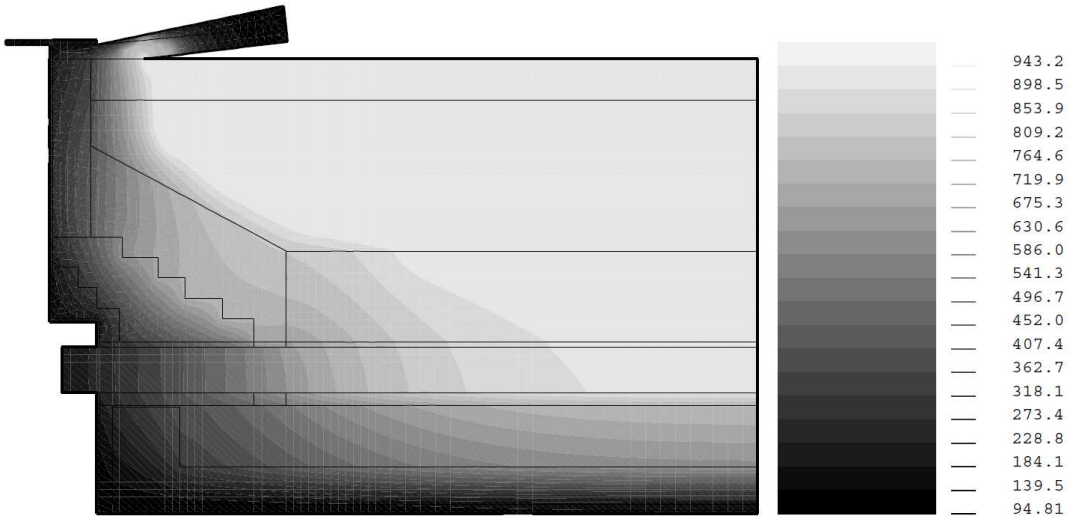


Figure 1.9 – Simulation result of isolines for the temperature in [7].

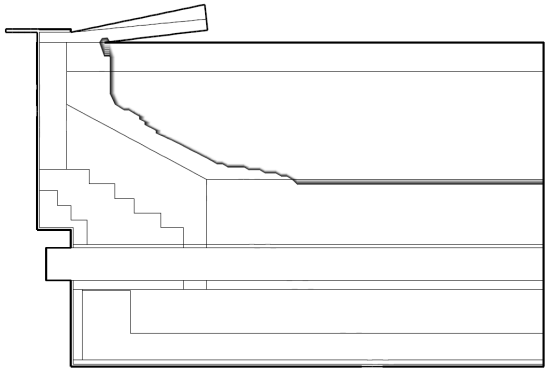


Figure 1.10 – Simulation result of ledge profile in [7].

All the models mentioned above didn't describe properly the thermal behaviour that happens in the bath, due to the fact that the bath layer was not included in the computational domain or regarded as solid. At the bath/ledge interface, the heat flux on the bath side was usually approximated by $h_b(T_b - T_m)$. As an input boundary condition, this approximation is significantly crucial, especially the bath/ledge heat transfer coefficient h_b is. There are some works dedicated to evaluate this coefficient [68, 58, 61]. An alternative way to avoid this approximation is to utilize the concept of inverse heat transfer method [52].

In the inverse method, the numerical calculation is combined with actual measurements. As aformentioned, the harsh condition inside the bath makes direct measurements difficult, but the "inverse" concept makes it possible to use the data measured outside the bath. For example in [4], the computation domain comprise two layers, sidewall and ledge, as shown in Fig. 1.11. P1, P2 and P3 are control points on the ledge front to decide its position and profile. Three thermocouples were set up inside the sidewall to measure the temperature for calibration. The temperature distribution was obtained by solving the following two-dimensional steady heat conduction equation through finite volume method:

$$\frac{\partial}{\partial x}\Big(k\frac{\partial T}{\partial x}\Big) + \frac{\partial}{\partial z}\Big(k\frac{\partial T}{\partial z}\Big) = 0 \tag{1.6}$$

The essence of this approach is an iterative procedure to minimize the discrepancy between the values obtained respectively by the measurement and numerical calculation at the locations where the thermocouples are. The coordinates of the control points are the arguments of a minimization equation. After each minimization, a new ledge front can be constructed according to the new positions of the control points. Then the computational domain is updated (moving mesh) and a new numerical simulation can be implemented. The iterations carries on until the calculated temperatures converge to measured values. In this approach, no heat flux was involved during calculation, so it avoided accounting for the latent heat. During the minimization, the interface position is the to-be-determined unknown.

Another type of inverse heat transfer method also covers the molten bath [40, 47, 41, 43, 42]. As shown in Fig. 1.12, by contrast with the model presented in [4], the heat flux was imposed on both left and right boundaries of the domain and the heat flux on the right boundary q''_{in} was taken as the to-be-determined unknown of the minimization. The thermocouples were also embedded in the sidewall to record transient temperature or thermal flux. This approach also comprises two alternately-solved parts, i.e. the direct calculation and inverse minimization. For the direct calculation, q''_{in} is known, the following enthalpy-based formulation of transient thermal conduction problem is solved by finite difference method based on fixed grids,

$$\rho C_p \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) - \delta h \frac{\partial \bar{\chi}}{\partial t} \tag{1.7}$$

where δh denotes the enthalpy; $\bar{\chi}$ is the liquid fraction in the mushy zone as shown in Fig. 1.12. The evolution of the latent heat is accounted for automatically in the enthalpy formulation

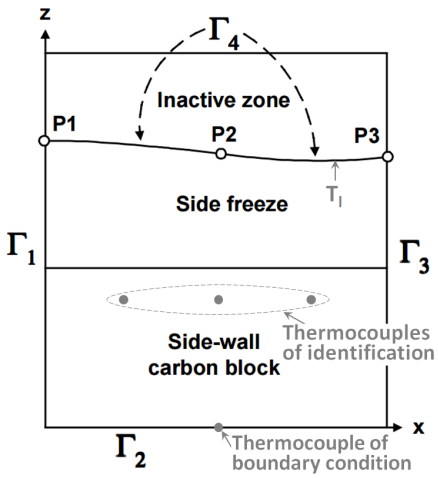


Figure 1.11 – Computational domain of [4].

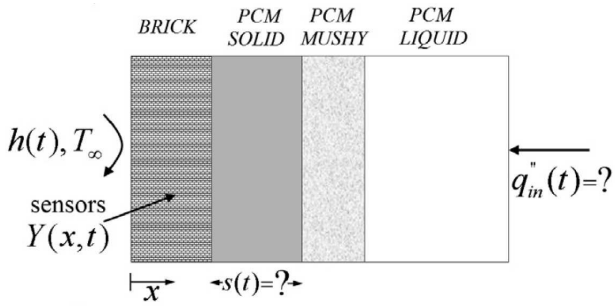


Figure 1.12 – The model setting up of the one-dimensional inverse heat transfer method in [40].

and the ledge front can be constructed according to the temperature result. For inverse minimization, the measured temperature or heat flux readings and the simulation results are utilized to estimate a new value of q''_{in}. It was reported that the minimization based on the heat flux measurements is more efficient [57, 40]. As an example, Fig. 1.13 displays the model setting up for the two-dimensional case through the inverse thermal conduction method in [47]. Fig. 1.14 presents the corresponding simulated ledge profile and comparison with the reference solution. It's reported that prediction errors increase as the ledge grows [47].

The main advantage of the inverse method can be thought of as that it attempts to provide the purely thermal model with real-time input thermal boundary conditions through real-time measurements with no need to simulate the hydrodynamics that happens in the molten bath. This might be attractive when computational capacity is limited. The advances of computer makes it possible to couple the thermal model with the fluid dynamic model [56], especially the increasingly mature commercial softwares aid the building of large-scale multi-physic models [39]. In [56], a steady-state three-dimensional model was built by solving three sets of equations alternately, i.e. non-linear convection-diffusion enthalpy-based equation with Joule effect,

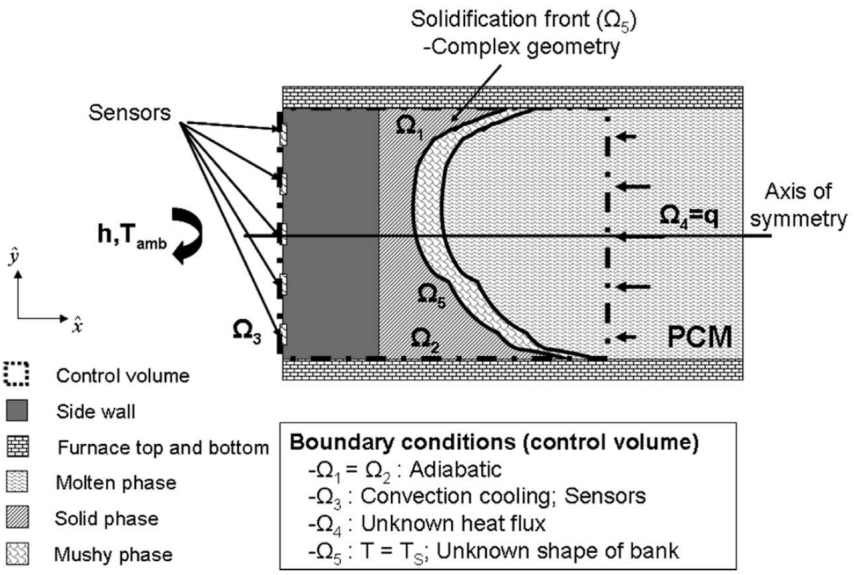


Figure 1.13 – The model setting up of the two-dimensional inverse heat transfer method in [47].

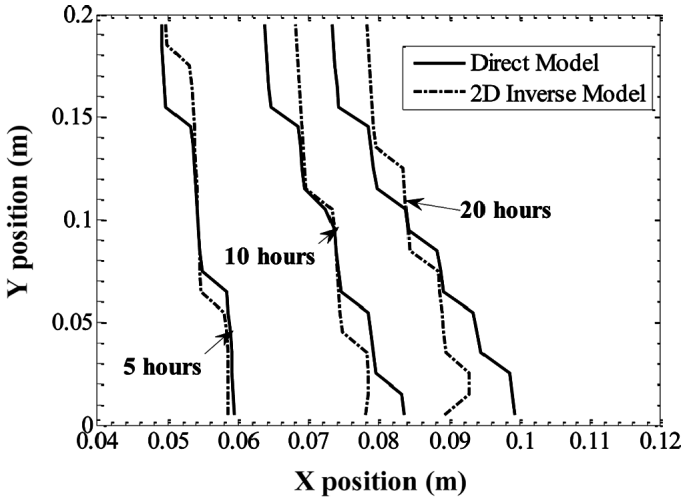


Figure 1.14 – The predicted ledge profile in [47].

Navier-Stokes equations and Maxwell equations. In dealing with free boundary problem for hydrodynamics, a liquid fraction variable defined as a function of temperature was utilized. Fig. 1.15 presents two horizontal slices of calculated ledge profile in terms of liquid fraction.

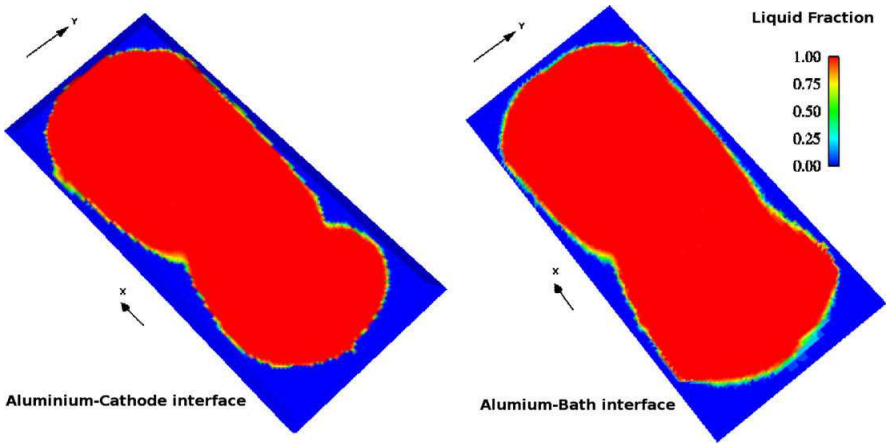


Figure 1.15 – Two horizontal slices of the calculated ledge profile in terms of liquid fraction in [56].

In summary, early research works predict ledge profile only based on the thermal balance model. Even though the influence of bath composition and other factors are analysed, yet they are taken into account merely as input parameters. Nowadays with the development of computing power, it has shown a trend of building a complex model by coupling more physicochemical processes.

1.3 Objectives and originality

During the Hall-Héroult process, the cell's internal environment is complex and mutable because it is subjected to various operational disturbances. The ledge profile is the outcome of the interactions among a couple of physicochemical processes. *The aim of the present work is to establish a two-dimensional numerical model for ledge profile prediction with considering the interactions among thermal behaviour, bath composition and bath motion inside the cell through XFEM and level set method.*

Assumptions used in building the numerical model are summarized below.

- The bath is only composed of cryolite and excess AlF_3 in chemistry, other additives and impurities are ignored. It means he bath can be regarded as a binary $NaF - AlF_3$ system.
- The side ledge is pure cryolite in solid state. Then the ledge formation can be expressed as

$$3NaF + AlF_3 \underset{\text{thaw}}{\overset{\text{freeze}}{\rightleftharpoons}} Na_3AlF_6 \tag{1.8}$$

- The computational domain is vertically two-dimensional and only includes ledge-bath two materials. It means the ledge front (solid-liquid interface) is taken as a free internal boundary.

- The ledge front is assumed to be planar. Columnar or dendritic structures are not considered.
- The bath flow is treated as being laminar and governed by the incompressible Stokes equations. The bubble driven flow and MHD calculation are beyond the scope of the present work.

The **originality** of the present work is embodied in the following several aspects.

- In the context of ledge profile prediction, either enthalpy method or temperature-based formulation with moving mesh techniques is adopted to model the phase change process in most of the published works. In the present work, XFEM combined with level set method enables the model to be built based on a fixed mesh while using temperature-based formulation. Compared with enthalpy method, the description of the interface through level set method is relatively isolated from the thermal problem. This isolation or partial independence makes the coupling relationships get clear, which facilitates modularization in programming. Level set method also preserves the capability of dealing with the interface with topological change.
- In the context of ledge profile prediction, the variation of the bath composition is seldom described, so its influence is usually neglected in the published works. As shown in Fig. 1.16, three interactive processes including thermal behaviour, bath composition and bath motion are taken into consideration to predict the ledge profile in the present work.

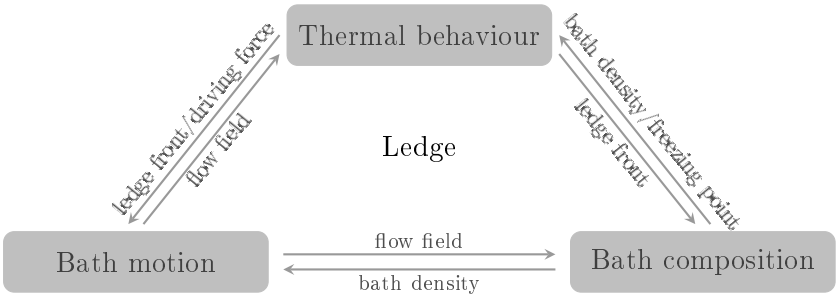


Figure 1.16 – Links among the interactive processes.

The influence of the bath composition variation is twofold:

– The melting point of bath is dependent on the bath composition. As shown in Fig. 1.17, the melting point can be expressed with respect to the weight percent of excess AlF_3. It corresponds to the following equation [64]:

$$T_m = 1011 + 0.50[AlF_3] - 0.13[AlF_3]^{2.2} \tag{1.9}$$

where $[AlF_3]\%$ denotes the weight percent of excess AlF_3.

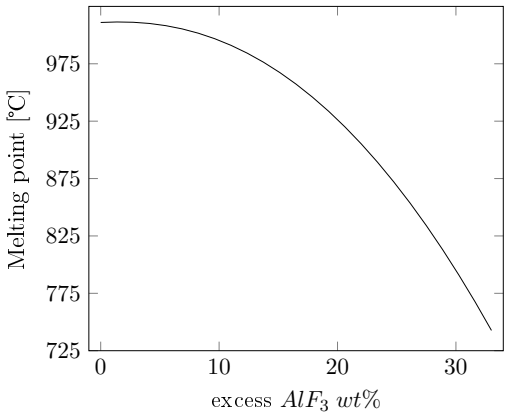


Figure 1.17 – Liquidus curve of $NaF - AlF_3$ system with respect to the weight percent of excess AlF_3.

– The bath density is dependent on bath composition. Fig. 1.18 presents the density of $NaF - AlF_3$ system with respect to the weight percent of excess AlF_3. The underlying equation is [60]

$$\rho_{1000} = \left[\left(1946 + 1113\chi_{AlF_3}\right)^{-3.6} + \left(\chi_{AlF_3}/859\right)^{3.6}\right]^{-1/3.6} \tag{1.10}$$

where ρ_{1000} denotes the bath density at 1000°C; χ_{AlF_3} is the molar fraction of AlF_3, and $\chi_{AlF_3} \approx \frac{2+3[AlF_3]\%}{8-3[AlF_3]\%}$.

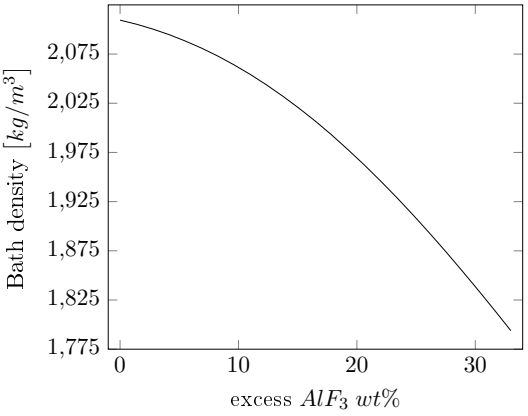


Figure 1.18 – Density of $NaF - AlF_3$ system with respect to the weight percent of excess AlF_3.

- Also in the context of applying XFEM/level set method to solve phase change problem, it's a fresh attempt to take into account the influence of fluid flow and species transport. XFEM has to deal with various discontinuities at the ledge front, including weak discontinuity in the temperature field and strong discontinuities in the species concentration, velocity and pressure fields. To the best knowledge of the authors, no attempt has been carried out to couple these three physics using XFEM in the literature.

- In modelling the melt flow, three driving forces are mainly investigated, including density jump between liquid and solid, temperature gradient and concentration gradient.

Remarks The programming work in the present study is based on the in-house code named FESh++ (Finite Element Shell in C++). This code was originally developed to solve thermo-electro-mechanical problem in the solid material [45, 23, 46, 18]. In recent years, the code has been further developed to involve time-dependent discontinuities and computational fluid dynamics in the present project and also another PhD project [48].

1.4 Organization of the thesis

The dissertation is organized as follows. In chapter 1, the motivation of the present project is discussed, followed by the review of the literature on numerical treatment of free boundary problems and progress of ledge profile simulation; in the end of this chapter, the objective and originality of this project are pointed out. In chapter 2, the mathematical model made up of a series of partial differential equations and the numerical methods including level set method and XFEM are briefly introduced. The numerical model building is divided into four steps, which are demonstrated one by one in the following chapters, respectively. In chapter 3, the one phase problem of coupling of fluid flow and species transport is solved by FEM. In chapter 4, the phase transition process coupled with species transport in a binary system is described by XFEM combined with level set method. Chapter 5 is dedicated to solving the Stefan problem of pure material with ensuing melt flow by XFEM combined with level set method. In chapter 6, all the three physical problems including phase transition, fluid flow and species transport are coupled; the model is applied to simulate the ledge solidification/melting process under certain scenarios. Finally, this project is concluded with the summary and outlook.

Bibliography

[1] A. G. Barantsev, V. V. Yurkov, V. C. Mann, T. V. Piskazhova, and K. F. Nikandrov. Model of process of electrolyses. *TMS-Light Metals*, pages 315–321, 2000.

[2] M. Bellet and V. D. Fachinotti. ALE method for solidification modelling. *Computer Methods in Applied Mechanics and Engineering*, 193(39):4355–4381, 2004.

[3] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601–620, 1999.

[4] P. Boily, L. I. Kiss, R. T. Bui, and P. Desclaux. Sensitivity analysis of the thermal detection of the ledge profile in an aluminium reduction cell. *TMS-Light Metals*, pages 1–7, 2001.

[5] V. Bojarevics and J. W. Evans. Mathematical modelling of Hall-Héroult pot instability and verification by measurements of anode current distribution. *TMS-Light Metals*, pages 783–788, 2015.

[6] J. N. Bruggeman. Pot heat balance fundamentals. In *6th Australasian Aluminum Smelting Workshop Proceedings*, pages 167–189, 1998.

[7] L. Consiglieri and M. C. Muniz. Existence of a solution for a free boundary problem in the thermoelectrical modelling of an aluminium electrolytic cell. *European Journal of Applied Mathematics*, 14(2):201–216, 2003.

[8] P. Coursol, G. Dufour, J. Coté, P. Chartrand, and P. Mackey. Application of thermodynamic models for better understanding and optimizing the Hall–Héroult process. *JOM*, 64(11):1326–1333, 2012.

[9] M. Cruchaga, L. Battaglia, M. Storti, and J. D'Elía. Numerical modeling and experimental validation of free surface flow problems. *Archives of Computational Methods in Engineering*, 23(1):139–169, 2016.

[10] J. Dolbow and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):131–150, 1999.

[11] J. Donea and A. Huerta. *Finite element methods for flow problems*. John Wiley & Sons, 2003.

[12] J. Donea, A. Huerta, J. Ph. Ponthot, and A. Rodriguez-Ferran. Arbitrary Lagrangian-Eulerian methods, Encyclopedia of computational mechanics. Vol 1. Ch 14, 2004.

[13] T. Drengstig. *On process model representation and AlF_3 dynamics of aluminium electrolysis cells*. PhD thesis, Norwegian University of Science and Technology, 1997.

[14] T. Drengstig, D. Ljungquist, and B. A. Foss. On the AlF_3 and temperature control of an aluminum electrolysis cell. *Modeling, Identification and Control*, 19(1):31–59, 1998.

[15] F. Duarte, R. Gormaz, and S. Natesan. Arbitrary Lagrangian-Eulerian method for Navier-Stokes equations with moving boundaries. *Computer Methods in Applied Mechanics and Engineering*, 193(45):4819–4836, 2004.

[16] M. Dupuis. Computation of aluminum reduction cell energy balance using ANSYS® finite element models. *TMS-Light Metals*, pages 409–418, 1998.

[17] Y. Feng, M. A. Cooksey, and M. P. Schwarz. CFD modelling of alumina mixing in aluminium reduction cells. *TMS-Light Metals*, pages 543–548, 2011.

[18] H. Fortin, N. Kandev, and M. Fafard. FEM analysis of voltage drop in the anode connector induced by steel stub diameter reduction. *Finite Elements in Analysis and Design*, 52:71–82, 2012.

[19] T. P. Fries and T. Belytschko. The extended/generalized finite element method: an overview of the method and its applications. *International Journal for Numerical Methods in Engineering*, 84(3):253–304, 2010.

[20] D. Fuster, G. Agbaglah, C. Josserand, S. Popinet, and S. Zaleski. Numerical simulation of droplets, bubbles and waves: state of the art. *Fluid Dynamics Research*, 41(6):065001, 2009.

[21] F. Gagnon, D. Ziegler, and M. Fafard. Electrochemical modelling using electroneutrality equation as a constraint. *Journal of Applied Electrochemistry*, 44(3):361–381, 2014.

[22] J. F. Gerbeau, T. Lelièvre, and C. Le Bris. Simulations of MHD flows with moving interfaces. *Journal of Computational Physics*, 184(1):163–191, 2003.

[23] P. Goulet. *Modélisation du comportement thermo-électro-mécanique des interfaces de contact d'une cuve de Hall-Héroult*. PhD thesis, Université Laval, 2004.

[24] K. Grjotheim. *Aluminium electrolysis: fundamentals of the Hall-Héroult process*. Aluminium-Verlag, 1982.

[25] V. Gusberti, D. S. Severo, B. J. Welch, and M. Skyllas-Kazacos. Modelling the aluminum smelting cell mass and energy balance - a tool based on the 1^{st} law of thermodynamics. In *Proceeding of 10th Australian Aluminium Smelting Technology Conference*, 2011.

[26] V. Gusberti, D. S. Severo, B. J. Welch, and M. Skyllas-Kazacos. Modeling the mass and energy balance of different aluminium smelting cell technologies. *TMS-Light Metals*, pages 929–934, 2012.

[27] F. H. Harlow and J. E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Physics of Fluids*, 8(12):2182–2189, 1965.

[28] W. E. Haupin. Calculating thickness of containing walls frozen from melt. *JOM*, 23(7):41–44, 1971.

[29] C. W. Hirt, A. A. Amsden, and J. L. Cook. An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *Journal of Computational Physics*, 14(3):227–253, 1974.

[30] C. W. Hirt and B. D. Nichols. Volume of fluid (VOF) method for the dynamics of free boundaries. *Journal of Computational Physics*, 39(1):201–225, 1981.

[31] T. Hofer. *Numerical simulation and optimization of the alumina distribution in an aluminium electrolysis pot.* PhD thesis, École Polytechnique Fédérale de Lausanne, 2011.

[32] G. Hou, J. Wang, and A. Layton. Numerical methods for fluid-structure interaction - a review. *Communications in Computational Physics*, 12(02):337–377, 2012.

[33] T. J. R. Hughes, W. K. Liu, and T. K. Zimmermann. Lagrangian-Eulerian finite element formulation for incompressible viscous flows. *Computer Methods in Applied Mechanics and Engineering*, 29(3):329–349, 1981.

[34] C. E. Kees, I. Akkerman, M. W. Farthing, and Y. Bazilevs. A conservative level set method suitable for variable-order approximations and unstructured meshes. *Journal of Computational Physics*, 230(12):4536–4558, 2011.

[35] L. I. Kiss and V. Dassylva-Raymond. Freeze thickness in the aluminum electrolysis cells. *TMS-Light Metals*, pages 431–436, 2008.

[36] H. Kohno and T. Tanahashi. Numerical analysis of moving interfaces using a level set method coupled with adaptive mesh refinement. *International Journal for Numerical Methods in Fluids*, 45(9):921–944, 2004.

[37] S. Kolås. Defining and verifying the "correlation line" in aluminum electrolysis. *JOM*, 59(5):55–60, 2007.

[38] S. Kolås and T. Støre. Bath temperature and AlF_3 control of an aluminium electrolysis cell. *Control Engineering Practice*, 17(9):1035–1043, 2009.

[39] S. Langlois, J. Rappaz, O. Martin, Y. Caratini, M. Flueck, A. Masserey, and G. Steiner. 3D coupled MHD and thermo-electrical modelling applied to AP technology pots. *TMS-Light Metals*, pages 769–775, 2015.

[40] M. LeBreux, M. Désilets, and M. Lacroix. Fast inverse prediction of phase change banks in high temperature furnaces with a Kalman filter coupled with a recursive least-square estimator. *International Journal of Heat and Mass Transfer*, 53(23):5250–5260, 2010.

[41] M. LeBreux, M. Désilets, and M. Lacroix. Control of the ledge thickness in high-temperature metallurgical reactors using a virtual sensor. *Inverse Problems in Science and Engineering*, 20(8):1215–1238, 2012.

[42] M. LeBreux, M. Désilets, and M. Lacroix. Prediction of the time-varying ledge profile inside a high-temperature metallurgical reactor with an unscented Kalman filter-based virtual sensor. *Numerical Heat Transfer, Part A: Applications*, 64(7):551–576, 2013.

[43] M. LeBreux, M. Désilets, and M. Lacroix. An unscented Kalman filter inverse heat transfer method for the prediction of the ledge thickness inside high-temperature metallurgical reactors. *International Journal of Heat and Mass Transfer*, 57(1):265–273, 2013.

[44] R. Loubère, P. H. Maire, M. Shashkov, J. Breil, and S. Galera. ReALE: a reconnection-based arbitrary-Lagrangian-Eulerian method. *Journal of Computational Physics*, 229(12):4724–4761, 2010.

[45] D. Marceau. *Modélisation du contact tridimensionnel avec frottement en grandes transformations et son application à l'étude des dispositifs d'ancrage multitorons*. PhD thesis, Université Laval, 2001.

[46] D. Marceau, P. Goulet, D. Richard, and M. Fafard. FESh++, une nouvelle approche orientée objet pour la simulation par éléments finis des problèmes multiphysiques. In *Actes du septièmes colloque en calcul des structures, Giens, France*, volume 2, pages 303–308, 2005.

[47] M. A. Marois, M. Désilets, and M. Lacroix. Prediction of a 2-D solidification front in high temperature furnaces by an inverse analysis. *Numerical Heat Transfer, Part A: Applications*, 59(3):151–166, 2011.

[48] D. Martin. *Multiphase modelling of melting/solidification with high density variations using XFEM*. PhD thesis, Université Laval, 2016.

[49] E. Olsson and G. Kreiss. A conservative level set method for two phase flow. *Journal of Computational Physics*, 210(1):225–246, 2005.

[50] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Springer Science & Business Media, 2006.

[51] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

[52] M. N. Ozisik. *Inverse heat transfer: fundamentals and applications*. CRC Press, 2000.

[53] J. E. Pilliod and E. G. Puckett. Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *Journal of Computational Physics*, 199(2):465–502, 2004.

[54] D. Richard, P. Goulet, M. Dupuis, and M. Fafard. Thermo-chemo-mechanical modeling of a Hall-Héroult cell thermal bake-out. *TMS-Light Metals*, 3:669–674, 2006.

[55] W. J. Rider and D. B. Kothe. Reconstructing volume tracking. *Journal of Computational Physics*, 141(2):112–152, 1998.

[56] Y. Safa, M. Flueck, and J. Rappaz. Numerical simulation of thermal problems coupled with magnetohydrodynamic effects in aluminium cell. *Applied Mathematical Modelling*, 33(3):1479–1492, 2009.

[57] M. Samai and T. Loulou. A comparative study of heat flux and temperature based objective functional to solve inverse heat conduction problems. *Numerical Heat Transfer, Part B: Fundamentals*, 56(1):75–104, 2009.

[58] D. S. Severo and V. Gusberti. A modelling approach to estimate bath and metal heat transfer coefficients. *TMS-Light Metals*, pages 309–314, 2009.

[59] A. D. Sneyd and A. Wang. Interfacial instability due to MHD mode coupling in aluminium reduction cells. *Journal of Fluid Mechanics*, 263:343–360, 1994.

[60] A. Solheim. The density of molten NaF-LiF-AlF_3-CaF_2-Al_2O_3 in aluminium electrolysis. *Aluminum Transactions*, 2(1):161–168, 2000.

[61] A. Solheim. Some aspects of heat transfer between bath and sideledge in aluminium reduction cells. *TMS-Light Metals*, pages 381–386, 2011.

[62] A. Solheim. Concentration gradients of individual anion species in the cathode boundary layer of aluminium reduction cells. *TMS-Light Metals*, pages 665–670, 2012.

[63] A. Solheim, S. T. Johansen, S. Rolseth, and J. Thonstad. Gas induced bath circulation in aluminium reduction cells. *Journal of Applied Electrochemistry*, 19(5):703–712, 1989.

[64] A. Solheim, S. Rolseth, E. Skybakmoen, L. Støen, Å. Sterten, and T. Støre. Liquidus temperatures for primary crystallization of cryolite in molten salt systems of interest for aluminum electrolysis. *Metallurgical and Materials Transactions B*, 27(5):739–744, 1996.

[65] A. Solheim and K. Tschöpe. Model for excessive cathode wear by a "carbon pump" at the cell bottom. *TMS-Light Metals*, pages 1257–1262, 2013.

[66] Y. Sui, H. Ding, and P. D. M. Spelt. Numerical simulations of flows with moving contact lines. *Annual Review of Fluid Mechanics*, 46:97–119, 2014.

[67] M. Sussman, P. Smereka, and S. Osher. A level set approach for computing solutions to incompressible two-phase flow. *Journal of Computational Physics*, 114(1):146–159, 1994.

[68] M. P. Taylor and B. J. Welch. Bath/freeze heat transfer coefficients: experimental determination and industrial application. *TMS-Light Metals*, pages 781–789, 1985.

[69] T. E. Tezduyar. Finite element methods for flow problems with moving boundaries and interfaces. *Archives of Computational Methods in Engineering*, 8(2):83–130, 2001.

[70] T. E. Tezduyar. Interface-tracking and interface-capturing techniques for finite element computation of moving boundaries and interfaces. *Computer Methods in Applied Mechanics and Engineering*, 195(23):2983–3000, 2006.

[71] J. Thonstad. *Aluminium electrolysis: fundamentals of the Hall-Héroult Process*. Aluminium-Verlag, 2001.

[72] J. Thonstad and S. Rolseth. Alternative electrolyte compositions for aluminium electrolysis. *Mineral Processing and Extractive Metallurgy*, 2013.

[73] R. von Kaenel, J. Antille, M. V. Romerio, and O. Besson. Impact of magnetohydrodynamic and bubbles driving forces on the alumina concentration in the bath of an Hall-Héroult cell. *TMS-Light Metals*, pages 585–590, 2013.

[74] C. C. Wei, J. Chen, B. J. Welch, and V. R. Voller. Modelling of dynamic ledge heat transfer. *TMS-Light Metals*, pages 309–316, 1997.

[75] K. Zhang, Y. Feng, P. Schwarz, Z. Wang, and M. Cooksey. Computational fluid dynamics (CFD) modeling of bubble dynamics in the aluminum smelting process. *Industrial & Engineering Chemistry Research*, 52(33):11378–11390, 2013.

Chapter 2

Mathematical models and numerical methods

2.1 Introduction

This chapter is dedicated to present the mathematical model and numerical methods involved in this project. The mathematical model refers to the series of partial differential equations governing the fluid flow, species transport and phase change phenomena. The derivation of the governing equations is given in section 2.2. The numerical methods mainly refer to the level set method and XFEM, which are used to capture the interface implicitly and deal with the discontinuities at the interface, respectively. Section 2.3 introduces the definition and initialization of the level set function, as well as how to construct the normal speed field to update the level set values. The XFEM and associated numerical integration are presented in section 2.4.

2.2 Governing equations

2.2.1 Fluid flow problem

2.2.1.1 Mass continuity equation

Mass conservation law is stated mathematically in Lagrangian description as

$$\frac{d}{dt}\int_{\Omega}\rho\, d\Omega = 0 \tag{2.1}$$

where ρ is density (kg/m^3) and $\frac{d}{dt}$ is the material derivative. According to Reynolds transport theorem (see appendix A.1), Eq. 2.1 can be converted into the following form:

$$\int_{\Omega}\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\vec{u})\ d\Omega = 0 \tag{2.2}$$

where $\vec{u}$ is the velocity field (m/s). Eq. 2.2 holds for any volume, so it can be further written in the differential form:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \tag{2.3}$$

Eq. 2.3 is the most universal form of mass continuity equation in Eulerian description, neglecting the issue of compressibility or incompressibility.

In the following several paragraphs, the incompressibility hypothesis of the fluid flow involved in the present study is justified. Physically speaking, incompressible flow refers to the flow in which the material density is constant. Thus, in Lagrangian description, which always coincides with the physical description, incompressibility can be expressed mathematically as [1]

$$\frac{d\rho}{dt} = 0 \tag{2.4}$$

It should be noted that it is the density of fluid parcel that keeps constant, not the density at a fixed position in space.

For a pure liquid flow problem, the flow is supposed to be incompressible. However, as aforementioned, in the present study, the fluid flow problem is coupled with thermal and species transport problems. That means the physical property of fluid is influenced by the temperature distribution and species concentration distribution to some extent. So density inside the electrolyte (molten bath, the only fluid involved in the present study) is non-uniform in space and mutable locally in time. Will this change the incompressibility property of the liquid flow?

First, the influence from species concentration distribution (bath composition) is analysed. The only one chemical reaction, as shown by Eq. 1.8, just takes place at the liquid boundary, i.e. the surface in contact with the solid ledge. That means no chemical reaction occurs in the bulk fluid. In Lagrangian description, the fluid parcel is a mass closed system. If tracing one fluid parcel as shown in Fig. 2.1, even though the species concentration distribution is non-uniform in space and mutable in time, the density of the fluid parcel remains constant, as long as its chemical composition does not change, i.e. no chemical reaction occurs inside, and the volume variation caused by mixing in between (of species) is negligible. Then in the present study, the variation of species concentration in the bulk fluid is regarded not to change the incompressibility property of fluid flow. The influence of the chemical reaction taking place at the ledge surface is accounted for in phase change problem.

Secondly, the temperature variation's influence on density is discussed. Since the fluid parcel is a closed system in terms of mass, but open in terms of energy, temperature variation is allowed within fluid parcel. Then, the fluid parcel tends to experience volume expansion or contraction due to the thermal effect. In the strict sense, the density no longer remains constant, that means the temperature variation leads to the fluid flow compressible. However, in the practical application of solving this kind of problem (natural convection flow or buoyancy-

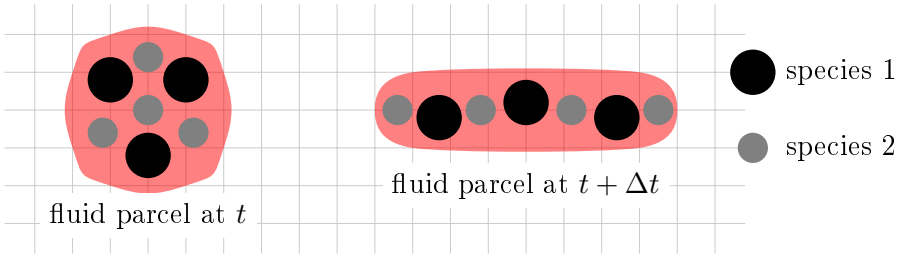


Figure 2.1 – Illustration of the species concentration variation's influence on the fluid density: as long as the chemical composition inside the fluid parcel does not change, i.e. no chemical reaction occurs, the material density remains constant.

driven flow caused by temperature gradient), Boussinesq approximation is always made for the sake of simplicity. The approximation has extreme accuracy for many such flows [10, 19]. In Boussinesq approximation, the density variation is negligible except in the gravity term where the density is multiplied by the gravitational acceleration. In the present study, the Boussinesq approximation is adopted to handle the thermal problem's influence on the fluid flow. As a result, the density, appearing anywhere including in the mass continuity equation but only except in the gravity term, is assumed constant either for the fluid parcel or at a fixed position in space.

Based on the analysis above, the fluid flow involved in the present study is regarded as incompressible flow. Then Eq. 2.3 can be divided into the following two equations:

$$\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho = 0 \tag{2.5}$$

$$\nabla \cdot \vec{u} = 0 \tag{2.6}$$

In the literature, it is common that ρ is taken as constant in Eulerian description, then Eq. 2.5 vanishes. In the present study, due to the influence from the variation of fluid composition, Eq. 2.5 remains and will be replaced by the species transport equation. The consistency in between will be justified in the later section 2.2.2. Eq. 2.6 always represents the incompressibility of fluid flow, and only based on this property, can some simplifications be made in the momentum equation.

2.2.1.2 Linear momentum equation

According to Newton's second law, the linear momentum equation can be expressed directly in Lagrangian description as

$$\frac{d}{dt} \int_{\Omega} \rho \vec{u} d\Omega = \sum_i \vec{F}_i \tag{2.7}$$

where $\vec{F}_i$ represents any external force exerting on the fluid parcel, including body forces and surface forces. Mathematically, the resultant force can be expresses as

$$\sum_i \vec{F}_i = \int_{\Omega} \vec{f} d\Omega + \int_{\Gamma} \underline{\underline{\sigma}} \cdot \vec{n} d\Gamma \tag{2.8}$$

where $\vec{f}$ is the body force per unit volume (N/m^3), such as specific weight (gravity term) $\rho\vec{g}$; $\underline{\underline{\sigma}}$ is the Cauchy stress tensor (N/m^2). Introducing Eq. 2.8 into Eq. 2.7 and converting with the aid of Reynolds transport equation, Eq. 2.7 reads in Eulerian description

$$\int_\Omega \left(\frac{\partial \rho \vec{u}}{\partial t} + \nabla \cdot (\rho \vec{u} \otimes \vec{u}) \right) d\Omega = \int_\Omega \vec{f} d\Omega + \int_\Gamma \underline{\underline{\sigma}} \cdot \vec{n} d\Gamma \tag{2.9}$$

where $\nabla \cdot (\rho \vec{u} \otimes \vec{u}) = (\nabla \cdot \vec{u}) \rho \vec{u} + \nabla (\rho \vec{u}) \cdot \vec{u}$, in which the first term on the right hand side vanishes for incompressible flow, the second term on the right hand side can be further written as $\nabla (\rho \vec{u}) \cdot \vec{u} = \vec{u} \otimes \nabla \rho \cdot \vec{u} + \rho \nabla \vec{u} \cdot \vec{u} = (\nabla \rho \cdot \vec{u}) \vec{u} + \rho (\vec{u} \cdot \nabla) \vec{u}$. Then Eq. 2.9 becomes

$$\int_\Omega \rho \frac{\partial \vec{u}}{\partial t} + \rho (\vec{u} \cdot \nabla) \vec{u} + \left(\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho \right) \vec{u} \; d\Omega = \int_\Omega \vec{f} d\Omega + \int_\Gamma \underline{\underline{\sigma}} \cdot \vec{n} d\Gamma \tag{2.10}$$

All the fluid involved in the present study are regarded as Newtonian fluid. The constitutive law for the Newtonian fluid can be written as [6]

$$\underline{\underline{\sigma}} = -p \underline{\underline{I}} + \underbrace{2\mu \underline{\underline{D}} + \lambda (\nabla \cdot \vec{u}) \underline{\underline{I}}}_{\underline{\underline{\tau}}} \tag{2.11}$$

where the stress tensor $\underline{\underline{\sigma}}$ is symmetric; p is pressure (N/m^2); $\underline{\underline{I}}$ is the identity tensor; $\underline{\underline{\tau}}$ is named viscous stress tensor; μ is dynamic viscosity ($N \cdot s/m^2$); $\underline{\underline{D}} = \frac{1}{2} \left(\nabla \vec{u} + \nabla \vec{u}^{\mathrm{T}} \right)$ is shear rate tensor ($1/s$); $\lambda = -\frac{2}{3}\mu$ is called the second coefficient of viscosity. For incompressible flow, the term containing λ vanishes, then Eq. 2.11 becomes

$$\underline{\underline{\sigma}} = -p \underline{\underline{I}} + \underbrace{2\mu \underline{\underline{D}}}_{\underline{\underline{\tau}}} \tag{2.12}$$

Introducing Eq. 2.12 into Eq. 2.10 and converting the stress term by the divergence theorem, Eq. 2.10 can be written in the following equivalent differential form:

$$\rho \frac{\partial \vec{u}}{\partial t} + \rho (\vec{u} \cdot \nabla) \vec{u} + \vec{u} \left(\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho \right) - \nabla \cdot \left(2\mu \underline{\underline{D}} \right) + \nabla p = \vec{f} \tag{2.13}$$

Eq. 2.13 is the conservative form of Newtonian incompressible Navier-Stokes momentum equation. The corresponding non-conservative form, most commonly seen in the literature, is

$$\rho \frac{\partial \vec{u}}{\partial t} + \rho (\vec{u} \cdot \nabla) \vec{u} - \nabla \cdot \left(2\mu \underline{\underline{D}} \right) + \nabla p = \vec{f} \tag{2.14}$$

2.2.1.3 Boussinesq approximation

Boussinesq approximation is commonly used to model the natural convection (buoyancy driven flow purely caused by temperature gradient) without having to use the compressible formulation of the Navier-Stokes equations.

In Boussinesq approximation, it's assumed that the density only in the gravity term is the real density, which changes with the temperature variation. For the sake of not changing the

previous notation, the real density is denoted by $\tilde{\rho}$, while the notation of density in the other terms remains the same, always denoted by ρ and can be called reference density. The real density consists of two parts as

$$\tilde{\rho} = \rho + \Delta\rho \tag{2.15}$$

where $\Delta\rho$ denotes the density variation caused by the temperature gradient. $\Delta\rho$ can be further written as

$$\Delta\rho = -\rho\alpha_p \underbrace{(T - T_{ref})}_{\Delta T} \tag{2.16}$$

where T_{ref} is the reference temperature, based on which the reference density ρ is evaluated; α_p is thermal expansion coefficient, defined as $\alpha_p = -\frac{1}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_p$ [10]. In the later applications in this thesis, the magnitude order of α_p always keeps at 10^{-4}, which implies $\alpha_p \Delta T \ll 1$ even with considerable temperature variation (we say 100 °C). Therefore, the density variation caused by the temperature gradient is so small that the Boussinesq approximation is always valid.

Introducing Eq. 2.15 and Eq. 2.16 into Eq. 2.14, one can obtain (no other forces except for buoyancy and gravity are considered)

$$\rho\frac{\partial\vec{u}}{\partial t} + \rho\left(\vec{u}\cdot\nabla\right)\vec{u} - \nabla\cdot\left(2\mu\underline{\underline{D}}\right) + \nabla p + \rho\alpha_p\left(T - T_{ref}\right)\vec{g} - \rho\vec{g} = 0 \tag{2.17}$$

If gravity is excluded, Eq. 2.17 can be written in the following pressure shifted form:

$$\rho\frac{\partial\vec{u}}{\partial t} + \rho\left(\vec{u}\cdot\nabla\right)\vec{u} - \nabla\cdot\left(2\mu\underline{\underline{D}}\right) + \nabla p + \rho\alpha_p\left(T - T_{ref}\right)\vec{g} = 0 \tag{2.18}$$

2.2.2 Species transport problem

In the present study, the only fluid (molten bath) is a multicomponent mixture and, as mentioned above, there is no chemical reaction occurring in the bulk fluid. Therefore, not only the mixture as a whole, but also each single species should be conserved in mass, which could be expressed as in Lagrangian description

$$\frac{d}{dt}\int_\Omega \hat{\rho}_i d\Omega = 0 \tag{2.19}$$

where $\hat{\rho}_i$ is the density of species i in the mixture, not as it being pure substance. Applying Reynolds transport theorem, Eq. 2.19 could be written in the following differential form of Eulerian description:

$$\frac{\partial\hat{\rho}_i}{\partial t} + \nabla\cdot\left(\hat{\rho}_i\vec{u}_i\right) = 0 \tag{2.20}$$

where $\vec{u}_i$ is the velocity of species i. For a multicomponent mixture, the sum of mass conservation equation of each single species (Eq. 2.20) should derive the total mass conservation

equation (Eq. 2.3). Based on this equality, the following two definitions are made [16]. The density of multicomponent mixture, identical with that in Eq. 2.3, is defined as

$$\rho \equiv \sum_i \hat{\rho}_i \tag{2.21}$$

and the velocity $\vec{u}$ in Eq. 2.3, mass-averaged velocity, is defined as

$$\vec{u} \equiv \frac{1}{\rho} \sum_i \hat{\rho}_i \vec{u}_i \tag{2.22}$$

In Eq. 2.20, $\hat{\rho}_i \vec{u}_i$, denoted by $\vec{N}_i$, is the absolute mass flux of species i with respect to the fixed reference frame. In order to utilize the flow information related to the mixture taken as a whole, which is obtained by solving the fluid flow problem as explained in section 2.2.1, the absolute mass flux of one species should be expressed by a relative mass flux $\vec{J}_i$ with respect to the mass-averaged velocity $\vec{u}$ as

$$\vec{N}_i = \underbrace{\hat{\rho}_i \left(\vec{u}_i - \vec{u}\right)}_{\vec{J}_i} + \hat{\rho}_i \vec{u} \tag{2.23}$$

Eq. 2.23 can be depicted physically as that the absolute mass flux of species i equals its flux relative to the whole mixture's motion plus the flux following with the mixture. In addition, it's easy to prove that the sum of the relative flux over all the species equals zero, i.e. $\sum_i \vec{J}_i = 0$. For a binary system, the relative mass flux could be evaluated by Fick's first law, as expressed below [16]:

$$\vec{J}_i = -\rho D \nabla \omega_i \tag{2.24}$$

where ω_i is the mass fraction of species i, defined as $\frac{\hat{\rho}_i}{\rho}$; D is the diffusion coefficient or diffusivity (m^2/s). Introducing Eq. 2.24 and Eq. 2.23 into Eq. 2.20 and utilizing the incompressibility constraint $\nabla \cdot \vec{u} = 0$, the mass conservation equation of each species becomes

$$\frac{\partial \rho \omega_i}{\partial t} + \vec{u} \cdot \nabla \left(\rho \omega_i\right) + \nabla \cdot \underbrace{\left(-\rho D \nabla \omega_i\right)}_{\vec{J}_i} = 0 \tag{2.25}$$

It is obvious, as the identities $\sum_i \omega_i = 1$ and $\sum_i \vec{J}_i = 0$ always hold, that the sum of Eq. 2.25 over all the species turns out to be Eq. 2.5. If Eq. 2.5 is taken as a constraint that the density of the mixture must satisfy, then Eq. 2.25 is a constraint more specific and more strict, which is on each species rather than on the mixture. It means with respect to Eq. 2.5, Eq. 2.25 is sufficient but not necessary. When the mass transport is taken into account as solving the multicomponent fluid dynamic problem, Eq. 2.5 must be replaced by Eq. 2.25. For a binary system, since $\sum_i \omega_i = 1$, the transport equation of one species is sufficient. Removing the subscript i in Eq. 2.25, it becomes

$$\frac{\partial \rho \omega}{\partial t} + \vec{u} \cdot \nabla \left(\rho \omega\right) - \nabla \cdot \left(\rho D \nabla \omega\right) = 0 \tag{2.26}$$

where ω is the solute mass fraction. Eq.2.26 is in the conservative form, as $\omega\left(\frac{\partial \rho}{\partial t}+\vec{u}\cdot\nabla\rho\right)=0$, it can be rewritten in the following non-conservative form:

$$\rho\frac{\partial \omega}{\partial t}+\rho\vec{u}\cdot\nabla\omega-\nabla\cdot(\rho D\nabla\omega)=0 \tag{2.27}$$

In species transport problem, the density of mixture ρ is a function of the unknown ω. If the volume is additive, i.e. the volume of the mixture is exactly equal to the sum of the volume of the two species before mixing and species 1 is taken as solute, the density could be expressed as [15]

$$\rho=\frac{1}{\frac{\omega}{\rho_1}+\frac{1-\omega}{\rho_2}}=\frac{\rho_1\rho_2}{\rho_1+(\rho_2-\rho_1)\,\omega} \tag{2.28}$$

where ρ_1 and ρ_2 are densities of species 1 and 2 as they are being pure material. In practically industrial application, the volume might be not additive. In that case, empirical formula obtained through experiments could be required for specific binary system. In the following, ω is sometimes referred to as salinity.

2.2.3 Stefan problem

2.2.3.1 General introduction

In Lagrangian description, for a given closed (mass fixed) system, the first law of thermodynamics, commonly referred to as the principle of energy conservation, can be stated as [2]

$$\Delta E_{st}^{tot}=Q-W \tag{2.29}$$

where ΔE_{st}^{tot} is the increase of the total energy stored in the system; Q is the net heat entering into the system; W is the net work done by the system. Fig. 2.2 shows the composition of energy, heat and work involved in Eq. 2.29 for a general case.

Before building the mathematical model for the Stefan problem, the following aspects must be noted,

- Essentially, the classic Stefan problem is a pure heat transfer problem, in which the variation of mechanical energy is negligible;
- For an incompressible flow, the flow work done by pressure is negligible;
- For the liquid flow, the viscous dissipation is negligible given that it is not of high speed and highly viscous;
- In the present study, it's assumed that no chemical reaction occurs in the bulk fluid. In each single phase, the only internal energy related is the sensible heat U_{sens};

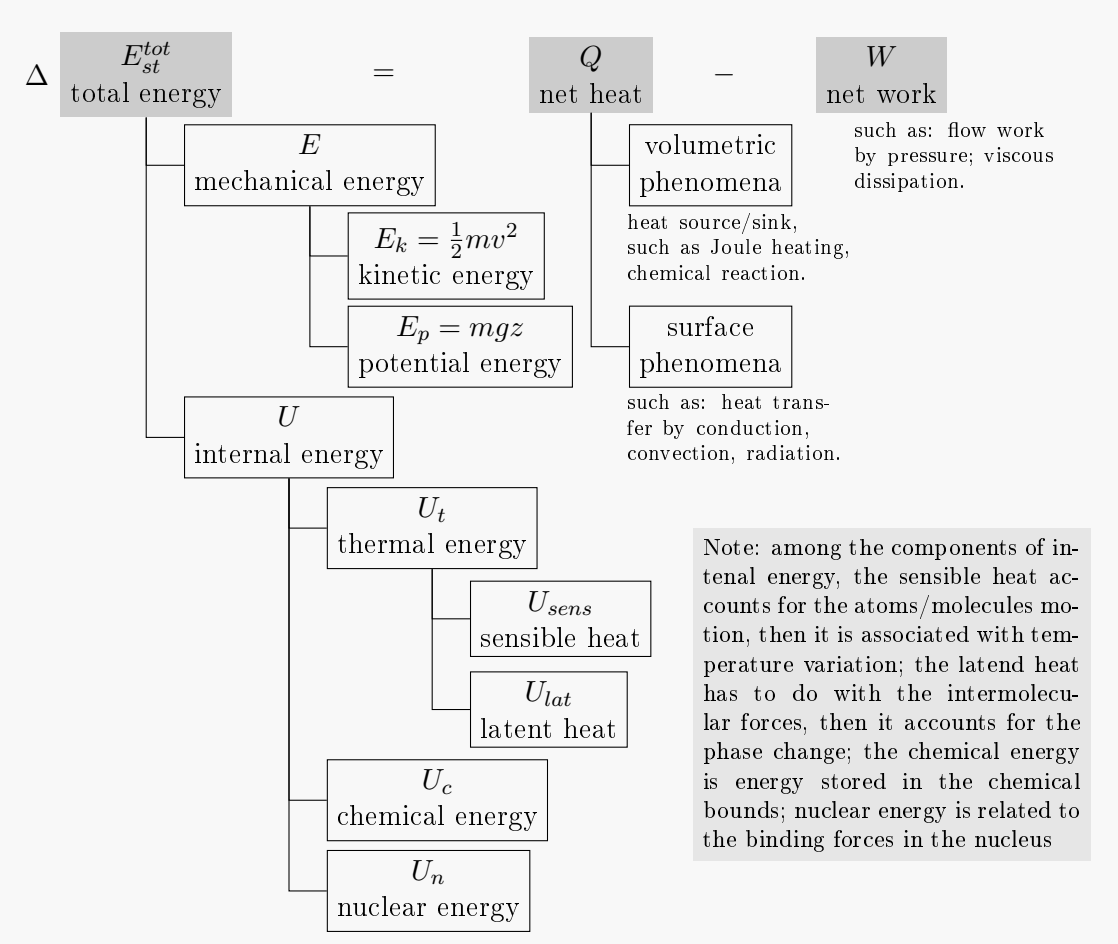


Figure 2.2 – The composition of internal energy, heat and work involved in Eq. 2.29 for a relatively general case.

- Chemical formula as shown by Eq. 1.8 is only used to show the quantitative relationship of the constituents between the solid ledge and the liquid bath. The real chemical formula in practice may be much more complicated than that shown by Eq. 1.8. Nevertheless, the chemical energy is neglected at the interface. Only the latent heat associated with phase change is taken into account as evaluating energy conservation at the interface.

Based on the above statements, energy conservation equations are built over single phase and at interface as phase change occurring, respectively.

2.2.3.2 Energy conservation equation over single phase material

Excluding the negligible portions mentioned above but the work done by the pressure, Eq. 2.29 can be written as [18]

$$\frac{d}{dt}\int_{\Omega}\rho e d\Omega = -\int_{\Gamma}\vec{q}\cdot\vec{n}d\Gamma - \int_{\Omega}p\nabla\cdot\vec{u}d\Omega + \int_{\Omega}sd\Omega \tag{2.30}$$

where e is the internal energy per unit mass (J/kg), actually only sensible heat is considered in here; $\vec{q}$ is the heat flux by conduction (W/m^2); s is the thermal source or sink (W/m^3). Utilizing Reynolds transport theorem and divergence theorem, Eq. 2.30 can be converted into the following Eulerian differential form:

$$\frac{\partial \rho e}{\partial t} + \nabla \cdot (\rho e \vec{u}) + \nabla \cdot \vec{q} + p \nabla \cdot \vec{u} - s = 0 \tag{2.31}$$

where $\vec{u}$ denotes the velocity field, it vanishes (zero value) in the solid phase. Eq.2.31 can be further expanded as

$$e\cancelto{0}{\left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u})\right)} + \rho\cancelto{\frac{de}{dt}}{\left(\frac{\partial e}{\partial t} + \vec{u} \cdot \nabla e\right)} + \nabla \cdot \vec{q} + \cancel{p \nabla \cdot \vec{u}}^{-\frac{p}{\rho}\frac{d\rho}{dt}} - s = 0 \tag{2.32}$$

After introducing the mass continuity law and definition of material derivative, Eq. 2.32 turns into

$$\rho \frac{de}{dt} + \nabla \cdot \vec{q} - \frac{p}{\rho}\frac{d\rho}{dt} - s = 0 \tag{2.33}$$

If the pressure is taken to be constant, the following equality holds:

$$-\frac{p}{\rho}\frac{d\rho}{dt} = p\rho \frac{d}{dt}\left(\frac{1}{\rho}\right) = \rho \frac{d}{dt}\left(\frac{p}{\rho}\right) \tag{2.34}$$

Then, Eq. 2.33 can be written as

$$\rho \frac{d}{dt}\left(e + \frac{p}{\rho}\right) + \nabla \cdot \vec{q} - s = 0 \tag{2.35}$$

where the quantity $e + \frac{p}{\rho}$ is the enthalpy per unit mass, which is denoted by h. To obtain the final energy conservation equation in terms of temperature, one state equation evaluating the enthalpy $h(T, p)$ and one constitutive law evaluating the flux by conduction $\vec{q}(T)$ are required. For constant pressure processes, the state equation to evaluate h can be expressed as [18]

$$\frac{dh}{dt} = \cancel{\left(\frac{\partial h}{\partial T}\right)_p}^{C_p}\frac{dT}{dt} + \left(\frac{\partial h}{\partial p}\right)_T \cancelto{0}{\frac{dp}{dt}} = C_p \frac{dT}{dt} \tag{2.36}$$

where C_p is the specific heat capacity at constant pressure ($J/\left(kg \cdot K\right)$). The conduction flux $\vec{q}$ is evaluated by Fourier's law as

$$\vec{q} = -k \nabla T \tag{2.37}$$

where k is the thermal conductivity ($W/\left(m \cdot K\right)$). Introducing Eq. 2.36 and Eq. 2.37 into Eq. 2.33, one can obtain

$$\rho C_p \frac{\partial T}{\partial t} + \rho C_p \vec{u} \cdot \nabla T - \nabla \cdot (k \nabla T) - s = 0 \tag{2.38}$$

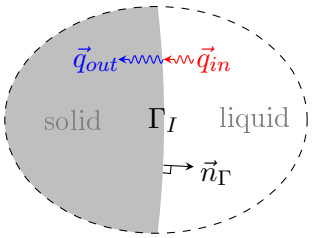


Figure 2.3 – Schematic view of region near the interface for a liquid/solid phase change process; as the heat outflow greater than the heat inflow at the interface (i.e. $|\vec{q}_{out} \cdot \vec{n}_\Gamma| > |\vec{q}_{in} \cdot \vec{n}_\Gamma|$), solidification occurs.

2.2.3.3 Energy conservation equation at interface as phase change occurring

During the phase change process of liquid/solid system, as shown in Fig. 2.3, the latent heat is released or absorbed at the interface. Therefore, the heat outflow at the solid side of the interface (denoted by Γ_I^-) is different from the heat inflow at the liquid side (denoted by Γ_I^+), i.e. heat flux jump (denoted by $[\![q]\!]$) at the interface exists.

In mathematics, the energy conservation at the interface can be expresses as [14]

$$\underbrace{\int_{\Gamma_I} [\![q]\!] d\Gamma}_{\int_{\Gamma_I^+} \vec{q}\cdot\vec{n}_\Gamma d\Gamma - \int_{\Gamma_I^-} \vec{q}\cdot\vec{n}_\Gamma d\Gamma} = \int_{\Gamma_I} \rho_s L V_I d\Gamma \tag{2.39}$$

where $\vec{n}_\Gamma$ is the unit liquid-pointing normal of the interface; L is the specific latent heat (J/kg); V_I is the interface normal speed (m/s); ρ_s is the density of the solid material. No matter the liquid phase and solid phase share the same density or not, the interface velocity always coincides with the velocity of solid front due to the immobility of the solid phase. In contrast, the mobility of the liquid phase will lead to a flow at the interface if the density is different between the two phases. In Eq. 2.39, V_I can be regarded either as interface velocity or as solid front velocity. For that reason, the density of solid is used as evaluating the amount of material that changes its phase.

Removing the integral operation and meanwhile evaluating the heat flux $\vec{q}$ by the Fourier's law, Eq. 2.39 can be written into the following differential form:

$$k_s \left.\frac{\partial T}{\partial n}\right|_{\Gamma_I^-} - k_l \left.\frac{\partial T}{\partial n}\right|_{\Gamma_I^+} = \rho_s L V_I \tag{2.40}$$

where subscripts s and l represent solid and liquid, respectively; $\frac{\partial}{\partial n}$ denotes derivative with respect to the normal direction.

2.2.4 Couplings at the interface

2.2.4.1 Between interface velocity and liquid velocity at the interface

The density may varies between different phases. As phase change occurs, variational density between liquid and solid phases would lead to volume expansion or contraction. It implies a fluid flow is induced in the liquid region. For instance, in the case $\rho_s > \rho_l$, as shown in Fig. 2.4, liquid compensation flow happens. The mass conservation law at the interface gives [9]

$$\rho_s V_I = \rho_l V_I - \rho_l \vec{u} \cdot \vec{n}_\Gamma \tag{2.41}$$

Then, the normal component of liquid velocity at the interface can be calculated by

$$\vec{u} \cdot \vec{n}_\Gamma = \left(1 - \frac{\rho_s}{\rho_l}\right) V_I \tag{2.42}$$

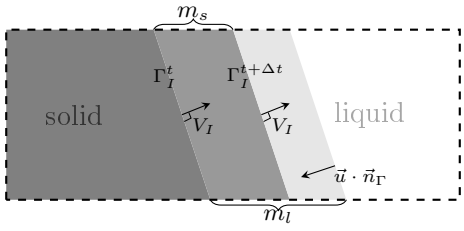


Figure 2.4 – Control volume analysis of phase change process in the case $\rho_s > \rho_l$: as solid front moves from Γ_I^t to $\Gamma_I^{t+\Delta t}$, liquid compensation happens with front velocity $\vec{u} \cdot \vec{n}_\Gamma$; $m_l = m_s$ denote the mass changing from liquid state to solid state.

2.2.4.2 Between interface velocity and concentration flux at the interface

For the multi-component solution, the chemical compositions on both sides of the interface are usually different. It means, as phase change occurs, some components would be released or absorbed by the interface. For a binary solution, the solute concentration at the interface must satisfy the following mass conservation law [9]:

$$-\rho_l D_l \left.\frac{\partial \omega}{\partial n}\right|_{\Gamma_I^+} + \rho_s D_s \left.\frac{\partial \omega}{\partial n}\right|_{\Gamma_I^-} = \rho_s V_I (\omega_l - \omega_s) \tag{2.43}$$

where the terms on the left hand side represent the solute mass flux jump. In practical applications, ω_s and ω_l are usually linked linearly by expression $\omega_s = k_p \omega_l$, where k_p is the partition coefficient determined according to the phase diagram [17].

2.3 Level set method

2.3.1 Definition

In level set method, instead of utilizing some makers to trace the discontinuity, the discontinuity is embedded into an one-dimension-raised scalar function, called level set function and

usually denoted by ϕ. The discontinuity is implicitly described as the zero-level of ϕ, which is defined as the shortest signed distance function. As shown in Fig. 2.5, considering a domain

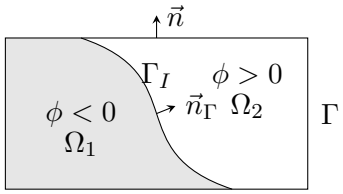


Figure 2.5 – Discontinuity description by level set method.

Ω divided by a sharp interface Γ_I such that $\Omega_1 \cup \Omega_2 = \Omega$ and $\Omega_1 \cap \Omega_2 = \Gamma_I$, the definition of level set function can be expressed as [13]

$$\phi(\vec{x},t) = sign\big((\vec{x}-\vec{x}_I)\cdot\vec{n}_\Gamma\big)\min_{\vec{x}_I\in\Gamma_I}\|\vec{x}-\vec{x}_I\| \qquad \vec{x}\in\Omega \tag{2.44}$$

where the unit normal of the interface $\vec{n}_\Gamma$ points from Ω_1 into Ω_2, such that

$$\phi(\vec{x},t)\begin{cases} <0 & \forall\vec{x}\in\Omega_1 \\ =0 & \forall\vec{x}\in\Gamma_I \\ >0 & \forall\vec{x}\in\Omega_2 \end{cases} \tag{2.45}$$

Besides locating the discontinuity, the level set method can provide other information related to the discontinuity profile, such as the unit normal $\vec{n}_\Gamma$ and the curvature κ:

$$\vec{n}_\Gamma = \frac{1}{\|\nabla\phi\|}\nabla\phi \tag{2.46}$$

$$\kappa = \nabla\cdot\vec{n}_\Gamma = \nabla\cdot\left(\frac{1}{\|\nabla\phi\|}\nabla\phi\right) \tag{2.47}$$

where $\|\nabla\phi\| = 1$ only holds as ϕ being signed distance function.

2.3.2 Initialization by the closest point projection (CPP) method

The initialization of level set function requires to calculate the shortest signed distance from each element node (it can be named slave point, denoted by $\vec{x}_s$) to the interface Γ_I. To this end, the interface should be firstly discretized in space. Each segment represents one element, denoted by Γ^e. Based on the newly built one-dimension-reduced mesh, one unique element node which is closest to the slave point can be found. Then the elements connected to that node are labelled as active elements. Looping over all the active elements, the point (called master point, denoted by $\vec{x}_m$) in each active element (called master segment, denoted by Γ^e_m) closest to the slave point is supposed to be identified. The parametric coordinate of the master point on the master segment is located by minimizing a squared distance function. The formulation of the minimization is often referred to as the closest point projection [11]. Fig. 2.6 presents schematically how the closest point projection works.

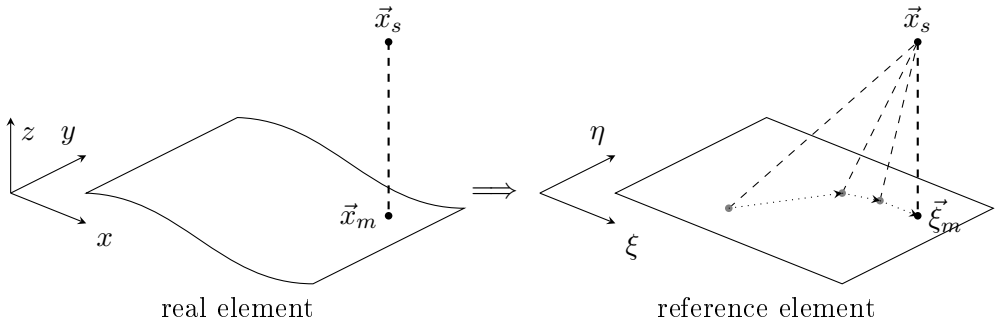


Figure 2.6 – Schematic diagram of closest point projection by Newton-Raphson method.

For any interface element with n^e nodes, the geometric approximation reads

$$\vec{x}(\xi) = [N]\{x\} \tag{2.48}$$

where

$$[N] = \begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 & \cdots & N_{n^e} & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 & \cdots & 0 & N_{n^e} & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 & \cdots & 0 & 0 & N_{n^e} \end{bmatrix} \tag{2.49a}$$

$$\{x\} = \langle x_1 \quad y_1 \quad z_1 \quad x_2 \quad y_2 \quad z_2 \quad \cdots \quad x_{n^e} \quad y_{n^e} \quad z_{n^e} \rangle^{\mathrm{T}} \tag{2.49b}$$

The squared distance function is defined as

$$d(\xi) := \frac{1}{2}(\vec{x} - \vec{x}_s) \cdot (\vec{x} - \vec{x}_s) \tag{2.50}$$

The minimum of the squared distance function occurs at the stationary point, where the gradient of the squared distance function is supposed to be zero, as expressed below:

$$\nabla\vec{x} \cdot (\vec{x} - \vec{x}_s) = 0 \tag{2.51}$$

or

$$\underbrace{\nabla\vec{x}^{\mathrm{T}}}_{2\times3}\underbrace{(\vec{x} - \vec{x}_s)}_{3\times1} = 0 \tag{2.52}$$

Eq. 2.51 or Eq. 2.52 is non-linear. In the present work, Newton-Raphson method is adopted to do linearisation. The solution of Eq. 2.51 or Eq. 2.52 can be found in appendix B.

In FESh++, during the initialization, the results at all nodes with DOF ϕ are stored in std::map<const FEShNode*, FEShVect3D> FEShDiscontinuity::*aMapNodeIntProjCoor*, which is filled by FEShDiscontinuity::computePhiInitialValue(). Note that if the foot of perpendicular falls outside of all the inversely connected elements, the element node on the newly built one-dimension-reduced mesh, which is closest to the slave point, is taken as the closest point to the slave point.

An example, that a circle-shaped discontinuity resides in the center of an unit square domain, is presented here. The radius of the circle is $\frac{\sqrt{2}}{6}$. Fig. 2.7 shows the mesh used to calculate

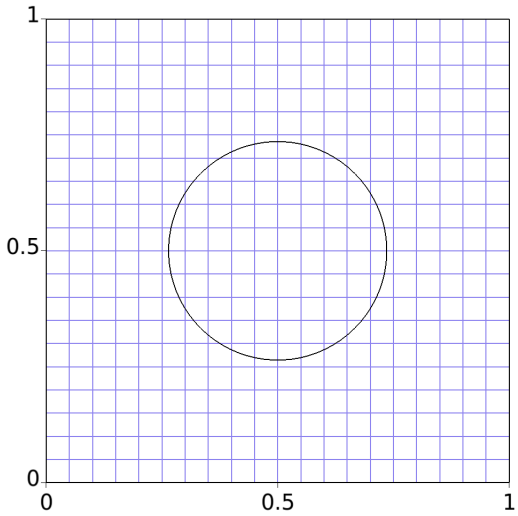
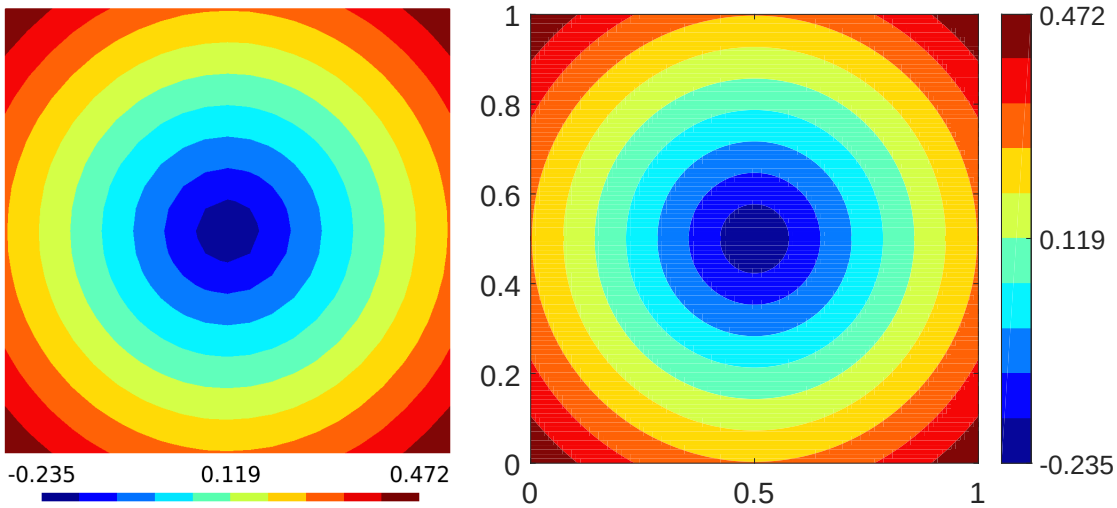


Figure 2.7 – The mesh used to calculate the shortest signed distance to the circle discontinuity by the CPP method.



(a) by CPP (b) analytical solution

Figure 2.8 – The shortest signed distance to the circle discontinuity.

the shortest signed distance to the circle discontinuity by the CPP method. Fig. 2.8a presents the result obtained by CPP method, which has a great agreement with the analytical solution as shown in Fig. 2.8b. The maximum relative error at element nodes is 0.35%.

Initialization of level set function through the CPP method mainly has the following two advantages.

- The discontinuity is not required to be described by a mathematical equation;
- The treatment is in the same manner no matter the dimensionality and no matter for straight lines or curves.

In FESh++, in the later steps after the initialization of ϕ, FEShDiscontinuity::*aMapNodeIntProjCoor*, which only contains results at the nodes belonging to the cut elements, is filled in FEShDiscontinuity::findNodeIntProjection(), following FEShDiscontinuity::findElementIntersection(). The algorithm to find out the foot of perpendicular from the nodes with DOF ϕ and also belonging to the cut elements to the discontinuity can be summarized as,

1. FEShBool lFound = false;
2. Loop over all the inversely connected elements (including elements on the periphery) of current node until lFound = true;
   a) Check up on the number of intersection points between current element and the discontinuity;
   b) If number of intersections == 1;
      i. lFound = true
      ii. Check if the intersection just happens on the node under evaluation;
      iii. If yes, the node under evaluation sits on the discontinuity;
      iv. If no, loop over all the elements containing this intersection, then only work on the element containing two intersections, calculate the distance from the node under evaluation to the segment built up by the two intersections through the package *geometry*[1]. Note that the foot of perpendicular can be outside of the segment;
   c) If number of intersections == 2, all the inversely connected elements with two intersections are examined. On each element, the distance from the node under evaluation to the segment built up by the two intersections is calculated through *geometry*. The foot of perpendicular can be outside of the segment. The shortest one is chosen and lFound = true;
   d) More than 2 intersections in one element is regarded as error.

2.3.3 Normal speed spreading

For the moving discontinuity, its motion is accounted for by the transport of level set function. The transport equation to update level set function reads [13]

$$\frac{\partial \phi}{\partial t} + \vec{v} \cdot \nabla \phi = 0 \tag{2.53}$$

where $\vec{v}$ denotes the external velocity field to advect ϕ. There are two situations according to how to define the external velocity field $\vec{v}$. In some problems, like two-phase flow, the interface moves along with the fluid field, then the fluid's velocity field can be directly used as $\vec{v}$ in Eq. 2.53. In that case, however, because the fluid particles in each phase may not move in the same pace with those at the interface Γ_I, ϕ would lose the property of being distance function quickly, therefore it requires reinitialization after each or a couple of time steps. In some other problems, like Stefan problem, the motion of the discontinuity doesn't adhere to the fluid field, which implies that the fluids' velocity field cannot be used to advect ϕ. Actually, in both situations, the external velocity field $\vec{v}$ can be constructed completely based on the interface velocity, such that the motion of ϕ everywhere can keep in step with that at the interface.

[1] http://www.softsurfer.com

Eq. 2.53 is often written in the following form:

$$\frac{\partial \phi}{\partial t} + F \parallel \nabla \phi \parallel = 0 \tag{2.54}$$

where $F = \vec{v} \cdot \frac{\nabla \phi}{\|\nabla \phi\|}$ is the normal speed field. Then, the problem turns into construction of F by spreading the interface normal speed V_I on the whole domain. The construction of F can be governed by [4]

$$sign(\phi) \nabla F \cdot \nabla \phi = 0 \tag{2.55}$$

with the input boundary condition:

$$F = V_I \qquad \text{on} \quad \Gamma_I \tag{2.56}$$

As indicated by Eq. 2.55, F is constructed to be orthogonal to ϕ, which minimizes the deviation of ϕ from the distance function [4]. Eq. 2.55 can be solved by finite element method with stabilization technique based on the identical mesh with level set function.

Since the discontinuity is not required to be aligned with the element edge, the imposition of boundary condition (Eq. 2.56) at the interface is not straightforward. For simplicity, Eq. 2.56 is applied directly to the nodes whose element is cut by Γ_I. In addition, when the discontinuity is open (not closed), it requires to apply Eq. 2.56 on the external boundary nodes satisfying the following condition to make the problem well-posed [4]:

$$(sign(\phi) \nabla \phi) \cdot \vec{n} < 0 \tag{2.57}$$

where $\vec{n}$ is the outward normal unit of the periphery of the domain.

The procedure to find out the boundary nodes and also their F value for the normal speed spreading in FESh++ is summarized below,

1. Loop over all the peripheral nodes to find out those satisfying Eq 2.57, i.e. the nodes denoted by empty red circle in Fig. 2.9. Store them in a container std::vector<const FEShNode*> *lBdyNodeF*;
2. Loop over all the nodes belonging to the cut elements, i.e. the nodes denoted by black-filled circle in Fig. 2.9, but not including those in step 1, to find out the projection point on the discontinuity and normal speed at that location. If the projection point is outside of the computational domain, the corresponding node is stored in the container *lBdyNodeF*;
3. The leading nodes, marked by black-filled circle and red empty circle simultaneously in Fig. 2.9, are selected out. On each leading node, loop over the inversely connected element (only on the periphery) of current leading node to find out the element containing location $\phi = 0$. Evaluate the normal speed at the location $\phi = 0$ and assign the result to the leading node, and also to the remaining nodes (on the periphery, marked by red empty circle);
4. On some internal nodes which are also stored in *lBdyNodeF* at step 2, the value on the leading node is applied.

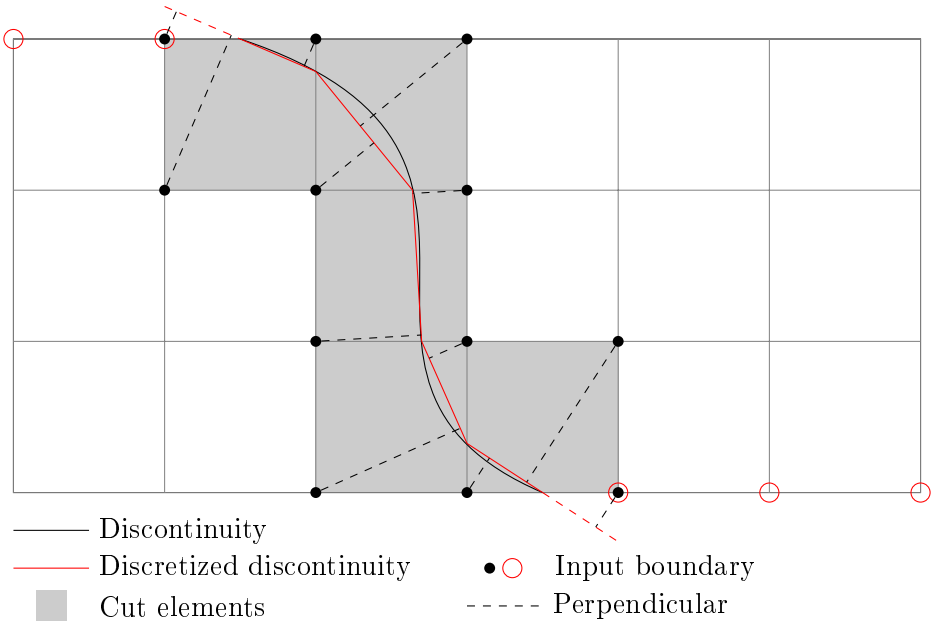


Figure 2.9 – Evaluation of boundary condition and identification of boundary nodes for the normal speed spreading based on the mesh made up of bi-linear quadrangular elements.

2.4 Extended finite element method

2.4.1 Approximation function

In contrast to the continuum-based finite element method (FEM), the extended finite element method (XFEM) exists to handle the discontinuity problem by enriching the approximation function locally. The XFEM approximation of a variable $u(\vec{x},t)$ with one discontinuity can be written as [8]

$$u^h(\vec{x},t) = \underbrace{\sum_{i\in\mathcal{N}} N_i(\vec{x})u_i(t)}_{\text{strd. FEM approx.}} + \underbrace{\sum_{i\in\mathcal{N}_e} M_i(\vec{x},t)a_i(t)}_{\text{enrichment}} \tag{2.58}$$

where u^h denotes the approximated value, N_i is the FEM shape function of node i, u_i is the unknown quantity at node i, M_i is the local enrichment function, a_i is the additional degree of freedom, $\mathcal{N}$ is the set of total nodes and $\mathcal{N}_e$ is the collection of enriched nodes. As shown in Fig.2.10, only the nodes belonging to the cut elements are enriched. The local enrichment

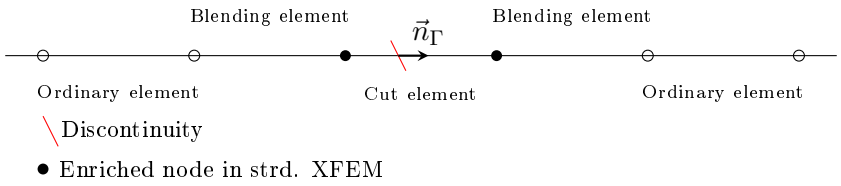


Figure 2.10 – One-dimensional mesh with discontinuity.

function can be further expanded as [8]

$$M_i(\vec{x},t) = N_i(\vec{x})\psi_i(\vec{x},t) \tag{2.59}$$

where N_i is also the FEM shape function but it is not necessary to take the same one with that in Eq.2.58, ψ_i is the global enrichment function of node i, which determines the discontinuity type of the approximation space. As shown in Fig. 2.11, the function that itself has a jump at

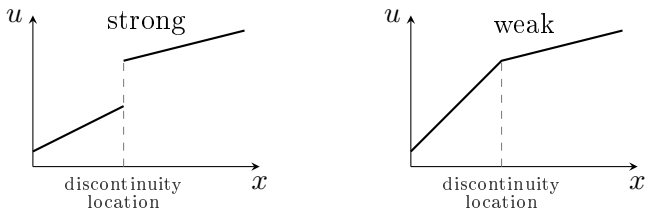


Figure 2.11 – Two types of discontinuity: strong and weak (or kink).

the interface is of strong discontinuity, while the function only whose gradient has a jump is of weak discontinuity. In general, the discontinuity type of a physical quantity is known a prior. Then it requires to choose proper enrichment function to build up the matched approximation space.

For strong discontinuity, the enrichment function usually adopts the sign-enrichment scheme (or Heaviside-enrichment scheme, almost no difference in between), which reads [8]

$$\psi_i(\vec{x},t) = sign(\phi(\vec{x},t)) - sign(\phi_i(t)) \tag{2.60}$$

where ϕ is the level set function. Fig. 2.12 presents the corresponding local enrichment functions in the cut element. Note that the local enrichment functions in sign-enrichment scheme vanish in the blending elements.

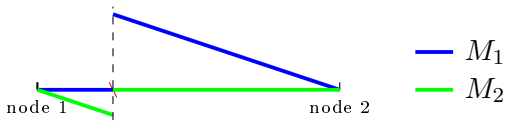


Figure 2.12 – Local enrichment functions of sign-enrichment scheme in the one-dimensional cut element.

For weak discontinuity, the abs-enrichment scheme below is usually employed [8]:

$$\psi_i(\vec{x},t) = |\phi(\vec{x},t)| - |\phi_i(t)| \tag{2.61}$$

As shown in Fig. 2.13, Eq. 2.61, often referred to as standard abs-enrichment scheme, can build an approximation space with an kink at the discontinuity location. However, due to the fact that the blending elements are only partially enriched, the property of partition of unity in the blending elements is no longer guaranteed, which is usually referred as blending element

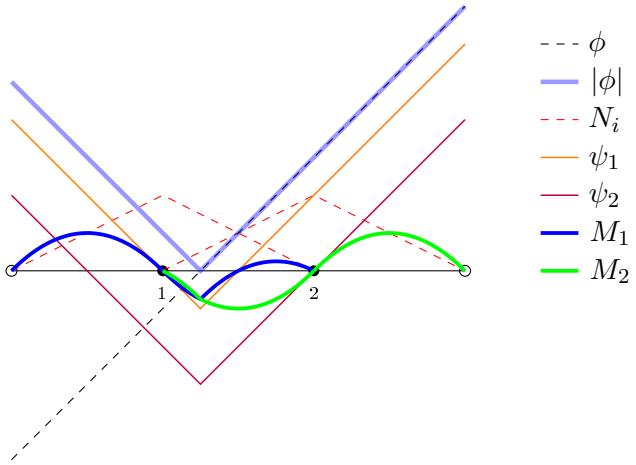


Figure 2.13 – Application of abs-enrichment scheme in cut and blending elements.

problem [5]. In the literature, there have been several remedies to overcome this problem. One of them is to use the following so-called modified abs-enrichment scheme [12]:

$$\psi(\vec{x},t) = N_j(\vec{x})|\phi_j(t)| - |N_j(\vec{x})\phi_j(t)| \tag{2.62}$$

where Einstein notation is applicable and the subscript j indicates the nodes that belong to the element containing point $\vec{x}$. As shown in Fig. 2.14, compared to the standard abs-enrichment

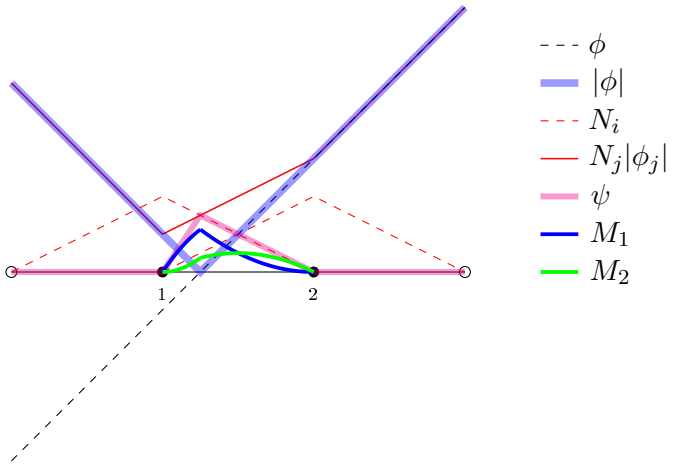


Figure 2.14 – Application of the modified abs-enrichment scheme in cut and blending elements.

scheme, the modified scheme has changed the global enrichment function dramatically such that the local enrichment functions in the blending elements disappear completely. Another solution of the blending element problem is to multiply the standard scheme by a ramping function R [7], as shown below:

$$\psi_j(\vec{x},t) = (|\phi(\vec{x},t)| - |\phi_j(t)|)R = (|\phi(\vec{x},t)| - |\phi_j(t)|)\sum_{i\in\mathcal{N}_e} N_i(\vec{x}) \tag{2.63}$$

where $j \in \mathcal{N}_{e,expanded}$, which is augmented $\mathcal{N}_e$ and includes all the nodes belonging to either cut elements or blending elements. Eq. 2.63 is often referred to as corrected abs-enrichment scheme. As shown in Fig. 2.15, the ramping function R only functions in the blending elements. The approximation space in the cut element is identical to that in the standard abs-enrichment scheme. In [3], the modified and corrected abs-enrichment schemes are compared

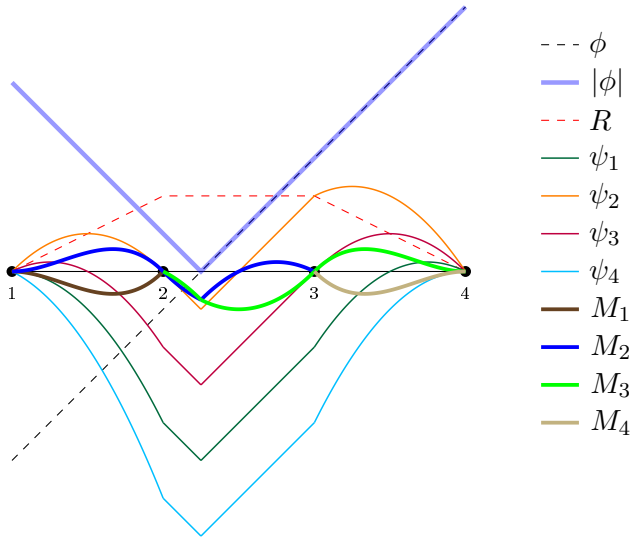


Figure 2.15 – Application of the corrected abs-enrichment scheme in cut and blending elements.

in terms of the convergence rate. It concludes that the modified scheme can yield suboptimal convergence rate while the corrected scheme can yield close-to-optimal convergence rate. Another way to deal with the weak discontinuity without blending element problem is to employ the sign-enrichment scheme, but it requires to enforce the continuity constraint at the interface.

2.4.2 Numerical integration

In the context of FEM, performing weak formulation and discretization on the governing equation would generate the following system in integral form:

$$\sum_e \int_{\Omega^e} f(\vec{x})\, d\vec{x} = 0 \tag{2.64}$$

Since the analytical integration of Eq. 2.64 is not always efficient or not even feasible, the standard practice is to do the mapping from the real element to the reference element first and then perform numerical integration in the reference element. This procedure can be described by the following equations:

$$\sum_e \int_{\Omega^e} f(\vec{x})\, d\vec{x} = \sum_e \int_{\Omega^e} f\left(\vec{\xi}\right) |J| d\vec{\xi} = \sum_e \left(\sum_i f\left(\vec{\xi}_i\right) |J| W_i \right)^e \tag{2.65}$$

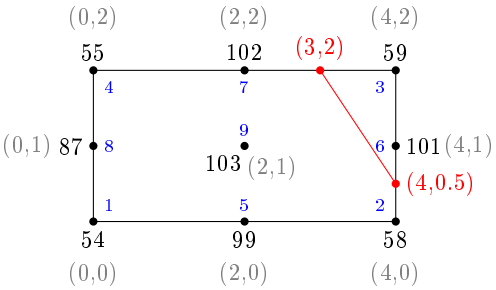
where $|J|$ denotes the determinant of Jacobian matrix J of the geometrical transformation; ξ_i indicates the location of the integration point i in the reference element; W_i is the weight at point i. In general, the numerical integration requires the integrand f to be smooth. But the real situation is that the solution space of XFEM in the cut elements is always not smooth. In order to make the integration algorithm still applicable in the cut elements, these elements must be split into pieces such that in each piece the integrand is smooth. After element splitting, each newly generated sub-element can be handled as the normal element in FEM. Then according to the mapping relationship, integration points for the original element and the corresponding weights can be obtained. For the cut element, numerical integration can be expressed as:

$$\begin{aligned}
&\int_{\Omega^e} f(\vec{x})\, d\vec{x} \\
&= \int_{\Omega^e} f\left(\vec{\xi}\right) |J| d\vec{\xi} \\
&= \sum_{e^{sub}} \int_{\Omega^{e^{sub}}} f\left(\vec{\xi}\right) |J| d\vec{\xi} \\
&= \sum_{e^{sub}} \int_{\Omega^{e^{sub}}} f\left(\vec{\xi}^{2nd}\right) |J||J^{2nd}| d\vec{\xi}^{2nd} \\
&= \sum_{e^{sub}} \left(\sum_j f\left(\vec{\xi}_j^{2nd}\right) |J||J^{2nd}| W_j^{2nd} \right)^{e^{sub}} \\
&= \sum_i f\left(\vec{\xi}_i\right) |J| W_i
\end{aligned} \tag{2.66}$$

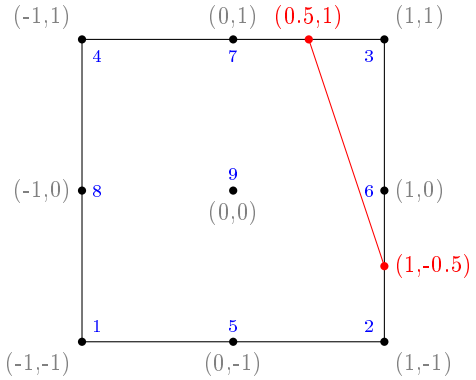
where the quantities with superscript $2nd$ represent quantities related to the second mapping.

In FESh++, the class named FEShIntgGroupEnr is derived from the class FEShIntgGroup to contain both standard integration algorithm and also that designed for the cut elements. Once the element is identified as cut element, FEShIntgEnr::generateIntg() is invoked to split the element into triangles with the aid of the package *Triangle*[2]. The procedure to obtain the integration points and corresponding weights in the cut element can be illustrated with Fig. 2.16. Suppose that one quadrangular element (9-node, Q_2) with connectivities {54, 58, 59, 55, 99, 101, 102, 87, 103}, as shown in Fig. 2.16a, is intersected with the discontinuity at two locations (4,0.5) and (3,2). In its reference element, a sub-mesh consisting of 4 triangular elements and 6 element nodes can be obtained, as shown in Fig. 2.16b and Fig. 2.16c. Based on the 3-point integration scheme for triangle presented in Fig. 2.16d, the integration points (denoted by $\times$ in Fig. 2.16c) and corresponding weights obtained through Eq. 2.66 are $\vec{\xi}_1(-\frac{3}{4}, -\frac{1}{3})$, $\vec{\xi}_2(0, \frac{2}{3})$, $\vec{\xi}_3(-\frac{3}{4}, \frac{2}{3})$, $\vec{\xi}_4(\frac{1}{3}, \frac{5}{12})$, $\vec{\xi}_5(-\frac{5}{12}, -\frac{7}{12})$, $\vec{\xi}_6(\frac{7}{12}, -\frac{1}{3})$, $\vec{\xi}_7(\frac{2}{3}, -\frac{11}{12})$, $\vec{\xi}_8(\frac{2}{3}, -\frac{2}{3})$, $\vec{\xi}_9(-\frac{1}{3}, -\frac{11}{12})$, $\vec{\xi}_{10}(-\frac{11}{12}, 0)$, $\vec{\xi}_{11}(\frac{11}{12}, \frac{3}{4})$, $\vec{\xi}_{12}(\frac{2}{3}, \frac{3}{4})$, $W_1 = \frac{1}{2}$; $W_2 = \frac{1}{2}$; $W_3 = \frac{1}{2}$; $W_4 = \frac{13}{24}$; $W_5 = \frac{13}{24}$; $W_6 = \frac{13}{24}$; $W_7 = \frac{1}{6}$; $W_8 = \frac{1}{6}$; $W_9 = \frac{1}{6}$; $W_{10} = \frac{1}{8}$; $W_{11} = \frac{1}{8}$; $W_{12} = \frac{1}{8}$.

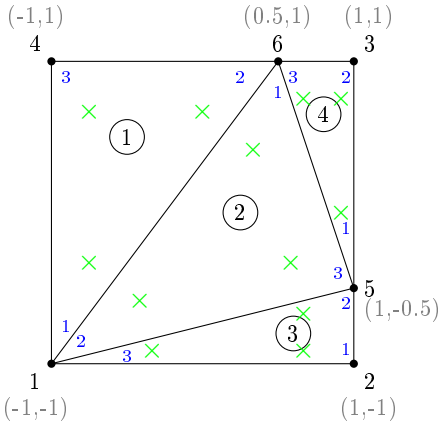
[2] https://www.cs.cmu.edu/~quake/triangle.html



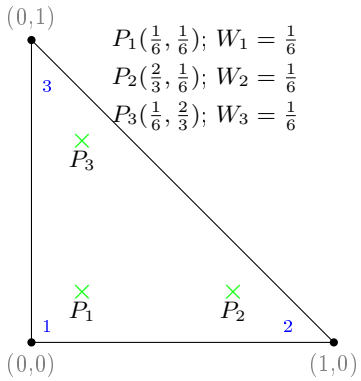
(a) One real cut element



(b) The reference element for the first mapping



(c) Newly generated elements in the reference element of the first mapping through *Triangle*



(d) The reference element with 3 integration points for the second mapping

Figure 2.16 – Triangulation of one sample cut element as well as identification of integration points and weights.

Bibliography

[1] G. K. Batchelor. *An introduction to fluid dynamics*. Cambridge University Press, 2000.

[2] T. L. Bergman, F. P. Incropera, and A. S. Lavine. *Fundamentals of heat and mass transfer*. John Wiley & Sons, 2011.

[3] K. W. Cheng and T. P. Fries. Higher-order XFEM for curved strong and weak discontinuities. *International Journal for Numerical Methods in Engineering*, 82(5):564–590, 2010.

[4] J. Chessa, P. Smolinski, and T. Belytschko. The extended finite element method (XFEM) for solidification problems. *International Journal for Numerical Methods in Engineering*, 53(8):1959–1977, 2002.

[5] J. Chessa, H. Wang, and T. Belytschko. On the construction of blending elements for local partition of unity enriched finite elements. *International Journal for Numerical Methods in Engineering*, 57(7):1015–1038, 2003.

[6] J. Donea and A. Huerta. *Finite element methods for flow problems*. John Wiley & Sons, 2003.

[7] T. P. Fries. A corrected XFEM approximation without problems in blending elements. *International Journal for Numerical Methods in Engineering*, 75(5):503–532, 2008.

[8] T. P. Fries and T. Belytschko. The extended/generalized finite element method: an overview of the method and its applications. *International Journal for Numerical Methods in Engineering*, 84(3):253–304, 2010.

[9] M. E. Glicksman, S. R. Coriell, and G. B. McFadden. Interaction of flows with the crystal-melt interface. *Annual Review of Fluid Mechanics*, 18(1):307–335, 1986.

[10] D. D. Gray and A. Giorgini. The validity of the Boussinesq approximation for liquids and gases. *International Journal of Heat and Mass Transfer*, 19(5):545–551, 1976.

[11] J. Kopačka, D. Gabriel, J. Plešek, and M. Ulbin. Assessment of methods for computing the closest point projection, penetration, and gap functions in contact searching problems. *International Journal for Numerical Methods in Engineering*, 105(11):803–833, 2016.

[12] N. Moës, M. Cloirec, P. Cartraud, and J. F. Remacle. A computational approach to handle complex microstructure geometries. *Computer Methods in Applied Mechanics and Engineering*, 192(28):3163–3177, 2003.

[13] S. Osher and R. P. Fedkiw. Level set methods: an overview and some recent results. *Journal of Computational physics*, 169(2):463–502, 2001.

[14] M. N. Ozisik. *Heat conduction*. John Wiley & Sons, 1993.

[15] G. S. H. Pau, J. B. Bell, K. Pruess, A. S. Almgren, M. J. Lijewski, and K. Zhang. High-resolution simulation and characterization of density-driven flow in CO_2 storage in saline aquifers. *Advances in Water Resources*, 33(4):443–455, 2010.

[16] J. C. Slattery. *Advanced transport phenomena*. Cambridge University Press, 1999.

[17] M. Theillard, F. Gibou, and T. Pollock. A sharp computational method for the simulation of the solidification of binary alloys. *Journal of Scientific Computing*, 63(2):330–354, 2015.

[18] S. Whitaker. *Fundamental principles of heat transfer*. Elsevier, 1976.

[19] R. K. Zeytounian. Joseph Boussinesq and his approximation: a contemporary view. *Comptes Rendus Mecanique*, 331(8):575–586, 2003.

Chapter 3

Finite element analysis of mass transport in a binary miscible solution of variable density

Résumé

Dans ce travail, un modèle bidimensionnel, basée sur la méthode des éléments finis et couplant les équations de Navier-Stokes et un problème de convection-diffusion, est développé pour étudier le transport de masse dans une solution composée de deux fluides immiscibles ou dans des solutions possédant des densités distinctes. Les non-linéarités sont induites par le terme convectif des équations de Navier-Stokes et la variation de la densité en fonction de la fraction massique. Les formes conservative et non-conservative de l'équation de convection-diffusion sont comparées en termes de la conservation de la masse. La méthode de substitution et l'algorithme de Newton-Raphson sont utilisés pour linéariser les systèmes d'équations et leurs performances sont comparées. Les simulations portent sur le problème dit « the lock exchange problem ».

Abstract

The subject concerned with in the present work is the mass transport phenomenon that happens in the solution made of two miscible fluids or solutions possessing distinct densities. The mass transport process is governed by two mechanisms, i.e. the relative motion in between species and the convection along with the flow of the mixture. In the present work, a two-dimensional finite element model is established to describe this process numerically. To this end, two sets of equations, including the incompressible Navier-Stokes equations and the transient convection-diffusion equation, are employed and coupled. The former governs the global mixture flow, while the latter accounts for the motion of one species, in which the

mass fraction of the denser species is taken as the unknown (mass-fraction-based). In variable density case, the conservative and non-conservative forms of the convection-diffusion equation are compared in terms of mass conservation. In addition, the coupling between these two sets of equations is fulfilled by two algorithms, corresponding to weak and strong couplings, respectively. Non-linearity arises from the convection term in Navier-Stokes equations and the density's dependency on the mass fraction. The linearisation methods such as substitution method and Newton-Raphson method are adopted and compared in terms of robustness and efficiency. All the related calculations are based on the lock exchange problem.

Keywords: finite element method, density driven, lock exchange, mass transport, binary miscible solution

3.1 Introduction

Mass transport mechanism for the bi-component system is fundamental to help us to get deep insight into a lot of natural phenomena, such as salinity driven ocean circulation, seawater intrusion into the coastal freshwater aquifer, transport of contaminants in groundwater, sediment flow in the lakes or rivers, and so on, as well as other engineering problems or scientific concerns, for example binary alloy casting, stratified flow, etc. All these phenomena and problems really matter either to environmental health or to industrial production. Due to the significance of this subject, it was widely studied by experimental and theoretical analysis [23, 20, 19]. Based on those achievements and with the development of computers, numerical simulations on this subject have got more attentions. In the literature, several types of mathematical models exist to describe mass transport that happens in a binary solution.

Problems, like seawater intrusion, solute transport in groundwater or oil recovery, are usually modelled as flow and transport in porous medium [9, 22, 15]. The porous media flow is usually characterised by Darcy's law, which can be regarded as the average version of Stokes equation [24]. Then, the mathematical model normally comprises Darcy's equation, continuity equation and solute transport equation. The coupling between flow field and solute transport is usually two-way. In specific applications, except for Darcy's equation, the continuity equation and solute transport equation may take different forms. For the continuity equation, as the density is variable rather than constant, the complete form applicable to the compressible case was used in some published works [9, 15], while some others still used the incompressibility constraint [22]. For the solute transport within this context, it mainly has three forms: i) the conservative mass-fraction-based convection-diffusion equation [9, 13], ii) the non-conservative mass-fraction-based convection-diffusion equation [15], and iii) the general convection-diffusion equation [22]. The first two types are non-linear while the third one is linear.

Excluding situations in porous media, the Navier-Stokes (NS) equations are usually adopted to account for the fluid dynamics. Then, the mathematical model might be constructed by

combining continuity equation, NS equations and solute transport equation. When the solute concentration variation impacts the solution density weakly, Boussinesq approximation is often utilized for simplification [25, 12, 26]. Under the assumption of Boussinesq approximation, since the density keeps constant except in the gravity term, the linear general convection-diffusion equation can be used for solute transport.

Without losing generality, the present work is more interested in the models without Boussinesq approximation. The mass transport in the binary solution can also be regarded as mixing of two fluids. In the published works, the studies on two fluids mixing are mainly based upon three configurations, i.e. the lock exchange flow [11, 3], the Rayleigh-Taylor instability flow [14] and the displacement flow driven by pressure [17, 16, 18]. Even though from the perspective of physics, all these works were dedicated to model two fluids mixing, the underlying mathematical models are slightly different. The difference mainly resides in the solute transport model. In works like [5, 6], the incompressible NS equations (including the incompressibility constraint) were used to describe the hydrodynamics, while the full continuity equation of the mixture accounts for the solute transport. Density of the mixture is also taken as a primitive variable besides the velocity and pressure. This model is only applicable to immiscible fluids, because only convection process is taken into account for solute transport. It's equivalent to the two-phase flow model to some extent, but the interface is not captured or tracked. In [3], a diffusion term is added artificially to avoid the development of discontinuities at the interface, but without rigorous derivation. In some other works [17, 16, 18, 7], the linear general convection-diffusion equation or pure convection equation taking the volume fraction as the unknown was employed to govern the solute transport. Rigorously speaking, for variable density flows, the volume fraction should not be taken as the concentration unknown, as the velocity field in NS equations is mass-averaged not volume-averaged [21]. Therefore, the mass-fraction-based solute transport equation is of more rigorous form in mathematics, but which is more often seen in the context of porous flow [10].

During the process of mixing two fluids, the mass transport is governed by two mechanisms. One is the relative motion in between species (diffusion process), while the other one is the convection along with the mixture flow (convection process). The present work takes both processes into account, which implies that the two fluids are miscible. The aim of the present work is to build a model by coupling the incompressible NS equations with the mass-fraction-based solute transport equation. The consistency (all quantities are mass-averaged) and completeness are demonstrated. Within the demonstration, the assertion that volume-preserving mixing with variable density locally doesn't change the incompressibility property of the solution is made. In addition, the coupling between hydrodynamics and solute transport is fulfilled by two algorithms, corresponding to weak and strong couplings, respectively. In the case of weak coupling, the hydrodynamic and solute transport problems are solved separately until both converge. The term "separately" refers to the fact that one problem is solved while

assuming all unknowns of the other one constant. In the case of strong coupling, these two problems are solved simultaneously and only one system of algebraic equations is generated. In both cases, the generated system is of non-linearity. This non-linearity mainly stems from two parts, i.e. the convection term in momentum equation and the density's dependency on solute concentration. For linearisation, substitution method and Newton-Raphson method are respectively adopted and also compared in terms of robustness and efficiency.

This paper is organized as follows. Section 3.2 is dedicated to elaborate the mathematical model. The numerical discretization by the finite element method (FEM) in both weak and strong coupling cases as well as linearisation by substitution method and Newton-Raphson method are presented in section 3.3. In section 3.4, the pure hydrodynamic and pure solute transport sub-models in the case of constant density are first verified by the Poiseuille flow between parallel plates and mixing of two parallel streams with different concentrations, respectively. Then, for the coupled case, calculations are conducted on the lock exchange problem. Two scenarios differing in the density contrast between the two species are worked with. Finally, section 3.5 concludes this study with a summary and outlook.

3.2 Mathematical models

The bi-component system concerned with in the present work refers to the mixture of two miscible fluids or solutions possessing distinct densities, as shown in Fig. 3.1. This system

Figure 3.1 – Initial state of two fluids mixing under the lock exchange configuration confined within a rectangular cavity Ω; $\rho_1 \neq \rho_2$; $\vec{g}$ denotes the gravitational acceleration.

can be analysed from different perspectives. Firstly, the mixture can be taken as a whole with variable density. In the Lagrangian description, a mass-closed fluid parcel is defined and traced. If the volume is assumed to be conserved during mixing, the density of the fluid parcel should always keep constant. According to the definition of incompressibility [1], it implies the mixing doesn't change the fact that the fluid is incompressible. Then, the mass and momentum conservation laws based on the mixture in the mixed formulation with primitive variables velocity $\vec{u}(\vec{x},t)$, pressure $p(\vec{x},t)$) and density $\rho(\vec{x},t)$ can be written as

$$\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho = 0 \tag{3.1a}$$

$$\nabla \cdot \vec{u} = 0 \tag{3.1b}$$

$$\rho \frac{\partial \vec{u}}{\partial t} + \rho \left(\vec{u} \cdot \nabla\right) \vec{u} - \nabla \cdot \left(2\mu \underline{\underline{D}}\right) + \nabla p - \rho \vec{g} = 0 \tag{3.1c}$$

where μ is the dynamic viscosity and $\underline{\underline{D}} = \frac{1}{2}\left(\nabla\vec{u} + \nabla\vec{u}^{\mathrm{T}}\right)$ is the shear rate tensor. Eq. 3.1a and Eq. 3.1b are obtained by splitting the continuity equation, and the latter is often referred to as incompressibility constraint.

Secondly, from the perspective of each individual species, the mass conservation law also holds. If no chemical reaction is involved, the continuity equation on each species reads

$$\frac{\partial\hat{\rho}_i}{\partial t} + \nabla\cdot(\hat{\rho}_i\vec{u}_i) = 0 \tag{3.2}$$

where $\hat{\rho}_i$ denotes mass concentration of species i in the mixture and $\vec{u}_i$ is the velocity of species i. The sum of Eq. 3.2 over species i would lead to the continuity equation of the mixture. It implies the following two identities:

$$\rho \equiv \sum_i \hat{\rho}_i \tag{3.3a}$$

$$\vec{u} \equiv \frac{1}{\rho}\sum_i \hat{\rho}_i\vec{u}_i \tag{3.3b}$$

Eq. 3.3 reveals the definitions of density and velocity for multi-component solution. Eq. 3.3b means that the velocity $\vec{u}$ is mass-averaged.

In Eq. 3.2, $\hat{\rho}_i\vec{u}_i$ represents the absolute mass flux of species i with respect to the fixed reference frame. The absolute mass flux can be divided into two portions: i) the flux relative to the mixture flow and ii) the flux that is keeping pace with the mixture flow. Therefore, it can be written as

$$\hat{\rho}_i\vec{u}_i = \hat{\rho}_i(\vec{u}_i - \vec{u}) + \hat{\rho}_i\vec{u} \tag{3.4}$$

The first term on the right hand side of Eq. 3.4 indicating the relative mass flux is often approximated by the Fick's law of diffusion in the binary solution, which reads [21]

$$\hat{\rho}_i(\vec{u}_i - \vec{u}) = -\rho D\nabla\omega_i \tag{3.5}$$

where D is the diffusion coefficient, it can be either isotropic or orthotropic; ω_i is the mass fraction of species i, defined as $\omega_i = \frac{\hat{\rho}_i}{\rho}$. Introducing Eq. 3.5 and Eq. 3.4 into Eq. 3.2, one obtains

$$\frac{\partial\rho\omega_i}{\partial t} + \vec{u}\cdot\nabla(\rho\omega_i) + \nabla\cdot(-\rho D\nabla\omega_i) = 0 \tag{3.6}$$

To get Eq. 3.6, the incompressibility constraint (Eq. 3.1b) has been applied. It's easy to prove that the summation of Eq. 3.6 over all the species would produce Eq. 3.1a, as the identity $\sum_i \omega_i = 1$ always holds. Mathematically, Eq. 3.6 is a stronger form relative to Eq. 3.1a, due to the fact that the former is a sufficient but not necessary condition of the latter. Therefore, Eq. 3.1a should be substituted by Eq. 3.6 without losing any accuracy and completeness. Since $\sum_i \omega_i = 1$, only one species transport equation is sufficient to describe the variation of the solution composition for a binary system. Dropping the subscript i in Eq. 3.6, one obtains

$$\frac{\partial\rho\omega}{\partial t} + \vec{u}\cdot\nabla(\rho\omega) + \nabla\cdot(-\rho D\nabla\omega) = 0 \tag{3.7}$$

where ω is usually the mass fraction of the denser species. ω is also be referred to as salinity in the following parts. Eq. 3.7 is in the conservative form, which can be further reduced into the following non-conservative form by utilizing Eq. 3.1a:

$$\rho\frac{\partial\omega}{\partial t} + \vec{u}\cdot\rho\nabla\omega + \nabla\cdot(-\rho D\nabla\omega) = 0 \tag{3.8}$$

Finally, the system is closed by the following equation of state [13]:

$$\rho = \frac{\rho_1\rho_2}{\rho_1 + (\rho_2 - \rho_1)\omega} \tag{3.9}$$

where ρ_1 and ρ_2 denote the density of the species at pure state. Eq. 3.9 holds in the case that there is no volume change upon mixing. In some practical applications, the relationship between mixture density ρ and mass fraction ω can be obtained through experiments.

In summary, the governing equations to describe the mass transport, that occurs in the binary solution with variable density, consist of Eq. 3.1b, Eq. 3.1c, any one of Eq. 3.7 and Eq. 3.8, and Eq. 3.9. To make a well-posed problem, all or part of the following initial and boundary conditions are also required to be provided:

$$\vec{u} = \vec{u}_0, \qquad \omega = \omega_0 \;\; ; \;\; \vec{x}\in\Omega \;\; , t = 0 \tag{3.10a}$$

$$\vec{u} = \vec{u}_D, \qquad \omega = \omega_D \;\; ; \;\; \vec{x}\in\Gamma_D, t\in(0,\infty) \tag{3.10b}$$

$$(-p\underline{\underline{I}} + 2\mu\underline{\underline{D}})\cdot\vec{n} = \vec{t}_N, \quad -\rho D\nabla\omega\cdot\vec{n} = q_{\omega,n}; \;\; \vec{x}\in\Gamma_N, t\in(0,\infty) \tag{3.10c}$$

where Γ_D and Γ_N represent the Dirichlet boundaries and Neumann boundaries, respectively. All the quantities on the right hand side of Eq. 3.10 are given value.

Remarks The species transport model established based on Eq. 3.7 or Eq. 3.8 is also applicable to the case governed by the general convection-diffusion equation in which the density is assumed to be constant and the molar concentration (mol/l or mol/m^3) is the unknown. The only modification in application is to assign unity to ρ, while ω is used to represent the molar concentration.

3.3 Finite element analysis

3.3.1 Weak formulation and time discretization

The respective weak formulations of Eq. 3.1c, Eq. 3.1b, Eq. 3.7 and Eq. 3.8 give

$$\begin{aligned}
W_{\vec{u}} &= \int_\Omega \delta\vec{u}\cdot\rho\frac{\partial\vec{u}}{\partial t}d\Omega + \int_\Omega \delta\vec{u}\cdot\rho\left(\vec{u}\cdot\nabla\right)\vec{u}d\Omega + \int_\Omega \underline{\underline{D}}(\delta\vec{u}) : 2\mu\underline{\underline{D}}(\vec{u})d\Omega \\
&\quad - \int_\Omega \nabla\cdot\delta\vec{u}\,pd\Omega - \int_\Omega \delta\vec{u}\cdot\rho\vec{g}d\Omega - \int_{\Gamma_N}\delta\vec{u}\cdot\vec{t}_N d\Gamma \qquad = 0 \qquad (3.11a)\\
W_p &= \int_\Omega -\delta p\nabla\cdot\vec{u}d\Omega \qquad = 0 \qquad (3.11b)\\
W_\omega^{consv} &= \int_\Omega \delta\omega(\omega\rho' + \rho)\frac{\partial\omega}{\partial t}d\Omega + \int_\Omega \delta\omega\vec{u}\cdot(\omega\rho'+\rho)\nabla\omega d\Omega \\
&\quad + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \int_{\Gamma_N}\delta\omega q_{\omega,n}d\Gamma \qquad = 0 \qquad (3.11c)\\
W_\omega &= \int_\Omega \delta\omega\rho\frac{\partial\omega}{\partial t}d\Omega + \int_\Omega \delta\omega\vec{u}\cdot\rho\nabla\omega d\Omega + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega \\
&\quad + \int_{\Gamma_N}\delta\omega q_{\omega,n}d\Gamma \qquad = 0 \qquad (3.11d)
\end{aligned}$$

where $\delta\vec{u}$, δp, and $\delta\omega$ are the test functions and ρ' is the density's first derivative with respect to the mass fraction ω. After the backward Euler method applied for time discretization, Eq. 3.11a, Eq. 3.11c and Eq. 3.11d turn into

$$\begin{aligned}
W_{\vec{u}} &= \int_\Omega \delta\vec{u}\cdot\rho\frac{1}{\Delta t}(\vec{u} - \vec{u}^{t-\Delta t})d\Omega + \int_\Omega \delta\vec{u}\cdot\rho\left(\vec{u}\cdot\nabla\right)\vec{u}d\Omega + \int_\Omega \underline{\underline{D}}(\delta\vec{u}) : 2\mu\underline{\underline{D}}(\vec{u})d\Omega \\
&\quad - \int_\Omega \nabla\cdot\delta\vec{u}\,pd\Omega - \int_\Omega \delta\vec{u}\cdot\rho\vec{g}d\Omega - \int_{\Gamma_N}\delta\vec{u}\cdot\vec{t}_N d\Gamma \qquad = 0 \qquad (3.12a)\\
W_\omega^{consv} &= \int_\Omega \delta\omega(\omega\rho' + \rho)\frac{1}{\Delta t}(\omega - \omega^{t-\Delta t})\omega d\Omega + \int_\Omega \delta\omega\vec{u}\cdot(\omega\rho'+\rho)\nabla\omega d\Omega \\
&\quad + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \int_{\Gamma_N}\delta\omega q_{\omega,n}d\Gamma \qquad = 0 \qquad (3.12b)\\
W_\omega &= \int_\Omega \delta\omega\rho\frac{1}{\Delta t}(\omega - \omega^{t-\Delta t})d\Omega + \int_\Omega \delta\omega\vec{u}\cdot\rho\nabla\omega d\Omega + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega \\
&\quad + \int_{\Gamma_N}\delta\omega q_{\omega,n}d\Gamma \qquad = 0 \qquad (3.12c)
\end{aligned}$$

where Δt is the time step size and the superscript $t-\Delta t$ denotes the previous time step. After time discretization, variables without superscript $t-\Delta t$ carry values of current time step by default.

3.3.2 Approximations in space

The Navier-Stokes equation in the mixed formulation would always result in an indefinite matrix system. It's known as saddle point problem. It requires that the finite elements for

velocity and pressure must be well matched to fulfill the inf-sup condition (LBB condition) [4]. One qualified element type and commonly used is Q2Q1 element as shown in Fig. 3.2. Under this configuration, the velocity $\vec{u}$ is approximated by bi-quadratic polynomials, while the pressure p is approximated by bi-linear polynomials. Just like the pressure, the approximation of mass fraction ω is also bi-linear.

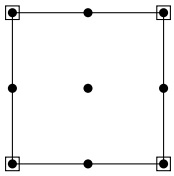


Figure 3.2 – Q2Q1 element; velocity DOFs reside on solid round nodes, while pressure and mass fraction DOFs on the empty squared nodes.

For a given element, unknowns are approximated in the matrix form as follows:

$$\vec{u} = [\bar{N}]\{u\} \tag{3.13a}$$

$$p = \langle N \rangle \{p\} \tag{3.13b}$$

$$\omega = \langle N \rangle \{\omega\} \tag{3.13c}$$

where

$$[N] = \begin{bmatrix} \bar{N}_1 & 0 & \bar{N}_2 & 0 & \cdots & \bar{N}_9 & 0 \\ 0 & \bar{N}_1 & 0 & \bar{N}_2 & \cdots & 0 & \bar{N}_9 \end{bmatrix} \tag{3.14a}$$

$$\langle N \rangle = \langle \; N_1 \quad N_2 \quad N_3 \quad N_4 \; \rangle \tag{3.14b}$$

$$\{u\} = \langle \; u_1 \quad v_1 \quad u_2 \quad v_2 \quad \cdots \quad u_9 \quad v_9 \; \rangle^{\mathrm{T}} \tag{3.14c}$$

$$\{p\} = \langle \; p_1 \quad p_2 \quad p_3 \quad p_4 \; \rangle^{\mathrm{T}} \tag{3.14d}$$

$$\{\omega\} = \langle \; \omega_1 \quad \omega_2 \quad \omega_3 \quad \omega_4 \; \rangle^{\mathrm{T}} \tag{3.14e}$$

3.3.3 Fully-discrete system

At the stage of linerization and spatial discretization, NS equations (including the incompressibility constraint) and the solute transport equation are regarded as independent in the case of weak coupling, which means, for example, the density in the momentum equation is not taken as a function of solute concentration for the time being. Then, two systems of algebraic equations would be generated, as shown by Eq. 3.15 in matrix form. But in the case of strong coupling, the hydrodynamics and solute transport are treated as a whole. Then, it only yields one system of algebraic equations, as shown by Eq. 3.16. Please note that in Eq. 3.15 and Eq. 3.16, terms with the superscript *consv* only exist for the situation when the conservative form of the species transport equation (Eq. 3.7) is used; sub-matrices with the subscript t are only generated by Newton-Raphson method; dropping the terms with the superscript *consv*

or with the subscript t, one can obtain the system generated by dealing with NS equations and the species transport equation in non-conservative form (Eq. 3.8) through the substitution method.

$$\begin{Bmatrix} R_{\vec{u}} \\ R_p \end{Bmatrix} + \begin{bmatrix} K_{\vec{u}\vec{u}} + \tilde{K}_{t,\vec{u}\vec{u}} & K_{\vec{u}p} \\ K_{\vec{u}p}^{\mathrm{T}} & 0 \end{bmatrix} \begin{Bmatrix} \Delta u \\ \Delta p \end{Bmatrix} = 0 \tag{3.15a}$$

$$\Big\{R_\omega + R_\omega^{consv}\Big\} + \Big[K_{\omega\omega} + \tilde{K}_{t,\omega\omega} + K_{\omega\omega}^{consv} + \tilde{K}_{t,\omega\omega}^{consv}\Big] \Big\{\Delta\omega\Big\} = 0 \tag{3.15b}$$

$$\begin{Bmatrix} R_{\vec{u}} \\ R_p \\ R_\omega + R_\omega^{consv} \end{Bmatrix} + \begin{bmatrix} K_{\vec{u}\vec{u}} + \tilde{K}_{t,\vec{u}\vec{u}} & K_{\vec{u}p} & \tilde{K}_{t,\vec{u}\omega} \\ K_{\vec{u}p}^{\mathrm{T}} & 0 & 0 \\ \tilde{K}_{t,\omega\vec{u}} + \tilde{K}_{t,\omega\vec{u}}^{consv} & 0 & K_{\omega\omega} + \tilde{K}_{t,\omega\omega} + K_{\omega\omega}^{consv} + \tilde{K}_{t,\omega\omega}^{consv} \end{bmatrix} \begin{Bmatrix} \Delta u \\ \Delta p \\ \Delta\omega \end{Bmatrix} = 0 \tag{3.16}$$

where

$$\begin{aligned} R_{\vec{u}} = &\sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \rho \left[\bar{N}\right] \frac{\{u\}^{i-1} - \{u\}^{t-\Delta t}}{\Delta t} + \Big(\rho^{i-1}[C]^{i-1} + \left[\check{B}\right]^{\mathrm{T}} 2\mu \left[\check{B}\right]\Big)\{u\}^{i-1} \\ &- \{B\}\langle N\rangle\{p\}^{i-1} - [\bar{N}]^{\mathrm{T}}\rho^{i-1}\{g\}d\Omega - \sum_e \int_{\Gamma_N^e} \left[\bar{N}\right]^{\mathrm{T}} \{t_N\}d\Gamma \end{aligned} \tag{3.17a}$$

$$R_p = \sum_e \int_{\Omega^e} -\{N\}\langle B\rangle\{u\}^{i-1}d\Omega \tag{3.17b}$$

$$\begin{aligned} R_\omega = &\sum_e \int_{\Omega^e} \{N\}\rho\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} + \{N\}\langle u\rangle[\bar{N}]^{\mathrm{T}}\rho^{i-1}[B]\{\omega\}^{i-1} \\ &+ [B]^{\mathrm{T}}\rho^{i-1}D[B]\{\omega\}^{i-1}d\Omega + \sum_e \int_{\Gamma_N^e} \{N\}q_{\omega,n}^{i-1}d\Gamma \end{aligned} \tag{3.17c}$$

$$\begin{aligned} R_\omega^{consv} = &\sum_e \int_{\Omega^e} \{N\}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} \\ &+ \{N\}\langle u\rangle[\bar{N}]^{\mathrm{T}}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}[B]\{\omega\}^{i-1}d\Omega \end{aligned} \tag{3.17d}$$

$$K_{\vec{u}\vec{u}} = \sum_e \int_{\Omega^e} [\bar{N}]^{\mathrm{T}} \frac{1}{\Delta t}\rho^{i-1}[\bar{N}] + \rho^{i-1}[C]^{i-1} + [\check{B}]^{\mathrm{T}} 2\mu[\check{B}]d\Omega \tag{3.17e}$$

$$\tilde{K}_{t,\vec{u}\vec{u}} = \sum_e \int_{\Omega^e} [\bar{N}]^{\mathrm{T}}\rho^{i-1}\Big([\bar{N}]_x\{u\}^{i-1}\langle\bar{N}_{10}\rangle + [\bar{N}]_y\{u\}^{i-1}\langle\bar{N}_{01}\rangle\Big)d\Omega \tag{3.17f}$$

$$K_{\vec{u}p} = \sum_e \int_{\Omega^e} -\{B\}\langle N\rangle d\Omega \tag{3.17g}$$

$$\tilde{K}_{t,\vec{u}\omega} = \sum_e \int_{\Omega^e} [\bar{N}]^{\mathrm{T}}(\rho')^{i-1} \left([\bar{N}]\frac{\{u\}^{i-1} - \{u\}^{t-\Delta t}}{\Delta t} - \{g\}\right)\langle N\rangle + (\rho')^{i-1}[C]^{i-1}\{u\}^{i-1}\langle N\rangle d\Omega \tag{3.17h}$$

$$K_{\omega\omega} = \sum_e \int_{\Omega^e} \{N\}\rho^{i-1}\frac{1}{\Delta t}\langle N\rangle + \{N\}\langle u\rangle^{i-1}[\bar{N}]^{\mathrm{T}}\rho^{i-1}[B] + [B]^{\mathrm{T}}\rho^{i-1}D[B]d\Omega \tag{3.17i}$$

$$\begin{aligned} \tilde{K}_{t,\omega\omega} = &\sum_e \int_{\Omega^e} \{N\}(\rho')^{i-1}\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t}\langle N\rangle \\ &+ \{N\}\langle u\rangle^{i-1}[\bar{N}]^{\mathrm{T}}(\rho')^{i-1}[B]\{\omega\}^{i-1}\langle N\rangle + [B]^{\mathrm{T}}(\rho')^{i-1}D[B]\{\omega\}^{i-1}\langle N\rangle d\Omega \end{aligned} \tag{3.17j}$$

$$K_{\omega\omega}^{consv} = \sum_e \int_{\Omega^e} \{N\} \left(\langle N \rangle \{\omega\}^{i-1} (\rho')^{i-1} \frac{1}{\Delta t} \langle N \rangle + \langle u \rangle^{i-1} [\bar{N}]^{\mathrm{T}} \langle N \rangle \{\omega\}^{i-1} (\rho')^{i-1} [B] \right) d\Omega \tag{3.17k}$$

$$\begin{aligned} \tilde{K}_{t,\omega\omega}^{consv} = & \sum_e \int_{\Omega^e} \{N\} \Big((\rho')^{i-1} + \langle N \rangle \{\omega\}^{i-1} (\rho'')^{i-1} \Big) \langle N \rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} \langle N \rangle \\ & + \{N\} \langle u \rangle^{i-1} [\bar{N}]^{\mathrm{T}} \Big((\rho')^{i-1} + \langle N \rangle \{\omega\}^{i-1} (\rho'')^{i-1} \Big) [B] \{\omega\}^{i-1} \langle N \rangle d\Omega \end{aligned} \tag{3.17l}$$

$$\tilde{K}_{t,\omega\vec{u}} = \sum_e \int_{\Omega^e} \{N\} \rho^{i-1} \langle \omega \rangle^{i-1} [B]^{\mathrm{T}} [\bar{N}] d\Omega \tag{3.17m}$$

$$\tilde{K}_{t,\omega\vec{u}}^{consv} = \sum_e \int_{\Omega^e} \{N\} \langle N \rangle \{\omega\}^{i-1} (\rho')^{i-1} \langle \omega \rangle^{i-1} [B]^{\mathrm{T}} [\bar{N}] d\Omega \tag{3.17n}$$

$\left[\check{B}\right]$ is called strain rate operator:

$$\left[\check{B}\right] = \begin{bmatrix} \bar{N}_{1,x} & 0 & \bar{N}_{2,x} & 0 & \cdots & \bar{N}_{9,x} & 0 \\ 0 & \bar{N}_{1,y} & 0 & \bar{N}_{2,y} & \cdots & 0 & \bar{N}_{9,y} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \frac{1}{2}\bar{N}_{2,y} & \frac{1}{2}\bar{N}_{2,x} & \cdots & \frac{1}{2}\bar{N}_{9,y} & \frac{1}{2}\bar{N}_{9,x} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \frac{1}{2}\bar{N}_{2,y} & \frac{1}{2}\bar{N}_{2,x} & \cdots & \frac{1}{2}\bar{N}_{9,y} & \frac{1}{2}\bar{N}_{9,x} \end{bmatrix} \tag{3.18}$$

$[B]$ is defined as gradient operator:

$$[B] = \nabla \langle N \rangle \tag{3.19}$$

$\langle B \rangle$ can be named divergence operator:

$$\langle B \rangle = \langle \bar{N}_{1,x} \quad \bar{N}_{1,y} \quad \bar{N}_{2,x} \quad \bar{N}_{2,y} \quad \cdots \quad \bar{N}_{9,x} \quad \bar{N}_{9,y} \rangle \tag{3.20}$$

$\langle \bar{N}_{10} \rangle$ is defined as

$$\langle \bar{N}_{10} \rangle = \langle \bar{N}_1 \quad 0 \quad \bar{N}_2 \quad 0 \quad \cdots \quad \bar{N}_9 \quad 0 \rangle \tag{3.21}$$

$\langle \bar{N}_{01} \rangle$ is defined as

$$\langle \bar{N}_{01} \rangle = \langle 0 \quad \bar{N}_1 \quad 0 \quad \bar{N}_2 \quad \cdots \quad 0 \quad \bar{N}_9 \rangle \tag{3.22}$$

$[C]$ is named convection operator:

$$[C]^{i-1} = [\bar{N}]^{\mathrm{T}} \Big([\bar{N}] \{u\}^{i-1} \Big)^{\mathrm{T}} \begin{Bmatrix} \partial_x \\ \partial_y \end{Bmatrix} [\bar{N}] \tag{3.23}$$

The superscript $i-1$ denotes the previous iteration step and ρ'' denotes the density's second derivative with respect to the mass fraction ω.

3.4 Numerical tests

3.4.1 Poiseuille flow between parallel plates

Poiseuille flow between parallel plates is taken as benchmark to test the pure hydrodynamic model. Poiseuille flow is induced by the pressure drop Δp between the inlet and outlet. As shown in Fig. 3.3, a parabola-shaped velocity profile forms when no-slip boundary conditions

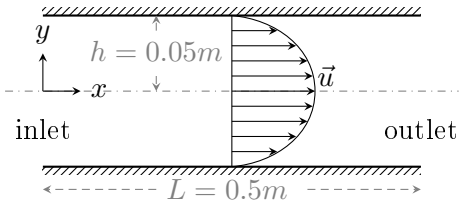


Figure 3.3 – Schematic diagram of Poiseuille flow.

are applied on the walls ($y = \pm h$). The separation distance between plates is denoted by $2 \times h$. The analytical solution of axial velocity's magnitude u reads [8]

$$u = -\frac{\Delta p}{2\mu L}(h^2 - y^2) \tag{3.24}$$

where $|y| \leq h$ and L denotes the distance that the pressure drop acts over. Fig. 3.4 presents the velocity and pressure distributions obtained by numerical simulation as $\rho = 1000kg/m^3$, $\mu = 0.1Pa \cdot s$, $\Delta p = -6Pa$. Fig. 3.5 displays the axial velocity magnitude at section $x = 0.225m$ and the comparison with the analytical solution. It can be seen that the numerical solution has good agreement with the analytical solution. The maximum relative error is 2.33%. The Reynolds number of the flow $\rho u_{\max} 2h/\mu$ is approximately 150 which indicates the flow is laminar.

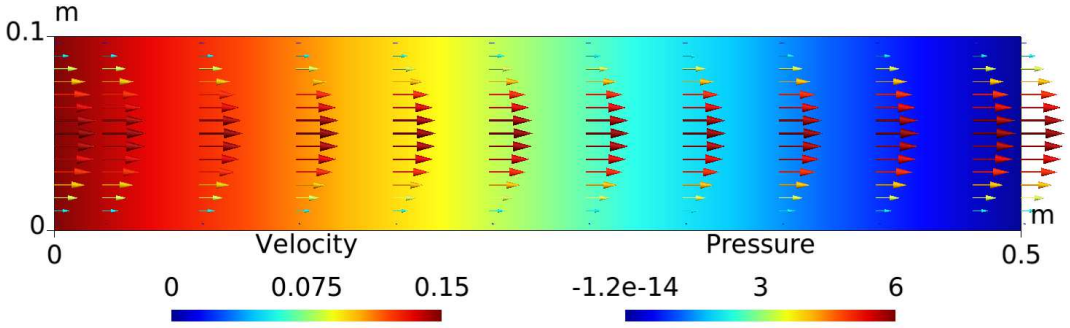


Figure 3.4 – Calculated velocity and pressure distributions after Poiseuille flow well-developed.

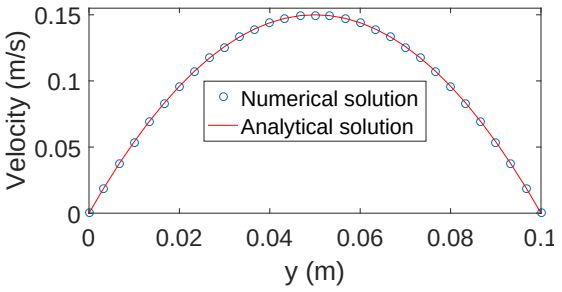


Figure 3.5 – The axial velocity magnitude at section $x = 0.225m$ and comparison with the analytical solution.

3.4.2 Mixing of two parallel streams with different concentrations

The species transport model with constant density is verified by the mixing problem of two parallel streams being of different concentrations. As shown in Fig. 3.6, two streams with

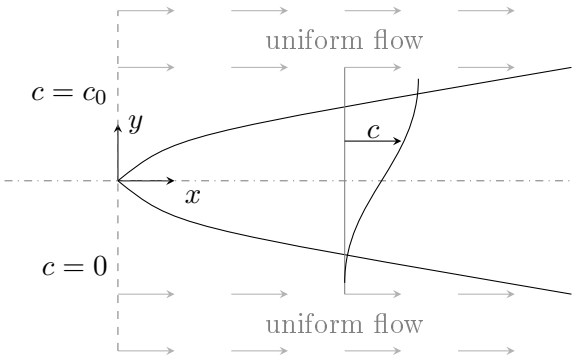


Figure 3.6 – Schematic diagram of the mixing of two parallel streams with different concentrations.

different concentrations start to mix at $x = 0m$. The boundary conditions of concentration (mol/m^3) can be expressed as

$$c(x,y) = \begin{cases} 0, & x = 0, y \leq 0 \\ c_0, & x = 0, y > 0 \end{cases} \tag{3.25}$$

The background flow field is axial and uniform of magnitude u. The transport process is governed by two mechanisms, i.e. longitudinal convection with flow and lateral diffusion caused by the concentration gradient. The analytical solution of concentration distribution reads [2]

$$c(x,y) = \frac{c_0}{2} erfc\left(\frac{y}{\sqrt{4D_y x/u}}\right) \tag{3.26}$$

where $erfc()$ is complementary error function and D_y is the lateral diffusivity. The numerical simulation is conducted within a rectangular domain in size of $160 \times 40m^2$, with $u = 0.1m/s$, $D_y = 0.02m^2/s$, $c_0 = 0.1mol/m^3$. Fig. 3.7 presents the concentration distribution after it reaches steady state, from which an obvious transitional zone can be seen. The simulation results along two sections ($x = 40.41m$ and $x = 80.95m$) are extracted to compare with the analytical solution, as shown in Fig. 3.8. It can be seen that they have great agreement. The maximum difference is about $1.1 \times 10^{-4} mol/m^3$.

3.4.3 Lock exchange flow

As shown in Fig. 3.1, two fluids of different densities, initially still and separated by a vertical "lock gate", are confined in a rectangular enclosure. Once the gate is taken away, due to the imbalanced hydrostatic pressure caused by non-uniform density distribution, the heavier fluid tends to start moving along the bottom while the lighter fluid flows in the opposite direction

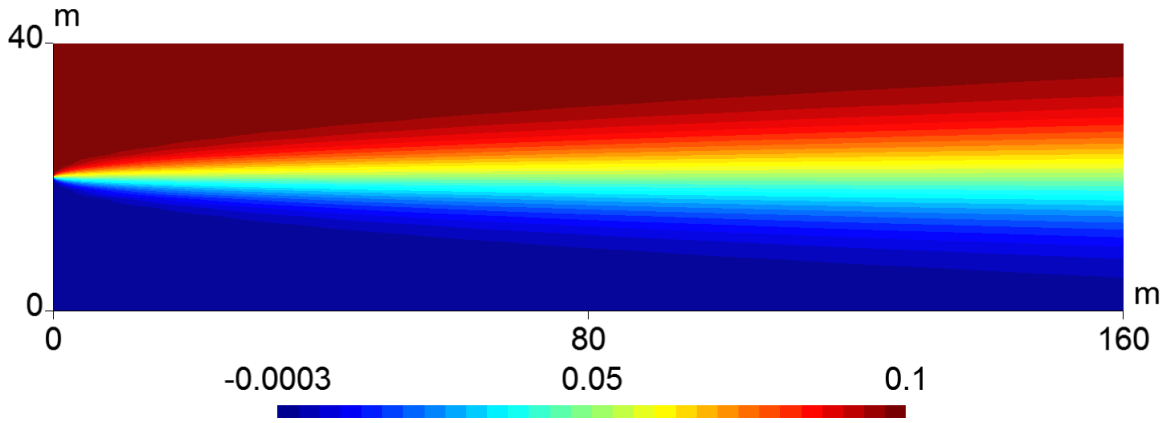


Figure 3.7 – The concentration distribution after the model reaching steady state.

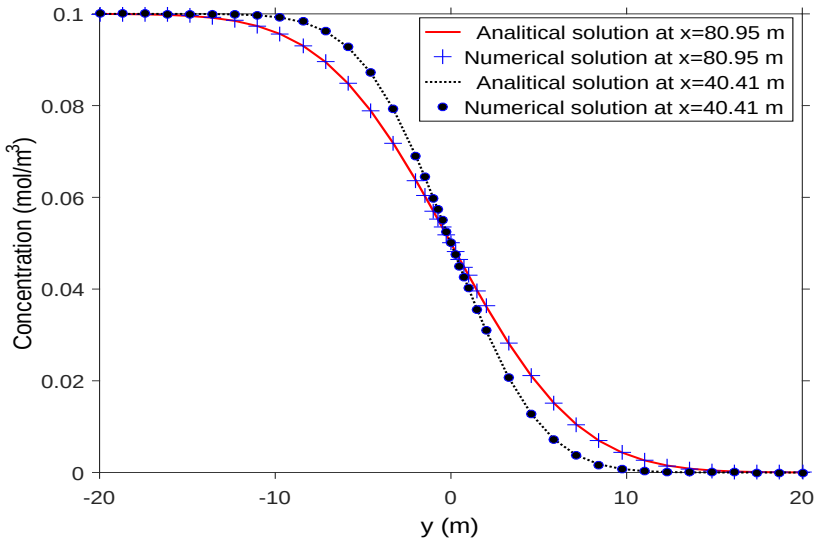


Figure 3.8 – Calculated and analytical concentrations at sections $x = 40.41m$ and $x = 80.95m$, respectively.

along the top. Meanwhile, mutual penetration also takes place between the two fluids. The mixing process will last for a while until it finally restores to still state again. In the final state, these two fluids are supposed to be completely mixed.

Numerical simulations of this mixing process are performed under two scenarios, which differ in the density contrast between the two fluids. In the first scenario, the density contrast takes up 10% relative to the denser one as $\rho_1 = 10kg/m^3$, $\rho_2 = 9kg/m^3$. In the second scenario, the density contrast takes up 0.5% just by increasing the density of the lighter one as $\rho_2 = 9.95kg/m^3$. All the other parameters are identical for both scenarios, such as $\mu = 0.01Pa \cdot s$, $D = 0.02m^2/s$, $g = -9.8m/s^2$, $\Delta t = 0.5s$ and the cavity is $2m \times 1m$ in size. No-slip conditions for velocities and zero-flux condition for salinity are applied on all solid walls. The pressure at right upper corner is set to be zero as a reference value. Initially,

each fluid is supposed to take up exactly half of the container in space. After well mixed, the theoretical mass fraction of denser fluid, i.e. $\frac{\rho_1}{\rho_1+\rho_2}$, can be easily calculated. This value is 0.526 in scenario one, 0.501 in scenario two.

Under each scenario, different options in the aspects such as coupling degree, linearization scheme, the form of the solute transport equation (conservative or non-conservative) are tested in terms of accuracy and efficiency. All the simulations last for 500 time steps, after which the mixing in both scenarios is done. The mixing is considered to be done when it fulfills the criteria that the flow speed reduces below the magnitude order $10^{-5}m/s$ and the mass fraction gets uniformly distributed under the precision with five decimal places. The convergence criterion of the iterations is the relative error between two consecutive solutions less than 10^{-6} for each unknown variable. The evolution of the flow pattern and mass fraction of the denser fluid for the two scenarios are presented in Fig. 3.9 and Fig. 3.10, respectively. At the very early stage of the mixing ($t = 0.5s$), the transitional belt of the mass fraction field starts to incline due to the induced circulation in scenario one; however, the transitional zone in scenario two almost keeps vertical due to the circulation formed in this case is weak and the diffusion process is overwhelmingly dominated. From the outset, the induced circulation will undergo a period of speeding up. Meanwhile, a S-shaped transitional zone, also discussed in [22], will form (see $t = 2s$ in Fig. 3.9 and $t = 19s$ in Fig. 3.10). The circulation reaches the highest speed $0.565m/s$ at $t = 4.5s$ in scenario one, while $0.0975m/s$ at $t = 19s$ in scenario two. Since then, the circulation has gradually slowed down. It also can be seen that the flow pattern induced by 0.5% density difference is monotonous compared to that by 10% density difference. For instance, two more sub-circulations start to form in the corners at $t = 8.5s$ in Fig. 3.9, then, they grow gradually and merge again as shown at $t = 13s$ and $t = 25s$; this process of merging and breaking up among circulations repeats several times. In both scenarios, the diffusion process mainly dominates at the later stage of the mixing as the circulation gets faint. The mixing for both scenarios consumes nearly the same time (500 time steps as $\Delta t = 0.5s$), which implies that the stronger circulation facilitates the mixing process on one hand, but on the other hand, it consumes time to make itself resume to the resting state.

Table 3.1 gives the summary of run times and relative errors relating to mass conservation under different situations. The run time refers to the time consumed based on Processor Intel® Core™ i7-3770 CPU @ 3.40GHz × 8. The relative error is calculated by

$$\text{Err.}\% = \frac{\text{theoretical value} - \text{simulated value after well mixed}}{\text{theoretical value}} \times 100\% \tag{3.27}$$

where all the values refer to the mass fraction of the denser fluid. As shown in Table 3.1, the Newton-Raphson method does not converge on velocity field either through strong coupling or weak coupling in the case with large density difference. The reason might be that, Newton-Raphson method requires better initially guessed value compared to the substitution method.

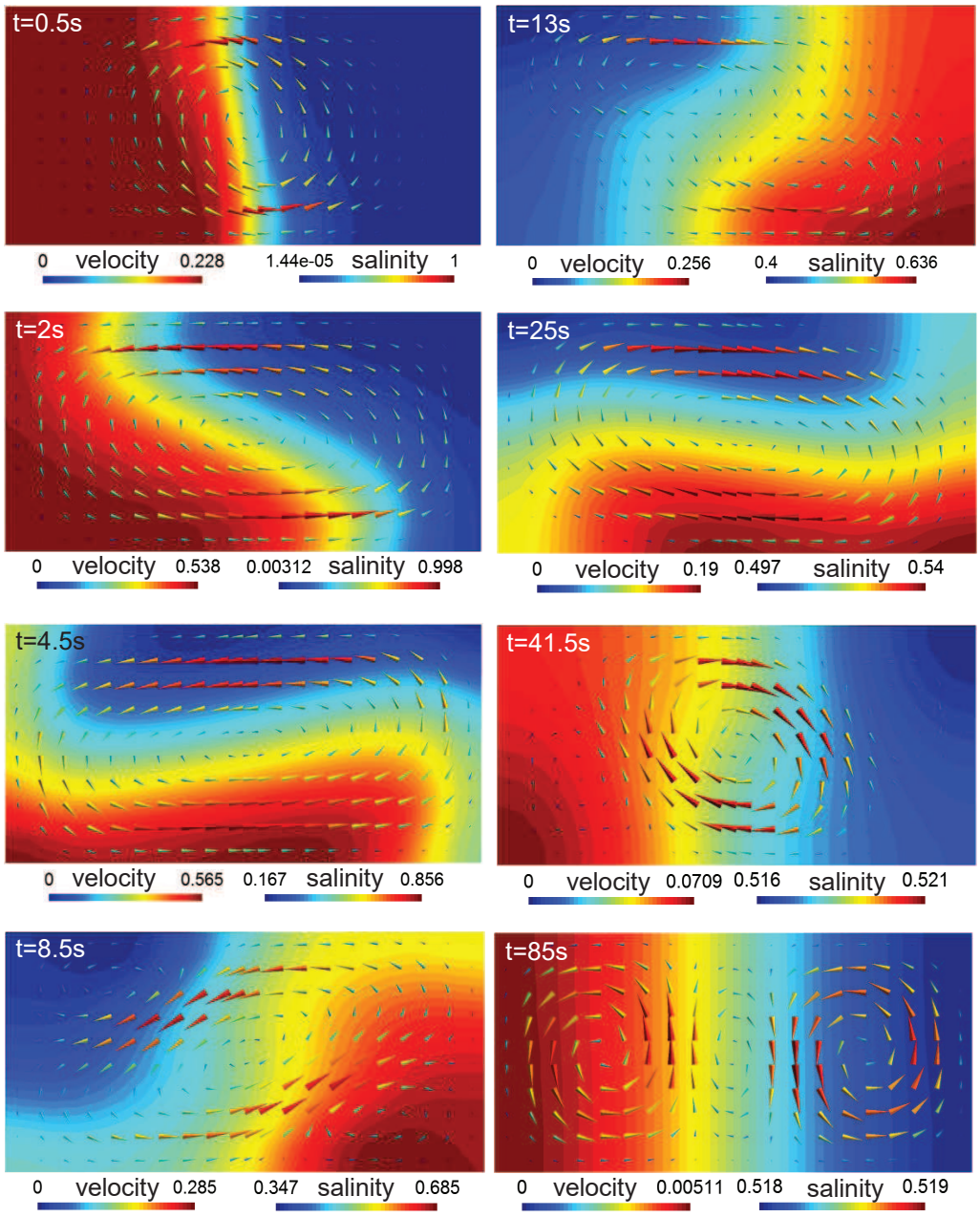


Figure 3.9 – Evolution of the flow pattern (velocity unit m/s) and mass fraction of the denser fluid for the case with 10% density distinct, substitution method, conservative form and strong coupling.

In the case with small density difference, the models by the Newton-Raphson method spend more time than those ones by the substitution method, since more time is required to calculate and assemble some extra terms as shown in Eq. 3.15 and Eq. 3.16. Speaking of models' run time, the most import factor is the coupling degree. As shown in Table 3.1, all the models by strong coupling are considerably time-saving compared to those by weak coupling, and they also have a very slightly better performance on mass conservation (3.16% vs 3.18%,

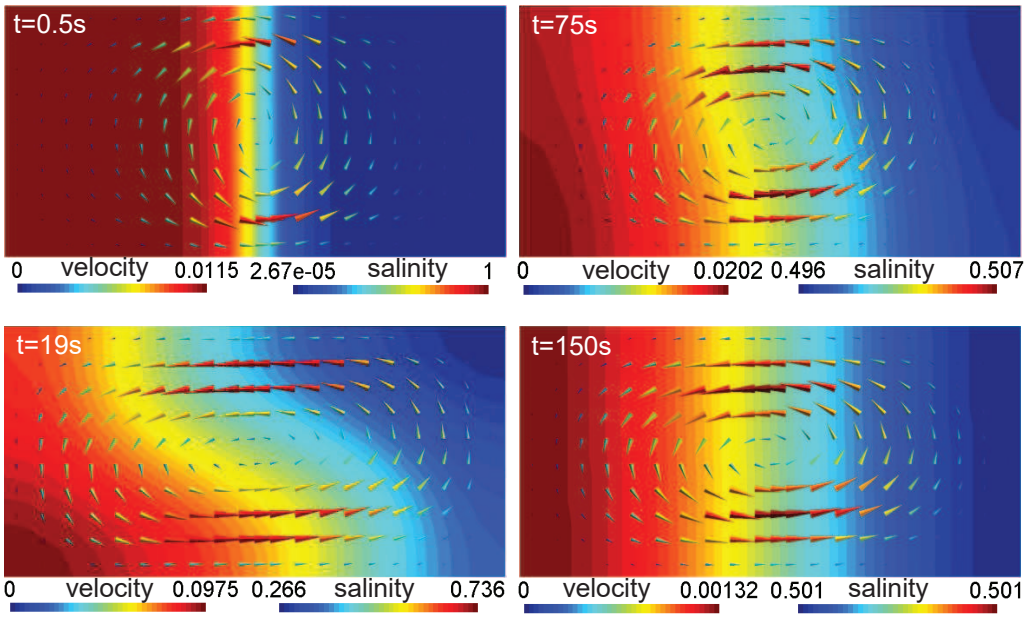


Figure 3.10 – Evolution of the flow pattern (velocity unit m/s) and mass fraction of the denser fluid for the case with 0.5% density distinct, Newton-Raphson method, conservative form and strong coupling.

Table 3.1 – Statistics of run times and errors with respect to mass conservation for lock exchange flow simulations under different situations

		Non-conservative				Conservative			
		weak coupling		strong coupling		weak coupling		strong coupling	
		Subst [a]	NR [b]	Subst	NR	Subst	NR	Subst	NR
$\rho_{gap}10\%$	t_{run} [c]	2.40	-[e]	0.10	-	2.53	-	0.11	-
	Err.[d]	3.18	-	3.16	-	1.45	-	1.42	-
$\rho_{gap}0.5\%$	t_{run}	2.23	2.39	0.07	0.13	2.53	2.71	0.07	0.14
	Err.	0.13	0.13	0.13	0.13	0.02	0.02	0.02	0.02

[a] Substitution method [b] Newton-Raphson method [c] Running time by hour [d] Relative error (%)
[e] Not converge

1.42% vs 1.45%) when the density difference is large. In the aspect of mass conservation, the models with solute transport equation in conservative form are always better mass-conserved compared to those with non-conservative form. The additional time consumed due to the extra conservative term in Eq. 3.15 and Eq. 3.16 are minor and acceptable.

3.5 Conclusions

The finite element model to depict the mass transport process taking place in a binary miscible solution with variable density has been presented. Particularly three aspects, including coupling degree (strong or weak), linearisation scheme (substitution method or Newton-Raphson

method), the form of solute transport equation (conservative or non-conservative), are examined by modelling the lock exchange problem in terms of computational efficiency or accuracy with respect to mass conservation. The strong coupling, substitution method and the solute transport equation in conservative form have a better performance compared to their counterparts. One thing should be pointed out here is that, although the strong coupling scheme is computationally efficient, it makes the links among physics complex or tangled to some extent. For the problem with multi-physics (more than two), the weak coupling scheme might enable each sub-problem relatively independent, which lets model's structure clearer and more suitable for modularization in practice. In addition, even though the present model is built for miscible fluids, it also apples to immiscible fluids problem theoretically if the diffusion term in the solute transport equation is removed. However, it might face a challenge that the instability problem caused by the discontinuities at the interface.

Acknowledgements

The authors would like to acknowledge the National Sciences and Engineering Research Council of Canada (NSERC) funding NSERC Discovery Grant (No. 36518) for supporting the present research. A part of the research presented in this paper was financed by the Fonds de Recherche du Québec-Nature et Technologie (FRQNT) by the intermediary of the Aluminium Research Centre-REGAL. The first author also acknowledges the financial support from the Chinese Scholarship Council (CSC).

Bibliography

[1] G. K. Batchelor. *An introduction to fluid dynamics*. Cambridge University Press, 2000.

[2] J. Bear. *Hydraulics of groundwater*. Courier Corporation, 2012.

[3] V. K. Birman, J. E. Martin, and E. Meiburg. The non-Boussinesq lock-exchange problem. Part 2. High-resolution simulations. *Journal of Fluid Mechanics*, 537:125–144, 2005.

[4] J. Donea and A. Huerta. *Finite element methods for flow problems*. John Wiley & Sons, 2003.

[5] Y. Fraigneau, J. L. Guermond, and L. Quartapelle. Approximation of variable density incompressible flows by means of finite elements and finite volumes. *Communications in Numerical Methods in Engineering*, 17(12):893–902, 2001.

[6] J. L. Guermond and A. Salgado. A splitting method for incompressible flows with variable density based on a pressure Poisson equation. *Journal of Computational Physics*, 228(8):2834–2846, 2009.

[7] Y. Hallez and J. Magnaudet. Effects of channel geometry on buoyancy-driven mixing. *Physics of Fluids*, 20(5):053306, 2008.

[8] J. Happel and H. Brenner. *Low Reynolds number hydrodynamics - with special applications to particulate media*, volume 1. Springer Science & Business Media, 2012.

[9] A. W. Herbert, C. P. Jackson, and D. A. Lever. Coupled groundwater flow and solute transport with fluid density strongly dependent upon concentration. *Water Resources Research*, 24(10):1781–1795, 1988.

[10] A. J. Landman and R. J. Schotting. Heat and brine transport in porous media: the Oberbeck-Boussinesq approximation revisited. *Transport in Porous Media*, 70(3):355–373, 2007.

[11] R. J. Lowe, J. W. Rottman, and P. F. Linden. The non-Boussinesq lock-exchange problem. Part 1. Theory and experiments. *Journal of Fluid Mechanics*, 537:101–124, 2005.

[12] T. J. McDougall, R. J. Greatbatch, and Y. Lu. On conservation equations in oceanography: how accurate are Boussinesq ocean models? *Journal of Physical Oceanography*, 32(5):1574–1584, 2002.

[13] G. S. H. Pau, J. B. Bell, K. Pruess, A. S. Almgren, M. J. Lijewski, and K. Zhang. High-resolution simulation and characterization of density-driven flow in CO_2 storage in saline aquifers. *Advances in Water Resources*, 33(4):443–455, 2010.

[14] A. R. Piriz, O. D. Cortazar, J. J. López Cela, and N. A. Tahir. The Rayleigh-Taylor instability. *American Journal of Physics*, 74(12):1095–1098, 2006.

[15] T. J. Povich, C. N. Dawson, M. W. Farthing, and C. E. Kees. Finite element methods for variable density flow and solute transport. *Computational Geosciences*, 17(3):529–549, 2013.

[16] K. C. Sahu, H. Ding, and O. K. Matar. Numerical simulation of non-isothermal pressure-driven miscible channel flow with viscous heating. *Chemical Engineering Science*, 65(10):3260–3267, 2010.

[17] K. C. Sahu, H. Ding, P. Valluri, and O. K. Matar. Linear stability analysis and numerical simulation of miscible two-layer channel flow. *Physics of Fluids*, 21(4):042104, 2009.

[18] K. C. Sahu, H. Ding, P. Valluri, and O. K. Matar. Pressure-driven miscible two-fluid channel flow with density gradients. *Physics of Fluids*, 21(4):043603, 2009.

[19] T. Séon, J. Znaien, D. Salin, J. P. Hulin, E. J. Hinch, and B. Perrin. Transient buoyancy-driven front dynamics in nearly horizontal tubes. *Physics of Fluids*, 19(12):123603, 2007.

[20] J. O. Shin, S. B. Dalziel, and P. F. Linden. Gravity currents produced by lock exchange. *Journal of Fluid Mechanics*, 521:1–34, 2004.

[21] J. C. Slattery. *Advanced transport phenomena*. Cambridge University Press, 1999.

[22] M. L. Szulczewski and R. Juanes. The evolution of miscible gravity currents in horizontal porous layers. *Journal of Fluid Mechanics*, 719:82–96, 2013.

[23] R. Taylor and R. Krishna. *Multicomponent mass transfer*, volume 2. John Wiley & Sons, 1993.

[24] S. Whitaker. Flow in porous media I: a theoretical derivation of Darcy's law. *Transport in Porous Media*, 1(1):3–25, 1986.

[25] Y. N. Young, H. Tufo, A. Dubey, and R. Rosner. On the miscible Rayleigh-Taylor instability: two and three dimensions. *Journal of Fluid Mechanics*, 447:377–408, 2001.

[26] R. K. Zeytounian. Joseph Boussinesq and his approximation: a contemporary view. *Comptes Rendus Mecanique*, 331(8):575–586, 2003.

Chapter 4

Numerical simulation of Stefan problem coupled with mass transport in a binary system through XFEM/level set method

Résumé

Cet article porte sur l'application de la méthode des éléments finis étendus à la simulation du phénomène de changement de phase couplé au transport d'espèces chimiques dans un système binaire. À cet effet, le problème de Stefan et l'équation de diffusion sont résolus pour décrire les distributions de température et de concentration, respectivement. La température est enrichie faiblement via le schéma « abs-enrichment ». La température de fusion, imposée à l'interface solide/liquide par le biais de la méthode de pénalité, dépend de la concentration. En raison de la discontinuité forte de la concentration, le schéma de type « sign-enrichment » est utilisé. L'évolution de l'interface solide/liquide est capturée par la technique de la fonction de niveaux (level-set). La robustesse et la précision du modèle sont démontrées à travers des simulations numériques.

Abstract

This paper deals with the application of the extended finite element method to simulate the phase change phenomenon in a binary system while considering the interaction with the mass transport of chemical species. To this end, the thermal conduction equation with the Stefan condition and the mass diffusion equation are solved to depict the temperature and the solute concentration distributions, respectively. In the heat transfer problem with phase change, the temperature field is weakly enriched using the corrected abs-enrichment scheme to avoid

the blending element problem. The melting temperature imposed by the penalty method at the solid/liquid interface is solute concentration dependent. In the mass transport problem, due to the strong discontinuity in the concentration field at the interface, the sign-enrichment scheme is used. A constant separating out concentration is enforced on the solid side of the interface also by the penalty method, while a mass flux is applied naturally on the liquid side. The phase interface is captured implicitly by the level set method, which is discretized on the same finite element mesh and therefore eliminates the need for remeshing techniques. The robustness and the accuracy of the model are demonstrated through numerical case study.

Keywords: extended finite element method, level set method, Stefan problem, mass transport problem

4.1 Introduction

The classical Stefan problem can trace back to the end of 19^{th} century [25], which was originally raised in relation with ice formation. Within the past more than one hundred years, the corresponding mathematical model (governing equations system) has been widely employed in various fields [11]. This extensive application in practice proves the realistic value of studying it. However, analytical solutions of Stefan problem only exist under certain very specific circumstances (simple geometry and boundary conditions). On most occasions, we have to resort to numerical methods.

Solving the Stefan problem through conventional numerical methods such as finite element method (FEM) always encounters two main challenges. The first one comes from the fact that, besides temperature, the position of phase front is a priori unknown as well, which makes Stefan problem fall into the category of free boundary value problems. The second challenge is that the rapid change (discontinuity) of physical quantities in the vicinity of phase front makes those conventional methods established on continuity hypothesis fail to perform appropriate approximation.

In the field of solving the Stefan problem numerically, there are several typical approaches such as enthalpy method, phase field method, moving mesh methods, conventional numerical methods based on the fixed mesh while utililizing front tracking/capturing techniques, etc. In enthalpy method [28, 29], the energy conservation equation in enthalpy formulation is adopted. Unlike temperature formulation, the enthalpy that measures the total heat can naturally account for the liberation or absorption of latent heat during the phase change process. This property makes it possible that only one equation is solved over the entire domain and tracking the interface is avoided. So the enthalpy method probably is the easiest scheme to implement among all the numerical methods to solve Stefan problem. But due to the fact that the interface cannot be described precisely, enthalpy method is more suitable to handle the case in which the phase transition happens over a mushy zone rather than at a sharp front. Another commonly

used method for phase change problem (especially for dendritic growth of binary system) is phase field method [8, 20]. In this method, an auxiliary variable (the phase field) with constant but distinct value in each phase and varying rapidly in between over a layer of certain thickness at the interface is introduced. One advantage of phase field method is that the Gibbs-Thomson effect is taken into account without the need to compute the interface curvature [3]. As mentioned above, Stefan problem is of free boundary problem, so some generic treatments designed to handle free boundary problems are also applicable to solve the Stefan problem. The following two types of solution belong to this situation. In moving mesh methods, the underlying mesh is continuously updated to maintain the interface aligned with element edges, such that the energy conservation equation can be solved separately in each phase [23]. The drawbacks of this kind of methods are obvious, such as mesh reconstruction is cumbersome and it's hard to manage the case when the interface possesses complex geometry or undergoes fierce topological change. By contrast with moving mesh methods, front tracking/capturing techniques combined with the conventional numerical methods could be used to solve Stefan problem based on a fixed mesh [4, 30]. However, in order to overcome the problem caused by discontinuity at the interface, local mesh refinement is required, which always increases the computational cost. For more detailed or systematic comparison among those numerical approaches discussed above, articles [13, 12] are recommended.

As a matter of fact, the discontinuity in the solution space is not well addressed by the aforementioned approaches. For that reason, the extended finite element method (XFEM) [1] as an alternative method was introduced to solve Stefan problem for the first time in 2002 [6, 16]. The XFEM has inherent capability on coping with the discontinuity problem by enriching the approximation space to conform with the solution space. In [16], the algorithm for one-dimensional case was implemented. Almost at the same time or maybe earlier, a generic algorithm combined with level set method to capture the interface for two-dimensional case was developed in [6]. The level set method holds great promise to describe free boundary from the geometric point of view, as it possesses advantages such as, easy to implement, capable to handle the case with topological change and able to give sufficient interface information (like unit normal and curvature). What's more, the combination of XFEM and level set method is natural and straightforward, as the level set function is utilized to construct the enrichment function of XFEM. The paper [6] also involved the aspects on spreading the interface normal velocity and imposing interface temperature condition (Dirichlet-type) by the penalty method. Since then, the combination of XFEM/level set method has been further developed under the framework of phase change problem [33, 22, 2, 24, 15]. Its accuracy and robustness were well demonstrated. In [33], the process of dendritic solidification of an undercooled melt was modelled with considering the Gibbs-Thompson effect at the interface. In [22], different ways (penalty method and Lagrange multiplier method) to impose the interface condition and different treatments in the blending elements were investigated. In [24], it focused on the one-dimensional algorithm for the physically nonlinear phase change problem with using Newton-

Raphson method for linearization. In [15], the Lagrange multiplier method was applied not only for enforcing the interface condition but also for the purpose to evaluate the heat flux jump at the interface by utilizing the Lagrange multiplier values.

To the best knowledge of the authors, all the published application studies of XFEM combined with level set method are just restricted to the phase change problem of pure substance. No research work has been published on the mixture (at least a binary system), in which some physical properties (like density, freezing point) are species concentration dependent. Given all that, the present study aims to develop a numerical model to simulate the phase change process of a binary mixture based on XFEM combined with level set method. To this end, the classical Stefan problem and mass transport problem are required to be solved simultaneously. From mathematical point of view, the entire model is composed of four sub-models as shown in Fig. 4.1. Fig. 4.1 also displays the links among the sub-models. It can be seen that the classical Stefan problem (shaded area) is two-way coupled with the mass transport problem. Note that, the present study only focuses on the phase change process with planar front, that means the situation with cellular and further dendritic interface is beyond the scope of this paper. In addition, the volume variation caused by the density difference between phases is also ignored.

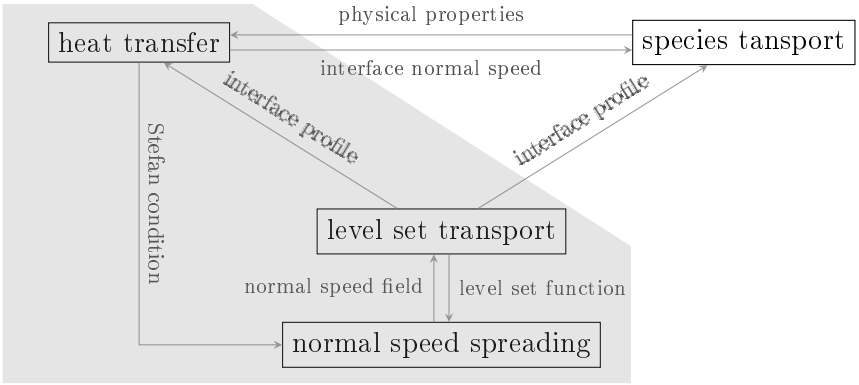


Figure 4.1 – Model's composition and the links among the sub-models; the shaded part accounts for the classical Stefan problem, which is two-way coupled with the mass transport problem.

This paper is organized as follows. Section 4.2 first introduces some basic concepts of XFEM; then, the governing equations and the numerical analysis of each sub-model are also presented in this section. In section 4.3, the model is employed to simulate the phase transition process of saline water ($NaCl - H_2O$ solution) to test the model's performance. At last, section 4.4 concludes this study with a summary and outlook.

4.2 Mathematical models and numerical analysis

4.2.1 XFEM approximation

Compared with the standard finite element method, the essential improvement of XFEM is that it enriches the approximation space locally according to the a priori known property (discontinuity) of the solution space. In XFEM, the approximation function of a variable $u(\vec{x},t)$ with one enrichment can be written as

$$u(\vec{x},t) = \underbrace{\sum_{i\in\mathcal{N}} N_i(\vec{x})u_i(t)}_{\text{strd. FEM approx.}} + \underbrace{\sum_{i\in\mathcal{N}_e} M_i(\vec{x},t)a_i(t)}_{\text{enrichment}} \tag{4.1}$$

where N_i is the standard finite element shape function of node i; u_i is the unknown quantity at node i; M_i is the local enrichment function; a_i is the additional degree of freedom; $\mathcal{N}$ is the set of total nodes; $\mathcal{N}_e$ is the collection of enriched nodes. Enriched nodes of standard XFEM are those nodes belonging to the cut elements as shown in Fig. 4.2. The local enrichment function can be further expanded as

$$M_i(\vec{x},t) = N_i(\vec{x})\psi_i(\vec{x},t) \tag{4.2}$$

where the shape function N_i mostly takes the same order with that in Eq. 4.1 but not necessarily; ψ_i is the global enrichment function of node i, which has different forms for different types of discontinuity.

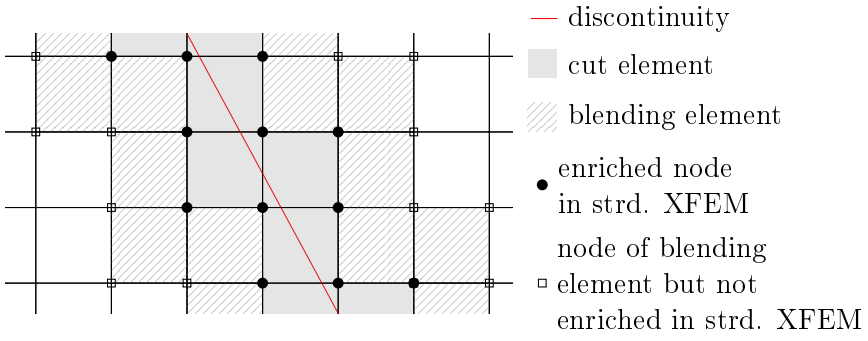


Figure 4.2 – Illustration of local enrichment for the mesh composed of bi-linear quadrangle elements.

In general, there are two types of discontinuity: strong and weak (singularity is not considered in the present study). A variable being of strong discontinuity means the variable undergoes a jump across the interface in its solution space. A weak discontinuity, also called kink, means the variable itself is continuous but its first-order derivative has a jump at the interface.

The global enrichment function plays the pivotal role on constructing the approximation space. For strong discontinuity, the global enrichment function always adopts the sign-enrichment

scheme (or Heaviside-enrichment scheme), which reads [10]

$$\psi_i(\vec{x},t) = sign(\phi(\vec{x},t)) - sign(\phi_i(t)) \tag{4.3}$$

where ϕ is the level set function, which is explained in section 4.2.4. Fig. 4.3 provides a visualized view of the local enrichment functions (M_i) corresponding to the sign-enrichment scheme for a certain given discontinuity in the bi-linear quadrangle reference element.

For weak discontinuity, there exist four typical types of enrichment schemes, as listed below.

1. Standard abs-enrichment scheme [10]

$$\psi_i(\vec{x},t) = |\phi(\vec{x},t)| - |\phi_i(t)| \tag{4.4}$$

2. Modified abs-enrichment scheme [17]

$$\psi(\vec{x},t) = N_j(\vec{x})|\phi_j(t)| - |N_j(\vec{x})\phi_j(t)| \tag{4.5}$$

   where Einstein notation is complied with and j indicates the node belonging to the element into which $\vec{x}$ falls. In this scheme, the global enrichment function ψ is not bound to any specific element node.

3. Corrected abs-enrichment scheme [9]

$$\psi_j(\vec{x},t) = (|\phi(\vec{x},t)| - |\phi_j(t)|)R = (|\phi(\vec{x},t)| - |\phi_j(t)|)\sum_{i\in\mathcal{N}_e} N_i(\vec{x}) \tag{4.6}$$

   where $j \in \mathcal{N}_{e,expanded}$, which stands for the set of nodes belonging to either cut elements or blending elements; R is the ramping function.

4. Sign-enrichment scheme (see Eq. 4.3)

It is well known that the standard abs-enrichment scheme would cause problems in the blending elements (see Fig. 4.2), for the reason that the blending element being only partially enriched makes the property of partition of unity lost [7]. The blending element problem is crucial for the performance and efficiency of the adopted abs-enrichment scheme. Both modified abs-enrichment scheme and corrected abs-enrichment scheme are developed to surmount this problem. Fig. 4.4 visualizes the above three different abs-enrichment schemes by picturing ψ of the central node in a 2×2 mesh, which contains 3 cut elements and 1 blending element. It can be seen that, compared with the standard abs-enrichment scheme, the modified scheme changes the global enrichment greatly, especially in the blending element it vanishes completely. By contrast, the corrected scheme only affects the global enrichment in the blending element with the aid of the ramping function. In [5], these two schemes were compared, with the conclusion that the corrected abs-enrichment is better than the modified abs-enrichment in terms of the convergence rate. The sign-enrichment scheme also can be applied for weak discontinuity. If

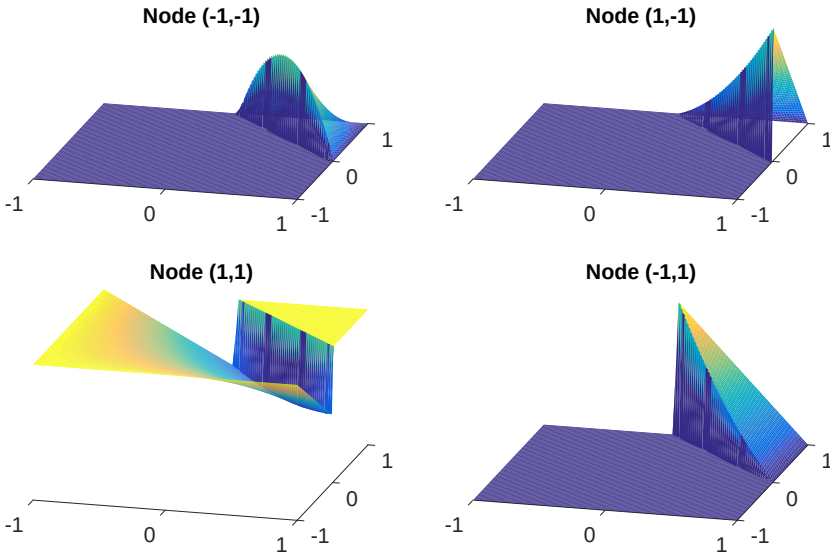


Figure 4.3 – Local enrichment functions of sign-enrichment scheme in a bi-linear quadrangle reference element with the discontinuity cutting through edge points (1,0) and (0,1).

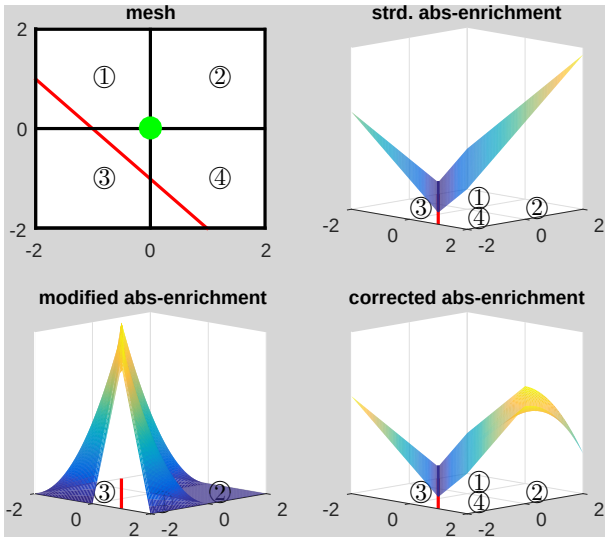


Figure 4.4 – Visualization of three types of abs-enrichment schemes by picturing ψ of the central node in a 2×2 mesh (the red line denotes the discontinuity and the green dot denotes the concerned central node).

so, the blending element problem can be avoided. But continuity at the interface must be enforced, which comes with more complexity.

In construction of the approximation space with weak discontinuity in the present study, the corrected abs-enrichment scheme is preferred due to its close-to-optimal convergence rate and relative easiness of implementation.

4.2.2 Heat transfer sub-model

4.2.2.1 Governing equations

As shown in Fig. 4.5, for a simple liquid-solid phase change problem, the computational domain $\Omega \subset \mathbb{R}^2$ is divided by the moving interface Γ_I into liquid region Ω_l and solid region Ω_s, such that $\Omega_l \cup \Omega_s = \Omega$ and $\Omega_l \cap \Omega_s = \Gamma_I$. The subscripts l and s denote variables in the liquid and solid phases, respectively. Γ_D denotes the exterior boundary of Dirichlet type, while Γ_N denotes the Neumann type. $\vec{n}$ represents the outward unit normal on the exterior boundary. $\vec{n}_\Gamma$ is the unit normal at the interface, by convention pointing from the solid to the liquid. The temperature at location $\vec{x}$ and at the time t is $T(\vec{x}, t)$.

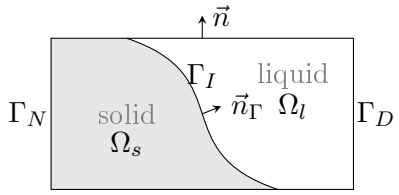


Figure 4.5 – Schematic of computational domain for phase change problem.

The evolution of temperature over the domain Ω is governed by the energy conservation equation below [6]:

$$\rho C_p \frac{\partial T}{\partial t} - \nabla \cdot (k \nabla T) - s = 0 \tag{4.7}$$

where ρ is the density; C_p is the specific heat; k is the thermal conductivity; s is the heat source. In the present study, the dynamic flow induced probably by the buoyancy force or liquid volume change is ignored. So the thermal convection mechanism is not taken into account. All the physical properties involved in Eq. 4.7 are discontinuous at the interface Γ_I and assumed to be temperature dependent in the stage of model building. Temperature at the interface Γ_I has to satisfy the following constraint (without considering Gibbs-Thompson effect),

$$T = T_m \tag{4.8}$$

where T_m is the melting point, which varies with the species concentration at the interface. Eq. 4.8 can be enforced through the penalty method. To make the problem well-posed, the following initial and external boundary conditions are also required,

$$T = T_0 \quad ; \qquad \vec{x} \in \Omega, \quad t = 0 \tag{4.9a}$$

$$T = T_D \quad ; \qquad \vec{x} \in \Gamma_D, t \in (0, \infty) \tag{4.9b}$$

$$-k \nabla T \cdot \vec{n} = q_{n,T} \; ; \qquad \vec{x} \in \Gamma_N, t \in (0, \infty) \tag{4.9c}$$

where T_0, T_D and $q_{n,T}$ are given value.

4.2.2.2 Weak formulation and discretization

With interface constraint (Eq. 4.8) imposed by the penalty method, the weak form of Eq. 4.7 can be written as

$$\begin{aligned} W_T = &\int_\Omega \delta T \rho C_p \frac{\partial T}{\partial t} d\Omega + \int_\Omega \nabla \delta T \cdot k \nabla T d\Omega - \int_\Omega \delta T s d\Omega + \int_{\Gamma_N} \delta T q_{n,T} d\Gamma \\ &+ \int_{\Gamma_I} \delta T \beta_T (T - T_m) d\Gamma = 0 \end{aligned} \tag{4.10}$$

where δT is the test function and β_T is the penalty parameter. Applying the Euler backward scheme to approximate the time derivative in Eq. 4.10, one obtains

$$\begin{aligned} W_T = &\frac{1}{\Delta t} \int_\Omega \delta T \rho C_p T d\Omega - \frac{1}{\Delta t} \int_\Omega \delta T \rho C_p T^{t-\Delta t} d\Omega + \int_\Omega \nabla \delta T \cdot k \nabla T d\Omega - \int_\Omega \delta T s d\Omega \\ &+ \int_{\Gamma_N} \delta T q_{n,T} d\Gamma + \int_{\Gamma_I} \delta T \beta_T (T - T_m) d\Gamma = 0 \end{aligned} \tag{4.11}$$

where Δt is the time step size and the superscript $t - \Delta t$ denotes the previous time step. Variables without the superscript $t - \Delta t$ carry the value of current time step by default.

Since the temperature is of weak discontinuity at the interface, the corrected abs-enrichment scheme is used to build the shape function. The specific form of Eq. 4.1 for temperature based on one element Ω^e containing n^e nodes can be written as

$$T = \langle N \rangle \{T\} \tag{4.12}$$

where

$$\langle N \rangle = \langle \; N_1 \quad N_2 \quad \cdots \quad N_{n^e} \quad M_1 \quad M_2 \quad \cdots \quad M_{n^e} \; \rangle \tag{4.13a}$$

$$\{T\} = \langle \; T_1 \quad T_2 \quad \cdots \quad T_{n^e} \quad \breve{T}_1 \quad \breve{T}_2 \quad \cdots \quad \breve{T}_{n^e} \; \rangle^{\mathrm{T}} \tag{4.13b}$$

where $\breve{T}$ denotes the additional degree of freedom of temperature.

After introducing Eq. 4.12, Eq. 4.11 can be written into the following discrete increment form:

$$\mathbf{R}_{\mathbf{M}}^{i-1} + \mathbf{R}_{\mathbf{K}}^{i-1} + \mathbf{R}_{\mathbf{F}}^{i-1} + \left(\frac{1}{\Delta t} \mathbf{M}_t^{i-1} + \mathbf{K}_t^{i-1} + \mathbf{F}_t^{i-1} \right) \{\Delta T\} = 0 \tag{4.14}$$

where

$$\mathbf{R}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \{N\} \left(\rho C_p\right)^{i-1} \langle N \rangle \frac{\{T\}^{i-1} - \{T\}^{t-\Delta t}}{\Delta t} d\Omega \tag{4.15a}$$

$$\mathbf{R}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} k^{i-1} [B] \{T\}^{i-1} d\Omega \tag{4.15b}$$

$$\mathbf{R}_{\mathbf{F}}^{i-1} = -\sum_e \int_{\Omega^e} \{N\} s^{i-1} d\Omega + \sum_e \int_{\Gamma_N^e} \{N\} q_{n,T}^{i-1} d\Gamma + \sum_e \int_{\Gamma_I^e} \{N\} \beta_T \Big(\langle N \rangle \{T\}^{i-1} - T_m \Big) d\Gamma \tag{4.15c}$$

$$\mathbf{M}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\} \Bigg(\left(\rho C_p\right)^{i-1} + \left(\rho' C_p + \rho C_p'\right)^{i-1} \langle N \rangle \Big(\{T\}^{i-1} - \{T\}^{t-\Delta t} \Big) \Bigg) \langle N \rangle d\Omega \tag{4.15d}$$

$$\mathbf{K}_t^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \Big(k^{i-1} [B] + \left(k'\right)^{i-1} [B] \{T\}^{i-1} \langle N \rangle \Big) d\Omega \tag{4.15e}$$

$$\mathbf{F}_t^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} \left(q'_{n,T}\right)^{i-1} \langle N \rangle d\Gamma - \sum_e \int_{\Omega^e} \{N\} \left(s'\right)^{i-1} \langle N \rangle d\Omega + \sum_e \int_{\Gamma_I^e} \{N\} \beta_T \langle N \rangle d\Gamma \tag{4.15f}$$

$\Delta T = T^i - T^{i-1}$ is the increment between two adjacent iteration steps and the primed variable stands for the derivative with respect to temperature. In the derivation of Eq. 4.14, the Newton-Raphson method is used for linearization, since all the physical properties (ρ, C_p, k, s and $q_{n,T}$) are assumed to be temperature dependent. Throughout this paper, all the interpolation functions (enriched or unenriched) are denoted by the vector $\langle N \rangle$, which is a dynamic container with mutable size; $[B] = \nabla \langle N \rangle$ is the gradient operator; superscript $i-1$ denotes the value at previous iteration step.

4.2.2.3 Evaluation of interface normal speed

The interface normal speed V_I is governed by the following Stefan condition at the interface [6]:

$$[\![q]\!] = \rho_s L V_I \tag{4.16}$$

where $[\![q]\!] = k_s \frac{\partial T}{\partial n} - k_l \frac{\partial T}{\partial n}$ is the heat flux jump in the normal direction at the interface, which is the only driving force of the interface propagation; L is the latent heat released or absorbed as the phase change occurs. Appropriate evaluation of the heat flux jump is critical. Theoretically, one can pick up one evaluation point on each side of the interface for the evaluation. It seems that the closer (e.g. on the order of magnitude 10^{-6} of the element size) the distance between the evaluation point and the interface is, more accurate result can be obtained. But actually it's not the case, because small oscillations that happen in the temperature field near the interface always lead to bad evaluation [14, 2]. For this reason, the solution proposed in [2] is adopted in the present study. As shown in Fig. 4.6, five points in each phase (the interface also included) are selected in the normal direction of the interface. δd is an adjustable parameter which indicates the distance between the furthest evaluation point and the interface. Based on the discrete data of temperature at those points, one can draw the least squares regression lines on both sides of the interface. The slope of the line in

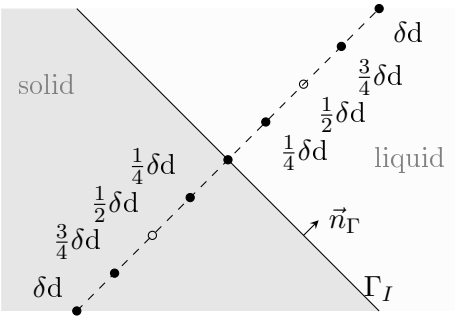


Figure 4.6 – Layout of the evaluation points for heat flux jump calculation.

each phase is used to approximate the temperature normal gradient. Then Eq. 4.16 would be rewritten as

$$V_I = \frac{1}{\rho L}\left(k_s \frac{2}{5}\frac{2T_m + T_s^{\frac{1}{4}\delta d} - T_s^{\frac{3}{4}\delta d} - 2T_s^{\delta d}}{\delta d} - k_l \frac{2}{5}\frac{2T_l^{\delta d} + T_l^{\frac{3}{4}\delta d} - T_l^{\frac{1}{4}\delta d} - 2T_m}{\delta d}\right) \tag{4.17}$$

4.2.3 Species transport sub-model

4.2.3.1 Governing equations

Theoretically, the density is neither uniform nor constant in the binary solution system obtained by mixing two chemical species. It always depends on the local solution composition. Given this fact, in order to preserve the capability to deal with the case involving a mutable density, the following mass conservation equation [19] is used

$$\frac{\partial \rho\omega}{\partial t} - \nabla \cdot (\rho D \nabla \omega) = 0 \tag{4.18}$$

where ω is the dimensionless mass fraction of the species that is regarded as solute, i.e. $\frac{\text{solute mass}}{\text{mixture mass}}$ in unit volume; D is the diffusion coefficient; the density of the mixture ρ is identical to that in Eq. 4.7. In most majority of the published works [27, 21, 31, 32, 26], the mass conservation equation took molar concentration c (mol/l) as the primitive variable. In such situation, the model built based on Eq. 4.18 can also be used by assigning unity to ρ; then, ω could represent molar concentration.

In the binary solution, the chemical composition in one phase tends to be distinct from that in the other phase. As the phase change takes place, one species would be absorbed or rejected by the phase front. This phenomenon would lead to a jump in the species concentration field at the interface. It means the unknown ω in Eq. 4.18 is of strong discontinuity. Then, the sign-enrichment scheme (see Eq. 4.3) should be employed to build the approximation space of ω.

Since ω is treated as being strongly discontinuous, some constraints are required to be enforced at the interface to make the problem well-posed. On the solid side of the interface (denoted

by Γ_{I^-}), the following Dirichlet-type condition is imposed:

$$\omega(\vec{x},t) = \omega_s, \qquad \vec{x} \in \Gamma_{I^-} \tag{4.19}$$

where ω_s refers to the value on the solidus line in phase diagram. On the liquid side of the interface (denoted by Γ_{I^+}), the following mass flux derived according to the mass conservation law is enforced as Neumann-type condition:

$$-\rho D\nabla\omega\cdot\vec{n}_{\Gamma_I} = \rho_s\big(\omega(\vec{x},t)-\omega_s\big)V_I, \qquad \vec{x}\in\Gamma_{I^+} \tag{4.20}$$

where ρ_s is the solid density at the interface. According to the phase diagram, the species concentration on the solid side ω_s is related to the species concentration on the liquid side $\omega(\vec{x},t)$ as $\vec{x}\in\Gamma_{I^+}$. This relationship is mostly approximated linearly through a partition coefficient k_p, i.e. $\omega_s = k_p\omega(\vec{x},t)$ [26]. The value of partition coefficient is physical problem dependent; for instance, in $NaCl-H_2O$ solution, the solid phase is composed of pure ice, indicating $k_p = 0$. Without loss of generality, the mathematical analysis is performed under the general case, i.e. ρ_s and ω_s are taken as unknown dependent.

Besides the constraints at the interface, the following initial and external boundary conditions are also required to make the equation system closed:

$$\omega = \omega_0 \quad ; \qquad \vec{x}\in\Omega, \quad t=0 \tag{4.21a}$$

$$\omega = \omega_D \quad ; \qquad \vec{x}\in\Gamma_D, t\in(0,\infty) \tag{4.21b}$$

$$-\rho D\nabla\omega\cdot\vec{n} = q_{n,\omega}\ ; \qquad \vec{x}\in\Gamma_N, t\in(0,\infty) \tag{4.21c}$$

where ω_0, ω_D and $q_{n,\omega}$ are given value.

4.2.3.2 Weak formulation and discretization

With the enforcement of the interface constraint Eq. 4.19 by the penalty method as well as Eq. 4.20 naturally added in, the weak form of Eq. 4.18 can be written as

$$\begin{aligned} W_\omega = &\int_\Omega \delta\omega\Big(\rho+\omega\rho'\Big)\frac{\partial\omega}{\partial t}d\Omega + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \int_{\Gamma_N}\delta\omega q_{n,\omega}d\Gamma \\ &+\int_{\Gamma_{I^+}}\delta\omega\rho_s\big(\omega-\omega_s\big)V_I d\Gamma + \int_{\Gamma_{I^-}}\delta\omega\beta_\omega(\omega-\omega_s)d\Gamma = 0 \end{aligned} \tag{4.22}$$

where $\delta\omega$ is the test function; the primed quantity represents the derivative with respect to ω; β_ω is the penalty parameter. Applying the Euler backward scheme to Eq. 4.22, one can obtain

$$\begin{aligned} W_\omega = &\frac{1}{\Delta t}\int_\Omega \delta\omega\Big(\rho+\omega\rho'\Big)\big(\omega-\omega^{t-\Delta t}\big)d\Omega + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \int_{\Gamma_N}\delta\omega q_{n,\omega}d\Gamma \\ &+\int_{\Gamma_{I^+}}\delta\omega\rho_s\big(\omega-\omega_s\big)V_I d\Gamma + \int_{\Gamma_{I^-}}\delta\omega\beta_\omega(\omega-\omega_s)d\Gamma = 0 \end{aligned} \tag{4.23}$$

The specific form of Eq. 4.1 for species concentration based on element Ω^e containing n^e nodes reads

$$\omega = \langle N \rangle \{\omega\} \tag{4.24}$$

where

$$\{\omega\} = \langle \ \omega_1 \quad \omega_2 \quad \cdots \quad \omega_{n^e} \quad \breve{\omega}_1 \quad \breve{\omega}_2 \quad \cdots \quad \breve{\omega}_{n^e} \ \rangle^{\mathrm{T}} \tag{4.25}$$

where $\breve{\omega}$ represents the additional degree freedom of species concentration; the shape function $\langle N \rangle$ is constructed based on the sign-enrichment scheme (Eq. 4.3), due to ω being of strong discontinuity. After introducing Eq. 4.24 and applying Newton-Raphson method for linearization, Eq. 4.23 can be written into the following discrete form:

$$\bar{\mathbf{R}}_{\mathbf{M}}^{i-1} + \bar{\mathbf{R}}_{\mathbf{K}}^{i-1} + \bar{\mathbf{R}}_{\mathbf{F}}^{i-1} + \left(\frac{1}{\Delta t} \bar{\mathbf{M}}_t^{i-1} + \bar{\mathbf{K}}_t^{i-1} + \bar{\mathbf{F}}_t^{i-1} \right) \{\Delta\omega\} = 0 \tag{4.26}$$

where

$$\bar{\mathbf{R}}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \{N\} \left(\rho + \omega\rho'\right)^{i-1} \langle N \rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{4.27a}$$

$$\bar{\mathbf{R}}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \left(\rho D\right)^{i-1} [B] \{\omega\}^{i-1} d\Omega \tag{4.27b}$$

$$\begin{aligned} \bar{\mathbf{R}}_{\mathbf{F}}^{i-1} = & \sum_e \int_{\Gamma_N^e} \{N\} q_{n,\omega}^{i-1} d\Gamma + \sum_e \int_{\Gamma_{I+}^e} \{N\} \rho_s^{i-1} \Big(\langle N \rangle \{\omega\}^{i-1} - \omega_s^{i-1} \Big) V_I d\Gamma \\ & + \sum_e \int_{\Gamma_{I-}^e} \{N\} \beta_\omega \Big(\langle N \rangle \{\omega\}^{i-1} - \omega_s^{i-1} \Big) d\Gamma \end{aligned} \tag{4.27c}$$

$$\bar{\mathbf{M}}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\} \left(\left(\rho + \omega\rho'\right)^{i-1} + \left(2\rho' + \omega\rho''\right)^{i-1} \langle N \rangle \Big(\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t} \Big) \right) \langle N \rangle d\Omega \tag{4.27d}$$

$$\bar{\mathbf{K}}_t^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \Big(\left(\rho D\right)^{i-1} [B] + \left(\rho' D + \rho D'\right)^{i-1} [B] \{\omega\}^{i-1} \langle N \rangle \Big) d\Omega \tag{4.27e}$$

$$\begin{aligned} \bar{\mathbf{F}}_t^{i-1} = & \sum_e \int_{\Gamma_N^e} \{N\} \left(q'_{n,\omega}\right)^{i-1} \langle N \rangle d\Gamma + \sum_e \int_{\Gamma_{I-}^e} \{N\} \beta_\omega \Big(1 - \left(\omega'_s\right)^{i-1} \Big) \langle N \rangle d\Gamma \\ & + \sum_e \int_{\Gamma_{I+}^e} \{N\} \left(\left(\rho'_s\right)^{i-1} \Big(\langle N \rangle \{\omega\}^{i-1} - \omega_s^{i-1} \Big) + \rho_s^{i-1} - \left(\rho_s \omega'_s\right)^{i-1} \right) \langle N \rangle V_I d\Gamma \end{aligned} \tag{4.27f}$$

In the derivation of Eq. 4.26, the physical quantities including ρ, D, $q_{\omega,n}$ and ω_s are assumed to depend on the species concentration.

4.2.4 Level set transport sub-model

4.2.4.1 Governing equations

In level set method, the unknown variable $\phi(\vec{x}, t)$ is initially defined as the shortest signed distance function away from the interface:

$$\phi(\vec{x}, t) = sign\big((\vec{x} - \vec{x}_I) \cdot \vec{n}_\Gamma \big) \min_{\vec{x}_I \in \Gamma_I} \|\vec{x} - \vec{x}_I\|, \qquad \vec{x} \in \Omega \tag{4.28}$$

The interface Γ_I corresponds to the set of points at which $\phi(\vec{x},t)=0$. The solid domain Ω_s is defined as the set of points that satisfy $\phi(\vec{x},t)<0$ while the points belonging to the liquid domain Ω_l satisfy $\phi(\vec{x},t)>0$. Then, interface capture is converted to a transport problem of $\phi(\vec{x},t)$. As the interface propagates, $\phi(\vec{x},t)$ must be correspondingly updated by the following conservation equation [18]:

$$\frac{\partial\phi}{\partial t}+\vec{v}\cdot\nabla\phi=0 \tag{4.29}$$

where $\vec{v}$ is the velocity field to advect ϕ. The unit normal vector of the interface can be obtained by solving the following equation:

$$\vec{n}_\Gamma=\frac{1}{\|\nabla\phi\|}\nabla\phi \tag{4.30}$$

Introducing Eq. 4.30 into Eq. 4.29, one obtains [18]

$$\frac{\partial\phi}{\partial t}+F\|\nabla\phi\|=0 \tag{4.31}$$

where $F=\vec{v}\cdot\vec{n}_\Gamma$ is the normal speed field, whose value at the interface should coincide with V_I calculated by Eq. 4.17 in the heat transfer sub-model. The construction of F is presented in section 4.2.5.

4.2.4.2 Weak formulation and discretization

Since Eq. 4.31 is a first order hyperbolic equation and the unknown $\phi(\vec{x},t)$ is smooth over Ω, the standard finite element method with Galerkin/least-square (GLS) stabilization technique [6] can be used to solve it. The weak form of Eq. 4.31 is

$$\int_\Omega\delta\phi\frac{\partial\phi}{\partial t}d\Omega+\int_\Omega\delta\phi F\|\nabla\phi\|d\Omega+\sum_e\int_{\Omega^e}\Big(F\nabla\delta\phi\cdot\frac{1}{\|\nabla\phi\|}\nabla\phi\Big)\tau_\phi^e\Big(\frac{\partial\phi}{\partial t}+F\|\nabla\phi\|\Big)d\Omega=0 \tag{4.32}$$

where $\delta\phi$ is the test function; τ_ϕ^e is the element stabilization parameter defined as $\tau_\phi^e=\frac{h^e}{|F^e|}$, in which h^e is the element shape factor and often takes the value of element size [6].

The approximation function of ϕ based on element Ω^e containing n^e nodes can be written as

$$\phi=\langle N\rangle\{\phi\} \tag{4.33}$$

where $\{\phi\}=\langle\ \phi_1\quad\phi_2\quad\cdots\quad\phi_{n^e}\ \rangle^{\mathrm{T}}$. After introducing Eq. 4.33 and performing the forward Euler time integration scheme, Eq. 4.32 can be written into the following discrete increment form:

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1}+\tilde{\mathbf{R}}_{\mathbf{C}}+\tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1}+\Big(\tilde{\mathbf{M}}+\tilde{\mathbf{M}}_{\mathbf{GLS}}\Big)\{\Delta\phi\}=0 \tag{4.34}$$

where

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1}=\sum_e\int_{\Omega^e}\{N\}\langle N\rangle\frac{\{\phi\}^{i-1}-\{\phi\}^{t-\Delta t}}{\Delta t}d\Omega \tag{4.35a}$$

$$\tilde{\mathbf{R}}_{\mathbf{C}} = \sum_e \int_{\Omega^e} \{N\} F \left\| [B]\{\phi\}^{t-\Delta t} \right\| d\Omega \tag{4.35b}$$

$$\tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1} = \sum_e \int_{\Omega^e} F[B]^{\mathrm{T}} \frac{[B]\{\phi\}^{t-\Delta t}}{\left\| [B]\{\phi\}^{t-\Delta t} \right\|} \tau_\phi^e \Big(\langle N \rangle \frac{\{\phi\}^{i-1} - \{\phi\}^{t-\Delta t}}{\Delta t} + F \left\| [B]\{\phi\}^{t-\Delta t} \right\| \Big) d\Omega \tag{4.35c}$$

$$\tilde{\mathbf{M}} = \sum_e \int_{\Omega^e} \frac{1}{\Delta t} \{N\} \langle N \rangle d\Omega \tag{4.35d}$$

$$\tilde{\mathbf{M}}_{\mathbf{GLS}} = \sum_e \int_{\Omega^e} F[B]^{\mathrm{T}} \frac{[B]\{\phi\}^{t-\Delta t}}{\left\| [B]\{\phi\}^{t-\Delta t} \right\|} \tau_\phi^e \langle N \rangle d\Omega \tag{4.35e}$$

4.2.5 Normal speed spreading sub-model

4.2.5.1 Governing equations

One obtains the normal speed only at the interface (V_I) by solving Eq. 4.17. The updating of ϕ by solving Eq. 4.31 requires to spread V_I over the entire domain to construct F. For this purpose, the following construction equation is used [6]:

$$sign(\phi) \nabla\phi \cdot \nabla F = 0 \tag{4.36}$$

with boundary condition at the interface:

$$F(\vec{x}) = V_I, \qquad \vec{x} \in \Gamma_I \tag{4.37}$$

Eq. 4.36 implies that F is constructed to be orthogonal to ϕ (see Fig. 4.7), so the property of ϕ being signed distance function would be conserved to some extent [6]. The term $sign(\phi)$ in Eq. 4.36 exists to make Eq. 4.37 become the influx boundary condition. When the interface does not form a closed path, the influx boundary should be augmented by some segments of the external boundary that satisfy the criterion $\vec{n} \cdot \big(sign(\phi) \nabla\phi(\vec{x}_\Gamma) \big) < 0$, as shown in Fig. 4.7. See [6] for more details.

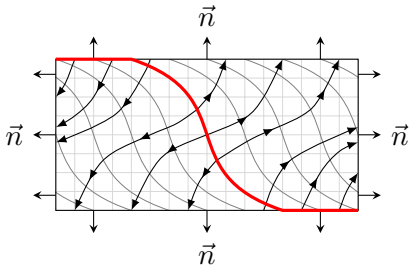


Figure 4.7 – Schematic drawing of spreading the normal speed field; grey curves represent the contours of ϕ created based on Fig. 4.5; black curves with arrows represent the characteristics of Eq. 4.36; the single red curve represents the augmented boundary.

4.2.5.2 Weak formulation and discretization

The standard finite element method with GLS stabilization scheme is adopted to solve Eq. 4.36. The weak formulation of Eq. 4.36 gives

$$\int_{\Omega} \delta F sign(\phi) \nabla\phi \cdot \nabla F d\Omega + \sum_{e} \int_{\Omega^e} \big(\nabla \delta F \cdot \nabla\phi\big) \tau_F^e \big(\nabla\phi \cdot \nabla F\big) d\Omega = 0 \tag{4.38}$$

where the stabilization parameter τ_F^e is defined as $\tau_F^e = \frac{h^e}{\|\nabla\phi\|}$. After discretization, the following discrete algebraic equation system is obtained:

$$\Big(\hat{\mathbf{K}} + \hat{\mathbf{K}}_{\mathbf{GLS}}\Big)\{F\} = 0 \tag{4.39}$$

where

$$\hat{\mathbf{K}} = \sum_{e} \int_{\Omega^e} \{N\} sign(\phi) \big(\nabla\phi\big)^{\mathrm{T}} [B] d\Omega \tag{4.40a}$$

$$\hat{\mathbf{K}}_{\mathbf{GLS}} = \sum_{e} \int_{\Omega^e} [B]^{\mathrm{T}} \nabla\phi \tau_F^e \big(\nabla\phi\big)^{\mathrm{T}} [B] d\Omega \tag{4.40b}$$

4.2.6 Computational procedure

The algorithm to solve the Stefan problem with considering the influence of mass transport in the binary system can be summarized as follow:

1. Identify the computational domain and initial position of the interface.
2. Initialize ϕ, T and ω.
3. Evaluate interface normal speed V_I by solving Eq. 4.17.
4. **Normal speed spreading sub-model**:
   a) identify the influx boundary and the Dirichlet-type condition;
   b) obtain F by solving Eq. 4.39.
5. **Level set transport sub-model**: update ϕ by solving Eq. 4.34.
6. **Heat transfer sub-model** and **Species transport sub-model** are solved alternately until both get converged. At each alternation, the relevant physical properties are updated. In each sub-model, the computation proceeds as below:
   a) loop over all the elements and at each element
      i. check if cut or not; if so, do splitting and update the integration points;
      ii. check if enriched or not; if so, update the enrichment function;
      iii. contribute the elementary matrices and residual vectors to Eq. 4.14 or Eq. 4.26;
   b) update T or ω by solving Eq. 4.14 or Eq. 4.26.
7. Go back to step 3 until the end.

4.3 Numerical tests

4.3.1 Generalities

$NaCl - H_2O$ solution could be the most common saline solution in our daily life, whose solidification process has been relatively well studied through numerical methods [21, 31, 32]. Among these published works, most majority worked only on 1D case. In this section, the 1D case whose results are available in the literature is taken as benchmark problem to verify the model of the present study. After that, the model is applied to deal with a 2D case in which the initial interface is inclined. The thermophysical data employed in the present modelling are the same as those in [21], as listed in Table. 4.1. It can be seen that the thermal diffusivity, denoted by α, is larger by two to three orders of magnitude from the mass diffusion coefficient. It implies that, compared with heat, the rejected species is more likely to get accumulated at the ice front. This phenomenon is confirmed in the following simulations. In addition, the freezing point is defined as a function of the concentration, $T_m = -1.86c$ [21], where c denotes the molar concentration in unit mol/l.

Table 4.1 – Thermophysical data employed for the $NaCl - H_2O$ solution

Items / Material	ρ (kg/m^3)	k $(W/m \cdot K)$	c_p $(J/kg \cdot K)$	α (m^2/s)	L (J/kg)	D (m^2/s)
Liquid	998	0.555	4220	1.32×10^{-7}	3.53×10^5	1.18×10^{-9}
Ice	912	2.25	1830	1.35×10^{-6}		/

4.3.2 1D case: the interface is vertical

Although the model in the present study is designed for 2D problems, 1D simulation can be achieved by starting it up with initial and boundary conditions that are uniform in the ignored dimension. The computational domain is $0.51mm \times 0.1mm$ in size. Fig. 4.8 displays the mesh and initial location of the interface. Initially, it is pure ice on the left side of the interface

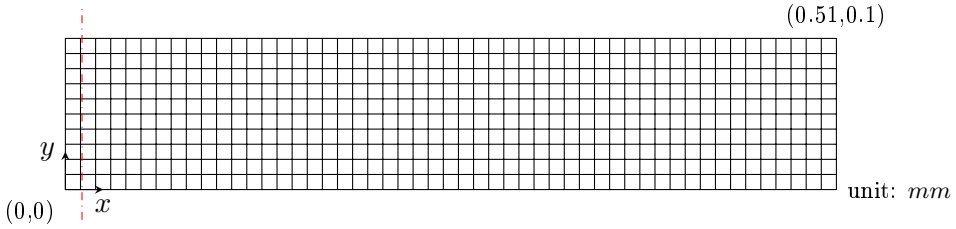


Figure 4.8 – The mesh containing 510 bi-linear quadrangle elements used for the 1D case and the initial interface location (at $x = 0.0102mm$) indicated by the red dash-dotted line.

while on the right it is filled with saline solution with uniform concentration $2 \times 10^{-3} mol/l$. To trigger the solidification process, a cooling function defined as $T = -5 \times 10^{-3} t - 0.0105$ is

enforced at $x = 0mm$. All the other sides are set to be insulated. The remaining parameters are set as $\Delta t = 0.01s$, $\delta d = 0.25h^e$, $\beta_T = \beta_\omega = 10^8$.

Fig. 4.9 shows the calculated temperature and concentration profiles at some moments as well as the comparison with those in [21]. All the drawings are in the logarithmic scale as that in

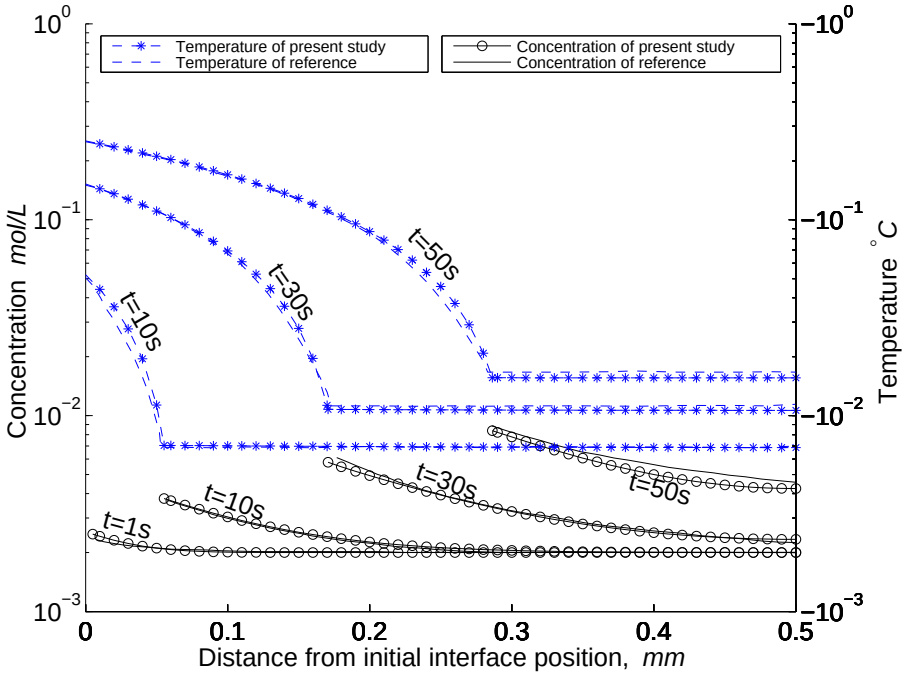


Figure 4.9 – Comparison of temperature and concentration profiles at some moments with those in [21]; the inflection points on temperature curves, aligned with the left tips of concentration curves, indicate the positions of the interface.

[21]. It can be seen that the results obtained by the present model have a good agreement with those in [21]. In Fig. 4.9, the inflection points in temperature curves, which are always aligned with the left tips of concentration curves, indicate the positions of the interface at each moment. As shown in Fig. 4.9, as the temperature on the left end drops in the way prescribed by the cooling function, the ice front propagates rightwards. With the propagation of ice front, salt accumulates on the front, which decreases the freezing point. Due to the difference of diffusivity for thermal and mass transport, the distribution of heat is much more uniform than that of salt concentration in the liquid region.

4.3.3 2D case: the initial interface is inclined

For the 2D case, the freezing or thawing of saline solution takes place in a rectangular domain of $1mm \times 0.5mm$ in size with an initially inclined solid-liquid interface. Fig. 4.10 displays the mesh and the initial location of the interface. A cooling function defined as $T = -5\times10^{-3}t - 0.1$ is enforced at $x = 0mm$ within the first 100s. In the remaining period of the simulation, the temperature at $x = 0mm$ is kept as constant, $T = -0.6$℃. On the right end, the temperature

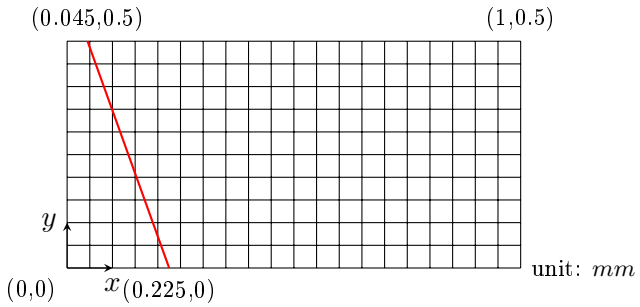


Figure 4.10 – The mesh containing 200 bi-linear quadrangle elements used for the 2D case and the initial interface location indicated by the red line.

is prescribed as 1.1℃ all the time. Both the top and bottom sides are set to be adiabatic. For the mass transport process, no mass flux is allowed crossing the exterior boundaries. Physically speaking, since the temperature on the right end is always higher than the freezing point, the simulation should finally reach a steady state that the propagation of the phase front comes to rest at some location inside the domain.

As shown in Fig. 4.10, the initial interface especially at the bottom is a little bit far from the left end. That requires a proper initial distribution of the temperature for a gentle start. For this purpose, a steady bi-material thermal problem, which can be regarded as a special case of Stefan problem with interface fixed, was modelled based on the boundary conditions described above. The obtained temperature distribution, as shown by the right sub-figure in Fig. 4.11 at $t = 0s$, was then employed as the initial condition for the transient simulation. The left sub-figure in Fig. 4.11 at $t = 0s$ presents the initial distribution of the salt concentration. Initially on the left side of the interface, it's pure ice; while on the right side, it's filled with saline solution of uniform concentration $2 \times 10^{-3} mol/l$. The remaining parameters are set as $\Delta t = 0.1s$, $\delta d = 0.25h^e$, $\beta_T = \beta_\omega = 10^8$.

The simulation takes $339.3s$ to reach the steady state. Fig. 4.11 shows the salt concentration (the left column of figures) and temperature (the right column of figures) distributions at some typical moments $t = 0s, 5s, 50s, 101.6s, 339.3s$. From these figures, we can see the evolutions of salt concentration, interface profile and temperature with time over the whole domain. By comparison, the interface locations can be easily recognized in the left column of figures rather than in those right ones, because of the strong discontinuity feature of the salt concentration at the interface. At the very early stage ($t = 5s$), the upper part of the interface moves faster than the lower part, which leads more salt accumulating at the upper part. As this trend continues, the interface almost becomes vertical at $t = 50s$, but the salt concentration on the top is still obviously larger than that at the lower part. With the ice front propagating rightwards, more salt accumulates at the ice front and its distribution along the interface gets more uniform. Compared with the temperature (representing the heat), the more obvious accumulation of the salt is due to the fact that the mass diffusivity is much smaller than

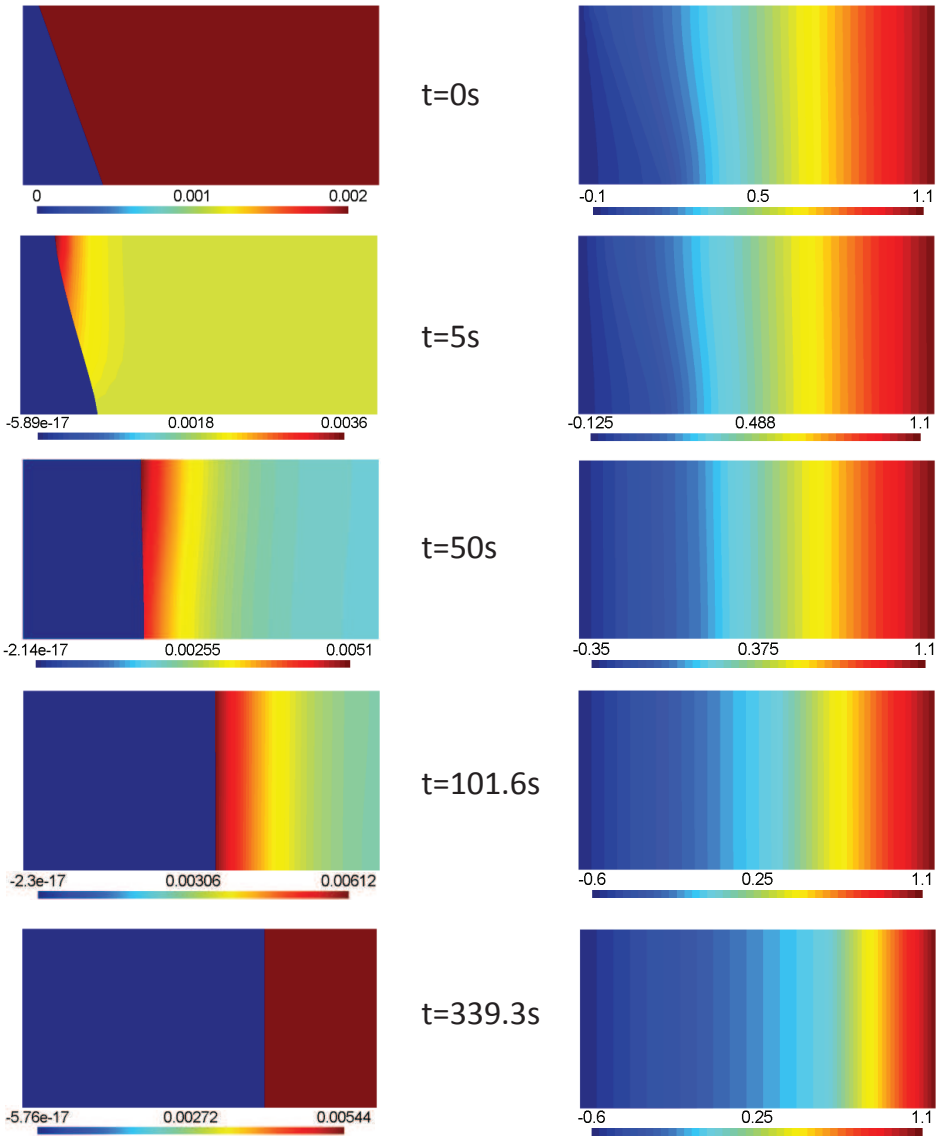
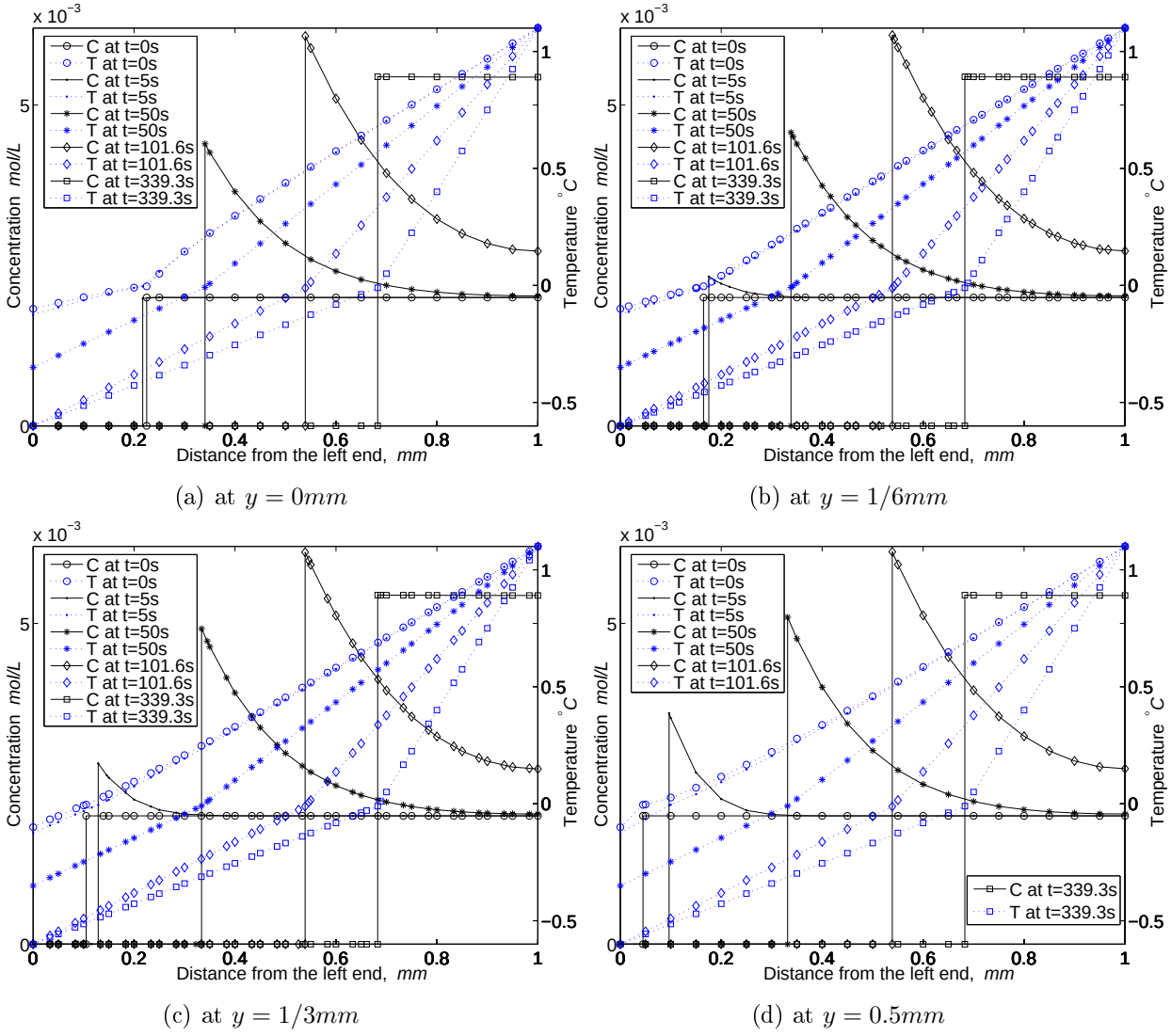


Figure 4.11 – Numerical results of salt concentration and temperature distributions at some moments; the left column of figures corresponding to the salt concentration, while the right column corresponding to the temperature.

the thermal diffusivity. At $t = 101.6s$, this salt accumulation reaches the maximum. Since then, the mass diffusion process in the liquid part starts to dominate compared with the front accumulation, because the motion of the interface gets slow with the descending of the heat jump at the interface.

The above description can also be verified in Fig. 4.12, which presents the evolution of salt concentration and temperature profiles along horizontal lines at $y = 0mm, 1/6mm, 1/3mm, 0.5mm$ extracted from the solution in Fig. 4.11. These figures reveal more details of what happens during this phase transition process. As shown in Fig. 4.12a, at the very early stage, the



(a) at $y = 0mm$ (b) at $y = 1/6mm$

(c) at $y = 1/3mm$ (d) at $y = 0.5mm$

Figure 4.12 – Evolution of temperature and salt concentration profiles at some cross sections; the inflection points on temperature curves, aligned with the breaking points on concentration curves, indicate the position of the interface.

ice front moves leftwards a little at the bottom of the domain, which implies the melting of a small amount of ice. The salt concentration in that region accordingly slightly drops to $1.986 \times 10^{-3} mol/l$ at $t = 5s$. The reason of the melting is that, the initial interface at the bottom is far from the left end, so initially the temperature gradient is relatively flat, which leads the heat flux flowing into solid part less than that running out from the liquid part. It also can be seen from Fig. 4.12 that, the interface, salt concentration and temperature profiles get almost uniform along the vertical direction after $t = 101.6s$, so it becomes 1D problem due to the vertically uniform-distributed boundary condition. Fig. 4.13 presents the interface profiles at the moments $t = 0s$, $5s$, $15s$, $20s$, $30s$, $50s$, $75s$, $101.6s$ and $339.3s$.

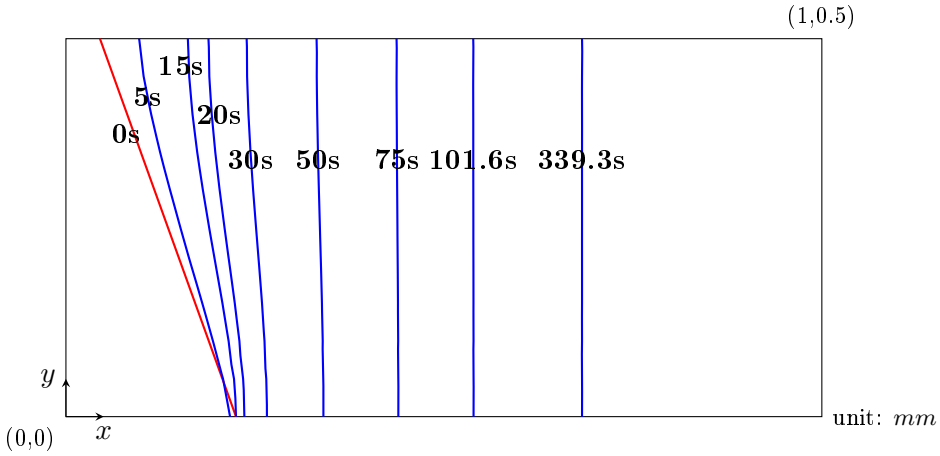


Figure 4.13 – The profiles of the interface at some moments.

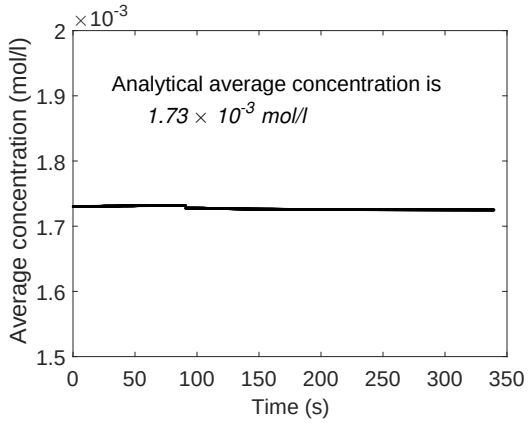


Figure 4.14 – The average salt concentration at each time step.

During the simulation, there is no mass flux crossing the exterior boundaries. It means the salt mass should always be conserved throughout. Initially, the salt only exists in the liquid solution with concentration $2 \times 10^{-3} mol/l$. The initial volume ratio that the liquid solution occupies over the entire domain is $0.4325 mm^2/0.5 mm^2 = 0.865$. So the average salt concentration with respect to the total volume in theory should always be $1.73 \times 10^{-3} mol/l$. Fig. 4.14 shows the average concentration at each time step. After calculation, the maximum relative error is 0.29%, which means the mass is well conserved.

4.4 Discussion and conclusion

In this work, a numerical model, based on XFEM and level set method, has been presented to describe the phase change process in the binary solution. The numerical tests demonstrate its accuracy and performance. With the help of XFEM and level set method, all the numerical tests were conducted with uniform and fixed mesh, which is attractive compared with those

methods requiring mesh local refinement or rebuilding. Within the framework of applying XFEM/level set method to solve phase change problem, it's a fresh attempt to take into account the mass transport process. In the mathematical analysis section, the mass transport sub-model is built based on a universal form of mass conservation equation, which is able to manage the case with variable density. This treatment gets it easy in the future to add the fluid flow model, to which the impact of variable density is important.

4.5 Acknowledgement

The authors would like to acknowledge the National Sciences and Engineering Research Council of Canada (NSERC) funding NSERC Discovery Grant (No. 36518) for supporting the present research. A part of the research presented in this paper was financed by the Fonds de Recherche du Québec-Nature et Technologie (FRQNT) by the intermediary of the Aluminium Research Centre-REGAL. The first author also acknowledges the financial support from the Chinese Scholarship Council (CSC).

Bibliography

[1] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601–620, 1999.

[2] M. K. Bernauer and R. Herzog. Implementation of an X-FEM solver for the classical two-phase Stefan problem. *Journal of Scientific Computing*, 52(2):271–293, 2012.

[3] W. J. Boettinger, J. A. Warren, C. Beckermann, and A. Karma. Phase-field simulation of solidification 1. *Annual Review of Materials Research*, 32(1):163–194, 2002.

[4] S. Chen, B. Merriman, S. Osher, and P. Smereka. A simple level set method for solving Stefan problems. *Journal of Computational Physics*, 135(1):8–29, 1997.

[5] K. W. Cheng and T. P. Fries. Higher-order XFEM for curved strong and weak discontinuities. *International Journal for Numerical Methods in Engineering*, 82(5):564–590, 2010.

[6] J. Chessa, P. Smolinski, and T. Belytschko. The extended finite element method (XFEM) for solidification problems. *International Journal for Numerical Methods in Engineering*, 53(8):1959–1977, 2002.

[7] J. Chessa, H. Wang, and T. BelytschkoM. On the construction of blending elements for local partition of unity enriched finite elements. *International Journal for Numerical Methods in Engineering*, 57:1015–1038, 2003.

[8] B. Echebarria, R. Folch, A. Karma, and M. Plapp. Quantitative phase-field model of alloy solidification. *Physical Review E*, 70(6):061604, 2004.

[9] T. P. Fries. A corrected XFEM approximation without problems in blending elements. *International Journal for Numerical Methods in Engineering*, 75(5):503–532, 2008.

[10] T. P. Fries and T. Belytschko. The extended/generalized finite element method: an overview of the method and its applications. *International Journal for Numerical Methods in Engineering*, 84(3):253–304, 2010.

[11] S. C. Gupta. *The classical Stefan problem: Basic concepts, modelling and analysis*. Elsevier, 2003.

[12] H. Hu and S. A. Argyropoulos. Mathematical modelling of solidification and melting: a review. *Modelling and Simulation in Materials Science and Engineering*, 4(4):371–396, 1996.

[13] E. Javierre, C. Vuik, F. J. Vermolen, and S. van der Zwaag. A comparison of numerical models for one-dimensional Stefan problems. *Journal of Computational and Applied Mathematics*, 192(2):445–459, 2006.

[14] H. Ji and J. E. Dolbow. On strategies for enforcing interfacial constraints and evaluating jump conditions with the extended finite element method. *International Journal for Numerical Methods in Engineering*, 61(14):2508–2535, 2004.

[15] D. Martin, H. Chaouki, J. L. Robert, M. Fafard, and D. Ziegler. A XFEM Lagrange multiplier technique for Stefan problems. *Frontiers in Heat and Mass Transfer*, 7(1), 2016.

[16] R. Merle and J. Dolbow. Solving thermal and phase change problems with the extended finite element method. *Computational Mechanics*, 28(5):339–350, 2002.

[17] N. Moës, M. Cloirec, P. Cartraud, and J. F. Remacle. A computational approach to handle complex microstructure geometries. *Computer Methods in Applied Mechanics and Engineering*, 192(28):3163–3177, 2003.

[18] S. Osher and R. P. Fedkiw. Level set methods: an overview and some recent results. *Journal of Computational Physics*, 169(2):463–502, 2001.

[19] T. J. Povich, C. N. Dawson, M. W. Farthing, and C. E. Kees. Finite element methods for variable density flow and solute transport. *Computational Geosciences*, 17(3):529–549, 2013.

[20] J. C. Ramirez, C. Beckermann, A. Karma, and H. J. Diepers. Phase-field modeling of binary alloy solidification with coupled heat and solute diffusion. *Physical Review E*, 69(5):051607, 2004.

[21] B. Rubinsky. Solidification processes in saline solutions. *Journal of Crystal Growth*, 62(3):513–522, 1983.

[22] L. Salvatori and N. Tosi. Stefan problem through extended finite elements: review and further investigations. *Algorithms*, 2(3):1177–1220, 2009.

[23] G. Segal, K. Vuik, and F. Vermolen. A conserving discretization for the free boundary in a two-dimensional Stefan problem. *Journal of Computational Physics*, 141(1):1–21, 1998.

[24] P. Stapór. The XFEM for nonlinear thermal and phase change problems. *International Journal of Numerical Methods for Heat & Fluid Flow*, 25(2):400–421, 2015.

[25] J. Stefan. Über die theorie der eisbildung, insbesondere über die eisbildung im polarmeere. *Annalen der Physik*, 278(2):269–286, 1891.

[26] M. Theillard, F. Gibou, and T. Pollock. A sharp computational method for the simulation of the solidification of binary alloys. *Journal of Scientific Computing*, 63(2):330–354, 2015.

[27] H. L. Tsai and B. Rubinsky. A "front tracking" finite element study on change of phase interface stability during solidification processes in solutions. *Journal of Crystal Growth*, 70(1-2):56–63, 1984.

[28] V. R. Voller, M. Cross, and N. C. Markatos. An enthalpy method for convection/diffusion phase change. *International Journal for Numerical Methods in Engineering*, 24(1):271–284, 1987.

[29] Q. Wang, B. Li, and M. Fafard. Effect of anode change on heat transfer and magneto-hydrodynamic flow in aluminum reduction cell. *JOM*, 68(2):610–622, 2016.

[30] S. W. J. Welch and J. Wilson. A volume of fluid based method for fluid flows with phase change. *Journal of Computational Physics*, 160(2):662–682, 2000.

[31] K. Wollhöver, C. Körber, M. W. Scheiwe, and U. Hartmann. Unidirectional freezing of binary aqueous solutions: an analysis of transient diffusion of heat and mass. *International Journal of Heat and Mass Transfer*, 28(4):761–769, 1985.

[32] Y. Yang and H. S. Udaykumar. Sharp interface Cartesian grid method III: solidification of pure materials and binary solutions. *Journal of Computational Physics*, 210(1):55–74, 2005.

[33] N. Zabaras, B. Ganapathysubramanian, and L. Tan. Modelling dendritic solidification with melt convection using the extended finite element method. *Journal of Computational Physics*, 218(1):200–227, 2006.

Chapter 5

Numerical simulation of Stefan problem with ensuing melt flow through XFEM/level set method

Résumé

Ce travail porte sur l'étude du phénomène de changement de phase et l'écoulement fluide qui en résulte. Cet écoulement est induit par le gradient de température et/ou par la différence de densité entre les deux phases. L'évolution de l'interface solide/liquide est capturée par la méthode de la fonction de niveaux tandis que les discontinuités à l'interface sont prises en compte via la méthode XFEM. L'équation de transfert de chaleur, combinée au problème de Stefan, et le problème de Stokes sont considérés à cet effet. Des contraintes telles que la vitesse du fluide et la température de fusion à l'interface sont imposées par la méthode de pénalité. Des simulations de solidification et de fusion montrent que la méthode XFEM prédit des résultats en bonne concordance avec ceux obtenus par éléments finis et les solutions analytiques.

Abstract

This work focuses on investigating the application of the extended finite element method (XFEM) and level set method in describing the interplay between the thermal behaviour and ensuing melt flow during the liquid-solid phase transition process of pure material. The flow that happens in the melt is supposed to be induced either by temperature gradient (buoyancy driven flow) or by density gap between the two phases. The problem at hand is characterized by the free-moving phase interface as well as discontinuities therein. The free interface is captured implicitly by the level set method, while the discontinuities are dealt with by XFEM. Two sets of partial differential equations, i.e. temperature-based energy conservation equation

supplemented by Stefan condition and Stokes equations with Boussinesq approximation, are solved by XFEM. The quantities to be determined including temperature, fluid velocity and pressure present different degrees of discontinuity at the interface. Accordingly, the corrected abs-enrichment scheme is applied to build the approximation function for temperature, while the sign-enrichment scheme for melt velocity and pressure. Constraints at the interface, such as interface temperature and melt velocity, are imposed by the penalty method. The accuracy of the numerical model is verified by three benchmark tests: flow past a circular cylinder, infinite corner solidification and tin melting front. The simulation results have a good agreement with the analytical solutions or those obtained by the conventional finite element method (FEM).

Keywords: extended finite element method, level set method, Stefan problem, buoyancy driven flow

5.1 Introduction

The phase transition problem of liquid-solid system is prevalent. It can be encountered in various scientific or engineering fields, such as metal casting, alloy formation, crystal growth, thermal engergy storage, cryopreservation, and so on. In general, the classical Stefan problem only takes the thermal conduction as the principal mechnism of heat transfer [24]. But actually during the phase change process, due to the thermal expansion effect, a buoyancy driven flow induced by the temperature gradient usually exists in the melt. In addition, if the density variation of the two phases are taken into account, there is also a forced flow originating from the interface due to mass conservation. It has been shown that the flow that happens in the melt has a strong impact on the interface morphology through influencing the heat transfer [12, 15].

Due to the prevalence of solidification/melting phenomena as well as because the impact of the melt flow is considerable, it has attracted many scholars to study the interactions in between by numerical modelling over the last several decades. Speaking of numerical modelling, one difficulty arises due to the fact that the problem at hand falls into the category of free boundary problem, which also usually comes with discontinuity issues. Interface locating and discontinuity issues are always a challenge to those conventional numerical methods developed based on the continuity hypothesis. In order to make the conventional methods still applicable, one common practice in the literature is to employ deforming mesh [20, 35, 2], in which the free interface is kept to align with element edges; then the governing equations can be solved separately in each phase by the conventional methods. By this means, the sharpness of the phase front is well preserved, but it also has some shortcomings, such as, it's difficult to deal with the free front that possesses complex shape or topological change; mesh rebuilding is sometimes cumbersome. There also exist a number of other approaches on this subject, like enthalpy method [36, 16, 30, 35, 29], phase field method [3] and effective heat capacity method

[31]. Among these approaches, the enthalpy method is most commonly used. In the enthalpy method, the energy conservation equation adopts the enthalpy formulation rather than the temperature-based formation. By the enthalpy formulation, the latent heat released/absorbed at the interface as phase change occurs can be accounted for naturally, such that the explicit tracking of the interface is circumvented. Then only one set of equations is solved over the whole domain and the interface location can be recovered a posteriori according to the temperature distribution. The advantage of this method is that it eliminates the requirement of tracking interface and imposing thermal interface condition. Nevertheless, there is still a need to account for hydrodynamic conditions at the interface. In addition, the enthalpy method is more suitable to handle the phase change problems with mushy zone rather than those with sharp interface, due to the fact that the enthalpy is of strong discontinuity at the interface in the sharp interface case.

Instead of using the approaches mentioned above, an attempt to employ an alternative method, i.e. XFEM combined with level set method, is made in the present work. By contrast with the conventional numerical methods, XFEM exists to cope with discontinuous problems. Since it was first developed in 1999 within the field of fracture mechanics [4, 10], XFEM has been extensively and successfully applied in many fields [14]. The level set method has also proven to be an accurate and robust way to capture various free boundaries (even topologically changed) based on fixed grids [25]. Combining XFEM with level set method is natural, since level set method not only is able to determine where ought to be enriched but also itself is utilized to construct the enrichment function in XFEM [34]. XFEM combined with level set method was first applied to solve the classical Stefan problem in 2002 [7, 18]. Based on the work done in [7, 18], this combined approach has been further investigated and developed on some detailed aspects but still confined within the framework of classical Stefan problem [27, 5, 9, 32, 22]. Beyond that, very few research works are found taking the influence of melt flow into account [38, 33, 23]. In [38], an externally forced flow described by volume-averaged momentum and continuity equations [39] was introduced to transport the heat, while the density difference between phases and buoyancy effect were regarded negligible. The sharp interface was smeared over a certain thickness, so that FEM was utilized to solve the hydrodynamic portion. In [33], the melt flow driven either by the density difference between the two phases or by an arbitrary flux at the interface was considered. But the velocity field is simply constructed by extending the liquid normal speed at the interface over the whole domain, so no hydrodynamic calculation was involved. In [23], the flow induced by the density difference between the two phases was modelled by solving Stokes equations. A FEM-void scheme (without any enrichment) was used in the approximation of velocity and pressure fields.

The aim of the present work is to model the solidification/melting process with ensuing melt flow through XFEM combined with level set method. To this end, temperature-based energy conservation equation with Stefan condition and Stokes equations with Boussinesq approxi-

mation are solved by XFEM alternately. To the best knowledge of the authors, no published research works have done this before. As shown in Fig. 5.1, the whole model can be divided into four sub-models. Fig. 5.1 also illustrates the dependencies and message exchange among these sub-models. Compared with the classical Stefan problem, new discontinuities at the interface associated with the velocity and pressure are added. Apparently the pressure is of strong discontinuity, but it depends for the velocity. If no-slip boundary condition is applied at the interface and density difference between the two phases is ignored, the velocity is of weak discontinuity, otherwise it is strongly discontinuous, because the velocity on the solid side is always zero. Without loss of generality, the velocity is regarded strongly discontinuous in the present work, then it requires to impose constraint on the liquid side of the interface.

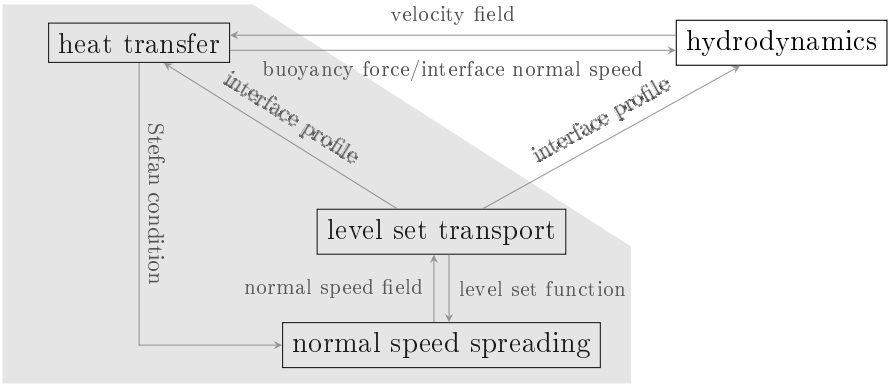


Figure 5.1 – The model's structure as well as the links among sub-models; the shaded part corresponds to the classical Stefan problem, which is two-way coupled with the hydrodynamic problem.

This paper is organized as follows. Section 5.2 introduces the governing equations, including energy conservation equations to describe the heat transfer, Stokes-Boussinesq equations for hydrodynamic flow, and equations involved in level set method to capture the interface. The numerical analysis through XFEM and FEM to solve the equations mentioned above is presented in section 5.3. Section 5.4 is devoted to validate the numerical model. Three benchmark tests are implemented to demonstrate its accuracy and robustness. At the end, section 5.5 concludes this paper.

5.2 Governing equations

5.2.1 Heat transfer

Description of the solidification/melting process in the present study is confined within a finite vertical domain $\Omega \subset \mathbb{R}^2$. As shown in Fig. 5.2, the computational domain Ω is divided by a free moving and sharp interface Γ_I into liquid region Ω_l and solid region Ω_s, such that $\Omega_l \cup \Omega_s = \Omega$ and $\Omega_l \cap \Omega_s = \Gamma_I$. The subscripts l and s throughout this paper are used to

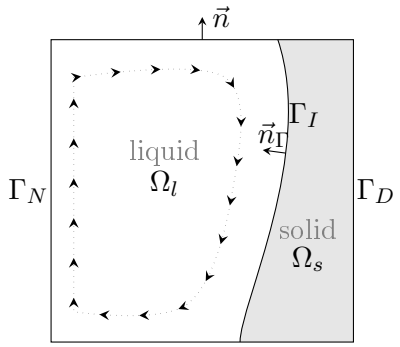


Figure 5.2 – Schematic of solidification/melting process with ensuing melt flow.

denote liquid and solid, respectively. The perimeter Γ consists of Neumann boundary Γ_N and Dirichlet boundary Γ_D, i.e. $\Gamma_N \cup \Gamma_D = \Gamma$ and $\Gamma_N \cap \Gamma_D = \emptyset$. $\vec{n}$ represents the outward unit normal of the exterior boundaries. $\vec{n}_\Gamma$ is the unit normal of the interface, pointing from solid into liquid by convention.

The energy conservation equation in terms of temperature $T(\vec{x}, t)$ formulation with both convection and conduction mechanisms being taken into account over the domain Ω, as well as initial and boundary conditions, reads

$$\rho C_p \frac{\partial T}{\partial t} + \rho C_p \vec{u} \cdot \nabla T - \nabla \cdot (k \nabla T) - s = 0; \qquad \vec{x} \in \Omega, t \in (0, \infty) \tag{5.1a}$$

$$T = T_D; \quad \vec{x} \in \Gamma_D, t \in (0, \infty) \tag{5.1b}$$

$$-k \nabla T \cdot \vec{n} = q_n; \quad \vec{x} \in \Gamma_N, t \in (0, \infty) \tag{5.1c}$$

$$T = T_0; \quad \vec{x} \in \Omega, t = 0 \tag{5.1d}$$

$$T = T_m; \quad \vec{x} \in \Gamma_I, t \in (0, \infty) \tag{5.1e}$$

where ρ is the density; C_p is the specific heat; $\vec{u}$ represents the velocity field of the heat-carrying medium, which is zero in the solid phase; k is the thermal conductivity; s is the heat source; T_D, q_n and T_0 are given value; and T_m is the melting point.

The phase change always comes with the liberation or absorption of latent heat at the interface Γ_I, where the energy conservation law can be expressed by the Stefan condition:

$$[\![q]\!] = \rho_s L V_I; \quad \vec{x} \in \Gamma_I, t \in (0, \infty) \tag{5.2}$$

where $[\![q]\!] = k_s \frac{\partial T}{\partial n}\Big|_{\Gamma_I^-} - k_l \frac{\partial T}{\partial n}\Big|_{\Gamma_I^+}$ is the heat flux jump in the normal direction of the interface, the superscripts $-$ and $+$ denote the solid side and liquid side, respectively; L represents the specific latent heat per unit mass; V_I is the normal speed of the interface, which has positive value as solidification occurs. The heat flux jump $[\![q]\!]$ is the only driving force to propogate the interface. The appropriate evaluation of $[\![q]\!]$ is crucial [19, 5], otherwise small numerical oscillations in the temperature field would be passed on, which finally leads to

wrong description of the interface profile. In the present study, the scheme proposed in [5] is employed. The basic idea of this scheme is to use the slope of the least squares regression line built by fitting a set of points distributed in the vicinity of the interface to approximate the temperature gradient. By doing so, the temperature is locally kind of smoothed, then small oscillations are supressed. For details, please see [5].

5.2.2 Hydrodynamics

The ensuing melt flow, as phase change takes place, is driven either by the buoyancy force or by the density jump at the interface. This flow is described by the incompressible Stokes equations with Boussinesq approximation in the present work. The governing equations with primitive variables velocity $\vec{u}(\vec{x},t)$ and pressure $p(\vec{x},t)$ in the pressure shifted form (the gravity term is excluded) can be written as

$$\rho\frac{\partial\vec{u}}{\partial t}-\nabla\cdot\left(2\mu\underline{\underline{D}}\right)+\nabla p+\rho\alpha_p\left(T-T_{ref}\right)\vec{g}=0; \qquad \vec{x}\in\Omega, t\in(0,\infty) \tag{5.3a}$$

$$\nabla\cdot\vec{u}=0; \qquad \vec{x}\in\Omega, t\in(0,\infty) \tag{5.3b}$$

$$\vec{u}=\vec{u}_0; \qquad \vec{x}\in\Omega, t=0 \tag{5.3c}$$

$$\vec{u}=\vec{u}_D; \qquad \vec{x}\in\Gamma_D, t\in(0,\infty) \tag{5.3d}$$

$$(-p\underline{\underline{I}}+2\mu\underline{\underline{D}})\cdot\vec{n}=\vec{t}_N; \qquad \vec{x}\in\Gamma_N, t\in(0,\infty) \tag{5.3e}$$

$$\vec{u}=0, p=0; \qquad \vec{x}\in\Omega_s, t\in(0,\infty) \tag{5.3f}$$

$$\vec{u}=\vec{u}_I; \qquad \vec{x}\in\Gamma_I^+, t\in(0,\infty) \tag{5.3g}$$

where μ is the dynamic viscosity; $\underline{\underline{D}}=\frac{1}{2}\left(\nabla\vec{u}+\nabla\vec{u}^{\mathrm{T}}\right)$ is the shear rate tensor; $\vec{u}_D$, $\vec{t}_N$ and $\vec{u}_0$ are given value; the last term on the left hand side of Eq. 5.3a accounts for the buoyancy force, in which T_{ref} denotes the reference temperature, α_p is the thermal expansion coefficient, and $\vec{g}$ is the gravitational acceleration; Eq. 5.3f states that the velocity and pressure in the solid region are set to be zero. $\vec{u}_I$ represents the melt velocity at the interface. If no-slip boundary condition is applied at the interface, $\vec{u}_I$ can be calculated by the following equation according to the mass conservation law [33]:

$$\vec{u}_I=\left(1-\frac{\rho_s}{\rho_l}\right)V_I\vec{n}_\Gamma \tag{5.4}$$

If the density jump between phases is ignored, the melt velocity at the interface is zero as indicated by Eq. 5.4.

5.2.3 Interface capture

5.2.3.1 Level set function and update

The level set function $\phi(\vec{x},t)$ is initially defined as the shortest signed distance to the interface as shown below [25]:

$$\phi=sign\big(\left(\vec{x}-\vec{x}_I\right)\cdot\vec{n}_\Gamma\big)\min_{\vec{x}_I\in\Gamma_I}\|\vec{x}-\vec{x}_I\|, \qquad \vec{x}\in\Omega \tag{5.5}$$

The entire domain Ω is endowed with distance property through Eq. 5.5. As indicated by Eq. 5.5, $\phi = 0$ means $\vec{x} \in \Gamma_I$, $\phi < 0$ means $\vec{x} \in \Omega_s$, and $\phi > 0$ means $\vec{x} \in \Omega_l$. Then, the interface propagation problem evolves into the transport problem of distance function. Once the level set function ϕ is initialized by Eq. 5.5, the update of ϕ is governed by the following transport equation [25]:

$$\frac{\partial \phi}{\partial t} + F\|\nabla \phi\| = 0 \quad \vec{x} \in \Omega, t \in (0, \infty) \tag{5.6}$$

where $F = \vec{v} \cdot \vec{n}_\Gamma = \vec{v} \cdot \frac{\nabla \phi}{\|\nabla \phi\|}$ is the normal component of velocity field $\vec{v}$ that transports the level set function.

5.2.3.2 Normal speed spreading

The normal speed field $F(\vec{x})$ in Eq. 5.6 can be constructed by spreading the interface normal speed over Ω as shown below [7]:

$$sign(\phi)\nabla \phi \cdot \nabla F = 0, \quad \vec{x} \in \Omega \tag{5.7a}$$

$$F(\vec{x}) = V_I, \quad \vec{x} \in \Gamma_I \tag{5.7b}$$

As indicated by Eq. 5.7a, F is constructed to be orthogonal to ϕ. In such a way, the distance property of ϕ can be retained to some extent. In order to start the spreading from the interface, the boundary condition represented by Eq. 5.7b acts as an influx condition by adding the component $sign(\phi)$ into Eq. 5.7a. See [7] for more details.

5.3 Numerical analysis

5.3.1 XFEM approximation

By contrast with FEM, the approximation space in XFEM is locally enriched to conform with the solution space where a priori known discontinuities exist. If only one discontinuity is considered, any variable $u(\vec{x}, t)$ can be approximated as

$$u(\vec{x}, t) = \underbrace{\sum_{i \in \mathcal{N}} N_i(\vec{x}) u_i(t)}_{\text{strd. FEM approx.}} + \underbrace{\sum_{i \in \mathcal{N}_e} M_i(\vec{x}, t) a_i(t)}_{\text{enrichment}} \tag{5.8}$$

where N_i is the interpolation function of node i; M_i is the local enrichment function; u_i and a_i are the ordinary and additional degree of freedoms (DOFs), respectively; $\mathcal{N}$ is the set of total nodes related with the variable under consideration; $\mathcal{N}_e$ is the set of enriched nodes. In the standard XFEM, $\mathcal{N}_e$ is composed of the nodes belonging to the cut elements as shown in Fig. 5.3.

It's known that the mixed formulation (velocity-pressure) of Stokes equations would lead to the saddle point problem. To make it solvable, one common practice is to make the approximating

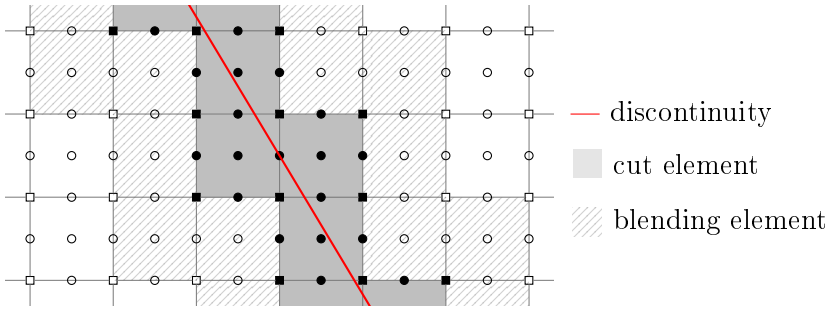


Figure 5.3 – Enrichment for the mesh comprises Q2Q1 elements; the element only with square nodes corresponds to bi-linear elements for matching pressure space (also the space of temperature, level set function, normal speed F and geometry); the element with all the nodes (square and rounded) corresponds to quadratic elements for matching velocity space; solid nodes represent those enriched in strd. XFEM.

polynomial of velocity one degree higher than that of pressure [11]. In the present study, Q2Q1 element is selected for the spacial discretization, as shown in Fig. 5.3. All the other variables, like temperature, level set function, normal speed F and geometry are approximated by the bi-linear polynomial, the same as pressure.

The local enrichment function M_i can be further written as

$$M_i(\vec{x},t) = N_i(\vec{x})\psi_i(\vec{x},t) \tag{5.9}$$

where ψ_i is the global enrichment function of node i. It determines discontinuity's characteristics of the approximation space. The scheme of ψ_i should be chosen carefully according to the discontinuity type (a priori known) of variables. In the present work, except the level set function and normal speed F which are continuous, the temperature is taken as being of weak discontinuity, while the velocity and pressure are treated as being strongly discontinuous. For the temperature, the corrected abs-enrichment scheme [13] as shown below is used:

$$\psi_i(\vec{x},t) = (|\phi(\vec{x},t) - \phi_i(t)|)\, R = (|\phi(\vec{x},t) - \phi_i(t)|) \sum_{j\in\mathcal{N}_e} N_j(\vec{x}) \tag{5.10}$$

where ϕ denotes the level set function; R is named ramping function. Besides the nodes belonging to cut elements, those nodes residing in blending elements (as shown in Fig. 5.3) also require to be enriched in this scheme. The ramping function R only takes effect in blending elements, since $\sum_{j\in\mathcal{N}_e} N_j(\vec{x})$ is always one in cut elements and zero in ordinary elements. This scheme effectively avoids problems caused in the blending element [8, 6]. For the velocity and pressure, the following sign-enrichment scheme is adopted:

$$\psi_i(\vec{x},t) = sign\,(\phi(\vec{x},t)) - sign\,(\phi_i(t)) \tag{5.11}$$

In this scheme, only nodes of cut elements are enriched.

5.3.2 Temperature-based energy conservation equation

The interface constraint represented by Eq. 5.1e is imposed by the penalty method. Following the procedure of finite element analysis, i.e. weak formulation, integration by parts, insertion of natural boundary conditions and backward Euler method for time integration, Eq. 5.1a gives

$$\begin{aligned} W_T = &\frac{1}{\Delta t}\int_\Omega \delta T \rho C_p T d\Omega - \frac{1}{\Delta t}\int_\Omega \delta T \rho C_p T^{t-\Delta t} d\Omega + \int_\Omega \delta T \rho C_p \vec{u}\cdot\nabla T d\Omega \\ &+ \int_\Omega \nabla\delta T \cdot k \nabla T d\Omega - \int_\Omega \delta T s d\Omega + \int_{\Gamma_N} \delta T q_n d\Gamma + \int_{\Gamma_I} \delta T \beta_T (T - T_m) d\Gamma = 0 \end{aligned} \tag{5.12}$$

where δT denotes the test function; β_T is the penalty parameter; Δt is the time step size; the superscript $t-\Delta t$ denotes the previous time step. Variables without superscript $t-\Delta t$ represent quantities at current time step by default.

The specific form of Eq. 5.8 for the temperature T can be written in the following matrix form based on one element Ω_e containing n^e nodes:

$$T = \langle N \rangle \{T\} \tag{5.13}$$

where

$$\langle N \rangle = \langle\ N_1 \quad N_2 \quad \cdots \quad N_{n^e} \quad M_1 \quad M_2 \quad \cdots \quad M_{n^e}\ \rangle \tag{5.14a}$$

$$\{T\} = \langle\ T_1 \quad T_2 \quad \cdots \quad T_{n^e} \quad \breve{T}_1 \quad \breve{T}_2 \quad \cdots \quad \breve{T}_{n^e}\ \rangle^{\mathrm{T}} \tag{5.14b}$$

$\breve{T}$ denotes the additional degree of freedom of temperature.

Substituting the temperature in Eq.5.12 by its approximation (Eq.5.13) and assembling, one can obtain the following system of algebraic equations:

$$\mathbf{R}_{\mathbf{M}}^{i-1} + \mathbf{R}_{\mathbf{C}}^{i-1} + \mathbf{R}_{\mathbf{K}}^{i-1} + \mathbf{R}_{\mathbf{F}}^{i-1} + \left(\frac{1}{\Delta t}\mathbf{M}_t^{i-1} + \mathbf{C}_t^{i-1} + \mathbf{K}_t^{i-1} + \mathbf{F}_t^{i-1}\right)\{\Delta T\} = 0 \tag{5.15}$$

where

$$\mathbf{R}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \{N\}(\rho C_p)^{i-1}\langle N\rangle \frac{\{T\}^{i-1} - \{T\}^{t-\Delta t}}{\Delta t} d\Omega \tag{5.16a}$$

$$\mathbf{R}_{\mathbf{C}}^{i-1} = \sum_e \int_{\Omega^e} \{N\}(\rho C_p)^{i-1}\vec{u}^{\mathrm{T}}[B]\{T\}^{i-1} d\Omega \tag{5.16b}$$

$$\mathbf{R}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} k^{i-1}[B]\{T\}^{i-1} d\Omega \tag{5.16c}$$

$$\mathbf{R}_{\mathbf{F}}^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} q_n^{i-1} d\Gamma - \sum_e \int_{\Omega^e} \{N\} s^{i-1} d\Omega + \sum_e \int_{\Gamma_I^e} \{N\}\beta_T \Big(\langle N\rangle\{T\}^{i-1} - T_m\Big) d\Gamma \tag{5.16d}$$

$$\mathbf{M}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\} \Bigg((\rho C_p)^{i-1} + \left(\rho' C_p + \rho C_p'\right)^{i-1} \langle N\rangle \Big(\{T\}^{i-1} - \{T\}^{t-\Delta t}\Big) \Bigg) \langle N\rangle d\Omega \tag{5.16e}$$

$$\mathbf{C}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\} \Big((\rho C_p)^{i-1} \vec{u}^{\mathrm{T}}[B] + \left(\rho' C_p + \rho C_p'\right)^{i-1} \vec{u}^{\mathrm{T}}[B]\{T\}^{i-1}\langle N\rangle \Big) d\Omega \tag{5.16f}$$

$$\mathbf{K}_t^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \Big(k^{i-1}[B] + \left(k'\right)^{i-1} [B]\{T\}^{i-1}\langle N\rangle \Big) d\Omega \tag{5.16g}$$

$$\mathbf{F}_t^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} \left(q_n'\right)^{i-1} \langle N\rangle d\Gamma - \sum_e \int_{\Omega^e} \{N\} \left(s'\right)^{i-1} \langle N\rangle d\Omega + \sum_e \int_{\Gamma_I^e} \{N\}\beta_T \langle N\rangle d\Gamma \tag{5.16h}$$

$[B] = \nabla\langle N\rangle$; variables with the superscript $i-1$ carry the value at previous iteration step; $\Delta T = T^i - T^{i-1}$ means temperature increment between adjacent iterations; the primed variables stand for the derivative with respect to the temperature. Without loss of generality, the physical properties, such as C_p, k, and ρ, are assumed to be temperature dependent. Eq. 5.15 is produced by applying the Newton-Raphson method for the linearization.

5.3.3 Stokes-Boussinesq equations

Imposing the interface constraint (Eq. 5.3g) by the penalty method, one obtains the following weak forms of Eq. 5.3a and Eq. 5.3b:

$$\begin{cases} W_{NS,\vec{u}} = \displaystyle\int_\Omega \delta\vec{u}\cdot\rho\frac{\partial\vec{u}}{\partial t}d\Omega + \int_\Omega \underline{\underline{D}}(\delta\vec{u}) : 2\mu\underline{\underline{D}}(\vec{u})d\Omega - \int_\Omega \nabla\cdot\delta\vec{u}\, p d\Omega \\ \qquad\qquad + \displaystyle\int_\Omega \delta\vec{u}\cdot\rho\alpha_p\left(T - T_{ref}\right)\vec{g}d\Omega - \int_{\Gamma_N} \delta\vec{u}\cdot\vec{t}_N d\Gamma + \int_{\Gamma_I^+} \delta\vec{u}\cdot\beta_u\left(\vec{u} - \vec{u}_I\right)d\Gamma = 0 \quad (5.17a) \\ W_{NS,p} = -\displaystyle\int_\Omega \delta p\nabla\cdot\vec{u}d\Omega = 0 \quad (5.17b) \end{cases}$$

where $\delta\vec{u}$ and δp are the test functions and β_u is the penalty parameter.

The specific forms of Eq. 5.8 for the velocity and pressure, respectively, can be written in the following matrix form:

$$\vec{u} = \left[\bar{N}\right]\{u\} \tag{5.18}$$

$$p = \langle N\rangle\{p\} \tag{5.19}$$

where

$$\left[\bar{N}\right] = \begin{bmatrix} \bar{N}_1 & 0 & \bar{N}_2 & 0 & \cdots & \bar{N}_{n^e} & 0 & \bar{M}_1 & 0 & \bar{M}_2 & 0 & \cdots & \bar{M}_{n^e} & 0 \\ 0 & \bar{N}_1 & 0 & \bar{N}_2 & \cdots & 0 & \bar{N}_{n^e} & 0 & \bar{M}_1 & 0 & \bar{M}_2 & \cdots & 0 & \bar{M}_{n^e} \end{bmatrix} \tag{5.20a}$$

$$\{u\} = \langle u_1 \quad v_1 \quad u_2 \quad v_2 \quad \cdots \quad u_{n^e} \quad v_{n^e} \quad \breve{u}_1 \quad \breve{v}_1 \quad \breve{u}_2 \quad \breve{v}_2 \quad \cdots \quad \breve{u}_{n^e} \quad \breve{v}_{n^e} \rangle^{\mathrm{T}} \tag{5.20b}$$

$$\{p\} = \langle \ p_1 \quad p_2 \quad \cdots \quad p_{n^e} \quad \breve{p}_1 \quad \breve{p}_2 \quad \cdots \quad \breve{p}_{n^e} \ \rangle^{\mathrm{T}} \tag{5.20c}$$

variables decorated with $\breve{}$ represent the additional degree of freedoms.

After introducing Eqs. 5.18 and 5.19 as well as applying the backward Euler method for time integration, Eq. 5.17 becomes

$$\begin{Bmatrix} \breve{\mathbf{R}}_{\mathbf{M}}^{i-1} + \breve{\mathbf{R}}_{\mathbf{K}}^{i-1} + \breve{\mathbf{R}}_{\mathbf{P}}^{i-1} + \breve{\mathbf{R}}_{\mathbf{F}}^{i-1} \\ \breve{\mathbf{R}}_{\mathbf{PV}}^{i-1} \end{Bmatrix} + \begin{bmatrix} \breve{\mathbf{M}} + \breve{\mathbf{K}} + \breve{\mathbf{F}} & \breve{\mathbf{P}} \\ \breve{\mathbf{P}}^{\mathrm{T}} & 0 \end{bmatrix} \begin{Bmatrix} \{\Delta u\} \\ \{\Delta p\} \end{Bmatrix} = 0 \tag{5.21}$$

where

$$\breve{\mathbf{R}}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \rho \left[\bar{N}\right] \frac{\{u\}^{i-1} - \{u\}^{t-\Delta t}}{\Delta t} d\Omega \tag{5.22a}$$

$$\breve{\mathbf{R}}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} \left[\breve{B}\right]^{\mathrm{T}} 2\mu \left[\breve{B}\right] \{u\}^{i-1} d\Omega \tag{5.22b}$$

$$\breve{\mathbf{R}}_{\mathbf{P}}^{i-1} = -\sum_e \int_{\Omega^e} \{B\}\langle N\rangle \{p\}^{i-1} d\Omega \tag{5.22c}$$

$$\breve{\mathbf{R}}_{\mathbf{PV}}^{i-1} = -\sum_e \int_{\Omega^e} \{N\}\langle B\rangle \{u\}^{i-1} d\Omega \tag{5.22d}$$

$$\begin{aligned} \breve{\mathbf{R}}_{\mathbf{F}}^{i-1} = & \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \rho\alpha_p \Big(T - T_{ref}\Big)\{g\} d\Omega + \sum_e \int_{\Gamma_I^{+,e}} \left[\bar{N}\right]^{\mathrm{T}} \beta_u \Big(\left[\bar{N}\right] \{u\}^{i-1} - \vec{u}_I \Big) d\Gamma \\ & - \sum_e \int_{\Gamma_N^e} \left[\bar{N}\right]^{\mathrm{T}} \{t_N\} d\Gamma d\Omega \end{aligned} \tag{5.22e}$$

$$\breve{\mathbf{M}} = \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \frac{1}{\Delta t} \rho \left[\bar{N}\right] d\Omega \tag{5.22f}$$

$$\breve{\mathbf{K}} = \sum_e \int_{\Omega^e} \left[\breve{B}\right]^{\mathrm{T}} 2\mu \left[\breve{B}\right] d\Omega \tag{5.22g}$$

$$\breve{\mathbf{P}} = -\sum_e \int_{\Omega^e} \{B\}\langle N\rangle d\Omega \tag{5.22h}$$

$$\breve{\mathbf{F}} = \sum_e \int_{\Gamma_I^{+,e}} \left[\bar{N}\right]^{\mathrm{T}} \beta_u \left[\bar{N}\right] d\Gamma \tag{5.22i}$$

$\left[\breve{B}\right]$ is called strain rate operator, defined as

$$\begin{bmatrix} \bar{N}_{1,x} & 0 & \cdots & \bar{N}_{n^e,x} & 0 & \bar{M}_{1,x} & 0 & \cdots & \bar{M}_{n^e,x} & 0 \\ 0 & \bar{N}_{1,y} & \cdots & 0 & \bar{N}_{n^e,y} & 0 & \bar{M}_{1,y} & \cdots & 0 & \bar{M}_{n^e,y} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \cdots & \frac{1}{2}\bar{N}_{n^e,y} & \frac{1}{2}\bar{N}_{n^e,x} & \frac{1}{2}\bar{M}_{1,y} & \frac{1}{2}\bar{M}_{1,x} & \cdots & \frac{1}{2}\bar{M}_{n^e,y} & \frac{1}{2}\bar{M}_{n^e,x} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \cdots & \frac{1}{2}\bar{N}_{n^e,y} & \frac{1}{2}\bar{N}_{n^e,x} & \frac{1}{2}\bar{M}_{1,y} & \frac{1}{2}\bar{M}_{1,x} & \cdots & \frac{1}{2}\bar{M}_{n^e,y} & \frac{1}{2}\bar{M}_{n^e,x} \end{bmatrix} \tag{5.23}$$

$\langle B\rangle$ can be named divergence operator, defined as

$$\langle \bar{N}_{1,x} \quad \bar{N}_{1,y} \quad \cdots \quad \bar{N}_{n^e,x} \quad \bar{N}_{n^e,y} \quad \bar{M}_{1,x} \quad \bar{M}_{1,y} \quad \cdots \quad \bar{M}_{n^e,x} \quad \bar{M}_{n^e,y} \rangle \tag{5.24}$$

5.3.4 Level set transport equation

Eq. 5.6 can be solved by FEM, given that the level set function $\phi(\vec{x},t)$ is smooth over the domain Ω. The weak form of Eq. 5.6, with the residual-based Galerkin/least-square (GLS) stabilization term [17, 7], can be written as

$$\int_{\Omega}\delta\phi\frac{\partial\phi}{\partial t}d\Omega+\int_{\Omega}\delta\phi F\|\nabla\phi\|d\Omega+\sum_{e}\int_{\Omega^e}\Big(F\nabla\delta\phi\cdot\frac{\nabla\phi}{\|\nabla\phi\|}\Big)\tau^e\Big(\frac{\partial\phi}{\partial t}+F\|\nabla\phi\|\Big)d\Omega=0 \quad (5.25)$$

where $\delta\phi$ is the test function; τ_{ϕ}^{e} is the element stabilization parameter defined as $\tau_{\phi}^{e}=\frac{h^e}{|F^e|}$ and h^e is the element shape factor. Eq. 5.25 is discretized in time by the forward Euler method. The same mesh as for the temperature and pressure fields are used to do the spatial discretization. After assembly, the final linear algebraic equation system reads

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1}+\tilde{\mathbf{R}}_{\mathbf{C}}+\tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1}+\Big(\tilde{\mathbf{M}}+\tilde{\mathbf{M}}_{\mathbf{GLS}}\Big)\{\Delta\phi\}=0 \quad (5.26)$$

where

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1}=\sum_{e}\int_{\Omega^e}\{N\}\langle N\rangle\frac{\{\phi\}^{i-1}-\{\phi\}^{t-\Delta t}}{\Delta t}d\Omega \quad (5.27a)$$

$$\tilde{\mathbf{R}}_{\mathbf{C}}=\sum_{e}\int_{\Omega^e}\{N\}F\left\|[B]\{\phi\}^{t-\Delta t}\right\|d\Omega \quad (5.27b)$$

$$\tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1}=\sum_{e}\int_{\Omega^e}F[B]^{\mathrm{T}}\frac{[B]\{\phi\}^{t-\Delta t}}{\|[B]\{\phi\}^{t-\Delta t}\|}\tau_{\phi}^{e}\Big(\langle N\rangle\frac{\{\phi\}^{i-1}-\{\phi\}^{t-\Delta t}}{\Delta t}+F\left\|[B]\{\phi\}^{t-\Delta t}\right\|\Big)d\Omega \quad (5.27c)$$

$$\tilde{\mathbf{M}}=\sum_{e}\int_{\Omega^e}\frac{1}{\Delta t}\{N\}\langle N\rangle d\Omega \quad (5.27d)$$

$$\tilde{\mathbf{M}}_{\mathbf{GLS}}=\sum_{e}\int_{\Omega^e}F[B]^{\mathrm{T}}\frac{[B]\{\phi\}^{t-\Delta t}}{\|[B]\{\phi\}^{t-\Delta t}\|}\tau_{\phi}^{e}\frac{1}{\Delta t}\langle N\rangle d\Omega \quad (5.27e)$$

5.3.5 Normal speed spreading equation

Likewise, FEM with GLS stabilization technique is utilized to solve Eq. 5.7 [7]. The corresponding weak formulation can be expressed as

$$\int_{\Omega}\delta F sign(\phi)\nabla\phi\cdot\nabla F d\Omega+\sum_{e}\int_{\Omega^e}\big(\nabla\delta F\cdot\nabla\phi\big)\tau_{F}^{e}\big(\nabla\phi\cdot\nabla F\big)d\Omega=0 \quad (5.28)$$

where δF is the test function; τ_{F}^{e} is the stabilization parameter, defined as $\frac{h^e}{\|\nabla\phi\|}$. The corresponding final algebraic equation system reads

$$\Big(\hat{\mathbf{K}}+\hat{\mathbf{K}}_{\mathbf{GLS}}\Big)\{\Delta F\}=0 \quad (5.29)$$

where

$$\hat{\mathbf{K}}=\sum_{e}\int_{\Omega^e}\{N\}sign(\phi)\big(\nabla\phi\big)^{\mathrm{T}}[B]d\Omega \quad (5.30a)$$

$$\hat{\mathbf{K}}_{\mathbf{GLS}}=\sum_{e}\int_{\Omega^e}[B]^{\mathrm{T}}\nabla\phi\tau_{F}^{e}\big(\nabla\phi\big)^{\mathrm{T}}[B]d\Omega \quad (5.30b)$$

5.3.6 Computational procedure

The usage of the present model, that has been built to simulate the solidification/melting process with considering the influence of ensuing melt flow, is supposed to follow the procedure summarized below:

1. Identify the computational domain and initial location of the interface.
2. Initialize ϕ, T, $\vec{u}$ and set reference value for p.
3. Evaluate interface normal speed V_I by Eq. 5.2.
4. **Normal speed spreading sub-model:**
   a) identify the influx boundaries and the Dirichlet-type conditions [7];
   b) obtain F by assembling and solving Eq. 5.29.
5. **Level set transport sub-model:** update ϕ by assembling and solving Eq. 5.26.
6. **Heat transfer sub-model** and **Hydrodynamics sub-model** are solved alternately, until both converge. In each single sub-model, it runs as below:
   a) loop over all elements
      i. check if cut or not; if so, do splitting and regenerate integration points;
      ii. check if enriched or not, then accordingly build the interpolation function;
      iii. assemble the discrete system Eq. 5.15 or Eq. 5.21;
   b) update T by solving Eq. 5.15 while $\vec{u}$ is assumed to be known; Eq. 5.21 is solved to update $\vec{u}$ and p while T is assumed to be known.
7. Go back to step 3 until the end.

5.3.7 Ill-conditioned problem

One drawback of XFEM formulation is that it tends to generate an ill-conditioned system [14, 21], especially when the nodes' support taking up in one phase is small and far as shown in Fig. 5.4. In this case, the iterative solvers tend not to converge. To alleviate this problem, the present work employs a technique, which is analogous to that proposed in [28], to remove the enrichment with small and far support in one phase. Removing the enrichment means that the originally enriched nodes would be treated as ordinary nodes. This technique is applied to variables with strong discontinuity. The procedure to find out the enriched nodes and excluding those with small and far support in one phase is demonstrated as below: on each cut element,

1. if the ratio of the area occupied by solid A_s to the area occupied by liquid A_l greater than a user-defined constant C_{block} ($C_{block} \gg 1$), the current element's nodes where $\phi > 0$ (nodes belonging to the liquid phase) are labelled as enriched;
2. else if $\frac{A_s}{A_l} < \frac{1}{C_{block}}$, the current element's nodes where $\phi < 0$ (nodes belonging to the solid phase) are labelled as enriched;

3. else all the nodes of current element are labelled as enriched.

As shown in Fig. 5.4, through the above procedure, the filled dot surrounded by a square is left to be not enriched.

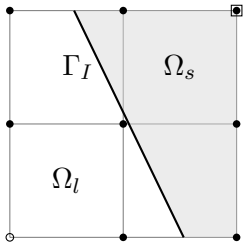


Figure 5.4 – Illustration of the case in which one node's support in one phase is small and far; the filled dots denote the enriched nodes; the filled dot surrounded by a square represents the node whose enrichment is blocked.

5.4 Numerical tests

5.4.1 Flow past a circular cylinder

The simulation of flow past a circular cylinder has always served as a benchmark problem to test the model established based on the conventional numerical methods. Let's consider a specific case that the flow occurs in a rectangular channel of $1.5m \times 0.5m$ in size, the cylinder is of radius $0.09m$ and its center locates at $(0.4m, 0.25m)$ as the left lower corner of the channel is taken as origin. The flow is driven by the pressure drop $\Delta p = -1Pa$ along the channel. No-slip boundary conditions are applied on the channel walls and the cylinder wall. Fig. 5.5 presents a typical triangular mesh in the context of FEM in such a way that cylinder wall acts as an external boundary. This problem can also be treated as fluid-structure interaction problem by bringing the cylinder into the computational domain. Then, it introduces discontinuities in the velocity and pressure fields at the cylinder wall. With the help of XFEM, an uniform and structured mesh without the internal hole, as shown in Fig. 5.6, will be feasible. To verify the XFEM hydrodynamic model of the present work, the same problem in steady case is modelled by FEM based on the mesh as shown in Fig. 5.5 and XFEM based on the mesh as shown in Fig. 5.6, respectively. In the XFEM modelling, the penalty parameter is set as $\beta_u = 10^8$ to impose the no-slip condition at the cylinder wall. The simulation results obtained by both methods over the whole domain and at the section $x = 0.4m$ are shown in figures 5.7 and 5.8, respectively. It can be seen that the XFEM results are in good agreement with those by FEM.

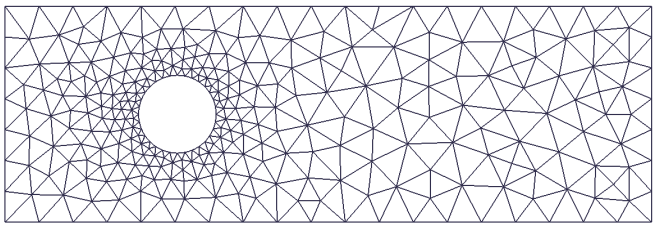


Figure 5.5 – A typical triangular mesh (containing 550 P2P1 elements) in the context of FEM.

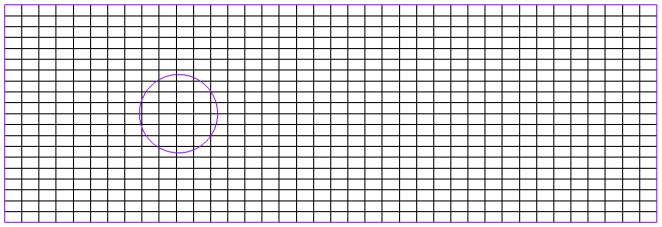


Figure 5.6 – The mesh (containing 760 Q2Q1 elements) used in XFEM as well as the discontinuity location (cylinder wall).

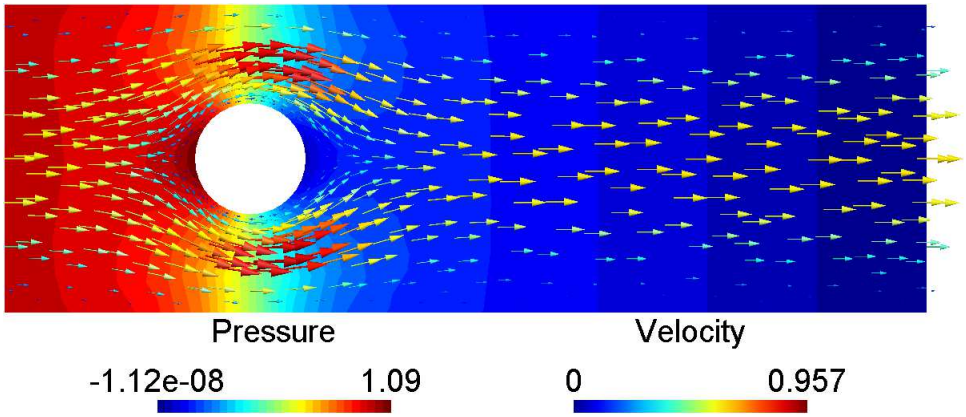
5.4.2 Infinite corner solidification

The solidification problem in an infinite corner has often been used to validate the classical two-dimensional Stefan model [7, 5]. As shown in Fig. 5.9, initially the quarter space $(x, y > 0)$ is filled with liquid of temperature T_{i}, which is greater than the melting point of the liquid T_m; the left and bottom walls are subject to a constant temperature T_w less than T_m, while the right and top boundaries are assumed to be insulated. For the case when both phases have the same thermal diffusivity α ($\alpha = \frac{k}{\rho C_p}$), the analytical solution of the interface location exists [26], which can be written as

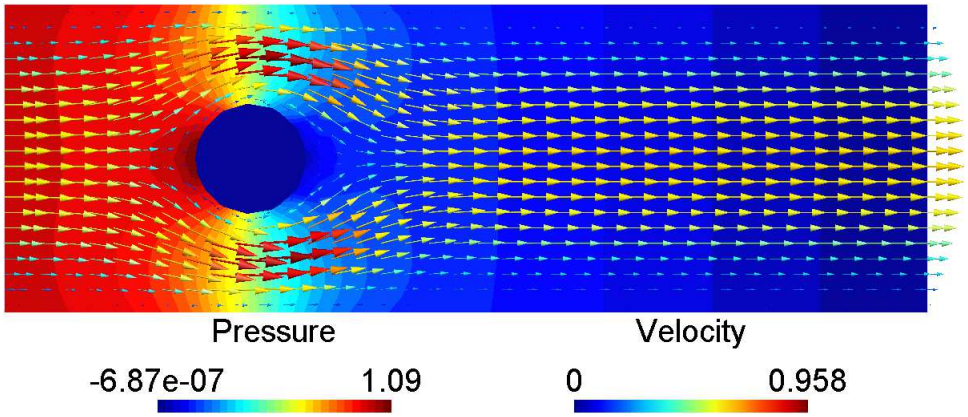
$$y^* = \left(a^b + \frac{c}{(x^*)^b - a^b} \right)^{1/b} \tag{5.31}$$

where $x^* = \frac{x}{\sqrt{4\alpha t}}$, $y^* = \frac{y}{\sqrt{4\alpha t}}$ denote the dimensionless coordinates; the constants $a = 0.70766$, $b = 5.02$ and $c = 0.159$ remain the same with those in [5].

The specific values shown in Fig. 5.9 refer to that in [7, 5], and also the material properties, including $\rho_l = \rho_s = 1kg/m^3$, $C_{p,l} = C_{p,s} = 1J/(kg \cdot K)$, $k_l = k_s = 1W/(m \cdot K)$, $L = 0.25J/kg$. In this application, a mesh containing 3136 bi-linear quadrangle elements is used and $\beta_T = 10^9$. The numerical results of the interface location at three time steps in comparison with the analytical solution Eq. 5.31 are presented in Fig. 5.10. The numerical results coincide well with the analytical solution.



(a) by FEM



(b) By XFEM

Figure 5.7 – Velocity and pressure distributions obtained by FEM and XFEM, respectively, over the whole domain.

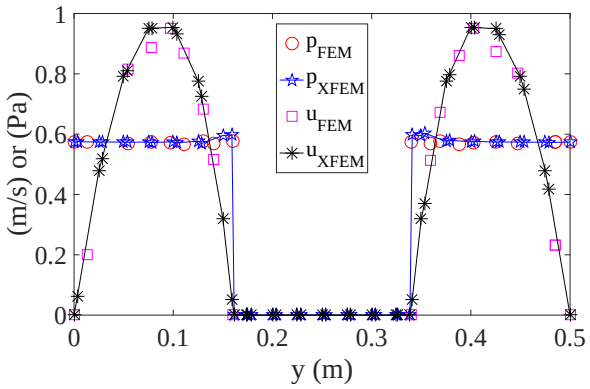


Figure 5.8 – Velocity's x-component u and pressure p distributions obtained by FEM and XFEM, respectively, at section $x = 0.4m$.

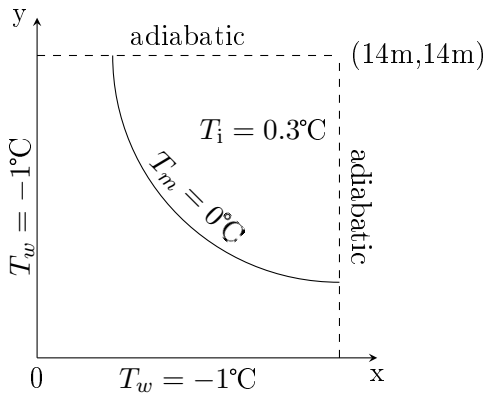


Figure 5.9 – Solidification problem in an infinite corner.

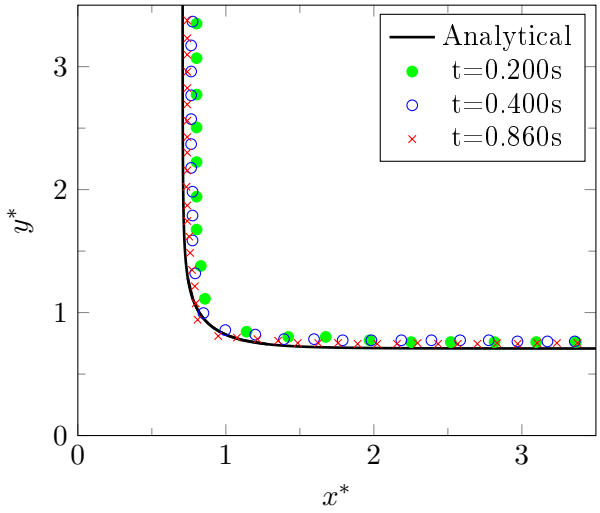


Figure 5.10 – Comparison between analytical and numerical solutions of the interface location for the corner solidification problem in the dimensionless coordinates at three different time steps.

5.4.3 Tin melting front

The tin melting front problem is a classical example in Comsol application gallery[1] based on the benchmark study in [37]. It describes the interaction between the heat transfer and the induced buoyancy flow as the metal tin melts. As shown in Fig. 5.11, the square cavity of $0.1m \times 0.1m$ in size is initially occupied by liquid tin as $x < 0.06m$ and solid tin as $x > 0.06m$. That means the initial liquid-solid interface is vertical and located at $x = 0.06m$. The temperature at the interface is always kept as the melting point $T = 505$℃. The left and right side-walls are subject to a constant higher temperature ($T_{\text{hot}} = 508$℃) and a constant lower temperature ($T_{\text{cold}} = 503$℃), respectively. The top and bottom are set to be insulated. For the sake of gentle start, a common practice is to take the steady solution, obtained by assuming the

[1] https://www.Comsol.com/model/tin-melting-front-6234

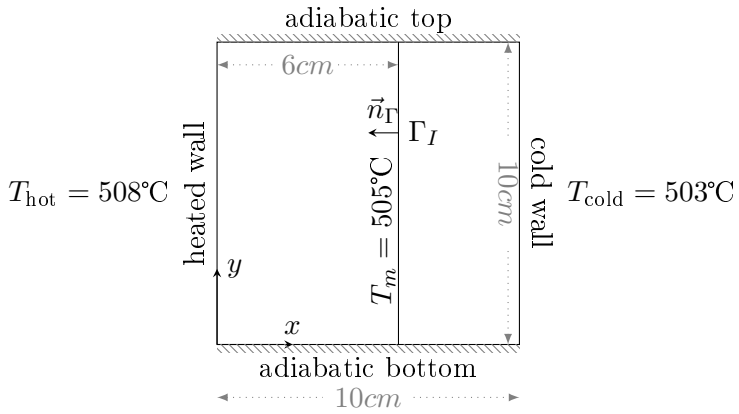


Figure 5.11 – Computational domain of the tin melting front problem and the thermal boundary conditions.

melting front fixed, as the initial temperature distribution for this transient process. Table

Table 5.1 – Thermophysical properties of tin

PARAMETER	VALUE
Density, ρ_s	$7500 kg/m^3$
Density, ρ_l	$7030 kg/m^3$ [1]
Specific heat capacity, C_p	$200 J/(kg \cdot K)$
Thermal conductivity, k	$60 W/(m \cdot K)$
Melting point, T_m	505℃
Latent heat of fusion, L	$60000 J/kg$
Dynamic viscosity, μ	$6 \times 10^{-3} kg/(m \cdot s)$
Reference temperature, T_{ref}	505℃
Thermal expansion coefficient, α_p	$2.67 \times 10^{-4} K^{-1}$

5.1 shows the thermophysical properties of tin used in the classical Comsol application except that in that application both phases have the same density $\rho = 7500 kg/m^3$.

Two scenarios are designed in the present work. Scenario 1 focuses on the influence of buoyancy flow on the heat transfer, in which the same density $\rho = 7500 kg/m^3$ is used for both phases. Regarding to the hydrodynamics, no-slip boundary conditions are applied on the periphery of the cavity and the interface, as shown in Fig. 5.12. Scenario 2 is to examine the influence of the enforced flow fed by the density jump between the two phases. In this case, different densities for liquid and solid phases are employed as shown in Table 5.1. The density gap in between reaches to 6.27% with respect to the heavier phase. Note that, under the initial and boundary conditions set above (in scenario 1) for the temperature, if no buoyancy flow or other flows exist, the system is just in the thermal equilibrium state. To break the equilibrium, a higher temperature ($T_{\text{hot}} = 530℃$) is applied on the left boundary. In addition, taking the density difference between phases into account is equivalent to allowing additional liquid tin to flow into the liquid phase from the interface. Due to the fact that the liquid tin is incompressible,

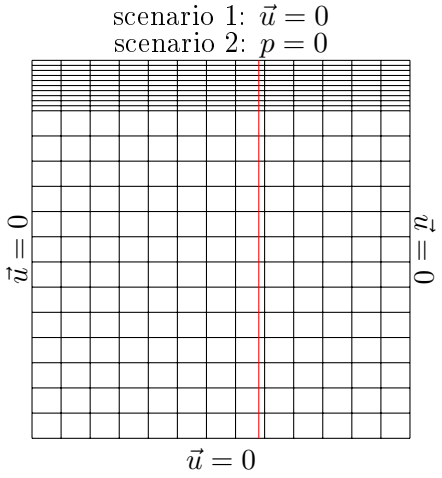


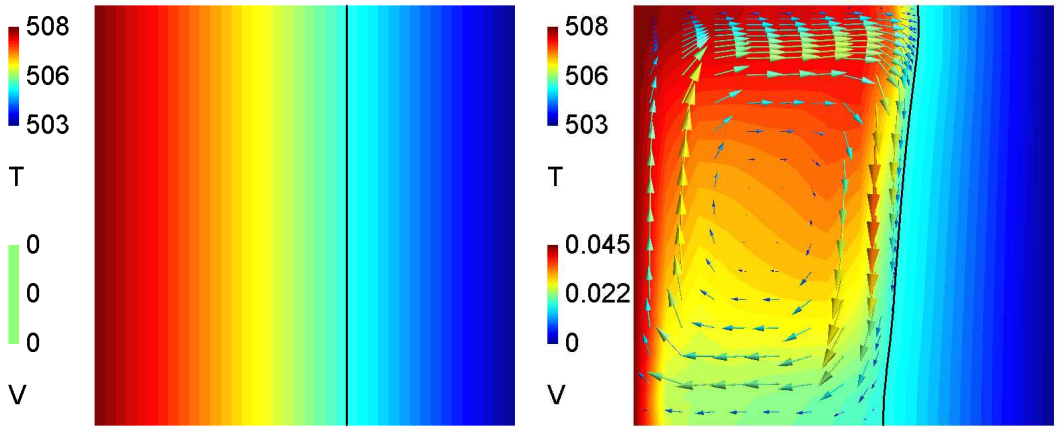
Figure 5.12 – The mesh (containing 299 quadrangle elements) used in both scenarios for tin melting front problem, the initial interface (red line) and boundary conditions for hydrodynamics.

"penetration" must be allowed somewhere on the periphery of the cavity. So in the second scenario, $p = 0$ is applied on the top of the cavity to leave the velocity there free. In both scenarios, the penalty parameter β_T is set to be 10^8 for the temperature, $\beta_u = 10^6$ for velocity and the blocking constant C_{block} is 50. Fig. 5.12 shows the mesh.

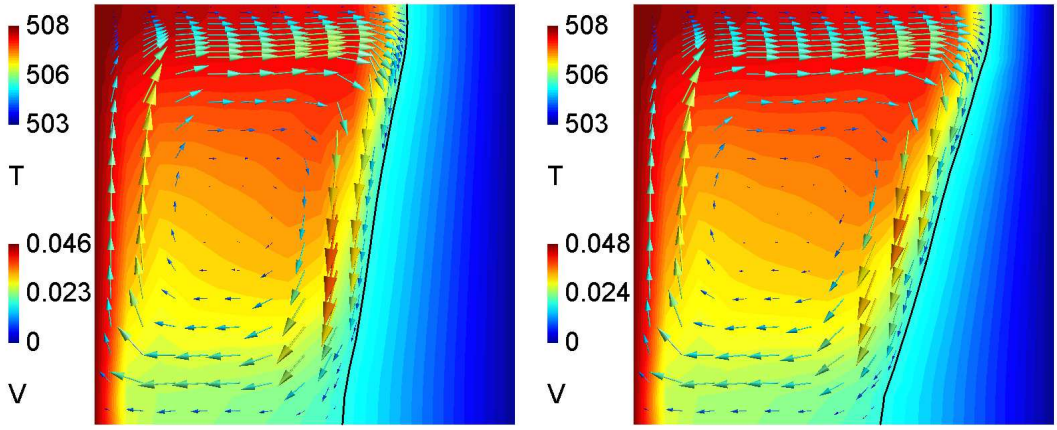
Fig. 5.13 presents the temperature distribution and melt flow pattern at several different time steps in scenario 1. It can be seen that the induced buoyancy flow has great impact on the temperature distribution. The flow enhances the rightward heat transfer along the top of the cavity. It increases the input heat flux at the upper position of the interface, which makes the original thermal balance there lost, then the upper solid tin first starts to melt. As the melt adjacent to the interface flows downwards, the heat taken from the top by the flow decreases gradually along the interface. It manifests a general tendency that the melting speed gets slower along the interface from top to bottom.

The temperature and velocity distributions at four time steps for scenario 2 is presented in Fig. 5.14. We can see that, as the melting takes place, a mass flux drawn from the melting front is triggered. The velocity of this flow at the interface is determined by the density difference between phases and the melting front speed (see Eq. 5.4). In the present case, the magnitude of the maximum flow velocity is near the order $10^{-5} m/s$. As shown in Fig. 5.14, the influence of such faint flow to the temperature distribution is almost negligible.

The comparisons between the results obtained by XFEM and Comsol in the two scenarios, including interface profiles, temperature and velocity magnitude distributions at certain time steps, are presented in figures 5.15-5.22. In Comsol, FEM combined with the mesh-deforming technique was used. It can be seen that the results obtained by XFEM have a good agreement



(a) $t = 0s$ (b) $t = 200s$



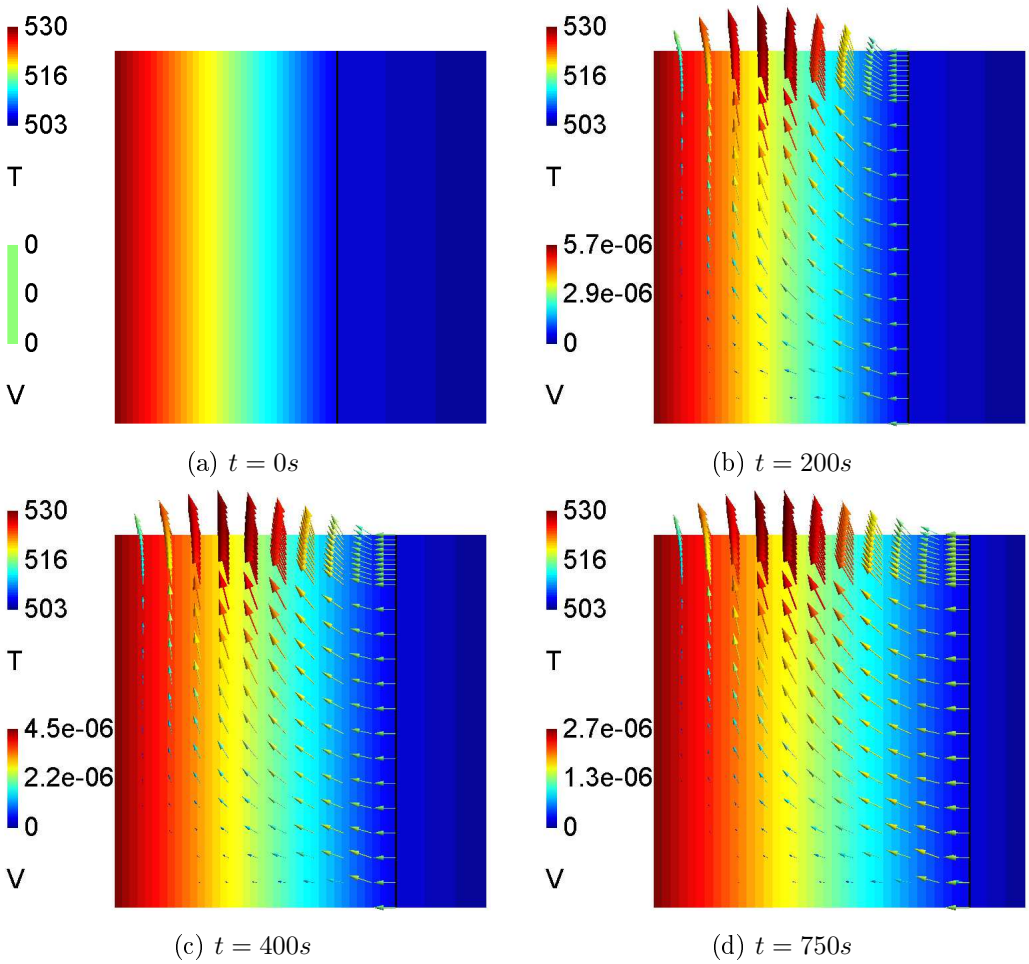
(c) $t = 385s$ (d) $t = 740s$

Figure 5.13 – Temperature and velocity distributions at four time steps for scenario 1 of tin melting front problem.

with those by Comsol.

5.5 Conclusions

A two-dimensional numerical model based on XFEM and level set method to describe the phase change process with ensuing melt flow has been presented. The model is characterized by describing the interaction between the heat transfer and the melt flow. The tin melting problem shows that the influence of the buoyancy flow induced by temperature gradient is obvious, while the melt flow caused by the density difference between solid and liquid phases (the density gap reaches to 6.27% with respect to the heavier one) is very faint. In the model, XFEM exhibits its flexibility in dealing with different types of discontinuity. Especially for the velocity field, the sign-enrichment scheme with the imposition of Dirichlet-type conditions at the interface works well for the weakly or strongly discontinuous situation. The enforcement of various physical constraints at the interface is straightforward by the penalty method. The



(a) $t = 0s$ (b) $t = 200s$

(c) $t = 400s$ (d) $t = 750s$

Figure 5.14 – Temperature and velocity distributions at four time steps for scenario 2 of tin melting front problem.

numerical tests show the accuracy of the present model.

Acknowledgements

The authors would like to acknowledge the National Sciences and Engineering Research Council of Canada (NSERC) funding NSERC Discovery Grant (No. 36518) for supporting the present research. A part of the research presented in this paper was financed by the Fonds de Recherche du Québec-Nature et Technologie (FRQNT) by the intermediary of the Aluminium Research Centre-REGAL. The first author also acknowledges the financial support from the Chinese Scholarship Council (CSC).

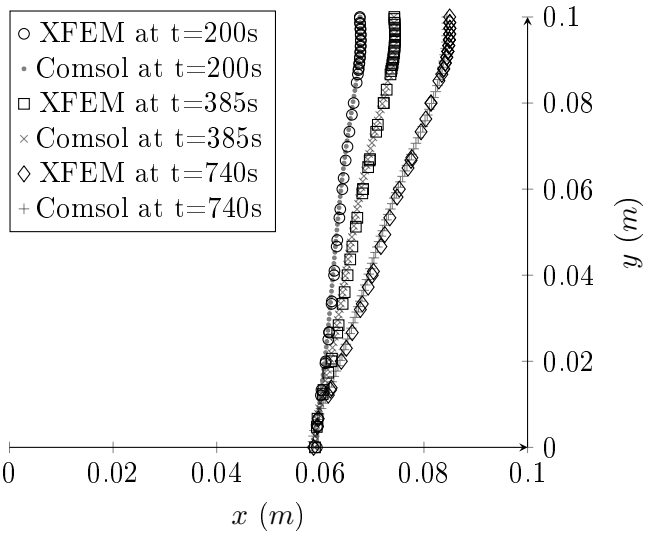
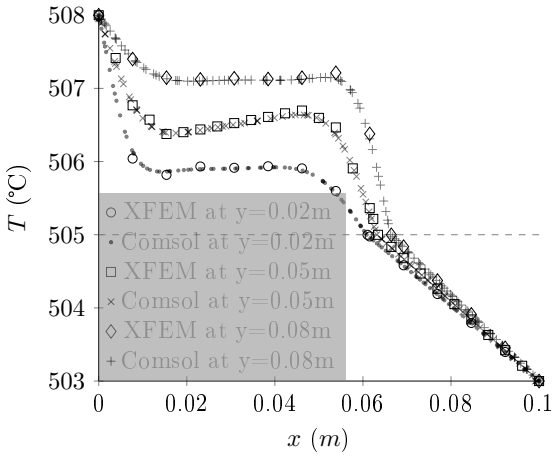
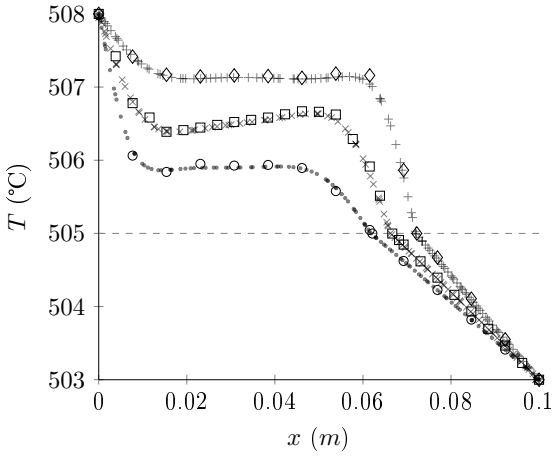


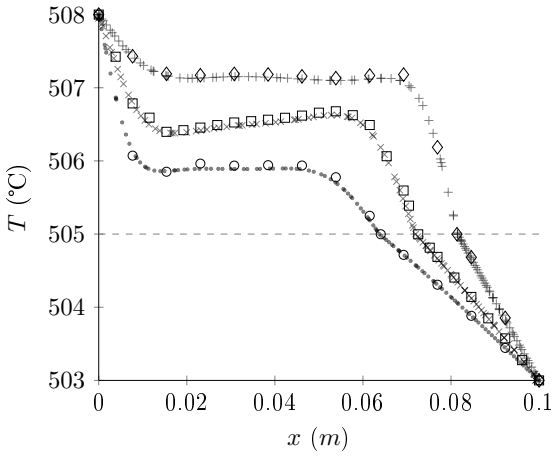
Figure 5.15 – Comparison between the interface profiles obtained by XFEM and Comsol respectively at three different time steps in scenario 1.



(a) at t = 200s

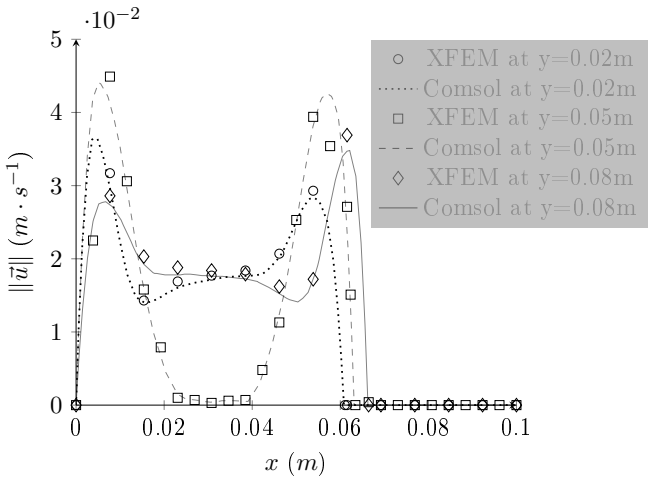


(b) at t=385s

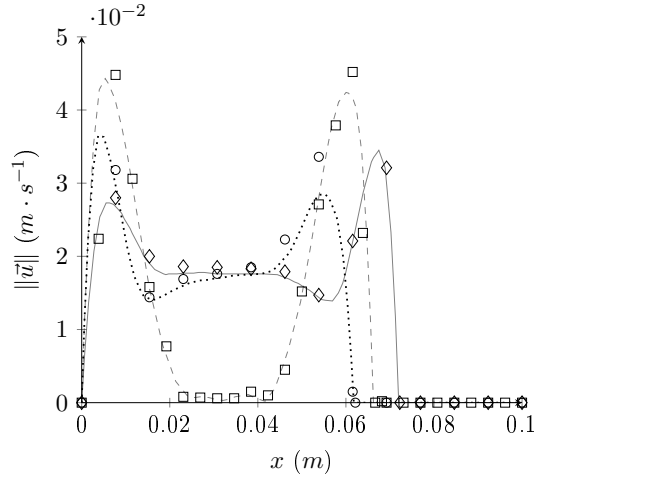


(c) at t=740s

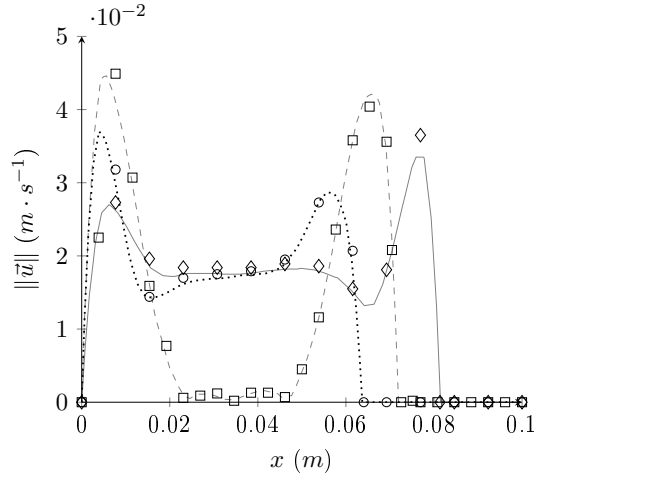
Figure 5.16 – Comparison between the temperature profiles obtained by XFEM and Comsol respectively at certain locations and time steps in scenario 1.



(a) at t=200s



(b) at t=385s



(c) at t=740s

Figure 5.17 – Comparison between the velocity magnitude obtained by XFEM and Comsol respectively at certain locations and time steps in scenario 1.

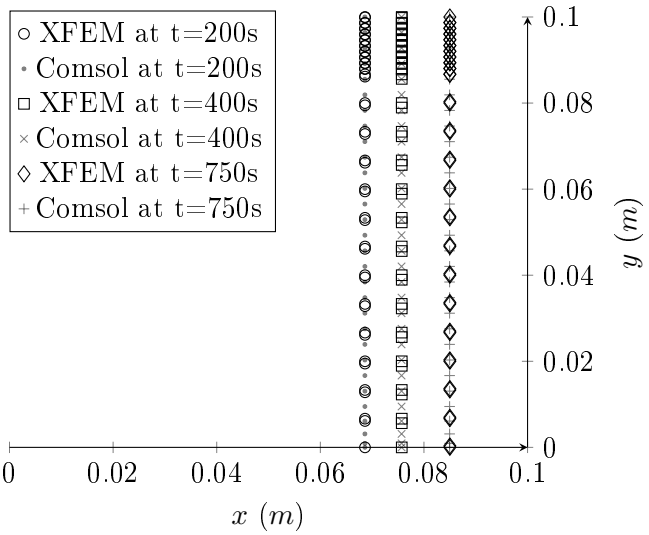
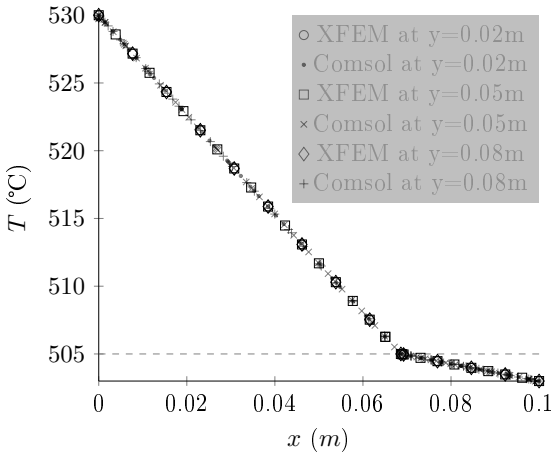
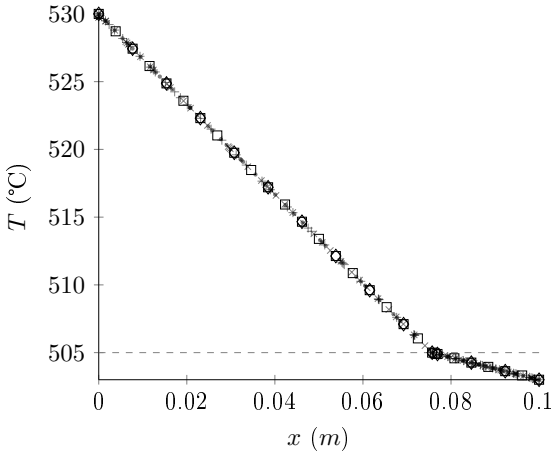


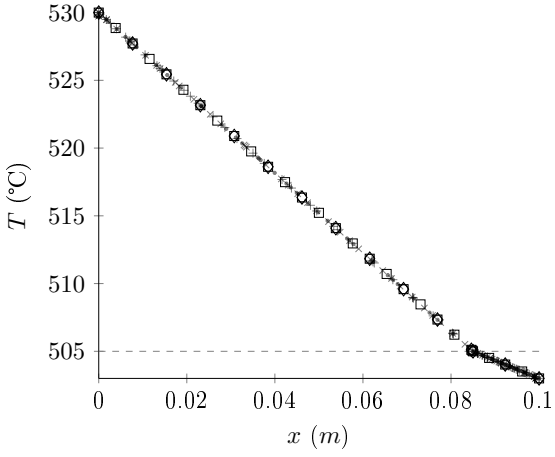
Figure 5.18 – Comparison between the interface profiles obtained by XFEM and Comsol respectively at three different time steps in scenario 2.



(a) at t = 200s

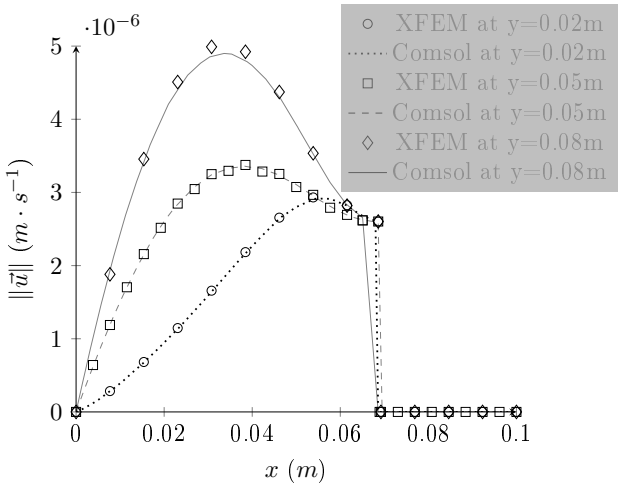


(b) at t=400s

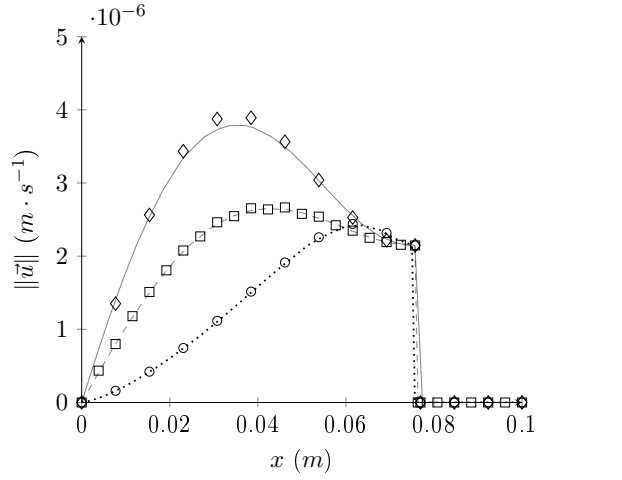


(c) at t=750s

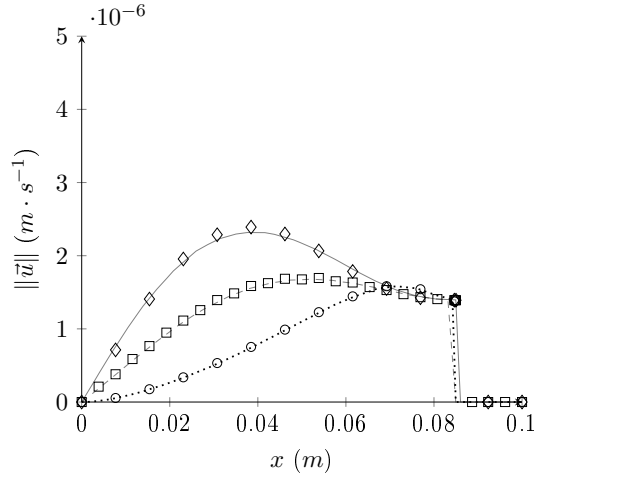
Figure 5.19 – Comparison between the temperature profiles obtained by XFEM and Comsol respectively at certain locations and time steps in scenario 2.



(a) at t=200s

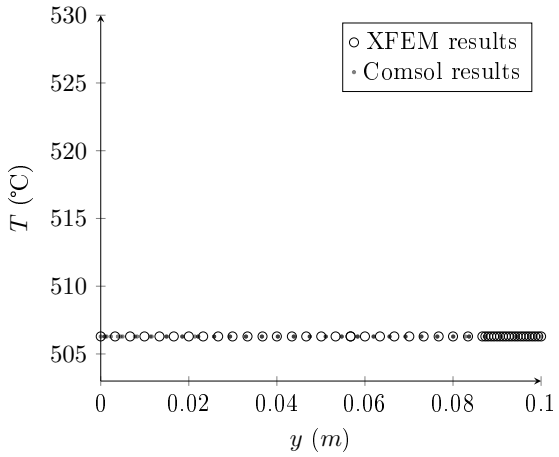


(b) at t=400s

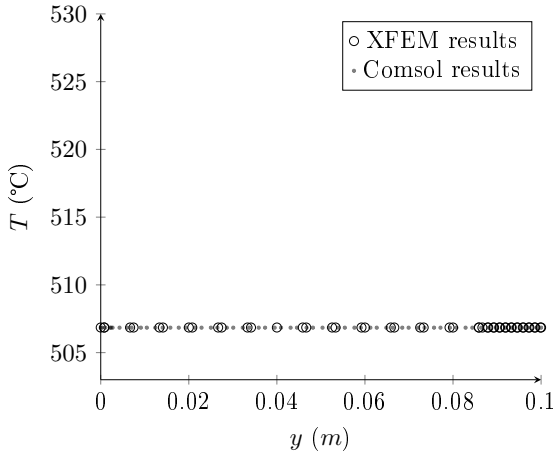


(c) at t=750s

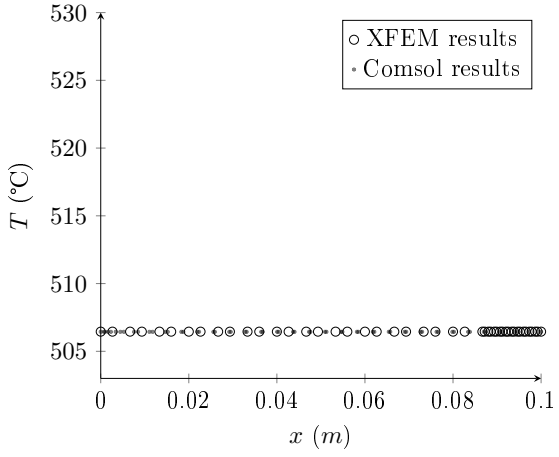
Figure 5.20 – Comparison between the velocity magnitude obtained by XFEM and Comsol respectively at certain locations and time steps in scenario 2.



(a) at t = 200s, x=0.065m

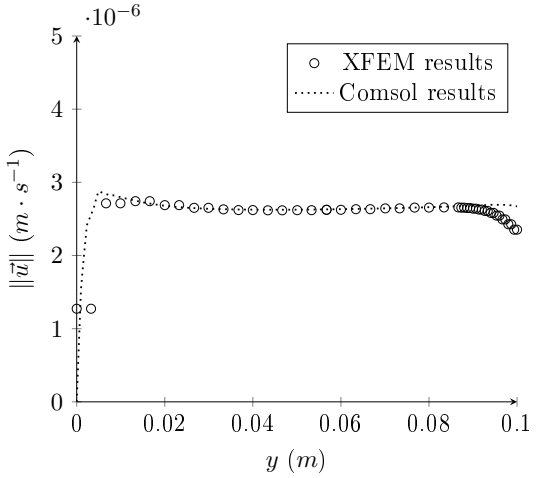


(b) at t=400s, y=0.07m

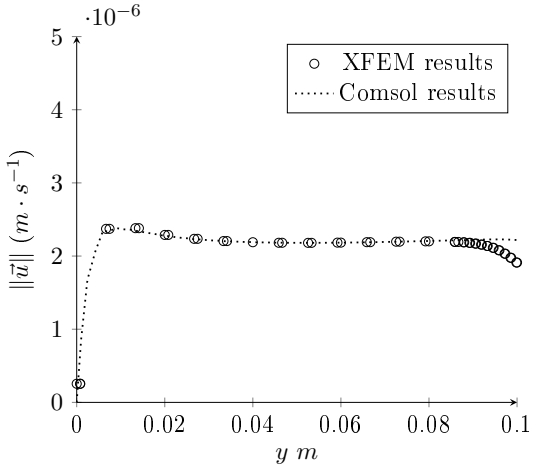


(c) at t=750s, x=0.08m

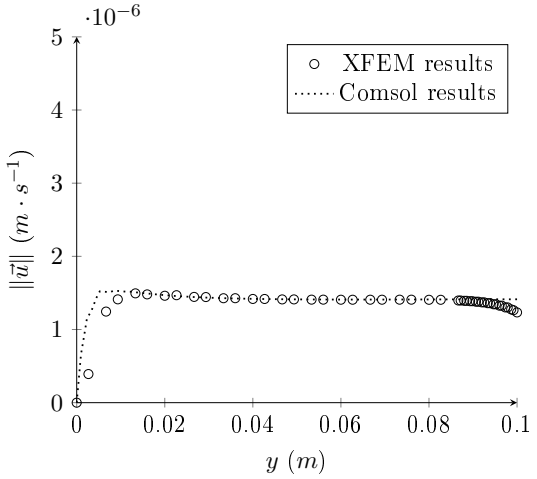
Figure 5.21 – Comparison between the temperature profiles obtained by XFEM and Comsol respectively along certain vertical sections at different time steps in scenario 2.



(a) at t=200s, x=0.065m



(b) at t=400s, x=0.07m



(c) at t=750s, x=0.08m

Figure 5.22 – Comparison between the velocity magnitude obtained by XFEM and Comsol respectively along certain vertical sections at different time steps in scenario 2; the decrease of velocity magnitude at the right end of XFEM results might be due to the no-slip condition at the interface.

Bibliography

[1] B. B. Alchagirov and A. M. Chochaeva. Temperature dependence of the density of liquid tin. *High Temperature*, 38(1):44–48, 2000.

[2] E. Bänsch, J. Paul, and A. Schmidt. An ALE finite element method for a coupled Stefan problem and Navier-Stokes equations with free capillary surface. *International Journal for Numerical Methods in Fluids*, 71(10):1282–1296, 2013.

[3] C. Beckermann, H. J. Diepers, I. Steinbach, A. Karma, and X. Tong. Modeling melt convection in phase-field simulations of solidification. *Journal of Computational Physics*, 154(2):468–496, 1999.

[4] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601–620, 1999.

[5] M. K. Bernauer and R. Herzog. Implementation of an X-FEM solver for the classical two-phase Stefan problem. *Journal of Scientific Computing*, 52(2):271–293, 2012.

[6] K. W. Cheng and T. P. Fries. Higher-order XFEM for curved strong and weak discontinuities. *International Journal for Numerical Methods in Engineering*, 82(5):564–590, 2010.

[7] J. Chessa, P. Smolinski, and T. Belytschko. The extended finite element method (XFEM) for solidification problems. *International Journal for Numerical Methods in Engineering*, 53(8):1959–1977, 2002.

[8] J. Chessa, H. Wang, and T. Belytschko. On the construction of blending elements for local partition of unity enriched finite elements. *International Journal for Numerical Methods in Engineering*, 57(7):1015–1038, 2003.

[9] A. Cosimo, V. Fachinotti, and A. Cardona. An enrichment scheme for solidification problems. *Computational Mechanics*, 52(1):17–35, 2013.

[10] J. Dolbow and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):131–150, 1999.

[11] J. Donea and A. Huerta. *Finite element methods for flow problems*. John Wiley & Sons, 2003.

[12] M. C. Flemings. Solidification processing. *Metallurgical Transactions*, 5(10):2121–2134, 1974.

[13] T. P. Fries. A corrected XFEM approximation without problems in blending elements. *International Journal for Numerical Methods in Engineering*, 75(5):503–532, 2008.

[14] T. P. Fries and T. Belytschko. The extended/generalized finite element method: an overview of the method and its applications. *International Journal for Numerical Methods in Engineering*, 84(3):253–304, 2010.

[15] M. E. Glicksman, S. R. Coriell, and G. B. McFadden. Interaction of flows with the crystal-melt interface. *Annual Review of Fluid Mechanics*, 18(1):307–335, 1986.

[16] N. Hannoun, V. Alexiades, and T. Z. Mai. Resolving the controversy over tin and gallium melting in a rectangular cavity heated from the side. *Numerical Heat Transfer: Part B: Fundamentals*, 44(3):253–276, 2003.

[17] T. J. R. Hughes, L. P. Franca, and G. M. Hulbert. A new finite element formulation for computational fluid dynamics: VIII. the Galerkin/least-squares method for advective-diffusive equations. *Computer Methods in Applied Mechanics and Engineering*, 73(2):173–189, 1989.

[18] H. Ji, D. Chopp, and J. E. Dolbow. A hybrid extended finite element/level set method for modeling phase transformations. *International Journal for Numerical Methods in Engineering*, 54(8):1209–1233, 2002.

[19] H. Ji and J. E. Dolbow. On strategies for enforcing interfacial constraints and evaluating jump conditions with the extended finite element method. *International Journal for Numerical Methods in Engineering*, 61(14):2508–2535, 2004.

[20] V. Kumar, F. Durst, and S. Ray. Modeling moving-boundary problems of solidification and melting adopting an arbitrary Lagrangian-Eulerian approach. *Numerical Heat Transfer, Part B: Fundamentals*, 49(4):299–331, 2006.

[21] C. Lang, D. Makhija, A. Doostan, and K. Maute. A simple and efficient preconditioning scheme for heaviside enriched XFEM. *Computational Mechanics*, 54(5):1357–1374, 2014.

[22] D. Martin, H. Chaouki, J. L. Robert, M. Fafard, and D. Ziegler. A XFEM Lagrange multiplier technique for Stefan problems. *Frontiers in Heat and Mass Transfer*, 7(1), 2016.

[23] D. Martin, H. Chaouki, J. L. Robert, D. Ziegler, and M. Fafard. Modelling of phase change with non-constant density using XFEM and a Lagrange multiplier. *Frontiers in Heat and Mass Transfer*, 7(1), 2016.

[24] A. M. Meirmanov. *The Stefan problem*, volume 3. Walter de Gruyter, 1992.

[25] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Springer Science & Business Media, 2006.

[26] K. A. Rathjen and L. M. Jiji. Heat conduction with melting or freezing in a corner. *Journal of Heat Transfer*, 93(1):101–109, 1971.

[27] L. Salvatori and N. Tosi. Stefan problem through extended finite elements: review and further investigations. *Algorithms*, 2(3):1177–1220, 2009.

[28] H. Sauerland and T. P. Fries. The stable XFEM for two-phase flows. *Computers & Fluids*, 87:41–49, 2013.

[29] S. S. Sebti, M. Mastiani, H. Mirzaei, A. Dadvand, S. Kashani, and S. A. Hosseini. Numerical study of the melting of nano-enhanced phase change material in a square cavity. *Journal of Zhejiang University-Science A*, 14(5):307–316, 2013.

[30] H. Shmueli, G. Ziskind, and R. Letan. Melting in a vertical cylindrical tube: numerical investigation and comparison with experiments. *International Journal of Heat and Mass Transfer*, 53(19):4082–4091, 2010.

[31] S. Singh and R. Bhargava. Numerical simulation of a phase transition problem with natural convection using hybrid FEM/EFGM technique. *International Journal of Numerical Methods for Heat & Fluid Flow*, 25(3):570–592, 2015.

[32] P. Stapór. The XFEM for nonlinear thermal and phase change problems. *International Journal of Numerical Methods for Heat & Fluid Flow*, 25(2):400–421, 2015.

[33] P. Stapór. A two-dimensional simulation of solidification processes in materials with thermo-dependent properties using XFEM. *International Journal of Numerical Methods for Heat & Fluid Flow*, 26(6):1661–1683, 2016.

[34] M. Stolarska, D. L. Chopp, N. Moës, and T. Belytschko. Modelling crack growth by level sets in the extended finite element method. *International Journal for Numerical Methods in Engineering*, 51(8):943–960, 2001.

[35] M. Ulvrová, S. Labrosse, N. Coltice, P. Råback, and P. J. Tackley. Numerical modelling of convection interacting with a melting and solidification front: application to the thermal evolution of the basal magma ocean. *Physics of the Earth and Planetary Interiors*, 206:51–66, 2012.

[36] V. R. Voller, M. Cross, and N. C. Markatos. An enthalpy method for convection/diffusion phase change. *International Journal for Numerical Methods in Engineering*, 24(1):271–284, 1987.

[37] F. Wolff and R. Viskanta. Solidification of a pure metal at a vertical wall in the presence of liquid superheat. *International Journal of Heat and Mass Transfer*, 31(8):1735–1744, 1988.

[38] N. Zabaras, B. Ganapathysubramanian, and L. Tan. Modelling dendritic solidification with melt convection using the extended finite element method. *Journal of Computational Physics*, 218(1):200–227, 2006.

[39] N. Zabaras and D. Samanta. A stabilized volume-averaging finite element method for flow in porous media and binary alloy solidification processes. *International Journal for Numerical Methods in Engineering*, 60(6):1103–1138, 2004.

Chapter 6

Numerical simulation of Stefan problem with ensuing melt flow and mass transport in the binary system through XFEM/level set method

Résumé

Ce travail a pour objectif le développement d'un modèle bidimensionnel pour simuler le phénomène de changement de phase avec prise en considération de l'effet de la composition chimique du milieu et du mouvement du fluide. À cet effet, les problèmes de Stefan, de Stokes et de transport d'espèces chimiques (système binaire) ont été couplés. La méthode XFEM est utilisée pour prendre en considération les discontinuités des différentes variables et la méthode de la fonction de niveaux est exploitée pour capturer l'évolution de l'interface solide/liquide. Le modèle ainsi développé a été appliqué à la simulation de l'évolution du profil du bain gelé dans le procédé de Hall-Héroult. Les résultats obtenus démontrent la robustesse du modèle, plus particulièrement lorsque la température de fusion dépend la composition chimique du bain.

Abstract

This work aims to establish a two-dimensional numerical model to describe the phase change process that happens in a binary system. During this process, the ensuing melt flow as well as the induced mass transport are supposed to be taken into account. To this end, the mathematical model made up of a series of partial differential equations including energy conservation equation in temperature-formulation, species transport equation and Stokes equations is solved by the extended finite element method (XFEM). The solid-liquid interface is captured implic-

itly by the level set method. In the vicinity of the interface, four physical quantities present different degrees of discontinuity. Temperature showing weak discontinuity is approximated by the corrected abs-enrichment scheme, while the other three including solute concentration, velocity and pressure are approximated by the sign-enrichment scheme. Constraints at the interface, such as melting point, mass flux and fluid velocity, are enforced by the penalty method. The melt flow can be induced either by the density difference between solid and liquid phases or by the buoyancy forces produced by the temperature gradient and compositional gradient. The Boussinesq approximation is only applied to deal with the density variation caused by temperature gradient, while the density in each term is set to be dependent on the melt composition to account for the influence of compositional gradient. The model is applied to simulate the ledge profile of the Hall–Héroult process under certain scenarios. It's shown that the model works well on describing the phase transition process of binary system, even though the melting point is sensitive to the melt composition.

Keywords: XFEM, level set method, phase change, buoyancy-driven flow, mass transport

6.1 Introduction

The classical Stefan problem generally refers to isothermal phase transition process of pure material, in which the melting point is constant and thermal conduction is the only mechanism to transfer the heat [22]. In comparison, the phase transition of bi-constituent material (binary system) is more complicated and characterized by the presence of preferential rejection or incorporation of the species at the interface due to the different solubility in liquid and solid phases [11]. This dilution or enrichment locally in species brings about two implications, that it affects both the local density and melting point, which are often sensitive to the local chemical composition in the binary system. In addition, if the influence of thermal expansion is also taken into account, the density variation will be subject simultaneously to both compositional gradient and temperature gradient. In that situation, a double-diffusive (thermosolutal) buoyancy flow is induced and has considerable impact on the heat transfer [2]. Therefore, it requires a comprehensive model to account for the interplay between the non-isothermal phase transition and the associated double-diffusive flow. In the past several decades, lots of research efforts have been dedicated to developing such comprehensive model due to its broad significance in many fields such as alloy casting, crystallization in magmas, formation/depletion of sea ice, and so on [21].

One prevalent model is the named continuum (single region) model [3, 4, 1, 34, 16, 13]. In this model, a mushy zone where liquid and solid coexist is also present to account for the non-isothermal phenomenon that the phase transition occurs over an extended temperature range. The model consists of momentum equation, continuity equation, energy conservation equation and species transport equation. This set of equations are obtained based on volume

averaging techniques and valid over the entire domain containing the liquid, solid and mushy zones. The energy conservation equation is in enthalpy formulation. Then, the need of tracking interface is eliminated. The parameters in different zones, as well as the permeability between zones (species partitioning), are defined in terms of solid/liquid fraction. The closure of the equation system is achieved by a suppplementary relationship between the solid/liquid fraction and the temperature distribution, i.e. a modified form of "lever rule" for binary phase diagrams. The continuum model has the virtue of being simple without tracking the interface. In [6, 19], the continuum model is even used to describe the phenomenon of detachment and drifting of solid crystals in the situation of hypereutectic solidification. The disadvantage of the continuum model is that it hardly offers precise details at the phase change area and appropriate constitutive relations are difficult to impose.

In contrast to continuum model, the other type is the two-phase model [37, 31], in which the conservation equations are solved in each phase based on moving mesh, or fixed mesh with interface capturing, along with the imposition of appropriate boundary conditions at the interface. In [37], a sharp-interface finite element model utilizing moving mesh is built to describe the dendritic solidification with thermosolutal flow. The interface is tracked explicitly by marker points. In [31], level set method is employed to implicitly capture the interface of dendritic growth of binary systems. According to the level set values, an artificial diffused solid-liquid interface is built. Within that diffused layer, the momentum, energy, species conservation equations are modified using volume averaging techniques, and then solved by the traditional finite element method. Compared to the continuum model, the two-phase model is superior in description of interaction between phases, but it also suffers inconvenience of remeshing or discontinuity problems at the interface.

In the present work, an attempt is made to establish a two-dimensional two-phase model to describe the phase change process of binary systems on a fixed mesh with the aid of extended finite element method (XFEM) and level set method. In dealing with the problems with non-smooth characteristics, the XFEM effectively eliminates the necessity of local mesh refinement and discontinuity smearing [12]. Especially the combination with the level set method, which provides XFEM with the discontinuity location and facilitates the construction of enrichment, has been successfully applied in various fields [15, 24]. The first introduction of the combination of XFEM and level set method to the classical Stefan problem can trace back to 2002 [8, 17]. Since then this combined model (XFEM/level set method) has been further applied and investigated under the framework of phase transition problem [23, 5, 10, 29, 36, 30], among which the most majority are still restricted to the classical Stefan problem. To the best knowledge of the authors, no research works utilizing the combination of XFEM and level set method to model the interplay between phase transition and the ensuing thermosolutal flow have been reported.

The present model is meant to be developed to simulate the ledge profile in the Hall–Héroult

process. Besides the melting point and density are sensitive to the melt composition, the Hall–Héroult system has another important feature that the density difference between solid and liquid phases is large [20]. In [20], the melt flow caused by this density difference is described, but the variation of melt composition is not considered. The influence of the melt flow caused by this density jump at the ledge front is further investigated in this work.

The following assumptions have been used to build the model in the present study:

1. The phase change front keeps clean and planar, that means formation of columnar or dendritic structures is beyond the scope of the present study, nevertheless the melting point is treated as changing with solute concentration.
2. The melt flow is described by the incompressible viscous Stokes equation.
3. In order to make the incompressibility of the melt always hold, it's assumed that the compositional variation only changes the local density but not the material density, meanwhile on the other hand, for the density variation caused by temperature gradient, the Boussinesq approximation is applied. In the literature, Boussinesq approximation is commonly used for both mechanisms (thermal and solutal effects).
4. Mass transport and flow are not allowed in the solid region.

The paper is organized as follows. In section 6.2, all the governing equations including heat transport equations, species transport equation, Stokes equations with Boussinesq approximation, as well as the definition and update equation of level set function are briefly introduced. Section 6.3 is dedicated to present the numerical analysis that how the governing equation are dealt with through XFEM and also some aspects on numerical implementation. In section 6.4, the established model is applied to simulate the ledge solidification/melting that occurs during the Hall-Héroult process under certain scenarios. Finally, the study is concluded by section 6.5.

6.2 Mathematical models

As shown in Fig. 6.1, for the liquid-solid phase transition problem, the liquid region Ω_l and solid region Ω_s, making up the computational domain $\Omega \subset \mathbb{R}^2$, are speared by the moving interface Γ_I such that $\Omega_l \cup \Omega_s = \Omega$ and $\Omega_l \cap \Omega_s = \Gamma_I$. Throughout the paper, the subscripts l and s are used to denote quantities in the liquid and solid state, respectively. By utilizing the level set method to capture the moving interface, a scalar function denoted by ϕ is defined as the signed distance function [8]:

$$\phi(\vec{x},t) = sign\big((\vec{x}-\vec{x}_I)\cdot\vec{n}_\Gamma\big)\min_{\vec{x}_I\in\Gamma_I}\|\vec{x}-\vec{x}_I\|, \qquad \vec{x}\in\Omega \tag{6.1}$$

such that the interface Γ_I is implicitly described as the zero-level of ϕ. $\vec{n}_\Gamma$ is the unit normal at the interface. By convention, $\vec{n}_\Gamma$ is defined as pointing from solid to liquid as shown in Fig.

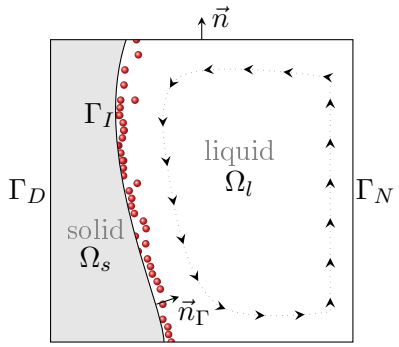


Figure 6.1 – Schematic diagram of phase change of a binary system with ensuing melt flow and species partitioning at the interface (one species represented by little spheres is rejected from the solid).

6.1. Then, the regions in different phase states are distinguished by

$$\phi(\vec{x},t)\begin{cases} <0, & \forall \vec{x}\in\Omega_s \\ =0, & \forall \vec{x}\in\Gamma_I \\ >0, & \forall \vec{x}\in\Omega_l \end{cases} \tag{6.2}$$

In general, to describe the phase transition with ensuing thermosolutal flow and species transport in a binary system, the mathematical model is composed of governing equations over the whole domain Ω, the governing equation at the interface Γ_I, constrains at the interface, as well as constitutive relations between the physical properties (density, melting point) and the melt composition. These components in detail are listed below [26, 8, 33].

1. Governing equations over the whole domain Ω include

   heat transport equation

$$\rho C_p\frac{\partial T}{\partial t}+\rho C_p\vec{u}\cdot\nabla T-\nabla\cdot(k\nabla T)-s=0 \tag{6.3}$$

   species transport equation

$$\frac{\partial \rho\omega}{\partial t}+\vec{u}\cdot\nabla(\rho\omega)+\nabla\cdot(-\rho D\nabla\omega)=0 \tag{6.4}$$

   pressure shifted Stokes equations with Boussinesq approximation

$$\rho\frac{\partial\vec{u}}{\partial t}-\nabla\cdot\left(2\mu\underline{\underline{D}}\right)+\nabla p+\rho\alpha_p\left(T-T_{ref}\right)\vec{g}-(\rho-\rho_0)\vec{g}=0 \tag{6.5a}$$

$$\nabla\cdot\vec{u}=0 \tag{6.5b}$$

   level set update equation

$$\frac{\partial\phi}{\partial t}+F\parallel\nabla\phi\parallel=0 \tag{6.6}$$

   normal speed spreading equation

$$sign(\phi)\nabla\phi\cdot\nabla F=0 \tag{6.7}$$

2. The governing equation at the interface Γ_I is

$$k_s \frac{\partial T}{\partial n}\bigg|_{\Gamma_I^-} - k_l \frac{\partial T}{\partial n}\bigg|_{\Gamma_I^+} = \rho_s L V_I \tag{6.8}$$

3. Constraints at the interface Γ_I include

$$T = T_m \tag{6.9a}$$

$$F = V_I \tag{6.9b}$$

$$\rho D \frac{\partial \omega}{\partial n}\bigg|_{\Gamma_I^-} - \rho D \frac{\partial \omega}{\partial n}\bigg|_{\Gamma_I^+} = \rho_s V_I (\omega - \omega_s) \tag{6.9c}$$

$$\vec{u}\Big|_{\Gamma_I^+} = \vec{u}_I = \left(1 - \frac{\rho_s}{\rho_l}\right) V_I \cdot \vec{n}_\Gamma \tag{6.9d}$$

4. Constitutive relations to close the model such as dependencies of the density and melting point on the solute concentration ($\rho(\omega)$ and $T_m(\omega)$).

where T denotes temperature; ω is the mass fraction of solute; $\vec{u}$ denotes fluid velocity; p denotes pressure; $F = \vec{v} \cdot \frac{\nabla \phi}{\|\nabla \phi\|}$ is the normal component of the velocity field $\vec{v}$ used to advect ϕ; ρ denotes density; ρ_0 is the density at initial composition and reference temperature; C_p is the specific heat capacity; k is the thermal conductivity; s is thermal source or sink; D is the mass diffusivity; $\underline{\underline{D}} = \frac{1}{2}\left(\nabla \vec{u} + \nabla \vec{u}^{\mathrm{T}}\right)$ represents shear rate tensor; μ is the dynamic viscosity; α_p is the thermal expansion coefficient; T_{ref} is the reference temperature; $\vec{g}$ is the gravitational acceleration; L is the latent heat; V_I is the normal speed of the interface; T_m is the melting point; superscripts $+$ and $-$ are used to denote quantities on the liquid side and the solid side of the interface, respectively.

In addition, a well-posed model also requires to be supplemented with the following initial and boundary conditions:

$$T = T_0, \quad \omega = \omega_0, \quad \vec{u} = \vec{u}_0; \qquad \vec{x} \in \Omega, t = 0 \tag{6.10a}$$

$$T = T_D, \quad \omega = \omega_D, \quad \vec{u} = \vec{u}_D; \qquad \vec{x} \in \Gamma_D, t \in (0, \infty) \tag{6.10b}$$

$$-k\nabla T \cdot \vec{n} = q_{T,n}, \quad -\rho D \nabla \omega \cdot \vec{n} = q_{\omega,n}, \quad (-p\underline{\underline{I}} + 2\mu \underline{\underline{D}}) \cdot \vec{n} = \vec{t}_N; \quad \vec{x} \in \Gamma_N, t \in (0, \infty) \tag{6.10c}$$

where Γ_D represents the exterior boundary of Dirichlet-type, while Γ_N the Neumann type; all the quantities on the right hand side of Eq. 6.10 are given value.

Remarks (1) For species transport, the code is developed based on the mass-fraction-based conservative formulation, i.e. Eq. 6.4. Eq. 6.4 exists to deal with the case in which density varies with the liquid composition. If the density variation is negligible for species transport, the same code also can be used by setting density (only in Eq. 6.4) to be constant, and

the unknown can be molar concentration. **(2)** The density ρ as well as the melting point T_m is a function of solute concentration, but independent on temperature T. The buoyancy force caused by temperature gradient is accounted for by the term $\rho\alpha_p\left(T-T_{ref}\right)\vec{g}$ in Eq. 6.5a. **(3)** Eq. 6.8 is often referred to as Stefan condition. The left-hand side of Eq. 6.8 representing the heat flux jump at the interface is the only driving force to move the phase front. In evaluating the heat flux jump to avoid small oscillations in the temperature field, the temperature gradient is approximated by the slope of the least squares regression line, which is built by fitting a set of points distributed in the vicinity of the interface. A parameter δd requires to be set to indicate the range within which the evaluation points are selected, please see [5] for details. Once V_I is obtained, it can be used as internal boundary condition (see Eq. 6.9b) to construct the normal speed field F (see Eq. 6.7) to advect the level set function ϕ. As shown in Eq. 6.7, F is constructed in the way to be orthogonal to ϕ, such that the property of ϕ being signed distance function can be conserved to a degree. **(4)** The species partition is governed by Eq. 6.9c. If species is not allowed to move in the solid region, i.e. $D_s = 0$, Eq. 6.9c becomes a boundary condition of Neumann type requiring to be imposed on the liquid side of the interface. **(5)** In the case that the influence of density difference between phases is taken into account, Eq. 6.9d should be imposed on the liquid side of the interface as a boundary condition of Dirichlet-type.

6.3 Numerical analysis

6.3.1 Approximation schemes

The mathematical model established above can be roughly divided into five sub-problems, governed by the equations 6.3 to 6.7, respectively. In these equations, there are mainly six unknown quantities involved, including temperature T, solute concentration ω, velocity $\vec{u}$, pressure p, level set function ϕ and normal speed F. All these unknowns manifest different degrees of non-smoothness at the interface except for ϕ and F. It means all the unknowns should be approximated by XFEM except that ϕ and F can be approximated by FEM. In XFEM, the approximation of any variable $u(\vec{x},t)$ with one enrichment can be written as [15]

$$u(\vec{x},t) = \underbrace{\sum_{i\in\mathcal{N}} N_i(\vec{x})u_i(t)}_{\text{strd. FEM approx.}} + \underbrace{\sum_{i\in\mathcal{N}_e} M_i(\vec{x},t)a_i(t)}_{\text{enrichment}} \tag{6.11}$$

where quantities with subscript i represent the values adhering to node i; N_i is the standard FEM shape function; M_i is the local enrichment function; u_i and a_i are the ordinary and additional degree of freedoms (DOFs), respectively; $\mathcal{N}$ is the set of total nodes related with the variable under consideration; $\mathcal{N}_e$ is the set of enriched nodes. In the standard XFEM, the enriched nodes are those belonging to the elements cut through by the discontinuity.

The Stokes equation in the mixed (velocity-pressure) formulation requires that the approximation spaces of velocity and pressure must be well matched to satisfy the inf-sup (LBB) condition. One commonly used qualified element type is Q2Q1. In the present work, the identical mesh composed of Q2Q1 elements is shared by all the sub-problems. Except that velocity is approximated by the bi-quadratic polynomial, all the other variables are approximated by the bi-linear polynomial. For a given element Ω^e, the approximations can be written in matrix form as

$$\vec{u} = [\bar{N}]\{u\} \tag{6.12a}$$
$$p = \langle N \rangle \{p\} \tag{6.12b}$$
$$T = \langle N \rangle \{T\} \tag{6.12c}$$
$$\omega = \langle N \rangle \{\omega\} \tag{6.12d}$$
$$\phi = \langle N \rangle \{\phi\} \tag{6.12e}$$
$$F = \langle N \rangle \{F\} \tag{6.12f}$$

where

$$[\bar{N}] = \begin{bmatrix} \bar{N}_1 & 0 & \bar{N}_2 & 0 & \cdots & \bar{N}_9 & 0 & \bar{M}_1 & 0 & \bar{M}_2 & 0 & \cdots & \bar{M}_9 & 0 \\ 0 & \bar{N}_1 & 0 & \bar{N}_2 & \cdots & 0 & \bar{N}_9 & 0 & \bar{M}_1 & 0 & \bar{M}_2 & \cdots & 0 & \bar{M}_9 \end{bmatrix} \tag{6.13a}$$

$$\{u\} = \langle u_1 \quad v_1 \quad u_2 \quad v_2 \quad \cdots \quad u_9 \quad v_9 \quad \breve{u}_1 \quad \breve{v}_1 \quad \breve{u}_2 \quad \breve{v}_2 \quad \cdots \quad \breve{u}_9 \quad \breve{v}_9 \rangle^{\mathrm{T}} \tag{6.13b}$$
$$\langle N \rangle = \langle \ N_1 \quad N_2 \quad \cdots \quad N_4 \quad M_1 \quad M_2 \quad \cdots \quad M_4 \ \rangle \tag{6.13c}$$
$$\{p\} = \langle \ p_1 \quad p_2 \quad \cdots \quad p_4 \quad \breve{p}_1 \quad \breve{p}_2 \quad \cdots \quad \breve{p}_4 \ \rangle^{\mathrm{T}} \tag{6.13d}$$
$$\{T\} = \langle \ T_1 \quad T_2 \quad \cdots \quad T_4 \quad \breve{T}_1 \quad \breve{T}_2 \quad \cdots \quad \breve{T}_4 \ \rangle^{\mathrm{T}} \tag{6.13e}$$
$$\{\omega\} = \langle \ \omega_1 \quad \omega_2 \quad \cdots \quad \omega_4 \quad \breve{\omega}_1 \quad \breve{\omega}_2 \quad \cdots \quad \breve{\omega}_4 \ \rangle^{\mathrm{T}} \tag{6.13f}$$
$$\{\phi\} = \langle \ \phi_1 \quad \phi_2 \quad \cdots \quad \phi_4 \ \rangle^{\mathrm{T}} \tag{6.13g}$$
$$\{F\} = \langle \ F_1 \quad F_2 \quad \cdots \quad F_4 \ \rangle^{\mathrm{T}} \tag{6.13h}$$

Remarks The size of matrices and vectors in Eq. 6.12 varies with the enrichment degree of current element, i.e. fully-enriched, partially-enriched or non-enriched. For example, the enriched components in Eq. 6.13, denoted by M or $\bar{M}$ in the shape function and decorated with $\breve{}$ in the nodal value vector, don't exist for the non-enriched elements.

The local enrichment function M_i can be further expanded as [15]

$$M_i(\vec{x}, t) = N_i(\vec{x})\psi_i(\vec{x}, t) \tag{6.14}$$

where ψ_i denotes the global enrichment function of node i. It's ψ_i that determines the discontinuity type of the approximation space. Therefore, the scheme of ψ_i should be chosen according to the discontinuity type in the solution space, which is known a priori. Since the

temperature T is weakly discontinuous, the corrected abs-enrichment scheme, proposed in [14], is used:

$$\psi_i(\vec{x},t) = \left(|\phi(\vec{x},t) - \phi_i(t)|\right) R \tag{6.15}$$

where $R = \sum_{j\in\mathcal{N}_e} N_j(\vec{x})$ acts as a ramping function. In the corrected abs-enrichment scheme, the nodes belonging to both fully-enriched elements (cut elements) and partially-enriched elements (blending elements) are enriched. By utilizing this scheme, the oscillation problems caused in the blending elements are effectively avoided [9, 7]. On the other hand, the solute concentration ω, velocity $\vec{u}$ and pressure p are strongly discontinuous. Therefore, the following sign-enrichment scheme [15] is used:

$$\psi_i(\vec{x},t) = sign\left(\phi(\vec{x},t)\right) - sign\left(\phi_i(t)\right) \tag{6.16}$$

In this scheme, only nodes of the cut elements are enriched.

6.3.2 Heat transport sub-model

The weak formulation of Eq. 6.3, together with imposing the interface constraint Eq. 6.9a by the penalty method, gives

$$\begin{aligned} W_T = &\int_\Omega \delta T \rho C_p \frac{\partial T}{\partial t} d\Omega + \int_\Omega \delta T \rho C_p \vec{u}\cdot\nabla T d\Omega + \int_\Omega \nabla\delta T \cdot k \nabla T d\Omega - \int_\Omega \delta T s d\Omega \\ &+ \int_{\Gamma_N} \delta T q_{T,n} d\Gamma + \int_{\Gamma_I} \delta T \beta_T (T - T_m) d\Gamma = 0 \end{aligned} \tag{6.17}$$

where δT denotes the test function; β_T is the penalty parameter. After a series of operations, such as time discretization by the backward Euler method, linearization by the Newton-Raphson method, spatial discretization and assembly, Eq. 6.17 becomes,

$$\mathbf{R}_\mathbf{M}^{i-1} + \mathbf{R}_\mathbf{C}^{i-1} + \mathbf{R}_\mathbf{K}^{i-1} + \mathbf{R}_\mathbf{F}^{i-1} + \left(\frac{1}{\Delta t}\mathbf{M}_t^{i-1} + \mathbf{C}_t^{i-1} + \mathbf{K}_t^{i-1} + \mathbf{F}_t^{i-1}\right)\{\Delta T\} = 0 \tag{6.18}$$

where

$$\mathbf{R}_\mathbf{M}^{i-1} = \sum_e \int_{\Omega^e} \{N\}(\rho C_p)^{i-1}\langle N\rangle \frac{\{T\}^{i-1} - \{T\}^{t-\Delta t}}{\Delta t} d\Omega \tag{6.19a}$$

$$\mathbf{R}_\mathbf{C}^{i-1} = \sum_e \int_{\Omega^e} \{N\}(\rho C_p)^{i-1}\vec{u}^{\mathrm{T}}[B]\{T\}^{i-1} d\Omega \tag{6.19b}$$

$$\mathbf{R}_\mathbf{K}^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} k^{i-1}[B]\{T\}^{i-1} d\Omega \tag{6.19c}$$

$$\mathbf{R}_\mathbf{F}^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} q_{T,n}^{i-1} d\Gamma + \sum_e \int_{\Gamma_I^e} \{N\}\beta_T\Big(\langle N\rangle\{T\}^{i-1} - T_m\Big) d\Gamma - \sum_e \int_{\Omega^e} \{N\} s^{i-1} d\Omega \tag{6.19d}$$

$$\mathbf{M}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\}\left((\rho C_p)^{i-1} + \left(\rho' C_p + \rho C_p'\right)^{i-1}\langle N\rangle\Big(\{T\}^{i-1} - \{T\}^{t-\Delta t}\Big)\right)\langle N\rangle d\Omega \tag{6.19e}$$

$$\mathbf{C}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\}\left((\rho C_p)^{i-1}\vec{u}^{\mathrm{T}}[B] + \left(\rho' C_p + \rho C_p'\right)^{i-1}\vec{u}^{\mathrm{T}}[B]\{T\}^{i-1}\langle N\rangle\right) d\Omega \tag{6.19f}$$

$$\mathbf{K}_t^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \Big(k^{i-1}[B] + (k')^{i-1}[B]\{T\}^{i-1}\langle N\rangle \Big) d\Omega \tag{6.19g}$$

$$\mathbf{F}_t^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} \left(q'_{T,n}\right)^{i-1} \langle N\rangle d\Gamma + \sum_e \int_{\Gamma_I^e} \{N\}\beta_T\langle N\rangle d\Gamma - \sum_e \int_{\Omega^e} \{N\} \left(s'\right)^{i-1} \langle N\rangle d\Omega \tag{6.19h}$$

Remarks Throughout the paper, the following notations hold. $[B]$ represents $\nabla\langle N\rangle$. Variables with superscript $i-1$ carry the value at previous iteration step. The primed quantities stand for the derivative with respect to the unknown for current sub-problem. $\Delta T = T^i - T^{i-1}$ means temperature increment between adjacent iterations, and the other unknowns do likewise. Δt is the time step size. The superscript $t-\Delta t$ denotes the previous time step. Variables without superscript $t-\Delta t$ represent quantities at current step by default.

6.3.3 Species transport sub-model

The weak formulation of Eq. 6.4 together with enforcing the interface constraint Eq. 6.9c gives

$$\begin{aligned} W_\omega = &\int_\Omega \delta\omega\Big(\rho + \omega\rho'\Big)\frac{\partial\omega}{\partial t} d\Omega + \int_\Omega \delta\omega\vec{u}\cdot\Big(\rho+\omega\rho'\Big)\nabla\omega d\Omega + \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega \\ &+ \int_{\Gamma_N} \delta\omega q_{\omega,n} d\Gamma + \int_{\Gamma_{I^+}} \delta\omega\rho_s\big(\omega-\omega_s\big)V_I d\Gamma = 0 \end{aligned} \tag{6.20}$$

where δw is the test function. After the same operations as in the section 6.3.2, Eq. 6.20 becomes

$$\bar{\mathbf{R}}_{\mathbf{M}}^{i-1} + \bar{\mathbf{R}}_{\mathbf{C}}^{i-1} + \bar{\mathbf{R}}_{\mathbf{K}}^{i-1} + \bar{\mathbf{R}}_{\mathbf{F}}^{i-1} + \left(\frac{1}{\Delta t}\bar{\mathbf{M}}_t^{i-1} + \bar{\mathbf{C}}_t^{i-1} + \bar{\mathbf{K}}_t^{i-1} + \bar{\mathbf{F}}_t^{i-1}\right)\{\Delta w\} = 0 \tag{6.21}$$

where

$$\bar{\mathbf{R}}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \{N\}\big(\rho+\omega\rho'\big)^{i-1}\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{6.22a}$$

$$\bar{\mathbf{R}}_{\mathbf{C}}^{i-1} = \sum_e \int_{\Omega^e} \{N\}\vec{u}^{\mathrm{T}}\big(\rho+\omega\rho'\big)^{i-1}[B]\{\omega\}^{i-1} d\Omega \tag{6.22b}$$

$$\bar{\mathbf{R}}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}}\big(\rho D\big)^{i-1}[B]\{\omega\}^{i-1} d\Omega \tag{6.22c}$$

$$\bar{\mathbf{R}}_{\mathbf{F}}^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} q_{\omega,n}^{i-1} d\Gamma + \sum_e \int_{\Gamma_{I^+}^e} \{N\}\rho_s^{i-1}\Big(\langle N\rangle\{\omega\}^{i-1} - \omega_s^{i-1}\Big)V_I d\Gamma \tag{6.22d}$$

$$\bar{\mathbf{M}}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\}\Big(\big(\rho+\omega\rho'\big)^{i-1} + \big(2\rho'+\omega\rho''\big)^{i-1}\langle N\rangle\Big(\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}\Big)\Big)\langle N\rangle d\Omega \tag{6.22e}$$

$$\bar{\mathbf{C}}_t^{i-1} = \sum_e \int_{\Omega^e} \{N\}\vec{u}^{\mathrm{T}}\Big(\big(\rho+\omega\rho'\big)^{i-1}[B] + \big(2\rho'+\omega\rho''\big)^{i-1}[B]\{\omega\}^{i-1}\langle N\rangle\Big) d\Omega \tag{6.22f}$$

$$\bar{\mathbf{K}}_t^{i-1} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}}\Big(\big(\rho D\big)^{i-1}[B] + \big(\rho' D + \rho D'\big)^{i-1}[B]\{\omega\}^{i-1}\langle N\rangle\Big) d\Omega \tag{6.22g}$$

$$\bar{\mathbf{F}}_t^{i-1} = \sum_e \int_{\Gamma_N^e} \{N\} \left(q'_{\omega,n}\right)^{i-1} \langle N\rangle d\Gamma + \sum_e \int_{\Gamma_{I^+}^e} \{N\}\Big(\rho'_s\big(\omega-\omega_s\big) + \rho_s - \rho_s\omega'_s\Big)^{i-1}\langle N\rangle V_I d\Gamma \tag{6.22h}$$

6.3.4 Stokes sub-model

The weak formulation of Eq. 6.5, together with enforcing the interface constraint Eq. 6.9d by the penalty method, gives

$$\begin{cases} W_{\vec{u}} = \int_\Omega \delta\vec{u}\cdot\rho\dfrac{\partial\vec{u}}{\partial t}d\Omega + \int_\Omega \underline{\underline{D}}(\delta\vec{u}) : 2\mu\underline{\underline{D}}(\vec{u})d\Omega - \int_\Omega \nabla\cdot\delta\vec{u}pd\Omega - \int_\Omega \delta\vec{u}\cdot(\rho-\rho_0)\vec{g}d\Omega \\ \quad + \int_\Omega \delta\vec{u}\cdot\rho\alpha_p\left(T-T_{ref}\right)\vec{g}d\Omega - \int_{\Gamma_N}\delta\vec{u}\cdot\vec{t}_N d\Gamma + \int_{\Gamma_I^+}\delta\vec{u}\cdot\beta_u\left(\vec{u}-\vec{u}_I\right)d\Gamma = 0 \quad (6.23a) \\ W_p = -\int_\Omega \delta p\nabla\cdot\vec{u}d\Omega = 0 \quad (6.23b) \end{cases}$$

where $\delta\vec{u}$ and δp are the test functions. After discretization (backward Euler method in time) and assembly, Eq. 6.23 becomes

$$\begin{Bmatrix} \check{\mathbf{R}}_{\mathbf{M}}^{i-1} + \check{\mathbf{R}}_{\mathbf{K}}^{i-1} + \check{\mathbf{R}}_{\mathbf{P}}^{i-1} + \check{\mathbf{R}}_{\mathbf{F}}^{i-1} \\ \check{\mathbf{R}}_{\mathbf{PV}}^{i-1} \end{Bmatrix} + \begin{bmatrix} \check{\mathbf{M}} + \check{\mathbf{K}} + \check{\mathbf{F}} & \check{\mathbf{P}} \\ \check{\mathbf{P}}^{\mathrm{T}} & 0 \end{bmatrix} \begin{Bmatrix} \{\Delta u\} \\ \{\Delta p\} \end{Bmatrix} = 0 \quad (6.24)$$

where

$$\check{\mathbf{R}}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \rho \left[\bar{N}\right] \frac{\{u\}^{i-1} - \{u\}^{t-\Delta t}}{\Delta t} d\Omega \quad (6.25a)$$

$$\check{\mathbf{R}}_{\mathbf{K}}^{i-1} = \sum_e \int_{\Omega^e} \left[\check{B}\right]^{\mathrm{T}} 2\mu \left[\check{B}\right] \{u\}^{i-1} d\Omega \quad (6.25b)$$

$$\check{\mathbf{R}}_{\mathbf{P}}^{i-1} = -\sum_e \int_{\Omega^e} \{B\}\langle\bar{N}\rangle\{p\}^{i-1} d\Omega \quad (6.25c)$$

$$\check{\mathbf{R}}_{\mathbf{PV}}^{i-1} = -\sum_e \int_{\Omega^e} \{\bar{N}\}\langle B\rangle\{u\}^{i-1} d\Omega \quad (6.25d)$$

$$\begin{aligned} \check{\mathbf{R}}_{\mathbf{F}}^{i-1} = &\sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \rho\alpha_p\Big(T-T_{ref}\Big)\{g\}d\Omega + \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} (\rho-\rho_0)\{g\}d\Omega \\ &+ \sum_e \int_{\Gamma_I^{+,e}} \left[\bar{N}\right]^{\mathrm{T}} \beta_u\Big(\left[\bar{N}\right]\{u\}^{i-1} - \vec{u}_I\Big)d\Gamma - \sum_e \int_{\Gamma_N^e} \left[\bar{N}\right]^{\mathrm{T}} \{t_N\}d\Gamma \end{aligned} \quad (6.25e)$$

$$\check{\mathbf{M}} = \sum_e \int_{\Omega^e} \left[\bar{N}\right]^{\mathrm{T}} \frac{1}{\Delta t}\rho\left[\bar{N}\right] d\Omega \quad (6.25f)$$

$$\check{\mathbf{K}} = \sum_e \int_{\Omega^e} \left[\check{B}\right]^{\mathrm{T}} 2\mu\left[\check{B}\right] d\Omega \quad (6.25g)$$

$$\check{\mathbf{P}} = -\sum_e \int_{\Omega^e} \{B\}\langle N\rangle d\Omega \quad (6.25h)$$

$$\check{\mathbf{F}} = \sum_e \int_{\Gamma_I^{+,e}} \left[\bar{N}\right]^{\mathrm{T}} \beta_u \left[\bar{N}\right] d\Gamma \quad (6.25i)$$

$\left[\check{B}\right]$ is called strain rate operator:

$$\begin{bmatrix} \bar{N}_{1,x} & 0 & \cdots & \bar{N}_{9,x} & 0 & \bar{M}_{1,x} & 0 & \cdots & \bar{M}_{9,x} & 0 \\ 0 & \bar{N}_{1,y} & \cdots & 0 & \bar{N}_{9,y} & 0 & \bar{M}_{1,y} & \cdots & 0 & \bar{M}_{9,y} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \cdots & \frac{1}{2}\bar{N}_{9,y} & \frac{1}{2}\bar{N}_{9,x} & \frac{1}{2}\bar{M}_{1,y} & \frac{1}{2}\bar{M}_{1,x} & \cdots & \frac{1}{2}\bar{M}_{9,y} & \frac{1}{2}\bar{M}_{9,x} \\ \frac{1}{2}\bar{N}_{1,y} & \frac{1}{2}\bar{N}_{1,x} & \cdots & \frac{1}{2}\bar{N}_{9,y} & \frac{1}{2}\bar{N}_{9,x} & \frac{1}{2}\bar{M}_{1,y} & \frac{1}{2}\bar{M}_{1,x} & \cdots & \frac{1}{2}\bar{M}_{9,y} & \frac{1}{2}\bar{M}_{9,x} \end{bmatrix} \quad (6.26)$$

$\langle B \rangle$ can be named divergence operator:

$$\langle \bar{N}_{1,x} \quad \bar{N}_{1,y} \quad \cdots \quad \bar{N}_{9,x} \quad \bar{N}_{9,y} \quad \bar{M}_{1,x} \quad \bar{M}_{1,y} \quad \cdots \quad \bar{M}_{9,x} \quad \bar{M}_{9,y} \rangle \tag{6.27}$$

6.3.5 Level set transport sub-model

Since ϕ is smooth over domain Ω, Eq. 6.6 can be solved directly by FEM. With the GLS scheme for stabilization and the forward Euler method for time integration, please see [8] for details, Eq. 6.6 becomes

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1} + \tilde{\mathbf{R}}_{\mathbf{C}}^{i-1} + \tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1} + \Big(\tilde{\mathbf{M}}^{i-1} + \tilde{\mathbf{M}}_{\mathbf{GLS}}^{i-1}\Big)\{\Delta\phi\} = 0 \tag{6.28}$$

where

$$\tilde{\mathbf{R}}_{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \{N\}\langle N\rangle \frac{\{\phi\}^{i-1} - \{\phi\}^{t-\Delta t}}{\Delta t} d\Omega \tag{6.29a}$$

$$\tilde{\mathbf{R}}_{\mathbf{C}}^{i-1} = \sum_e \int_{\Omega^e} \{N\} F \left\| [B]\{\phi\}^{t-\Delta t} \right\| d\Omega \tag{6.29b}$$

$$\tilde{\mathbf{R}}_{\mathbf{GLS}}^{i-1} = \sum_e \int_{\Omega^e} F[B]^{\mathrm{T}} \frac{[B]\{\phi\}^{t-\Delta t}}{\|[B]\{\phi\}^{t-\Delta t}\|} \tau_\phi^e \Big(\langle N\rangle \frac{\{\phi\}^{i-1} - \{\phi\}^{t-\Delta t}}{\Delta t} + F \left\| [B]\{\phi\}^{t-\Delta t} \right\| \Big) d\Omega \tag{6.29c}$$

$$\tilde{\mathbf{M}}^{i-1} = \sum_e \int_{\Omega^e} \frac{1}{\Delta t} \{N\}\langle N\rangle d\Omega \tag{6.29d}$$

$$\tilde{\mathbf{M}}_{\mathbf{GLS}}^{i-1} = \sum_e \int_{\Omega^e} F[B]^{\mathrm{T}} \frac{[B]\{\phi\}^{t-\Delta t}}{\|[B]\{\phi\}^{t-\Delta t}\|} \tau_\phi^e \frac{1}{\Delta t} \langle N\rangle d\Omega \tag{6.29e}$$

τ_ϕ^e is the stabilization parameter (see [8]).

6.3.6 Normal speed spreading sub-model

Likewise, Eq. 6.7 is solved by FEM with the GLS stabilization technique [8]. The weak formulation and the space discretization lead to the following set of algebraic equations:

$$\Big(\hat{\mathbf{K}} + \hat{\mathbf{K}}_{\mathbf{GLS}}\Big)\{\Delta F\} = 0 \tag{6.30}$$

where

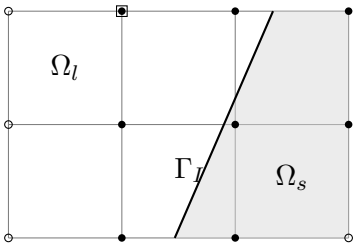
$$\hat{\mathbf{K}} = \sum_e \int_{\Omega^e} \{N\} sign(\phi) \big(\nabla\phi\big)^{\mathrm{T}} [B] d\Omega \tag{6.31a}$$

$$\hat{\mathbf{K}}_{\mathbf{GLS}} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}} \nabla\phi \tau_F^e \big(\nabla\phi\big)^{\mathrm{T}} [B] d\Omega \tag{6.31b}$$

τ_F^e is the stabilization parameter (see [8]).

6.3.7 Treatment of the calculation related to hydrodynamics and mass transport in the solid phase

The calculation related to hydrodynamics and mass transport in the solid region requires extra attention. In general, the transport of mass and momentum is always negligible in the solid. To this end, the ordinary components in velocity, pressure and concentration fields in the solid region are treated as Dirichlet boundary conditions in the present study, but the additional DOFs always keep active. At each time step according to the given interface location, as shown in Fig. 6.2, the ordinary DOFs on the nodes belonging to Ω_s are labelled as Dirichlet boundary and assigned with constant value.



◦ Ordinary node
• Enriched node
◙ Candidate node whose enrichment might be blocked

Figure 6.2 – A typical mesh in XFEM; shaded region represents the solid phase.

Remarks When the area ratio between the two parts separated by the phase front within one element is too small or too large, XFEM formulation would lead to ill-conditioned system [15, 18]. To alleviate this problem, the present study does not enrich the nodes that are originally supposed to but actually influenced faintly by the enrichment. As shown in Fig. 6.2, the enriched node surrounded by a square is likely to fall into this situation. If the area ratio is greater than a user-defined constant C_{block} ($C_{block} \gg 1$) or less than $\frac{1}{C_{block}}$, this node will just be taken as an ordinary one. An analogous treatment can be found in [25], in which the enrichment is blocked by assigning zero values to the additional DOFs.

6.3.8 Computational procedure

As shown in Fig. 6.3, the numerical model is mainly composed of five sub-models, and supplemented with the calculation of interface normal speed V_I by Eq. 6.8. The computational procedure can be outlined as below:

1. Initialize ϕ once computational domain and initial interface location are identified; meanwhile, T, ω, $\vec{u}$ and p are initialized.
2. Evaluate V_I according to current distribution of T by Eq. 6.8.
3. Spread V_I over Ω to obtain F by solving Eq. 6.30.

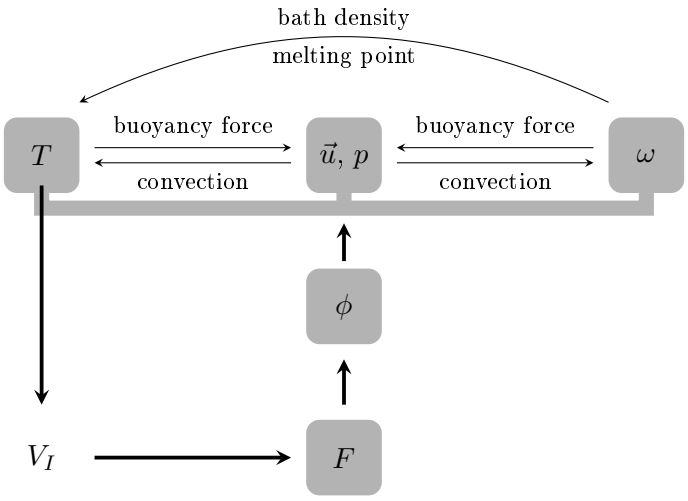


Figure 6.3 – The numerical model's structure (each sub-model is represented by its own unknowns) and the links among the sub-problems.

4. Update ϕ by solving Eq. 6.28; all the calculations get into the next time step.
5. Solve Eq. 6.18, Eq. 6.21 and Eq. 6.24 one by one based on the new interface location; one sub-problem is solved based on the latest solutions from the other two sub-problems; this process is repeated until all these three sub-models converge.
6. Go back to Step 2.

6.4 Numerical simulation of ledge profile

6.4.1 General description

During the Hall-Héroult process of producing aluminium, the side ledge plays a crucial role in maintaining the internal working condition of the cell at stable state. The thermal behaviour inside the cell influenced by the bath composition and liquid motion determines the ledge profile. The molten bath can be assumed to be composed of $3NaF \cdot AlF_3$ (cryolite) and excess AlF_3, while the solid ledge is taken as pure cryolite. In practice at the normal running condition, the cryolite ratio (CR) is around 2.3; the corresponding mass fraction of excess AlF_3 is 0.1085. Table 6.1 presents the material properties used in the following simulations. All the parameter values are from the literature, except that the coefficients related to each diffusion process in the liquid phase including k_l, D_l and μ_l are taken as free parameters. The values assigned to the free parameters are chosen to avoid the convection-dominant instability problem.

6.4.2 Melting case

As shown in Fig. 6.4, the computational domain is $0.1m \times 0.1m$ in size. The initial ledge front is vertical and located at $x = 0.06m$. Initially, the region as $x < 0.06m$ is filled with liquid

Table 6.1 – Thermophysical properties of ledge and bath

Parameters		Solid	Liquid
ρ	$[kg/m^3]$	2850[35]	$\left(\left(1946+1113\frac{2+3\omega}{8-3\omega}\right)^{-3.6}+\left(\frac{2+3\omega}{859(8-3\omega)}\right)^{3.6}\right)^{-1/3.6}+7.73$ [27]
C_p	$[J/(kg\cdot K)]$	1850[32]	1760[32]
k	$[W/(m\cdot K)]$	1.08[35]	160
μ	$[kg/(m\cdot s)]$	-	1×10^{-2}
D	$[m^2/s]$	-	4×10^{-5}
T_{ref}	[℃]	-	992 *
α_p	$[K^{-1}]$	-	4.7×10^{-4} **
T_m	[℃]	$1011+50\omega-3265\omega^{2.2}$ [28]	
L	$[J/kg]$	5.1×10^5 [35]	

* T_{ref} is set to be close to the melting point at the initial bath composition of $\omega = 0.1085$

** Average value calculated by density formula in [27]

bath being of uniform concentration of 0.1085 in terms of mass fraction of excess AlF_3. Fig. 6.4 also presents boundary conditions used in this case. For hydrodynamic calculation, no-slip boundary conditions are applied on the periphery of the entire domain and also on the ledge front; in addition, even though the density is distinct in different phases, the flow triggered by this density jump at the ledge front is omitted in this case, i.e. $\vec{u}_I = 0$. For species transport sub-model, the mass flux (species penetration) is not allowed on the external boundaries but is acceptable at the ledge front. For the thermal transport sub-model, temperature at the interface is always prescribed as the melting point T_m; the left and right side-walls are subject to higher temperature ($T = 1011$℃) and lower temperature ($T = 990$℃), respectively; the top and bottom are set to be insulated. For the sake of gentle start, the steady solution of the temperature, obtained under the same model settings but assuming the melting front fixed, is taken as the initial condition for the transient thermal problem. The hydrodynamic calculation starts with static state and the pressures are zero throughout the domain at initial time step. The pressure at $(0m, 0.1m)$ being $p = 0$ is taken as reference value throughout the simulation. The assignment of other model parameters is done as $C_{block} = 50$, $\Delta t = 1s$, $\delta d = 0.5h^e$, $\beta_T = 1\times10^8$, $\beta_u = 1\times10^6$.

According to the source of driving force, simulations based on the following three different scenarios are performed:

1. only temperature-gradient driven,
2. only concentration-gradient driven,
3. double-gradient driven.

In scenario 1, the calculation doesn't include solving the species transport sub-model. It means the influence of bath composition's variation is not taken into account. A constant melting point $T_m = 992$℃ is enforced at the ledge front in this scenario. In the second scenario, only the Boussinesq term in Eq. 6.5a is omitted. So the melt flow is driven purely by the species

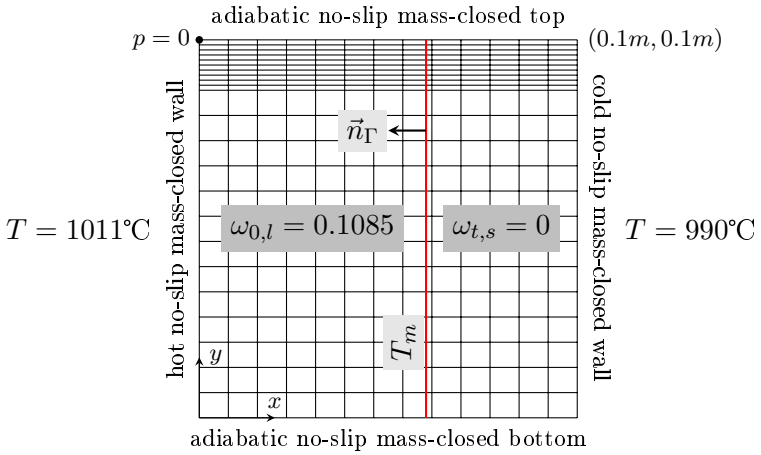
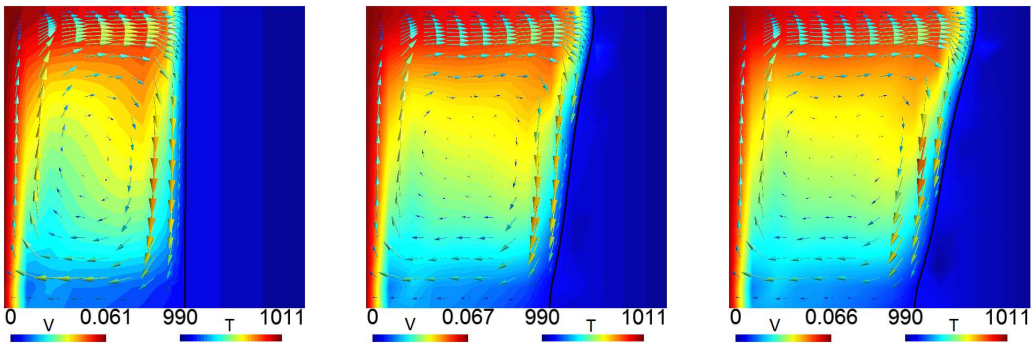


Figure 6.4 – Mesh used in the ledge melting case as well as the initial and boundary conditions; the red line represents the initial ledge front.

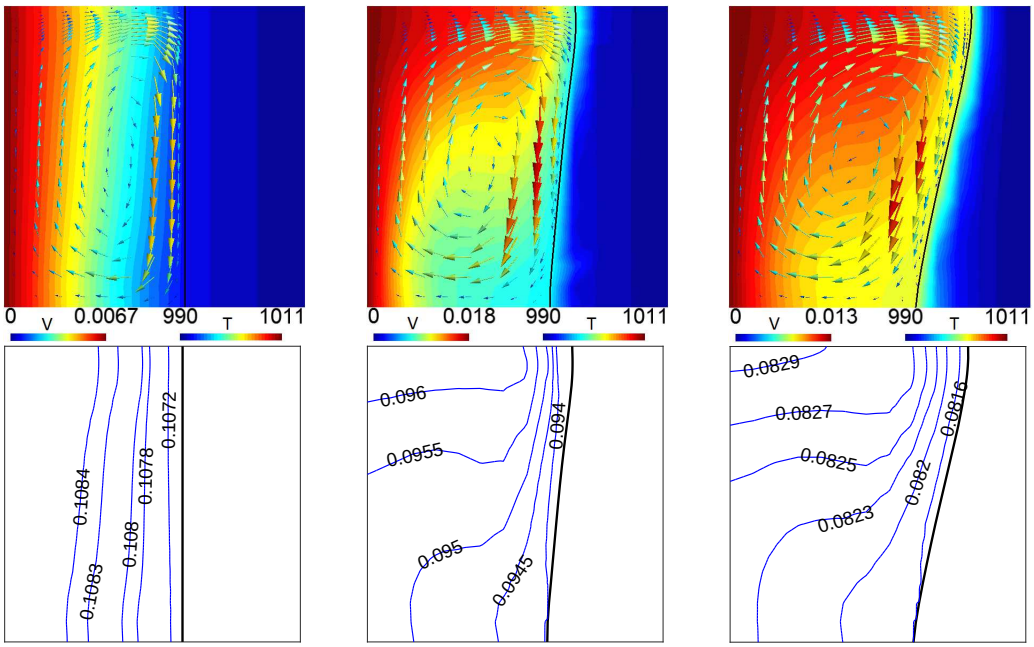
concentration gradient. In scenario 3, the melt flow is driven by both temperature gradient and concentration gradient.

The figures 6.5 to 6.7 present the temperature and velocity distributions at some time steps in each scenario, the latter two figures also present the concentration distributions of excess AlF_3. Observing the velocity magnitude shown in the sub-figure at $t = 5s$ of each scenario, we can conclude that the melt flow is mainly driven by the temperature gradient at the early stage. That's because the temperature gradient always exists but the concentration gradient won't form until the melting process start to take place. It can be seen from figures 6.6 and 6.7, the concentration of excess AlF_3 is gradually diluted in the melt as the ledge front propagates. The concentration dilution phenomenon originates from the nearby of the ledge front and spreads over the melt by diffusion and convection mechanisms. In this melting case, both temperature gradient and concentration gradient generate a downward driving force near the ledge front. The circulation flow induced by this downward driving force carries more heat into ledge front at the upper region. Therefore, the upper ledge melts faster. Under the present model settings, the driving force produced by the temperature gradient is stronger than that produced by the concentration gradient by comparing the velocity magnitude shown in the figures 6.5 to 6.7. The ledge melting speed is affected by the intensity of the circulation in the melt. As shown in Fig. 6.8, which exhibits the ledge front profiles at $t = 85s$ in each scenario, the ledge melting speed is the largest in scenario 1, then followed by scenario 3. The reason is that, the melting point is low and constant in the first scenario, while in the other two scenarios, the melting point is raised by the concentration dilution phenomenon (see the expression of T_m in Table 6.1). The variation of the melting point also can be seen in Fig. 6.9, which shows temperature profiles along the horizontal section of $y = 0.05m$ at different time steps in each scenario. The temperature gradient in scenario 1 is the largest, which generates the most intense circulation in these three scenarios.



(a) $t = 5s$ (b) $t = 60s$ (c) $t = 85s$

Figure 6.5 – Temperature and velocity distributions at some time steps for scenario 1 of ledge melting case.

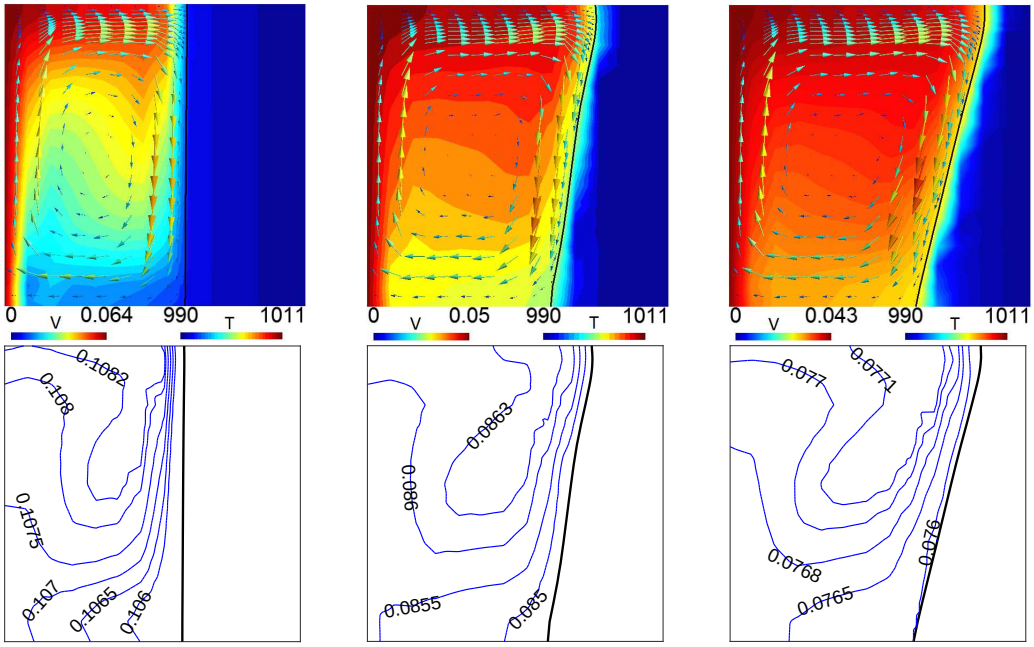


(a) $t = 5s$ (b) $t = 100s$ (c) $t = 310s$

Figure 6.6 – Temperature, velocity (above) and concentration (below) distributions at some time steps for scenario 2 of ledge melting case.

6.4.3 Solidification case

In this case, the supplementary flow caused by the density change between the two phases at the interface during the phase transition process is mainly investigated. To highlight the impact of this supplementary flow, the buoyancy force induced either by the temperature gradient or by the concentration gradient is ignored. As shown in Fig. 6.10, which shows the mesh and the boundary conditions, the initial ledge front is located at $x = 0.025m$, an outward heat flux increasing with coordinate y is imposed on the boundary $x = 0m$, and the temperature is prescribed as 995℃ on the boundary $x = 0.1m$ throughout the simulations.



(a) $t = 5s$ (b) $t = 100s$ (c) $t = 200s$

Figure 6.7 – Temperature, velocity (above) and concentration (below) distributions at some time steps for scenario 3 of ledge melting case.

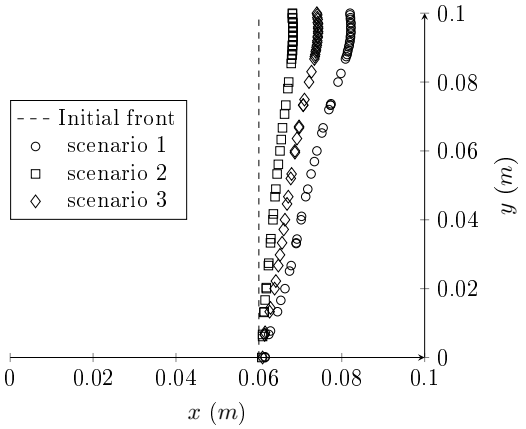
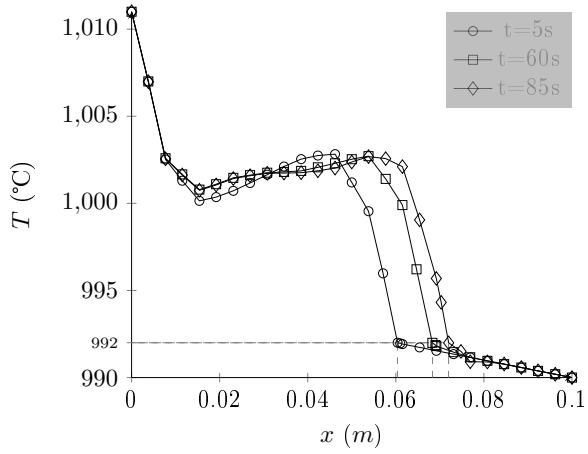


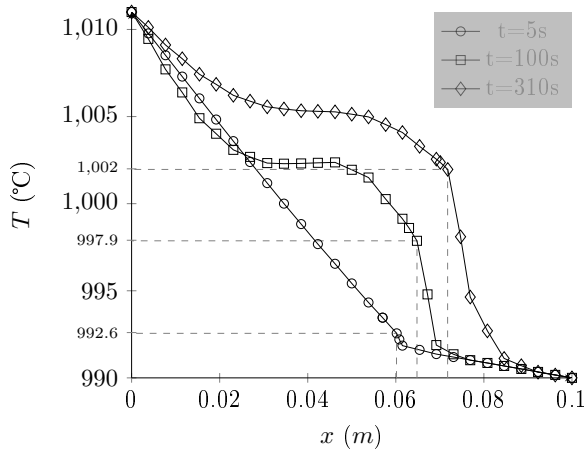
Figure 6.8 – Interface profile at $t = 85s$ in different scenarios of ledge melting case.

For the hydrodynamic calculation, the fluid velocity at the interface obtained by Eq. 6.9d is imposed through the penalty method, while the traction is set to be free on the top boundary to allow fluid penetration. All the other model settings are identical to that in the ledge melting case, but except $\rho_l = 2064kg/m^3$, $D_l = 10^{-6}m^2/s$, $k_l = 0.8W/(m \cdot K)$ and $\Delta t = 10s$ in this case. The density jump at the interface reaches up to $\frac{\rho_s - \rho_l}{\rho_s} \approx 28\%$.

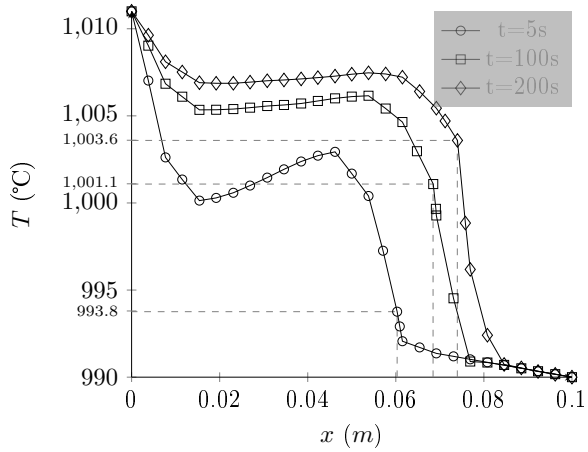
The velocity, temperature and concentration distributions at different time steps are shown in Fig. 6.11. It can be seen that, the outward heat flux imposed at $x = 0m$ produces temperature distribution in the solid region decreasing along y axis, while the temperature



(a) scenario 1



(b) scenario 2



(c) scenario 3

Figure 6.9 – Temperature profiles along the horizontal section of $y = 0.05m$ at different time steps in the scenarios of ledge melting case; the corresponding melting points and interface locations are marked out by the dashed lines.

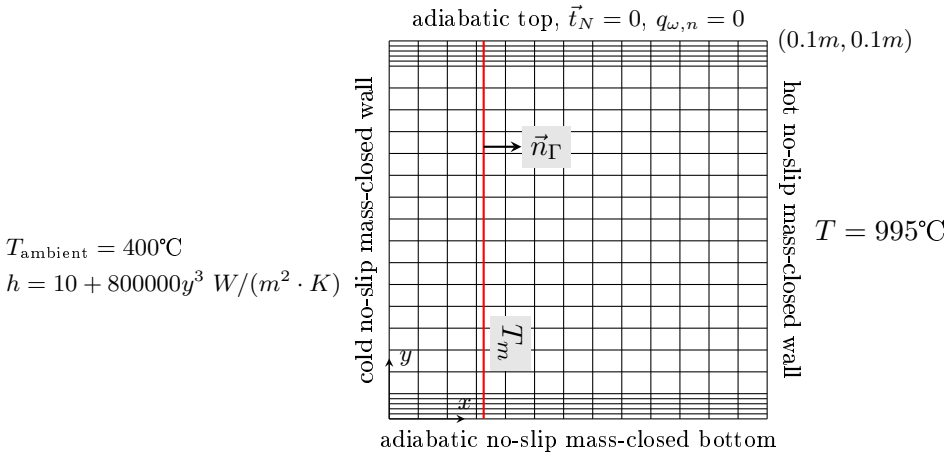
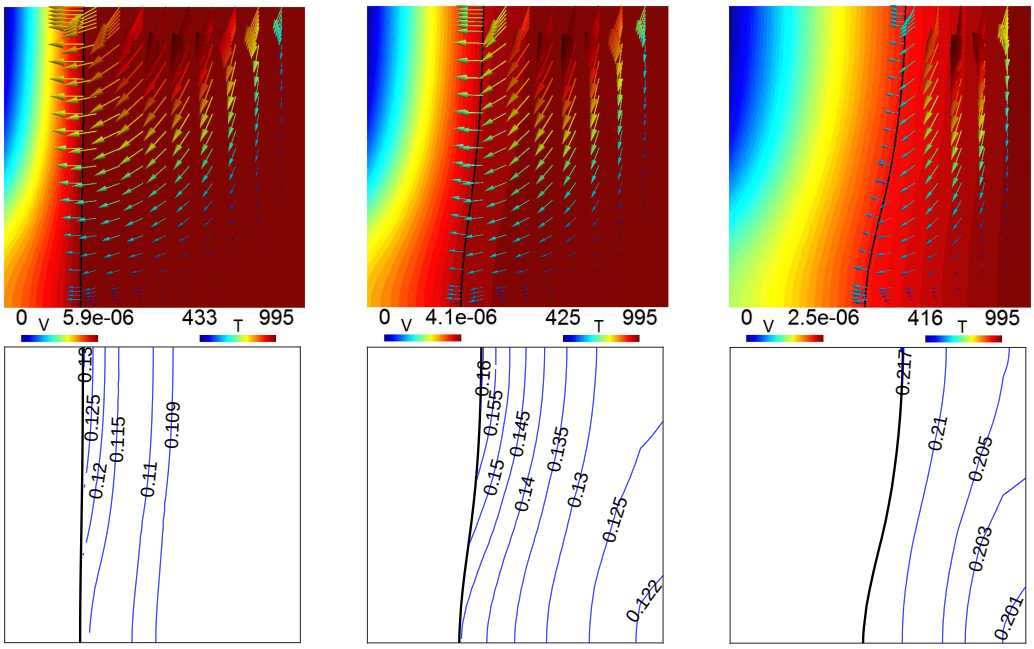


Figure 6.10 – Mesh used in the ledge solidification case and boundary conditions; the red line represents the initial ledge front.



(a) $t = 2min$ (b) $t = 30min$ (c) $t = 126min$

Figure 6.11 – Temperature, velocity (above) and concentration (below) distributions at different time steps in ledge solidification case; the bold solid line indicates the ledge front.

gradient is increasing. This temperature pattern makes the solidification take place and its speed increase along y axis. As the ledge front propagates, due to the density jump between different phases at the interface, volume contraction induces a supplementary melt flow going towards the ledge front. The local melt velocity is perpendicular to the ledge front, because no-slip condition is enforced at the interface. According to Eq. 6.9d, the melt velocity at the interface is proportional to the local normal speed of the interface. Another consequence of

the solidification is the rejection of the excess AlF_3 from the solid into the melt. The excess AlF_3 accumulates at the ledge front as the ledge front moves. This accumulation is intensified where the speed of ledge front is large. The concentration accumulation at the interface leads to the decrease of the melting point (see the expression of T_m in Table 6.1). As shown in Fig. 6.12, which exhibits the temperature profiles along the horizontal section of $y = 0.08m$ at different time steps, the melting point decreases dramatically during the solidification process.

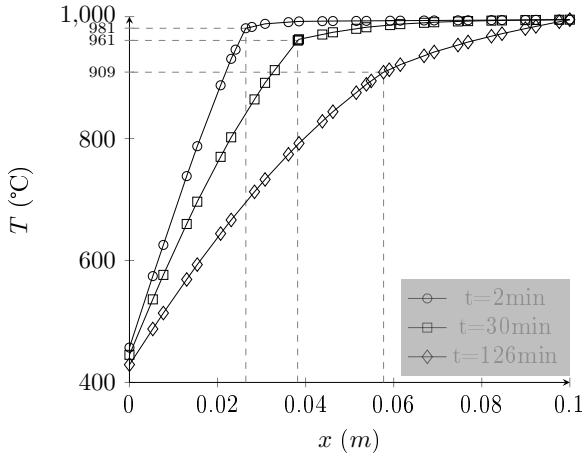


Figure 6.12 – Temperature profiles along the horizontal section of $y = 0.08m$ at different time steps in ledge solidification case; the corresponding melting points and interface locations are marked out by the dashed lines.

In theory, the supplementary flow exists to fulfil the mass conservation law under the constraint of incompressibility. Therefore, the inflow mass at the top boundary should be identical to the increment of the total mass. As shown in Fig. 6.13, the total mass (solid and liquid) is increasing as the solidification process goes on, and the increment of the total mass has a good agreement to the inflow mass at each time step.

In order to figure out the extent to which the solidification process is influenced by the melt flow caused by the density jump between phases, a simulation based on the same model settings but without melt flow is also done. Fig. 6.14 presents the interface profiles at different time steps in both simulations. It's shown that the impact of this melt flow is almost negligible. However, it just means that the velocity magnitude of this melt flow is too small to affect the transport of the heat and chemical species. Actually in the real cell, no bath is filled constantly from the top to supplement the melt, so the volume reduction by this large density jump between phases should be considerable. The simulation of volume reduction requires the top boundary to be treated as another free boundary, which is beyond the scope of the present work.

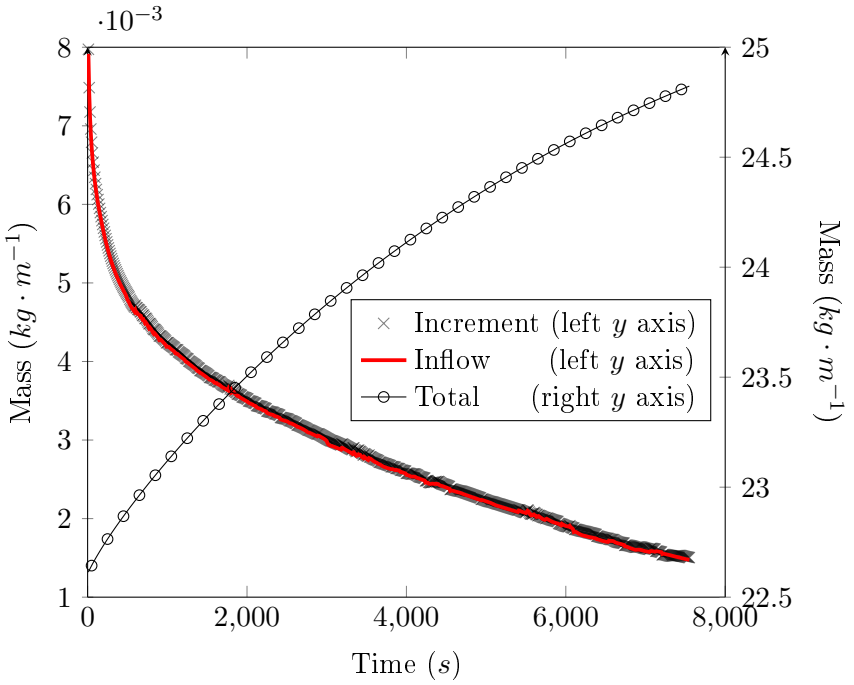


Figure 6.13 – The total mass, the mass increment and the inflow mass through the top boundary at each time step.

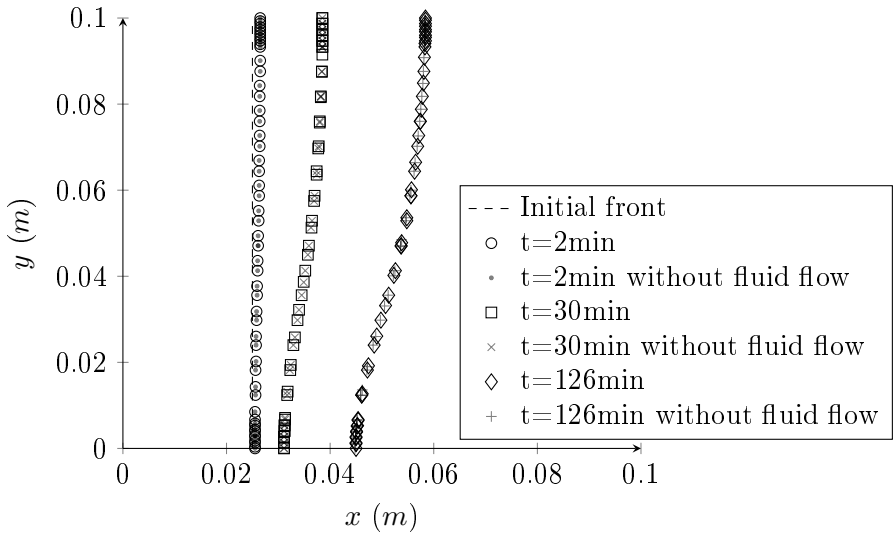


Figure 6.14 – Interface profiles at different time steps in two scenarios of ledge solidification case.

6.5 Conclusions

A two-dimensional numerical model based on XFEM and level set method to describe the phase change process with ensuing melt flow and mass transport in the binary system has been presented. The model is characterized by simulating this multi-physical problem based on a

fixed mesh, and that the density in each term and the melting point are dependent on the melt composition. The application of the present model to predict the ledge profile under certain scenarios shows its accuracy and robustness. In predicting the ledge profile under current model settings, it's shown that, the influence of the thermosolutal flow is obvious, while the melt flow induced by the density jump between solid and liquid phases is too faint to affect the heat and species transport, even though the density difference between the two phases reaches to $\sim 28\%$ with respect to the heavier one. In this model, more discontinuities emerge compared with the classical Stefan problem, XFEM shows its flexibility to deal with different types of discontinuity. In the future work, solving the complete Navier-Stokes equations and adding stabilization techniques to circumvent convection-dominant instability problem and the saddle point problem requires to be done to improve the capability of the present model.

Acknowledgements

The authors would like to acknowledge the National Sciences and Engineering Research Council of Canada (NSERC) funding NSERC Discovery Grant (No. 36518) for supporting the present research. A part of the research presented in this paper was financed by the Fonds de Recherche du Québec-Nature et Technologie (FRQNT) by the intermediary of the Aluminium Research Centre-REGAL. The first author also acknowledges the financial support from the Chinese Scholarship Council (CSC).

Bibliography

[1] C. Beckermann and R. Viskanta. Double-diffusive convection during dendritic solidification of a binary mixture. *PhysicoChemical Hydrodynamics*, 10(2):195–213, 1988.

[2] C. Beckermann and R. Viskanta. An experimental study of melting of binary mixtures with double-diffusive convection in the liquid. *Experimental Thermal and Fluid Science*, 2(1):17–26, 1989.

[3] W. D. Bennon and F. P. Incropera. A continuum model for momentum, heat and species transport in binary solid-liquid phase change systems-I. Model formulation. *International Journal of Heat and Mass Transfer*, 30(10):2161–2170, 1987.

[4] W. D. Bennon and F. P. Incropera. A continuum model for momentum, heat and species transport in binary solid-liquid phase change systems-II. Application to solidification in a rectangular cavity. *International Journal of Heat and Mass Transfer*, 30(10):2171–2187, 1987.

[5] M. K. Bernauer and R. Herzog. Implementation of an X-FEM solver for the classical two-phase Stefan problem. *Journal of Scientific Computing*, 52(2):271–293, 2012.

[6] P. R. Chakraborty and P. Dutta. Study of freckles formation during directional solidification under the influence of single-phase and multiphase convection. *Journal of Thermal Science and Engineering Applications*, 5(2):021004, 2013.

[7] K. W. Cheng and T. P. Fries. Higher-order XFEM for curved strong and weak discontinuities. *International Journal for Numerical Methods in Engineering*, 82(5):564–590, 2010.

[8] J. Chessa, P. Smolinski, and T. Belytschko. The extended finite element method (XFEM) for solidification problems. *International Journal for Numerical Methods in Engineering*, 53(8):1959–1977, 2002.

[9] J. Chessa, H. Wang, and T. Belytschko. On the construction of blending elements for local partition of unity enriched finite elements. *International Journal for Numerical Methods in Engineering*, 57(7):1015–1038, 2003.

[10] A. Cosimo, V. Fachinotti, and A. Cardona. An enrichment scheme for solidification problems. *Computational Mechanics*, pages 1–19, 2013.

[11] S. H. Davis. *Theory of solidification*. Cambridge University Press, 2001.

[12] J. Dolbow and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):131–150, 1999.

[13] H. El, M. F. Yousef, and J. M. Khodadadi. Numerical simulation of the effect of the size of suspensions on the solidification process of nanoparticle-enhanced phase change materials. *Journal of Heat Transfer*, 135(5):052901, 2013.

[14] T. P. Fries. A corrected XFEM approximation without problems in blending elements. *International Journal for Numerical Methods in Engineering*, 75(5):503–532, 2008.

[15] T. P. Fries and T. Belytschko. The extended/generalized finite element method: an overview of the method and its applications. *International Journal for Numerical Methods in Engineering*, 84(3):253–304, 2010.

[16] J. C. Heinrich and D. R. Poirier. Convection modeling in directional solidification. *Comptes Rendus Mecanique*, 332(5-6):429–445, 2004.

[17] H. Ji, D. Chopp, and J. E. Dolbow. A hybrid extended finite element/level set method for modeling phase transformations. *International Journal for Numerical Methods in Engineering*, 54(8):1209–1233, 2002.

[18] C. Lang, D. Makhija, A. Doostan, and K. Maute. A simple and efficient preconditioning scheme for heaviside enriched XFEM. *Computational Mechanics*, 54(5):1357–1374, 2014.

[19] A. Mahato and A. Kumar. Modeling transport phenomena of ice slurry in an ice forming unit. *International Journal of Refrigeration*, 69:205–222, 2016.

[20] D. Martin, H. Chaouki, J. L. Robert, D. Ziegler, and M. Fafard. Modelling of phase change with non-constant density using XFEM and a Lagrange multiplier. *Frontiers in Heat and Mass Transfer*, 7(1), 2016.

[21] E. J. Pickering. Macrosegregation in steel ingots: the applicability of modelling and characterisation techniques. *ISIJ International*, 53(6):935–949, 2013.

[22] J. Prüss, J. Saal, and G. Simonett. Existence of analytic solutions for the classical Stefan problem. *Mathematische Annalen*, 338(3):703–755, 2007.

[23] L. Salvatori and N. Tosi. Stefan problem through extended finite elements: review and further investigations. *Algorithms*, 2(3):1177–1220, 2009.

[24] H. Sauerland and T. P. Fries. The extended finite element method for two-phase and free-surface flows: a systematic study. *Journal of Computational Physics*, 230(9):3369–3390, 2011.

[25] H. Sauerland and T. P. Fries. The stable XFEM for two-phase flows. *Computers & Fluids*, 87:41–49, 2013.

[26] J. C. Slattery. *Advanced transport phenomena*. Cambridge University Press, 1999.

[27] A. Solheim. The density of molten NaF-LiF-AlF_3-CaF_2-Al_2O_3 in aluminium electrolysis. *Aluminum Transactions*, 2(1):161–168, 2000.

[28] A. Solheim, S. Rolseth, E. Skybakmoen, L. Støen, Å. Sterten, and T. Støre. Liquidus temperatures for primary crystallization of cryolite in molten salt systems of interest for aluminum electrolysis. *Metallurgical and Materials Transactions B*, 27(5):739–744, 1996.

[29] P. Stapór. The XFEM for nonlinear thermal and phase change problems. *International Journal of Numerical Methods for Heat & Fluid Flow*, 25(2):400–421, 2015.

[30] P. Stapór. A two-dimensional simulation of solidification processes in materials with thermo-dependent properties using XFEM. *International Journal of Numerical Methods for Heat & Fluid Flow*, 26(6):1661–1683, 2016.

[31] L. Tan and N. Zabaras. A level set simulation of dendritic solidification of multi-component alloys. *Journal of Computational Physics*, 221(1):9–40, 2007.

[32] M. P. Taylor, W. D. Zhang, V. Wills, and S. Schmid. A dynamic model for the energy balance of an electrolysis cell. *Chemical Engineering Research and Design*, 74(8):913–933, 1996.

[33] M. Theillard, F. Gibou, and T. Pollock. A sharp computational method for the simulation of the solidification of binary alloys. *Journal of Scientific Computing*, 63(2):330–354, 2015.

[34] V. R. Voller, A. D. Brent, and C. Prakash. The modelling of heat, mass and solute transport in solidification systems. *International Journal of Heat and Mass Transfer*, 32(9):1719–1731, 1989.

[35] C. C. Wei, J. Chen, B. J. Welch, and V. R. Voller. Modelling of dynamic ledge heat transfer. *TMS-Light Metals*, pages 309–316, 1997.

[36] N. Zabaras, B. Ganapathysubramanian, and L. Tan. Modelling dendritic solidification with melt convection using the extended finite element method. *Journal of Computational Physics*, 218(1):200–227, 2006.

[37] P. Zhao, J. Heinrich, and D. Poirier. Effect of variable density in dendritic solidification. In *42nd AIAA Aerospace Sciences Meeting and Exhibit*, page 628, 2004.

Chapter 7

Conclusion

7.1 Summary

7.1.1 On the aspect of the source code building

Much of the work in this project is to manipulate the source code. The in-house software FESh++ provides a great object-oriented programming environment to build finite element models. The new features (like new physical processes, new materials, new constitutive laws, new discontinuities, to name but a few) involved in the present work are added in terms of various *classes*. The addition of new *classes* and the manipulation of every *class* are straightforward, as long as one has a deep comprehension and clear picture in the mind of the code's structure and hierarchy. That means FESh++ supports the development of new features very much.

FESh++ is characterized by the capability to deal with the non-linear and multi-physical problems. As discussed in chapter 3, either for the non-linearity that happens in each physical problem (e.g. caused by the unknown-dependent physical properties) or for the coupling degree among different physics, there are two schemes for selection. Regarding to non-linearity issue, the linearisation tools include Newton-Raphson method and substitution method. Appendix C presents the contribution of each formulation term to the final discrete system based on these two algorithms, respectively. In the real application, the Newton-Raphson method is first implemented due to its faster convergence speed in theory; if it fails, the substitution method will be the second choice. Regarding to the issue of coupling degree, the strong coupling means that all the physical fields are addressed simultaneously and only one discrete system is produced; while in the weak coupling scheme, the physical processes are addressed alternately, i.e. one process is solved by utilizing the latest solutions of other process until all converge. As discussed in chapter 3, although the strong coupling scheme is computationally efficient (less total time consumption), it might make different physical processes tangled to each other too much, and the larger the final system of algebraic equations is, the higher it

sets requirements for the linear solver. The scheme of strong coupling might be used for the problem with two physics involved. For the problem with multi-physics (more than two), the weak coupling scheme would enable each sub-problem relatively independent, which leads to a model with clear structure and well modularized.

At present, FESh++ can deal with problems of being either continuous (by FEM) or discontinuous (weak or strong by XFEM). The discontinuous one is taken as the special case of continuous problems, then the *classes* related to the discontinuous case are built as the children *classes* of the normal (continuous) ones, such that both cases (continuous and discontinuous) can share their common parts and the old *classes* are compatible to the newly-added discontinuity without any modification. In summary, FESh++ is well-organized in structure and one has a large flexibility in both development and application of FESh++.

7.1.2 On the aspect of the numerical model

A two-dimensional numerical model to describe the phase change with ensuing melt flow and species transport in the binary system has been presented in this thesis. To overcome the difficulty caused by the presence of free moving internal boundary and discontinuities at the free boundary, XFEM combined with level set method is used such that simulations can be implemented over a fixed mesh while the interface keeps sharp. It's attractive because working with a fixed mesh avoids the cumbersome remeshing and the level set method holds the potential to deal with free interface with topological change (splitting or merging). The price of working with a fixed mesh is that it requires the extra effort to enforce the boundary conditions (constraints) at the interface. There are two types of boundary conditions involved. The Neumann-type ones can be accounted for naturally during the weak formulation, please refer to how the species flux is imposed at the liquid side of the interface. The Dirichlet-type ones such as the melting point and liquid velocity at the interface are imposed by the penalty method in the present work. The constraints enforcement through the penalty method works well with the penalty parameter β ranging from 10^6 to 10^8. Unlike the Lagrange multiplier method, the implementation of the penalty method is straightforward without changing the structure of the final matrix system.

The final model involves three physical problems including phase change, fluid flow and species transport. In the course of building the final model, three subjects listed below are investigated by coupling each two of these three problems. The accuracy of the corresponding sub-models is verified by the analytical solutions or those obtained by the conventional methods. The three subjects corresponding respectively to the chapters 3 to 5 include:

1. one-phase density driven flow

   The mixing of two miscible fluids with different densities is described by solving the coupled problem involving fluid flow and species transport through the conventional finite element method. The use of mass-fraction-based convection-diffusion equation to

govern the species' motion in the variable density case is justified. It also shows that this equation in conservative form is a little better than that in non-conservative form in mass conservation.

2. Stefan problem without convection mechanism in the binary system
   The phase change process without convection mechanism in the binary system is described by solving the coupled problem involving heat and species transports through XFEM combined with level set method. In the field of XFEM's application, it's a fresh attempt to consider the influence of melt composition while solving the Stefan problem.
3. Stefan problem with ensuing melt flow in pure material
   The phase change process with ensuing melt flow, which can be driven by the thermal buoyancy force or density jump between the two phases, in pure material is described by solving the coupled problem involving fluid flow and heat transport through XFEM combined with level set method. In the field of XFEM's application, it's also a fresh attempt to consider the influence of the ensuing melt flow while solving the Stefan problem.

The final model to describe the phase change with ensuing melt flow and species transport is applied to the bath/ledge system under certain scenarios. The present model works well in dealing with the case in which the melting point and melt density is sensitive to the solute concentration. The influences of three driving forces are mainly investigated. Under the current model setting, the influences of buoyancy forces caused by thermal and solute gradients are obvious, while the melt flow induced by the density jump between phases is too faint to affect the heat and species transport, even though the density difference reaches to $\sim 28\%$. But it doesn't mean the influence of this density difference is not important, because it will cause considerable volume contraction in the melt which is beyond the present project.

XFEM has exhibited its accuracy and flexibility in dealing with different types of discontinuity. The weak discontinuity in the temperature field is well approximated by the corrected abs-enrichment scheme without blending element problems, while the strong discontinuities in the species concentration and pressure fields are approximated by the sign-enrichment scheme. The discontinuity type of the velocity field depends on the constraints at the interface. Nevertheless, the sign-enrichment scheme with the imposition of Dirichlet-type conditions at the interface works well for either the weakly or strongly discontinuous situation.

One drawback of XFEM formulation is that it tends to yield an ill-conditioned system when the nodes' support taking up in one phase is small and far. As shown in the one-dimensional example presented in appendix D, the condition number increases dramatically as the discontinuity approaches to the element node. To alleviate this problem, the present work tend to convert the originally enriched nodes with small and far support in one phase back to ordinary ones. It shows that after this treatment, the condition number increases slightly and the

accuracy of the solutions is good as the discontinuity close to the element node.

In the present work, the level set method used to capture the free interface implicitly also shows its advantages in combining with XFEM. The level set values can be utilized not only to reconstruct the interface and its unit normal easily, but also to find out the enriched elements, to build the enrichment functions, as well as to define the material properties.

7.1.3 On the aspect of modelling ledge freezing/melting

The present model takes the influence of bath composition and bath motion into account while describing the freezing/melting of the ledge. Regarding to the bath composition, it is represented by the content of excess AlF_3 based on the assumption that the bath is taken as a binary system. This assumption is reasonable, because in industrial production the cryolite ratio (CR) which is just based on the same binary system is often used to indicate the bath composition. However, the present model is not merely limited to this situation, it can also be applied to investigate the influence of any species (additives or impurities) as long as the relevant constitutive relations are given. Regarding to the bath motion, the driving forces caused by the thermal gradient, solute gradient or density jump between phases are mainly investigated. But the present model also holds the potential to consider other external forces like Lorentz force and gas bubbles' drag force by adding them into the momentum equation.

7.2 Future work

The work to build a numerical model based on XFEM combined with level set method to describe the phase change with ensuing melt flow and species transport in the ledge/bath system has been done. In order to enhance the model's performance and extend its capability, future work can be done from the following aspects:

- Stabilization is needed. The instability of the present model is mainly caused by two facts, including convection-dominated problems whose severity is often indicated by Peclet number or Reynolds number and the saddle point problem in Stokes or Navier-Stokes mixed formulation. To overcome the instability problem related to the convection-domination, the streamline-upwind/Petrov-Galerkin (SUPG, also called streamline diffusion method) and Galerkin/least-squares (GLS) methods are commonly used. In both methods, weighted residuals with a so-called stabilization parameter τ are added into the weak form without giving up consistency and degrading accuracy. The solvability of saddle point problems requires the Ladyženskaja-Babuška-Brezzi condition to be fulfilled or circumvented. In the present work, the LBB condition is satisfied by choosing a specific combination of interpolation functions for velocity and pressure fields, i.e. Q2Q1 element is used. But the stability of the modified combination pair such as Q2Q1 in XFEM is an open problem. An alternative way is to circumvent the LBB condition

by adding appropriate residual-based perturbation terms, such that any combination of interpolations is supposed to be stable. This kind of method includes the pressure-stabilizing/Petrov-Galerkin (PSPG) and GLS methods.

- Reinitialization is needed in level set method. In the present model, the normal speed field F used to transport the level set function ϕ is constructed by spreading the speed of the interface to be orthogonal to ϕ, such that the property of ϕ being signed distance function is well preserved in the situation only planar interface (low curvature) involved. For the case in which the interface has large curvature, the orthogonality of F to ϕ will not be well guaranteed, then the update will make ϕ lose the property of being distance function. Therefore, ϕ requires to be reinitialized after each or several time steps.
- A more complex model can be developed. The accurate prediction of the ledge profile relies on the comprehensive description of the associated processes. The phrase "comprehensive description" not only refers to adding more physical phenomena such as electrics, magnetohydrodynamics and gas bubbles, also means including a larger domain in space or extending the present model into three dimensions. Building a such complex model (involving more discontinuities), which is really a giant and challenging project represents the development direction of the present model.

Appendix A

Tools

A.1 Reynolds transport theorem

In fluid mechanics, two distinct approaches, i.e. Lagrangian and Eulerian descriptions, are commonly used to describe physical processes from the mathematical point of view. The Lagrangian description is based on the concept **fluid parcel** (also called material element of fluid), which is a collection of matter of fixed identity and follows the flow. The fluid parcel is a closed deformable system, which means only the energy flux is allowed to cross the boundary, but the mass flux is not. Within the Lagrangian reference frame, the time change rate of a quantity is denoted by $\frac{d}{dt}$, known as material derivative (or substantial derivative, total derivative). The Eulerian description centres on the concept **control volume**, which is a field description. The control volume is an completely open system and allows mass and energy to pass through the boundary. The time change rate of a quantity within the Eulerian reference frame is called local derivative, denoted by $\frac{\partial}{\partial t}$.

The basic conservation laws of natural physics can apply directly in the Lagrangian description, since it works with mass-closed system. Nevertheless, the established differential equations are not in the convenient form to work with. The strategy is to convert the differential equations obtained by the Lagrangian description into the equivalent form in the Eulerian description through the bridge in between, known as Reynolds transport theorem. In calculus, the Reynolds transport theorem is the generalization of the Leibniz integral rule in three dimensions. It can be used to move the time derivative into the integral operator as time derivative of an integrated quantity being calculated.

Let $\mathbf{B}$ represents any extensive property, scalar or vector. The corresponding intensive prop-

erty is denoted by $\mathbf{b}$. Reynolds transport theorem can be expressed mathematically as

$$\begin{aligned}\frac{dB}{dt} &= \frac{d}{dt}\left(\int_{\Omega(t)} \mathbf{b} d\Omega\right) \\ &= \int_{\Omega(t)} \frac{\partial \mathbf{b}}{\partial t} + \nabla \cdot (\mathbf{b} \otimes \vec{u})\, d\Omega \\ &= \int_{\Omega(t)} \frac{\partial \mathbf{b}}{\partial t} d\Omega + \int_{\Gamma(t)} (\vec{u} \cdot \vec{n})\, \mathbf{b} d\Gamma\end{aligned} \tag{A.1}$$

where Ω is the control volume, Γ is the bounding surface of the control volume Ω, $\vec{u}$ is the velocity vector, and $\vec{n}$ is the outward unit normal vector at the surface. In words, the physical interpretation of Eq. A.1 reads

Time change rate of any extensive property in the mass-closed system = *Integral of local partial time change rate of the intersive property in the open system* + *Net flux of the intensive property over the surface of the open system*

Worthing noting is that the Reynolds transport theorem is only required when the conversion from the Lagrangian description to the Eulerian description happens to the extensive property. For an intensive property, the conversion can be accomplished simply by the definition of material derivative:

$$\frac{d\mathbf{b}}{dt} = \frac{\partial \mathbf{b}}{\partial t} + \vec{u} \cdot \nabla \mathbf{b} \tag{A.2}$$

A.2 Solver in FESh++

Within the framework of finite element method (including the extended finite element method), supposing u represents any unknown variable to be determined, we can always obtain the following algebraic equation system after weak formulation and discretization:

$$\begin{aligned}W &= \delta u^{\mathrm{T}} R(u) \\ &= \delta u^{\mathrm{T}} \left(K(u)\, u - F\right)\end{aligned} \tag{A.3}$$

where W represents weak formulation; $R(u)$ is the residual; $K(u)$ is the coefficient matrix, which is unknown dependent for non-linear problems; F is the load-related column, which is unknown independent. In the present study, mainly the following two methods are employed to solve Eq. A.3.

- Substitution method

$$\begin{aligned}W\left(u^i\right) &= \delta u^{\mathrm{T}} R\left(u^i\right) \\ &= \delta u^{\mathrm{T}} \Big(K\left(u^{i-1}\right)\left(u^{i-1} + \Delta u\right) - F\Big) \\ &= \delta u^{\mathrm{T}} \Big(R\left(u^{i-1}\right) + K\left(u^{i-1}\right)\Delta u\Big)\end{aligned} \tag{A.4}$$

- Newton-Raphson method

$$\begin{aligned} W\left(u^i\right) &= \delta u^{\mathrm{T}} R\left(u^i\right) \\ &= \delta u^{\mathrm{T}}\Big(R\left(u^{i-1}\right) + \Delta R\Big) \\ &= \delta u^{\mathrm{T}}\Big(R\left(u^{i-1}\right) + \frac{\partial R}{\partial u}\Big|_{u^{i-1}} \Delta u\Big) \\ &= \delta u^{\mathrm{T}}\Big(R\left(u^{i-1}\right) + K_t\left(u^{i-1}\right)\Delta u\Big) \end{aligned} \tag{A.5}$$

where the superscript i represents the iteration step and Δ represents the increment.

Two measures are used to check the solution converged or not, as shown below.

- The absolute value of residual must approach zero maximally, which can be expresses as

$$|R\left(u^i\right)| < 10^{-m} \tag{A.6}$$

  where m is a positive integer.

- The solution difference between two successive iterations must approach zero maximally. It can be fulfilled by evaluating the following relative error:

$$\Big|\frac{u^i - u^{i-1}}{u^{i-1}}\Big| < 10^{-m} \tag{A.7}$$

A.3 Jacobian matrix and determinant

Jacobian matrix $[J]$ is used for transformation of derivative from with respect to the real element (global coordinate system $x-y-z$) to with respect to the reference element (local parametric coordinate system $\xi-\eta-\zeta$). **Jacobian determinant** $|J|$ is used for transformation of domain integration from the real element to the reference element. Generally speaking, the severer the distortion of the real element is in comparison with the ideally shaped element, the worse is the transformation. Take bi-linear quadrangle element as example, as long as the element shape is structural (rectangle or parallelogram), the Jacobian determinant is uniformly distributed within the element. In other cases, the Jacobian determinant varies within the element, so it depends on the location where it's evaluated. Actually, the Jacobian determinant reveals the **local stretching** of the parametric coordinate space required to fit onto the global coordinate space. The variation intensity of Jacobian determinant reveals the degree of distortion.

For any real element,

1. in three dimensions

$$[J] = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{bmatrix} = \begin{Bmatrix} \frac{\partial}{\partial \xi} \\ \frac{\partial}{\partial \eta} \\ \frac{\partial}{\partial \zeta} \end{Bmatrix} \begin{Bmatrix} x & y & z \end{Bmatrix} = \underset{3\times n^e}{\begin{bmatrix} N_{1,\xi} & N_{2,\xi} & \dots & N_{n^e,\xi} \\ N_{1,\eta} & N_{2,\eta} & \dots & N_{n^e,\eta} \\ N_{1,\zeta} & N_{2,\zeta} & \dots & N_{n^e,\zeta} \end{bmatrix}} \underset{n^e\times 3}{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_{n^e} & y_{n^e} & z_{n^e} \end{bmatrix}} \tag{A.8}$$

   where $x = \langle N \rangle \{x\}$, $y = \langle N \rangle \{y\}$, $z = \langle N \rangle \{z\}$.

2. in two dimensions

$$[J] = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \end{bmatrix} = \begin{Bmatrix} \frac{\partial}{\partial \xi} \\ \frac{\partial}{\partial \eta} \end{Bmatrix} \begin{Bmatrix} x & y & z \end{Bmatrix} = \underset{2\times n^e}{\begin{bmatrix} N_{1,\xi} & N_{2,\xi} & \dots & N_{n^e,\xi} \\ N_{1,\eta} & N_{2,\eta} & \dots & N_{n^e,\eta} \end{bmatrix}} \underset{n^e\times 3}{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_{n^e} & y_{n^e} & z_{n^e} \end{bmatrix}} \tag{A.9}$$

3. in one dimension

$$[J] = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \end{bmatrix} = \begin{Bmatrix} \frac{\partial}{\partial \xi} \end{Bmatrix} \begin{Bmatrix} x & y & z \end{Bmatrix} = \underset{1\times n^e}{\begin{bmatrix} N_{1,\xi} & N_{2,\xi} & \dots & N_{n^e,\xi} \end{bmatrix}} \underset{n^e\times 3}{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_{n^e} & y_{n^e} & z_{n^e} \end{bmatrix}} \tag{A.10}$$

A.4 Tensor calculus

1. First derivative of scalar with respect to vector
   The first derivative of a scalar f with respect to a vector $\vec{x}$ is called the gradient of f.

$$\frac{d}{d\vec{x}} f = \langle \begin{array}{cccc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{array} \rangle^{\mathrm{T}} \tag{A.11}$$

2. Second derivative of scalar with respect to vector
   The second derivative of a scalar f with respect to a vector $\vec{x}$ is called the Hessian matrix of f.

$$\frac{d^2}{d\vec{x}^2} f = \begin{bmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial {x_2}^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial {x_n}^2} \end{bmatrix} \tag{A.12}$$

3. First derivative of vector with respect to vector
   The first derivative of a vector $\vec{f}$ with respect to another vector $\vec{x}$ is called the Jacobian of $\vec{f}$.

$$\frac{d}{d\vec{x}}\vec{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{A.13}$$

4. Gradient of vector in Cartesian coordinates

$$\nabla\vec{a} = \begin{bmatrix} a_{1,x} & a_{1,y} & a_{1,z} \\ a_{2,x} & a_{2,y} & a_{2,z} \\ a_{3,x} & a_{3,y} & a_{3,z} \end{bmatrix} \tag{A.14}$$

5. Divergence of tensor in Cartesian coordinates

$$\nabla \cdot \underline{\underline{A}} = \begin{Bmatrix} A_{11,x} + A_{12,y} + A_{13,z} \\ A_{21,x} + A_{22,y} + A_{23,z} \\ A_{31,x} + A_{32,y} + A_{33,z} \end{Bmatrix} \tag{A.15}$$

6. Outer product or Dyadic multiplication or Tensor product of two vectors

$$\vec{a} \otimes \vec{b} = \{a\}\langle b\rangle \tag{A.16}$$

7. Gradient of the product of a scalar and a vector

$$\nabla\left(a\vec{b}\right) = \vec{b} \otimes \nabla a + a\nabla\vec{b} \tag{A.17}$$

8. Divergence of outer product of two vectors

$$\begin{aligned} \nabla \cdot \left(\vec{a} \otimes \vec{b}\right) &= \vec{a}\left(\nabla \cdot \vec{b}\right) + \nabla\vec{a} \cdot \vec{b} \\ &= \vec{a}\left(\nabla \cdot \vec{b}\right) + \left(\vec{b} \cdot \nabla\right)\vec{a} \end{aligned} \tag{A.18}$$

9. Outer product of two vector, then dot product with a vector

$$\left(\vec{a} \otimes \vec{b}\right) \cdot \vec{c} = \left(\vec{b} \cdot \vec{c}\right)\vec{a} \tag{A.19}$$

10. Double-dot product of two tensors

$$\underline{\underline{A}} : \underline{\underline{B}} = A_{ij}B_{ij} = tr\left(\underline{\underline{A}}^{\mathrm{T}} \otimes \underline{\underline{B}}\right) = tr\left(A_{ki}B_{kj}\right) = \underline{\underline{B}} : \underline{\underline{A}} \tag{A.20}$$

    If $\underline{\underline{B}}$ is symmetric, i.e. $B_{ij} = B_{ji}$, we have

$$\underline{\underline{A}}^{\mathrm{T}} : \underline{\underline{B}} = A_{ji}B_{ij} = A_{ji}B_{ji} = \underline{\underline{A}} : \underline{\underline{B}} \tag{A.21}$$

11. Dot product of a tensor with a vector
    In general, $\underline{\underline{A}} \cdot \vec{b} \neq \vec{b} \cdot \underline{\underline{A}}$, otherwise the tensor $\underline{\underline{A}}$ is symmetric.

12. $\vec{a} \cdot \underline{\underline{B}} \cdot \vec{c} = \vec{a} \cdot \left(\underline{\underline{B}} \cdot \vec{c}\right)$

Appendix B

Solution of the squared distance gradient function

For the three-dimensional case, the variation of Eq. 2.51 or Eq. 2.52 can be written as

$$
\begin{aligned}
&\Delta\vec{R} \\
&= \nabla(\frac{d\vec{x}}{d\vec{\xi}}\Delta\vec{\xi})\cdot(\vec{x}-\vec{x}_s) + \nabla\vec{x}\cdot(\frac{d\vec{x}}{d\vec{\xi}}\Delta\vec{\xi}) \\
&= \nabla\left(\begin{bmatrix} \frac{\partial x}{\partial\xi} & \frac{\partial x}{\partial\eta} \\ \frac{\partial y}{\partial\xi} & \frac{\partial y}{\partial\eta} \\ \frac{\partial z}{\partial\xi} & \frac{\partial z}{\partial\eta} \end{bmatrix}\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix}\right)\begin{Bmatrix}x-x_s\\ y-y_s\\ z-z_s\end{Bmatrix} + \begin{bmatrix} \frac{\partial x}{\partial\xi} & \frac{\partial y}{\partial\xi} & \frac{\partial z}{\partial\xi} \\ \frac{\partial x}{\partial\eta} & \frac{\partial y}{\partial\eta} & \frac{\partial z}{\partial\eta} \end{bmatrix}\left(\begin{bmatrix} \frac{\partial x}{\partial\xi} & \frac{\partial x}{\partial\eta} \\ \frac{\partial y}{\partial\xi} & \frac{\partial y}{\partial\eta} \\ \frac{\partial z}{\partial\xi} & \frac{\partial z}{\partial\eta} \end{bmatrix}\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix}\right) \\
&= \begin{bmatrix} \frac{\partial}{\partial\xi}\left(\frac{\partial x}{\partial\xi}\Delta\xi+\frac{\partial x}{\partial\eta}\Delta\eta\right) & \frac{\partial}{\partial\eta}\left(\frac{\partial x}{\partial\xi}\Delta\xi+\frac{\partial x}{\partial\eta}\Delta\eta\right) \\ \frac{\partial}{\partial\xi}\left(\frac{\partial y}{\partial\xi}\Delta\xi+\frac{\partial y}{\partial\eta}\Delta\eta\right) & \frac{\partial}{\partial\eta}\left(\frac{\partial y}{\partial\xi}\Delta\xi+\frac{\partial y}{\partial\eta}\Delta\eta\right) \\ \frac{\partial}{\partial\xi}\left(\frac{\partial z}{\partial\xi}\Delta\xi+\frac{\partial z}{\partial\eta}\Delta\eta\right) & \frac{\partial}{\partial\eta}\left(\frac{\partial z}{\partial\xi}\Delta\xi+\frac{\partial z}{\partial\eta}\Delta\eta\right) \end{bmatrix}^{\mathrm{T}}\begin{Bmatrix}x-x_s\\ y-y_s\\ z-z_s\end{Bmatrix} \\
&\quad + \begin{bmatrix} \left(\frac{\partial x}{\partial\xi}\right)^2+\left(\frac{\partial y}{\partial\xi}\right)^2+\left(\frac{\partial z}{\partial\xi}\right)^2 & \frac{\partial x}{\partial\xi}\frac{\partial x}{\partial\eta}+\frac{\partial y}{\partial\xi}\frac{\partial y}{\partial\eta}+\frac{\partial z}{\partial\xi}\frac{\partial z}{\partial\eta} \\ \frac{\partial x}{\partial\eta}\frac{\partial x}{\partial\xi}+\frac{\partial y}{\partial\eta}\frac{\partial y}{\partial\xi}+\frac{\partial z}{\partial\eta}\frac{\partial z}{\partial\xi} & \left(\frac{\partial x}{\partial\eta}\right)^2+\left(\frac{\partial y}{\partial\eta}\right)^2+\left(\frac{\partial z}{\partial\eta}\right)^2 \end{bmatrix}\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix} \qquad \text{(B.1a)} \\
&= \begin{Bmatrix} \left(\frac{\partial^2 x}{\partial\xi^2}\Delta\xi+\frac{\partial^2 x}{\partial\eta\partial\xi}\Delta\eta\right)(x-x_s)+\left(\frac{\partial^2 y}{\partial\xi^2}\Delta\xi+\frac{\partial^2 y}{\partial\eta\partial\xi}\Delta\eta\right)(y-y_s)+\left(\frac{\partial^2 z}{\partial\xi^2}\Delta\xi+\frac{\partial^2 z}{\partial\eta\partial\xi}\Delta\eta\right)(z-z_s) \\ \left(\frac{\partial^2 x}{\partial\xi\partial\eta}\Delta\xi+\frac{\partial^2 x}{\partial\eta^2}\Delta\eta\right)(x-x_s)+\left(\frac{\partial^2 y}{\partial\xi\partial\eta}\Delta\xi+\frac{\partial^2 y}{\partial\eta^2}\Delta\eta\right)(y-y_s)+\left(\frac{\partial^2 z}{\partial\xi\partial\eta}\Delta\xi+\frac{\partial^2 z}{\partial\eta^2}\Delta\eta\right)(z-z_s) \end{Bmatrix} \\
&\quad + \begin{bmatrix} \left(\frac{\partial x}{\partial\xi}\right)^2+\left(\frac{\partial y}{\partial\xi}\right)^2+\left(\frac{\partial z}{\partial\xi}\right)^2 & \frac{\partial x}{\partial\xi}\frac{\partial x}{\partial\eta}+\frac{\partial y}{\partial\xi}\frac{\partial y}{\partial\eta}+\frac{\partial z}{\partial\xi}\frac{\partial z}{\partial\eta} \\ \frac{\partial x}{\partial\eta}\frac{\partial x}{\partial\xi}+\frac{\partial y}{\partial\eta}\frac{\partial y}{\partial\xi}+\frac{\partial z}{\partial\eta}\frac{\partial z}{\partial\xi} & \left(\frac{\partial x}{\partial\eta}\right)^2+\left(\frac{\partial y}{\partial\eta}\right)^2+\left(\frac{\partial z}{\partial\eta}\right)^2 \end{bmatrix}\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix} \\
&= \begin{bmatrix} \frac{\partial^2 x}{\partial\xi^2}(x-x_s)+\frac{\partial^2 y}{\partial\xi^2}(y-y_s)+\frac{\partial^2 z}{\partial\xi^2}(z-z_s) & \frac{\partial^2 x}{\partial\eta\partial\xi}(x-x_s)+\frac{\partial^2 y}{\partial\eta\partial\xi}(y-y_s)+\frac{\partial^2 z}{\partial\eta\partial\xi}(z-z_s) \\ \frac{\partial^2 x}{\partial\xi\partial\eta}(x-x_s)+\frac{\partial^2 y}{\partial\xi\partial\eta}(y-y_s)+\frac{\partial^2 z}{\partial\xi\partial\eta}(z-z_s) & \frac{\partial^2 x}{\partial\eta^2}(x-x_s)+\frac{\partial^2 y}{\partial\eta^2}(y-y_s)+\frac{\partial^2 z}{\partial\eta^2}(z-z_s) \end{bmatrix}
\end{aligned}
$$

$$+\left[\begin{array}{cc}\left(\frac{\partial x}{\partial \xi}\right)^2+\left(\frac{\partial y}{\partial \xi}\right)^2+\left(\frac{\partial z}{\partial \xi}\right)^2 & \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta}+\frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial z}{\partial \xi}\frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \xi}+\frac{\partial y}{\partial \eta}\frac{\partial y}{\partial \xi}+\frac{\partial z}{\partial \eta}\frac{\partial z}{\partial \xi} & \left(\frac{\partial x}{\partial \eta}\right)^2+\left(\frac{\partial y}{\partial \eta}\right)^2+\left(\frac{\partial z}{\partial \eta}\right)^2\end{array}\right]\Bigg)\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix}$$

The final increment form can be written as

$$\begin{aligned}0=&\begin{Bmatrix}\frac{\partial x}{\partial \xi}(x-x_s)+\frac{\partial y}{\partial \xi}(y-y_s)+\frac{\partial z}{\partial \xi}(z-z_s)\\ \frac{\partial x}{\partial \eta}(x-x_s)+\frac{\partial y}{\partial \eta}(y-y_s)+\frac{\partial z}{\partial \eta}(z-z_s)\end{Bmatrix}^{i-1}+\Bigg(\begin{bmatrix}\left(\frac{\partial x}{\partial \xi}\right)^2+\left(\frac{\partial y}{\partial \xi}\right)^2+\left(\frac{\partial z}{\partial \xi}\right)^2 & \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta}+\frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial z}{\partial \xi}\frac{\partial z}{\partial \eta}\\ \frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \xi}+\frac{\partial y}{\partial \eta}\frac{\partial y}{\partial \xi}+\frac{\partial z}{\partial \eta}\frac{\partial z}{\partial \xi} & \left(\frac{\partial x}{\partial \eta}\right)^2+\left(\frac{\partial y}{\partial \eta}\right)^2+\left(\frac{\partial z}{\partial \eta}\right)^2\end{bmatrix}\\ &+\begin{bmatrix}\frac{\partial^2 x}{\partial \xi^2}(x-x_s)+\frac{\partial^2 y}{\partial \xi^2}(y-y_s)+\frac{\partial^2 z}{\partial \xi^2}(z-z_s) & \frac{\partial^2 x}{\partial \eta \partial \xi}(x-x_s)+\frac{\partial^2 y}{\partial \eta \partial \xi}(y-y_s)+\frac{\partial^2 z}{\partial \eta \partial \xi}(z-z_s)\\ \frac{\partial^2 x}{\partial \xi \partial \eta}(x-x_s)+\frac{\partial^2 y}{\partial \xi \partial \eta}(y-y_s)+\frac{\partial^2 z}{\partial \xi \partial \eta}(z-z_s) & \frac{\partial^2 x}{\partial \eta^2}(x-x_s)+\frac{\partial^2 y}{\partial \eta^2}(y-y_s)+\frac{\partial^2 z}{\partial \eta^2}(z-z_s)\end{bmatrix}\Bigg)^{i-1}\begin{Bmatrix}\Delta\xi\\ \Delta\eta\end{Bmatrix}\end{aligned} \tag{B.2}$$

For the two-dimensional case, the variation of Eq. 2.51 or Eq. 2.52 can be written as

$$\begin{aligned}&\Delta R\\ &=\nabla\left(\frac{d\vec{x}}{d\xi}\Delta\xi\right)\cdot(\vec{x}-\vec{x}_s)+\nabla\vec{x}\cdot\left(\frac{d\vec{x}}{d\xi}\Delta\xi\right)\\ &=\begin{Bmatrix}\frac{\partial}{\partial \xi}\left(\frac{\partial x}{\partial \xi}\Delta\xi\right)\\ \frac{\partial}{\partial \xi}\left(\frac{\partial y}{\partial \xi}\Delta\xi\right)\\ \frac{\partial}{\partial \xi}\left(\frac{\partial z}{\partial \xi}\Delta\xi\right)\end{Bmatrix}^{\mathrm{T}}\begin{Bmatrix}x-x_s\\ y-y_s\\ z-z_s\end{Bmatrix}+\left\langle\begin{array}{ccc}\frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi}\end{array}\right\rangle\begin{Bmatrix}\frac{\partial x}{\partial \xi}\Delta\xi\\ \frac{\partial y}{\partial \xi}\Delta\xi\\ \frac{\partial z}{\partial \xi}\Delta\xi\end{Bmatrix}\\ &=\left(\frac{\partial^2 x}{\partial \xi^2}(x-x_s)+\frac{\partial^2 y}{\partial \xi^2}(y-y_s)+\frac{\partial^2 z}{\partial \xi^2}(z-z_s)+\left(\frac{\partial x}{\partial \xi}\right)^2+\left(\frac{\partial y}{\partial \xi}\right)^2+\left(\frac{\partial z}{\partial \xi}\right)^2\right)\Delta\xi\end{aligned} \tag{B.3}$$

The final increment form reads

$$\begin{aligned}0=&\left(\frac{\partial x}{\partial \xi}(x-x_s)+\frac{\partial y}{\partial \xi}(y-y_s)+\frac{\partial z}{\partial \xi}(z-z_s)\right)^{i-1}\\ &+\left(\frac{\partial^2 x}{\partial \xi^2}(x-x_s)+\frac{\partial^2 y}{\partial \xi^2}(y-y_s)+\frac{\partial^2 z}{\partial \xi^2}(z-z_s)+\left(\frac{\partial x}{\partial \xi}\right)^2+\left(\frac{\partial y}{\partial \xi}\right)^2+\left(\frac{\partial z}{\partial \xi}\right)^2\right)^{i-1}\Delta\xi\end{aligned} \tag{B.4}$$

Appendix C

Solution contribution analysis based on each term

C.1 Naming conventions of each term

$$\underbrace{\rho\frac{\partial \vec{u}}{\partial t}}_{\text{NS time term}} + \underbrace{\rho\left(\vec{u}\cdot\nabla\right)\vec{u}}_{\text{NS convection term}} + \underbrace{-\nabla\cdot\left(2\mu\underline{\underline{D}}\right)}_{\text{NS viscous term}} + \underbrace{\nabla p}_{\text{NS pressure term}} + \underbrace{\rho\alpha_p\left(T-T_{ref}\right)\vec{g}}_{\text{NS buoyancy term}} + \underbrace{-\rho\vec{g}}_{\text{NS gravity term}} = 0 \tag{C.1a}$$

$$\underbrace{\nabla\cdot\vec{u}}_{\text{NS solenoid term}} = 0 \tag{C.1b}$$

$$\underbrace{\rho\frac{\partial \omega}{\partial t}}_{\text{ST time term}} + \underbrace{\vec{u}\cdot\rho\nabla\omega}_{\text{ST convection term}} - \underbrace{\nabla\cdot\left(D\rho\nabla\omega\right)}_{\text{ST diffusion term}} + \underbrace{\omega\frac{\partial \rho}{\partial t}}_{\text{ST time consv term}} + \underbrace{\vec{u}\cdot\omega\nabla\rho}_{\text{ST convec consv term}} = 0 \tag{C.2a}$$

$$\underbrace{\rho C_p\frac{\partial T}{\partial t}}_{\text{Heat capacity term}} + \underbrace{\rho C_p\vec{u}\cdot\nabla T}_{\text{Heat convection term}} + \underbrace{-\nabla\cdot\left(k\nabla T\right)}_{\text{Heat conduction term}} + \underbrace{-s}_{\text{Heat source term}} = 0 \tag{C.3a}$$

$$\underbrace{T-T_m}_{\text{Heat penalty term}} = 0 \tag{C.3b}$$

$$\underbrace{\frac{\partial \phi}{\partial t}}_{\text{LS time term}} + \underbrace{F\|\nabla\phi\|}_{\text{LS convection term}} = 0 \tag{C.4a}$$

$$\underbrace{\sum_e\int_{\Omega^e}\left(F\nabla\delta\phi\cdot\frac{\nabla\phi}{\|\nabla\phi\|}\right)\tau^e\left(\frac{\partial\phi}{\partial t}+F\|\nabla\phi\|\right)d\Omega}_{\text{LS GLS term}} \tag{C.4b}$$

$$\underbrace{sign(\phi)\nabla\phi\cdot\nabla F}_{\text{Spread ortho term}} = 0 \tag{C.5a}$$

$$\underbrace{\sum_e\int_{\Omega^e}\left(\nabla\delta F\cdot\nabla\phi\right)\tau^e\left(\nabla\phi\cdot\nabla F\right)d\Omega}_{\text{Spread GLS term}} \tag{C.5b}$$

C.2 NS time term

$$\rho\frac{\partial\vec{u}}{\partial t} \tag{C.6}$$

The weak form is

$$W_{\text{NS time}} = \int_{\Omega}\delta\vec{u}\cdot\rho\frac{\partial\vec{u}}{\partial t}d\Omega \tag{C.7}$$

After time discretization by backward Euler method, it becomes

$$W_{\text{NS time}} = \int_{\Omega}\delta\vec{u}\cdot\rho\frac{1}{\Delta t}\left(\vec{u}-\vec{u}^{t-\Delta t}\right)d\Omega \tag{C.8}$$

On one element, Eq. C.8 can be discretized in space as

$$\begin{aligned} W^{e}_{\text{NS time}} &= \langle\delta u\rangle\int_{\Omega^e}[N]^{\mathrm{T}}\rho\frac{1}{\Delta t}[N]\left(\{u\}-\{u\}^{t-\Delta t}\right)d\Omega \\ &= \langle\delta u\rangle\int_{\Omega^e}[N]^{\mathrm{T}}\rho\frac{1}{\Delta t}[N]\,d\Omega\{u\} - \langle\delta u\rangle\int_{\Omega^e}[N]^{\mathrm{T}}\rho\frac{1}{\Delta t}[N]\,\{u\}^{t-\Delta t}d\Omega \end{aligned} \tag{C.9}$$

where

$$[N] = \begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 & \cdots & N_{n^e} & 0 & 0 & M_1 & 0 & 0 & M_2 & 0 & 0 & \cdots & M_{n^e} & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & \cdots & N_{n^e} & 0 & 0 & M_1 & 0 & 0 & M_2 & 0 & 0 & \cdots & M_{n^e} & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 & 0 & 0 & \cdots & N_{n^e} & 0 & 0 & M_1 & 0 & 0 & M_2 & 0 & 0 & \cdots & M_{n^e} \end{bmatrix} \tag{C.10}$$

- Contribution to the coefficient matrix by substitution method

$$[K]_{\text{NS time}} = -\sum_{e}\int_{\Omega^e}[N]^{\mathrm{T}}\rho\frac{1}{\Delta t}[N]\,d\Omega \tag{C.11}$$

- Contribution to the tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS time}} = -\sum_{e}\int_{\Omega^e}[N]^{\mathrm{T}}\rho\frac{1}{\Delta t}[N]\,d\Omega \tag{C.12}$$

- Contribution to the residual vector

$$\{R\}_{\text{NS time}} = \sum_{e}\int_{\Omega^e}[N]^{\mathrm{T}}\rho\,[N]\,\frac{\{u\}^{i-1}-\{u\}^{t-\Delta t}}{\Delta t}d\Omega \tag{C.13}$$

- Contribution to the tangential coefficient matrix, corresponding to the coupling with mass transport problem, by Newton-Raphson method

$$\left[K_{t,\vec{u}\omega}\right]_{\text{NS time}} = -\sum_{e}\int_{\Omega^e}[N]^{\mathrm{T}}(\rho')^{i-1}[N]\,\frac{\{u\}^{i-1}-\{u\}^{t-\Delta t}}{\Delta t}\langle N\rangle d\Omega \tag{C.14}$$

C.3 NS convection term

$$\rho\left(\vec{u}\cdot\nabla\right)\vec{u} \tag{C.15}$$

The corresponding weak formulation is

$$\begin{aligned} W_{\text{NS convection}} &= \int_{\Omega}\delta\vec{u}\cdot\rho\left(\vec{u}\cdot\nabla\right)\vec{u}d\Omega \\ &= \int_{\Omega}\delta u\rho\Big(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\Big)u \\ &\quad +\delta v\rho\Big(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\Big)v \\ &\quad +\delta w\rho\Big(u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\Big)wd\Omega \end{aligned} \tag{C.16}$$

where u,v and w are the scalar axial components of $\vec{u}$ in Cartesian coordinate system. The spatial discretization in matrix form on one element reads

$$\begin{aligned} W^{e}_{\text{NS convection}} &= \int_{\Omega^e}\langle\delta\hat{u}\rangle\{N\}\rho\langle N\rangle\Big(\{\hat{u}\}\frac{\partial}{\partial x}+\{\hat{v}\}\frac{\partial}{\partial y}+\{\hat{w}\}\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{u}\} \\ &\quad +\langle\delta\hat{v}\rangle\{N\}\rho\langle N\rangle\Big(\{\hat{u}\}\frac{\partial}{\partial x}+\{\hat{v}\}\frac{\partial}{\partial y}+\{\hat{w}\}\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{v}\} \\ &\quad +\langle\delta\hat{w}\rangle\{N\}\rho\langle N\rangle\Big(\{\hat{u}\}\frac{\partial}{\partial x}+\{\hat{v}\}\frac{\partial}{\partial y}+\{\hat{w}\}\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{w}\}d\Omega \\ &= \langle\delta\hat{u}\rangle\int_{\Omega^e}\Big[\hat{C}\Big]d\Omega\{\hat{u}\}+\langle\delta\hat{v}\rangle\int_{\Omega^e}\Big[\hat{C}\Big]d\Omega\{\hat{v}\}+\langle\delta\hat{w}\rangle\int_{\Omega^e}\Big[\hat{C}\Big]d\Omega\{\hat{w}\} \\ \text{or} &= \langle\delta u\rangle\int_{\Omega^e}\rho\left[C\right]d\Omega\left\{u\right\} \end{aligned} \tag{C.17}$$

where

$$[C] = [N]^{\mathrm{T}}\Big(\left[N\right]\left\{u\right\}\Big)^{\mathrm{T}}\begin{Bmatrix}\partial_x\\ \partial_y\\ \partial_z\end{Bmatrix}[N] \tag{C.18a}$$

$$\Big[\hat{C}\Big] = \{N\}\rho\langle N\rangle\Big(\{\hat{u}\}\frac{\partial}{\partial x}+\{\hat{v}\}\frac{\partial}{\partial y}+\{\hat{w}\}\frac{\partial}{\partial z}\Big)\langle N\rangle \tag{C.18b}$$

$$\{\hat{u}\} = \langle u_1 \quad u_2 \quad \cdots \quad u_{n^e} \quad \breve{u}_1 \quad \breve{u}_2 \quad \cdots \breve{u}_{n^e} \rangle^{\mathrm{T}} \tag{C.18c}$$

$$\{\hat{v}\} = \langle v_1 \quad v_2 \quad \cdots \quad v_{n^e} \quad \breve{v}_1 \quad \breve{v}_2 \quad \cdots \breve{v}_{n^e} \rangle^{\mathrm{T}} \tag{C.18d}$$

$$\{\hat{w}\} = \langle w_1 \quad w_2 \quad \cdots \quad w_{n^e} \quad \breve{w}_1 \quad \breve{w}_2 \quad \cdots \breve{w}_{n^e} \rangle^{\mathrm{T}} \tag{C.18e}$$

- Contribution to coefficient matrix by substitution method
  - Indexed by $\big(\langle\delta\hat{u}\rangle, \{\hat{u}\}\big)$

$$[K]_{\text{NS convection},uu} = -\sum_{e}\int_{\Omega^e}\Big[\hat{C}\Big]^{i-1}d\Omega \tag{C.19}$$

– Indexed by $\big(\langle\delta\hat{v}\rangle, \{\hat{v}\}\big)$

$$[K]_{\text{NS convection},vv} = -\sum_e \int_{\Omega^e} \left[\hat{C}\right]^{i-1} d\Omega \tag{C.20}$$

– Indexed by $\big(\langle\delta\hat{w}\rangle, \{\hat{w}\}\big)$

$$[K]_{\text{NS convection},ww} = -\sum_e \int_{\Omega^e} \left[\hat{C}\right]^{i-1} d\Omega \tag{C.21}$$

where $\left[\hat{C}\right]^{i-1} = \{N\}\rho\langle N\rangle\Big(\{\hat{u}\}^{i-1}\frac{\partial}{\partial x} + \{\hat{v}\}^{i-1}\frac{\partial}{\partial y} + \{\hat{w}\}^{i-1}\frac{\partial}{\partial z}\Big)\langle N\rangle$

- Contribution to tangential coefficient matrix by Newton-Raphson method
  The variation of weak form, $\Delta W_{\text{NS convection}}$, with respect to its components gives

$$\begin{aligned}
&\Delta W_{\text{NS convection}} = \\
&\int_\Omega \delta u\rho\Big(u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\Big)\Delta u + \underbrace{\delta u\rho\Big(\frac{\partial}{\partial x}\Big)u\Delta u + \delta v\rho\Big(\frac{\partial}{\partial x}\Big)v\Delta u + \delta w\rho\Big(\frac{\partial}{\partial x}\Big)w\Delta u}_{[N]^{\mathrm{T}}\rho[N]_x\{u\}\big\langle\ \langle N\rangle_i \quad 0 \quad 0\ \big\rangle\{u\}} \\
&+ \delta v\rho\Big(u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\Big)\Delta v + \underbrace{\delta u\rho\Big(\frac{\partial}{\partial y}\Big)u\Delta v + \delta v\rho\Big(\frac{\partial}{\partial y}\Big)v\Delta v + \delta w\rho\Big(\frac{\partial}{\partial y}\Big)w\Delta v}_{[N]^{\mathrm{T}}\rho[N]_y\{u\}\big\langle\ 0 \quad \langle N\rangle_i \quad 0\ \big\rangle\{u\}} \\
&+ \delta w\rho\Big(u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\Big)\Delta w + \underbrace{\delta u\rho\Big(\frac{\partial}{\partial z}\Big)u\Delta w + \delta v\rho\Big(\frac{\partial}{\partial z}\Big)v\Delta w + \delta w\rho\Big(\frac{\partial}{\partial z}\Big)w\Delta w}_{[N]^{\mathrm{T}}\rho[N]_z\{u\}\big\langle\ 0 \quad 0 \quad \langle N\rangle_i\ \big\rangle\{u\}}
\end{aligned} \tag{C.22}$$

  The following part shows the contribution to each matrix block.

– Indexed by $\big(\langle\delta\hat{u}\rangle, \{\hat{u}\}\big)$

$$[K_t]_{\text{NS convection},uu} = -\sum_e \int_{\Omega^e} \left[\hat{C}\right]^{i-1} + \{N\}\rho\Big(\frac{\partial}{\partial x}\Big)\langle N\rangle\{\hat{u}\}^{i-1}\langle N\rangle d\Omega \tag{C.23}$$

– Indexed by $\big(\langle\delta\hat{v}\rangle, \{\hat{v}\}\big)$

$$[K_t]_{\text{NS convection},vv} = -\sum_e \int_{\Omega^e} \left[\hat{C}\right]^{i-1} + \{N\}\rho\Big(\frac{\partial}{\partial y}\Big)\langle N\rangle\{\hat{v}\}^{i-1}\langle N\rangle d\Omega \tag{C.24}$$

– Indexed by $\big(\langle\delta\hat{w}\rangle, \{\hat{w}\}\big)$

$$[K_t]_{\text{NS convection},ww} = -\sum_e \int_{\Omega^e} \left[\hat{C}\right]^{i-1} + \{N\}\rho\Big(\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{w}\}^{i-1}\langle N\rangle d\Omega \tag{C.25}$$

– Indexed by $\big(\langle\delta\hat{u}\rangle, \{\hat{v}\}\big)$

$$[K_t]_{\text{NS convection},uv} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial y}\Big)\langle N\rangle\{\hat{u}\}^{i-1}\langle N\rangle d\Omega \tag{C.26}$$

- Indexed by $\big(\langle\delta\hat{u}\rangle, \{\hat{w}\}\big)$

$$[K_t]_{\text{NS convection},uw} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{u}\}^{i-1}\langle N\rangle d\Omega \tag{C.27}$$

- Indexed by $\big(\langle\delta\hat{v}\rangle, \{\hat{u}\}\big)$

$$[K_t]_{\text{NS convection},vu} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial x}\Big)\langle N\rangle\{\hat{v}\}^{i-1}\langle N\rangle d\Omega \tag{C.28}$$

- Indexed by $\big(\langle\delta\hat{v}\rangle, \{\hat{w}\}\big)$

$$[K_t]_{\text{NS convection},vw} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial z}\Big)\langle N\rangle\{\hat{v}\}^{i-1}\langle N\rangle d\Omega \tag{C.29}$$

- Indexed by $\big(\langle\delta\hat{w}\rangle, \{\hat{u}\}\big)$

$$[K_t]_{\text{NS convection},wu} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial x}\Big)\langle N\rangle\{\hat{w}\}^{i-1}\langle N\rangle d\Omega \tag{C.30}$$

- Indexed by $\big(\langle\delta\hat{w}\rangle, \{\hat{v}\}\big)$

$$[K_t]_{\text{NS convection},wv} = -\sum_e \int_{\Omega^e} \{N\}\rho\Big(\frac{\partial}{\partial y}\Big)\langle N\rangle\{\hat{w}\}^{i-1}\langle N\rangle d\Omega \tag{C.31}$$

- Contribution to residual vector
  - Indexed by $\langle\delta\hat{u}\rangle$

$$\{R\}_{\text{NS convection},u} = \sum_e \int_{\Omega^e} \Big[\hat{C}\Big]^{i-1} d\Omega\{\hat{u}\}^{i-1} \tag{C.32}$$

  - Indexed by $\langle\delta\hat{v}\rangle$

$$\{R\}_{\text{NS convection},v} = \sum_e \int_{\Omega^e} \Big[\hat{C}\Big]^{i-1} d\Omega\{\hat{v}\}^{i-1} \tag{C.33}$$

  - Indexed by $\langle\delta\hat{w}\rangle$

$$\{R\}_{\text{NS convection},w} = \sum_e \int_{\Omega^e} \Big[\hat{C}\Big]^{i-1} d\Omega\{\hat{w}\}^{i-1} \tag{C.34}$$

- Contribution to tangential coefficient matrix, corresponding to the coupling with mass transport problem, by Newton-Raphson method

$$\big[K_{t,\vec{u}\omega}\big]_{\text{NS convection}} = -\sum_e \int_{\Omega^e} (\rho')^{i-1}[C]^{i-1}\{u\}^{i-1}\langle N\rangle d\Omega \tag{C.35}$$

In the practical implementation, the corresponding matrices and vectors are expanded and nested to adjust to the following solution structure based on one element with n^e nodes:

$$\{u\} = \langle u_1 \quad v_1 \quad w_1 \quad u_2 \quad v_2 \quad w_2 \quad \cdots \quad u_{n^e} \quad v_{n^e} \quad w_{n^e} \quad \breve{u}_1 \quad \breve{v}_1 \quad \breve{w}_1 \quad \breve{u}_2 \quad \breve{v}_2 \quad \breve{w}_2 \quad \cdots \quad \breve{u}_{n^e} \quad \breve{v}_{n^e} \quad \breve{w}_{n^e} \rangle^{\mathrm{T}} \tag{C.36}$$

An Illustration

Based on the linear triangle element as shown in Fig. C.1, Table C.1 shows the tangential matrix related to the convection term.

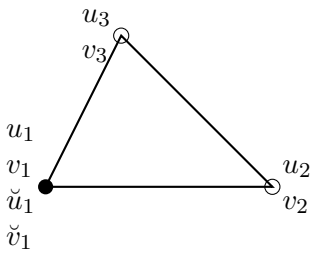


Figure C.1 – One sample triangle linear element used to illustrate the matrices and vectors related to the NS convection term; Node 1 is enriched.

Table C.1 – The tangential matrix (integral operator omitted) related to NS convection term

	u_1	v_1	u_2	v_2	u_3	v_3	$\breve{u}_1$	$\breve{v}_1$
δu_1	$\hat{C}_{11}+$ $\rho N_1 N_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_1 N_1\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{12}+$ $\rho N_1 N_2\langle N_x\rangle\{\hat{u}\}$	$\rho N_1 N_2\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{13}+$ $\rho N_1 N_3\langle N_x\rangle\{\hat{u}\}$	$\rho N_1 N_3\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{14}+$ $\rho N_1 M_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_1 M_1\langle N_y\rangle\{\hat{u}\}$
δv_1	$\rho N_1 N_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{11}+$ $\rho N_1 N_1\langle N_y\rangle\{\hat{v}\}$	$\rho N_1 N_2\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{12}+$ $\rho N_1 N_2\langle N_y\rangle\{\hat{v}\}$	$\rho N_1 N_3\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{13}+$ $\rho N_1 N_3\langle N_y\rangle\{\hat{v}\}$	$\rho N_1 M_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{14}+$ $\rho N_1 M_1\langle N_y\rangle\{\hat{v}\}$
δu_2	$\hat{C}_{21}+$ $\rho N_2 N_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_2 N_1\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{22}+$ $\rho N_2 N_2\langle N_x\rangle\{\hat{u}\}$	$\rho N_2 N_2\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{23}+$ $\rho N_2 N_3\langle N_x\rangle\{\hat{u}\}$	$\rho N_2 N_3\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{24}+$ $\rho N_2 M_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_2 M_1\langle N_y\rangle\{\hat{u}\}$
δv_2	$\rho N_2 N_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{21}+$ $\rho N_2 N_1\langle N_y\rangle\{\hat{v}\}$	$\rho N_2 N_2\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{22}+$ $\rho N_2 N_2\langle N_y\rangle\{\hat{v}\}$	$\rho N_2 N_3\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{23}+$ $\rho N_2 N_3\langle N_y\rangle\{\hat{v}\}$	$\rho N_2 M_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{24}+$ $\rho N_2 M_1\langle N_y\rangle\{\hat{v}\}$
δu_3	$\hat{C}_{31}+$ $\rho N_3 N_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_3 N_1\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{32}+$ $\rho N_3 N_2\langle N_x\rangle\{\hat{u}\}$	$\rho N_3 N_2\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{33}+$ $\rho N_3 N_3\langle N_x\rangle\{\hat{u}\}$	$\rho N_3 N_3\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{34}+$ $\rho N_3 M_1\langle N_x\rangle\{\hat{u}\}$	$\rho N_3 M_1\langle N_y\rangle\{\hat{u}\}$
δv_3	$\rho N_3 N_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{31}+$ $\rho N_3 N_1\langle N_y\rangle\{\hat{v}\}$	$\rho N_3 N_2\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{32}+$ $\rho N_3 N_2\langle N_y\rangle\{\hat{v}\}$	$\rho N_3 N_3\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{33}+$ $\rho N_3 N_3\langle N_y\rangle\{\hat{v}\}$	$\rho N_3 M_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{34}+$ $\rho N_3 M_1\langle N_y\rangle\{\hat{v}\}$
$\delta \breve{u}_1$	$\hat{C}_{41}+$ $\rho M_1 N_1\langle N_x\rangle\{\hat{u}\}$	$\rho M_1 N_1\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{42}+$ $\rho M_1 N_2\langle N_x\rangle\{\hat{u}\}$	$\rho M_1 N_2\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{43}+$ $\rho M_1 N_3\langle N_x\rangle\{\hat{u}\}$	$\rho M_1 N_3\langle N_y\rangle\{\hat{u}\}$	$\hat{C}_{44}+$ $\rho M_1 M_1\langle N_x\rangle\{\hat{u}\}$	$\rho M_1 M_1\langle N_y\rangle\{\hat{u}\}$
$\delta \breve{v}_1$	$\rho M_1 N_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{41}+$ $\rho M_1 N_1\langle N_y\rangle\{\hat{v}\}$	$\rho M_1 N_2\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{42}+$ $\rho M_1 N_2\langle N_y\rangle\{\hat{v}\}$	$\rho M_1 N_3\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{43}+$ $\rho M_1 N_3\langle N_y\rangle\{\hat{v}\}$	$\rho M_1 M_1\langle N_x\rangle\{\hat{v}\}$	$\hat{C}_{44}+$ $\rho M_1 M_1\langle N_y\rangle\{\hat{v}\}$

C.4 NS viscous term

$$-\nabla\cdot\left(2\mu\underline{\underline{D}}\right) \tag{C.37}$$

The corresponding weak formulation is

$$\begin{aligned}
W_{\text{NS viscous}} &= -\int_{\Omega}\delta\vec{u}\cdot\nabla\cdot\left(2\mu\underline{\underline{D}}\right)d\Omega \\
&= \int_{\Omega}\nabla\delta\vec{u}:2\mu\underline{\underline{D}}-\nabla\cdot\left(\delta\vec{u}\cdot2\mu\underline{\underline{D}}\right)d\Omega \\
&= \int_{\Omega}\frac{1}{2}\left(\nabla\delta\vec{u}+\nabla\delta\vec{u}\right):2\mu\underline{\underline{D}}d\Omega-\oint_{\Gamma}\left(\delta\vec{u}\cdot2\mu\underline{\underline{D}}\right)\cdot\vec{n}d\Gamma \\
&= \int_{\Omega}\frac{1}{2}\left(\nabla\delta\vec{u}+\left(\nabla\delta\vec{u}\right)^{\mathrm{T}}\right):2\mu\underline{\underline{D}}d\Omega-\int_{\Gamma_N}\delta\vec{u}\cdot\left(2\mu\underline{\underline{D}}\cdot\vec{n}\right)d\Gamma \\
&= \int_{\Omega}\delta\underline{\underline{D}}:2\mu\underline{\underline{D}}d\Omega-\int_{\Gamma_N}\delta\vec{u}\cdot\left(2\mu\underline{\underline{D}}\cdot\vec{n}\right)d\Gamma
\end{aligned} \tag{C.38}$$

The boundary term in Eq. C.38 serves as part of the Neumann boundary condition, turn to Section C.9. The shear rate tensor $\underline{\underline{D}}=\frac{1}{2}\left(\nabla\vec{u}+\nabla\vec{u}^{\mathrm{T}}\right)$ can be written in the following fully expanded form:

$$\begin{bmatrix}
\frac{\partial u}{\partial x} & \frac{1}{2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right) \\
\frac{1}{2}\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right) & \frac{\partial v}{\partial y} & \frac{1}{2}\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right) \\
\frac{1}{2}\left(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z}\right) & \frac{1}{2}\left(\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z}\right) & \frac{\partial w}{\partial z}
\end{bmatrix} \tag{C.39}$$

Then, $\underline{\underline{D}}$ can be reconstructed into the column vector as shown below:

$$\left\langle\ \frac{\partial u}{\partial x}\quad\frac{\partial v}{\partial y}\quad\frac{\partial w}{\partial z}\quad\frac{1}{2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\quad\frac{1}{2}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\quad\frac{1}{2}\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)\quad\frac{1}{2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\quad\frac{1}{2}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\quad\frac{1}{2}\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)\ \right\rangle^{\mathrm{T}} \tag{C.40}$$

The spatial discretization of the body term in Eq. C.38 on one element reads

$$W^{e}_{\text{NS viscous}}=\langle\delta u\rangle\int_{\Omega^e}[B]^{\mathrm{T}}\,2\mu\,[B]\,d\Omega\{u\} \tag{C.41}$$

where $[B]$ is named strain rate operator. Let's consider the situation in two dimensions, $[B]$ is constructed as shown in Eq. C.42 based on one element with n^e nodes.

$$
\begin{bmatrix}
N_{1,x} & 0 & 0 & N_{2,x} & 0 & 0 & \cdots & N_{n^e,x} & 0 & 0 & M_{1,x} & 0 & 0 & M_{2,x} & 0 & 0 & \cdots & M_{n^e,x} & 0 & 0 \\
0 & N_{1,y} & 0 & 0 & N_{2,y} & 0 & \cdots & 0 & N_{n^e,y} & 0 & 0 & M_{1,y} & 0 & 0 & M_{2,y} & 0 & \cdots & 0 & M_{n^e,y} & 0 \\
0 & 0 & N_{1,z} & 0 & 0 & N_{2,z} & \cdots & 0 & 0 & N_{n^e,z} & 0 & 0 & M_{1,z} & 0 & 0 & M_{2,z} & \cdots & 0 & 0 & M_{n^e,z} \\
\frac{1}{2}N_{1,y} & \frac{1}{2}N_{1,x} & 0 & \frac{1}{2}N_{2,y} & \frac{1}{2}N_{2,x} & 0 & \cdots & \frac{1}{2}N_{n^e,y} & \frac{1}{2}N_{n^e,x} & 0 & \frac{1}{2}M_{1,y} & \frac{1}{2}M_{1,x} & 0 & \frac{1}{2}M_{2,y} & \frac{1}{2}M_{2,x} & 0 & \cdots & \frac{1}{2}M_{n^e,y} & \frac{1}{2}M_{n^e,x} & 0 \\
\frac{1}{2}N_{1,z} & 0 & \frac{1}{2}N_{1,x} & \frac{1}{2}N_{2,z} & 0 & \frac{1}{2}N_{2,x} & \cdots & \frac{1}{2}N_{n^e,z} & 0 & \frac{1}{2}N_{n^e,x} & \frac{1}{2}M_{1,z} & 0 & \frac{1}{2}M_{1,x} & \frac{1}{2}M_{2,z} & 0 & \frac{1}{2}M_{2,x} & \cdots & \frac{1}{2}M_{n^e,z} & 0 & \frac{1}{2}M_{n^e,x} \\
0 & \frac{1}{2}N_{1,z} & \frac{1}{2}N_{1,y} & 0 & \frac{1}{2}N_{2,z} & \frac{1}{2}N_{2,y} & \cdots & 0 & \frac{1}{2}N_{n^e,z} & \frac{1}{2}N_{n^e,y} & 0 & \frac{1}{2}M_{1,z} & \frac{1}{2}M_{1,y} & 0 & \frac{1}{2}M_{2,z} & \frac{1}{2}M_{2,y} & \cdots & 0 & \frac{1}{2}M_{n^e,z} & \frac{1}{2}M_{n^e,y} \\
\frac{1}{2}N_{1,y} & \frac{1}{2}N_{1,x} & 0 & \frac{1}{2}N_{2,y} & \frac{1}{2}N_{2,x} & 0 & \cdots & \frac{1}{2}N_{n^e,y} & \frac{1}{2}N_{n^e,x} & 0 & \frac{1}{2}M_{1,y} & \frac{1}{2}M_{1,x} & 0 & \frac{1}{2}M_{2,y} & \frac{1}{2}M_{2,x} & 0 & \cdots & \frac{1}{2}M_{n^e,y} & \frac{1}{2}M_{n^e,x} & 0 \\
\frac{1}{2}N_{1,z} & 0 & \frac{1}{2}N_{1,x} & \frac{1}{2}N_{2,z} & 0 & \frac{1}{2}N_{2,x} & \cdots & \frac{1}{2}N_{n^e,z} & 0 & \frac{1}{2}N_{n^e,x} & \frac{1}{2}M_{1,z} & 0 & \frac{1}{2}M_{1,x} & \frac{1}{2}M_{2,z} & 0 & \frac{1}{2}M_{2,x} & \cdots & \frac{1}{2}M_{n^e,z} & 0 & \frac{1}{2}M_{n^e,x} \\
0 & \frac{1}{2}N_{1,z} & \frac{1}{2}N_{1,y} & 0 & \frac{1}{2}N_{2,z} & \frac{1}{2}N_{2,y} & \cdots & 0 & \frac{1}{2}N_{n^e,z} & \frac{1}{2}N_{n^e,y} & 0 & \frac{1}{2}M_{1,z} & \frac{1}{2}M_{1,y} & 0 & \frac{1}{2}M_{2,z} & \frac{1}{2}M_{2,y} & \cdots & 0 & \frac{1}{2}M_{n^e,z} & \frac{1}{2}M_{n^e,y}
\end{bmatrix}
\begin{Bmatrix}
u_1 \\ v_1 \\ w_1 \\ u_2 \\ v_2 \\ w_2 \\ \cdots \\ u_{n^e} \\ v_{n^e} \\ w_{n^e} \\ \breve{u}_1 \\ \breve{v}_1 \\ \breve{w}_1 \\ \breve{u}_2 \\ \breve{v}_2 \\ \breve{w}_2 \\ \cdots \\ \breve{u}_{n^e} \\ \breve{v}_{n^e} \\ \breve{w}_{n^e}
\end{Bmatrix}
\tag{C.42}
$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS viscous}} = -\sum_e \int_{\Omega^e} [B]^{\mathrm{T}}\, 2\mu\, [B]\, d\Omega \tag{C.43}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS viscous}} = -\sum_e \int_{\Omega^e} [B]^{\mathrm{T}}\, 2\mu\, [B]\, d\Omega \tag{C.44}$$

- Contribution to residual vector

$$\{R\}_{\text{NS viscous}} = \sum_e \int_{\Omega^e} [B]^{\mathrm{T}}\, 2\mu\, [B]\, \{u\}^{i-1} d\Omega \tag{C.45}$$

C.5 NS pressure term

$$\nabla p \tag{C.46}$$

The corresponding weak formulation is

$$\begin{aligned} W_{\text{NS pressure}} &= \int_\Omega \delta\vec{u} \cdot \nabla p d\Omega \\ &= -\int_\Omega \nabla \cdot \delta\vec{u}\, p d\Omega + \int_\Omega \nabla \cdot (\delta\vec{u}\, p)\, d\Omega \\ &= -\int_\Omega \nabla \cdot \delta\vec{u}\, p d\Omega + \oint_\Gamma (\delta\vec{u}\, p) \cdot \vec{n} d\Gamma \\ &= -\int_\Omega \nabla \cdot \delta\vec{u}\, p d\Omega + \int_{\Gamma_N} \delta\vec{u} \cdot \left(p\underline{\underline{I}} \cdot \vec{n}\right) d\Gamma \end{aligned} \tag{C.47}$$

The boundary term in Eq. C.47 also serves as part of the Neumann boundary condition, turn to C.9 for details. The spatial discretization of the body term in Eq. C.47 in matrix form on one element reads

$$W^e_{\text{NS pressure}} = -\langle \delta u \rangle \int_{\Omega^e} \{B\}\langle N\rangle d\Omega \{p\} \tag{C.48}$$

where $\{B\} = \langle\ N_{1,x} \quad N_{1,y} \quad N_{2,x} \quad N_{2,y} \quad \cdots \quad N_{9,x} \quad N_{9,y}\ \rangle$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS pressure}} = \sum_e \int_{\Omega^e} \{B\}\langle N\rangle d\Omega \tag{C.49}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS presssure}} = \sum_e \int_{\Omega^e} \{B\}\langle N\rangle d\Omega \tag{C.50}$$

- Contribution to residual vector

$$\{R\}_{\text{NS pressure}} = -\sum_e \int_{\Omega^e} \{B\}\langle N\rangle \{p\}^{i-1} d\Omega \tag{C.51}$$

C.6 NS solenoid term

$$-\nabla\cdot\vec{u} \tag{C.52}$$

The corresponding weak formulation is

$$W_{\text{NS solenoid}} = -\int_{\Omega}\delta p\nabla\cdot\vec{u}d\Omega \tag{C.53}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{NS solenoid}} = -\langle\delta p\rangle\int_{\Omega^e}\{N\}\langle B\rangle d\Omega\{u\} \tag{C.54}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS solenoid}} = \sum_{e}\int_{\Omega^e}\{N\}\langle B\rangle d\Omega \tag{C.55}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS solenoid}} = \sum_{e}\int_{\Omega^e}\{N\}\langle B\rangle d\Omega \tag{C.56}$$

- Contribution to residual vector

$$\{R\}_{\text{NS solenoid}} = -\sum_{e}\int_{\Omega^e}\{N\}\langle B\rangle\{u\}^{i-1}d\Omega \tag{C.57}$$

C.7 NS gravity term

$$-\rho\vec{g} \tag{C.58}$$

The corresponding weak formulation is

$$W_{\text{NS gravity}} = -\int_{\Omega}\delta\vec{u}\cdot\rho\vec{g}d\Omega \tag{C.59}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{NS gravity}} = -\langle\delta u\rangle\int_{\Omega^e}[N]^{\mathrm{T}}\rho\{g\}d\Omega \tag{C.60}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS gravity}} = 0 \tag{C.61}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS gravity}} = 0 \tag{C.62}$$

- Contribution to residual vector

$$\{R\}_{\text{NS gravity}} = -\sum_e \int_{\Omega^e} [N]^{\mathrm{T}} \rho\{g\} d\Omega \tag{C.63}$$

- Contribution to tangential coefficient matrix, corresponding to the coupling with mass transport problem, by Newton-Raphson method

$$\left[K_{t,\vec{u}\omega}\right]_{\text{NS gravity}} = \sum_e \int_{\Omega^e} [N]^{\mathrm{T}} (\rho')^{i-1} \{g\}\langle N\rangle d\Omega \tag{C.64}$$

C.8 NS buoyancy term

$$\rho\alpha_p \left(T - T_{ref}\right) \vec{g} \tag{C.65}$$

The weak form is

$$W_{\text{NS buoyancy}} = \int_\Omega \delta\vec{u} \cdot \rho\alpha_p \left(T - T_{ref}\right) \vec{g} d\Omega \tag{C.66}$$

On one element, it can be discretized as

$$W^e_{\text{NS buoyancy}} = \langle \delta u\rangle \int_{\Omega^e} [N]^{\mathrm{T}} \rho\alpha_p \Big(\langle N\rangle\{T\} - T_{ref}\Big)\{g\} d\Omega \tag{C.67}$$

- Contribution to the coefficient matrix by substitution method

$$[K]_{\text{NS buoyancy}} = 0 \tag{C.68}$$

- Contribution to the tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS buoyancy}} = 0 \tag{C.69}$$

- Contribution to the residual vector

$$\{R\}_{\text{NS buoyancy}} = \sum_e \int_{\Omega^e} [N]^{\mathrm{T}} \rho\alpha_p \Big(\langle N\rangle\{T\} - T_{ref}\Big)\{g\} d\Omega \tag{C.70}$$

C.9 NS Neumann term

The corresponding weak formulation is

$$\begin{aligned} W_{\text{NS Neumann}} &= \int_{\Gamma_N} \delta\vec{u} \cdot \left(p\underline{\underline{I}} \cdot \vec{n}\right) d\Gamma - \int_{\Gamma_N} \delta\vec{u} \cdot \left(2\mu\underline{\underline{D}} \cdot \vec{n}\right) d\Gamma \\ &= -\int_{\Gamma_N} \delta\vec{u} \cdot \Big(\left(-p\underline{\underline{I}} + 2\mu\underline{\underline{D}}\right) \cdot \vec{n} \Big) d\Gamma \\ &= -\int_{\Gamma_N} \delta\vec{u} \cdot \vec{t}_N d\Gamma \end{aligned} \tag{C.71}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{NS Neumann}} = -\langle \delta u \rangle \int_{\Gamma^e_N} [N]^{\mathrm{T}} \{t_N\} d\Gamma \tag{C.72}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS Neumann}} = 0 \tag{C.73}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS Neumann}} = 0 \tag{C.74}$$

- Contribution to residual vector

$$\{R\}_{\text{NS Neumann}} = -\sum_e \int_{\Gamma^e_N} [N]^{\mathrm{T}} \{t_N\} d\Gamma \tag{C.75}$$

C.10 NS penalty term

$$\vec{u} - \vec{u}_I = 0 \tag{C.76}$$

The interface constraint at the interface is imposed by penalty method. The variational formulation can be expressed as

$$\Pi = \frac{1}{2}\beta \left(\vec{u} - \vec{u}_I\right)^2 \tag{C.77}$$

The variational derivative of Π yields

$$\delta\Pi = \delta\vec{u} \cdot \beta \left(\vec{u} - \vec{u}_I\right) \tag{C.78}$$

Then, the weak form can be written as

$$W_{\text{NS penalty}} = \int_{\Gamma_I^+} \delta\vec{u} \cdot \beta \left(\vec{u} - \vec{u}_I\right) d\Gamma \tag{C.79}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{NS penalty}} = \langle \delta u \rangle \int_{\Gamma_I^{+,e}} [N]^{\mathrm{T}} \beta \Big([N] \{u\} - \vec{u}_I \Big) d\Gamma \tag{C.80}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{NS penalty}} = -\sum_e \int_{\Gamma_I^{+,e}} [N]^{\mathrm{T}} \beta [N] \, d\Gamma \tag{C.81}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{NS penalty}} = -\sum_e \int_{\Gamma_I^{+,e}} [N]^{\mathrm{T}} \beta [N] \, d\Gamma \tag{C.82}$$

- Contribution to residual vector

$$\{R\}_{\text{NS penalty}} = \sum_e \int_{\Gamma_I^{+,e}} [N]^{\mathrm{T}} \beta \Big([N] \{u\}^{i-1} - \vec{u}_I \Big) d\Gamma \tag{C.83}$$

C.11 ST time term

$$\rho \frac{\partial \omega}{\partial t} \tag{C.84}$$

The weak formulation reads

$$W_{\text{ST time}} = \int_\Omega \delta\omega \rho \frac{\partial \omega}{\partial t} d\Omega \tag{C.85}$$

After time discretization by backward Euler method, it becomes

$$W_{\text{ST time}} = \int_\Omega \delta\omega \rho \frac{\omega - \omega^{t-\Delta t}}{\Delta t} d\Omega \tag{C.86}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{ST time}} = \langle \delta\omega \rangle \int_{\Omega^e} \{N\} \rho \langle N \rangle \frac{\{\omega\} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.87}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST time}} = -\sum_e \int_{\Omega^e} \{N\} \rho^{i-1} \frac{1}{\Delta t} \langle N \rangle d\Omega \tag{C.88}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned} [K_t]_{\text{ST time}} = &-\sum_e \int_{\Omega^e} \{N\} \rho^{i-1} \frac{1}{\Delta t} \langle N \rangle \\ &+ \{N\} (\rho')^{i-1} \langle N \rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} \langle N \rangle d\Omega \end{aligned} \tag{C.89}$$

- Contribution to residual vector

$$\{R\}_{\text{ST time}} = \sum_e \int_{\Omega^e} \{N\} \rho \langle N \rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.90}$$

C.12 ST time consv term

$$\omega \frac{\partial \rho}{\partial t} \tag{C.91}$$

The weak formulation reads

$$\begin{aligned} W_{\text{ST time consv}} &= \int_{\Omega} \delta\omega\omega \frac{\partial \rho}{\partial t} d\Omega \\ &= \int_{\Omega} \delta\omega\omega\rho' \frac{\partial \omega}{\partial t} d\Omega \end{aligned} \tag{C.92}$$

After time discretization by backward Euler method, it becomes

$$W_{\text{ST time consv}} = \int_{\Omega} \delta\omega\omega\rho' \frac{\omega - \omega^{t-\Delta t}}{\Delta t} d\Omega \tag{C.93}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{ST time consv}} = \langle \delta\omega \rangle \int_{\Omega^e} \{N\}\langle N\rangle\{\omega\}\rho'\langle N\rangle \frac{\{\omega\} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.94}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST time consv}} = -\sum_{e} \int_{\Omega^e} \{N\}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1} \frac{1}{\Delta t} \langle N\rangle d\Omega \tag{C.95}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned} [K_t]_{\text{ST time consv}} = -\sum_{e} \int_{\Omega^e} &\{N\}\Big((\rho')^{i-1} + \langle N\rangle\{\omega\}^{i-1}(\rho'')^{i-1}\Big)\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t}\langle N\rangle \\ &+ \{N\}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\frac{1}{\Delta t}\langle N\rangle d\Omega \end{aligned} \tag{C.96}$$

- Contribution to residual vector

$$\{R\}_{\text{ST time consv}} = \sum_{e} \int_{\Omega^e} \{N\}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\langle N\rangle \frac{\{\omega\}^{i-1} - \{\omega\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.97}$$

C.13 ST convection term

$$\vec{u} \cdot \rho \nabla \omega \tag{C.98}$$

The weak formulation reads

$$W_{\text{ST convection}} = \int_{\Omega} \delta\omega \vec{u} \cdot \rho \nabla \omega d\Omega \tag{C.99}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{ST convection}} = \langle \delta\omega \rangle \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \rho \nabla \langle N\rangle \{\omega\} d\Omega \tag{C.100}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST convection}} = -\sum_e \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \rho^{i-1} \nabla \langle N\rangle d\Omega \tag{C.101}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{ST convection}} = -\sum_e \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \Big(\rho^{i-1} \nabla \langle N\rangle + (\rho')^{i-1} \nabla \langle N\rangle \{\omega\}^{i-1} \langle N\rangle\Big) d\Omega \tag{C.102}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to residual vector

$$\{R\}_{\text{ST convection}} = \sum_e \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \rho^{i-1} \nabla \langle N\rangle \{\omega\}^{i-1} d\Omega \tag{C.103}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to tangential coefficient matrix, corresponding to the coupling with hydrodynamic problem, by Newton-Raphson method

$$[K_{t,w\vec{u}}]_{\text{ST convection}} = -\sum_e \int_{\Omega^e} \{N\} \rho^{i-1} \langle \omega\rangle^{i-1} \big(\nabla \langle N\rangle\big)^{\mathrm{T}} [N] d\Omega \tag{C.104}$$

C.14 ST convec consv term

$$\vec{u} \cdot \omega \nabla \rho \tag{C.105}$$

The weak formulation reads

$$\begin{aligned} W_{\text{ST convec consv}} &= \int_\Omega \delta\omega \vec{u} \cdot \omega \nabla \rho d\Omega \\ &= \int_\Omega \delta\omega \vec{u} \cdot \omega \rho' \nabla \omega d\Omega \end{aligned} \tag{C.106}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{ST convec consv}} = \langle \delta\omega\rangle \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \langle N\rangle \{\omega\} \rho' \nabla \langle N\rangle \{\omega\} d\Omega \tag{C.107}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST convec consv}} = -\sum_e \int_{\Omega^e} \{N\}\langle u\rangle [N]^{\mathrm{T}} \langle N\rangle \{\omega\}^{i-1} (\rho')^{i-1} \nabla \langle N\rangle d\Omega \tag{C.108}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned}[K_t]_{\text{ST convec consv}} = &-\sum_e \int_{\Omega^e} \{N\}\langle u\rangle[N]^{\mathrm{T}}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\nabla\langle N\rangle \\ &+ \{N\}\langle u\rangle[N]^{\mathrm{T}}\Big((\rho')^{i-1} + \langle N\rangle\{\omega\}^{i-1}(\rho'')^{i-1}\Big)\nabla\langle N\rangle\{\omega\}^{i-1}\langle N\rangle d\Omega\end{aligned} \tag{C.109}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to residual vector

$$\{R\}_{\text{ST convec consv}} = \sum_e \int_{\Omega^e} \{N\}\langle u\rangle[N]^{\mathrm{T}}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\nabla\langle N\rangle\{\omega\}^{i-1} d\Omega \tag{C.110}$$

  If strongly coupled with NS equations, $\langle u\rangle = \langle u\rangle^{i-1}$.

- Contribution to tangential coefficient matrix, corresponding to the coupling with hydrodynamic problem, by Newton-Raphson method

$$[K_{t,\omega\vec{u}}]_{\text{ST convec consv}} = -\sum_e \int_{\Omega^e} \{N\}\langle N\rangle\{\omega\}^{i-1}(\rho')^{i-1}\langle\omega\rangle^{i-1}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}[N] d\Omega \tag{C.111}$$

C.15 ST diffusion term

$$-\nabla\cdot(\rho D\nabla\omega) \tag{C.112}$$

The weak formulation reads

$$\begin{aligned}W_{\text{ST diffusion}} &= -\int_\Omega \delta\omega\nabla\cdot(\rho D\nabla\omega)d\Omega \quad \rightarrow\textbf{integration by parts} \\ &= \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega - \int_\Omega \nabla\cdot(\delta\omega\rho D\nabla\omega)d\Omega \quad \rightarrow\textbf{divergence theorem} \\ &= \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega - \oint_\Gamma \delta\omega\rho D\nabla\omega\cdot\vec{n}d\Gamma \\ &= \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \int_{\Gamma_D} \cancelto{0}{\delta\omega}(-\rho D\nabla\omega\cdot\vec{n})d\Gamma + \int_{\Gamma_N}\delta\omega(-\rho D\nabla\omega\cdot\vec{n})d\Gamma \\ &= \int_\Omega \nabla\delta\omega\cdot\rho D\nabla\omega d\Omega + \underbrace{\int_{\Gamma_N}\delta\omega q_{\omega,n}d\Gamma}_{\textbf{ST Neumann term}}\end{aligned} \tag{C.113}$$

The treatment of **ST Neumann term** is in section C.16. For **ST diffusion term**, the spatial discretization in matrix form on one element reads

$$W^e_{\text{ST diffusion}} = \langle\delta\omega\rangle\int_{\Omega^e}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\rho D\nabla\langle N\rangle\{\omega\}d\Omega \tag{C.114}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST diffusion}} = -\sum_e\int_{\Omega^e}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\rho^{i-1}D\nabla\langle N\rangle d\Omega \tag{C.115}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{ST diffusion}} = -\sum_e \int_{\Omega^e} \left(\nabla\langle N\rangle\right)^{\mathrm{T}} \left(\rho^{i-1} D \nabla\langle N\rangle + (\rho')^{i-1} D \nabla\langle N\rangle \{\omega\}^{i-1} \langle N\rangle\right) d\Omega \quad \text{(C.116)}$$

- Contribution to residual vector

$$\{R\}_{\text{ST diffusion}} = \sum_e \int_{\Omega^e} \left(\nabla\langle N\rangle\right)^{\mathrm{T}} \rho^{i-1} D \nabla\langle N\rangle \{\omega\}^{i-1} d\Omega \quad \text{(C.117)}$$

C.16 ST Neumann term

The weak form can be written as

$$W_{\text{ST Neumann}} = \int_{\Gamma_N} \delta\omega q_{\omega,n} d\Gamma \quad \text{(C.118)}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{ST Neumann}} = \langle\delta\omega\rangle \int_{\Gamma^e_N} \{N\} q_{\omega,n} d\Gamma \quad \text{(C.119)}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{ST Neumann}} = 0 \quad \text{(C.120)}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{ST Neumann}} = -\sum_e \int_{\Gamma^e_N} \{N\} \left(q'_{\omega,n}\right)^{i-1} \langle N\rangle d\Gamma \quad \text{(C.121)}$$

- Contribution to residual vector

$$\{R\}_{\text{ST Neumann}} = \sum_e \int_{\Gamma^e_N} \{N\} q^{i-1}_{\omega,n} d\Gamma \quad \text{(C.122)}$$

C.17 Heat capacity term

$$\rho C_p \frac{\partial T}{\partial t} \quad \text{(C.123)}$$

After weak formulation, it gives

$$W_{\text{Heat capacity}} = \int_\Omega \delta T \rho C_p \frac{\partial T}{\partial t} d\Omega \quad \text{(C.124)}$$

After time discretization by backward Euler method, it becomes

$$W_{\text{Heat capacity}} = \int_\Omega \delta T \rho C_p \frac{1}{\Delta t} T d\Omega - \int_\Omega \delta T \rho C_p \frac{1}{\Delta t} T^{t-\Delta t} d\Omega \quad \text{(C.125)}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{Heat capacity}} = \langle\delta T\rangle \int_{\Omega^e} \{N\} \rho C_p \frac{1}{\Delta t} \langle N\rangle \{T\} d\Omega - \langle\delta T\rangle \int_{\Omega^e} \{N\} \rho C_p \frac{1}{\Delta t} \langle N\rangle \{T\}^{t-\Delta t} d\Omega \quad \text{(C.126)}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat capacity}} = -\sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\frac{1}{\Delta t}\langle N\rangle d\Omega \tag{C.127}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned}[K_t]_{\text{Heat capacity}} = &-\sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\frac{1}{\Delta t}\langle N\rangle \\ &+ \{N\}\Big(\rho' C_p + \rho C_p'\Big)^{i-1}\langle N\rangle\frac{\{T\}^{i-1} - \{T\}^{t-\Delta t}}{\Delta t}\langle N\rangle d\Omega\end{aligned} \tag{C.128}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat capacity}} = \sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\langle N\rangle\frac{\{T\}^{i-1} - \{T\}^{t-\Delta t}}{\Delta t}d\Omega \tag{C.129}$$

C.18 Heat convection term

$$\rho C_p \vec{u}\cdot\nabla T \tag{C.130}$$

After weak formulation, it gives

$$W_{\text{Heat convection}} = \int_\Omega \delta T \rho C_p \vec{u}\cdot\nabla T d\Omega \tag{C.131}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{Heat convection}} = \langle\delta T\rangle \int_{\Omega^e} \{N\}\rho C_p\langle u\rangle \left[N\right]^{\mathrm{T}} \nabla\langle N\rangle\{T\} d\Omega \tag{C.132}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat convection}} = -\sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\langle u\rangle \left[N\right]^{\mathrm{T}} \nabla\langle N\rangle d\Omega \tag{C.133}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned}[K_t]_{\text{Heat convection}} = &-\sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\langle u\rangle \left[N\right]^{\mathrm{T}} \nabla\langle N\rangle \\ &+ \{N\}\Big(\rho' C_p + \rho C_p'\Big)^{i-1}\langle u\rangle \left[N\right]^{\mathrm{T}} \nabla\langle N\rangle\{T\}^{i-1}\langle N\rangle d\Omega\end{aligned} \tag{C.134}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat convection}} = \sum_e \int_{\Omega^e} \{N\}\rho^{i-1}C_p^{i-1}\langle u\rangle \left[N\right]^{\mathrm{T}} \nabla\langle N\rangle\{T\}^{i-1} d\Omega \tag{C.135}$$

C.19 Heat conduction term

$$-\nabla\cdot(k\nabla T) \tag{C.136}$$

The corresponding weak formulation is

$$\begin{aligned}
W_{\text{Heat conduction}} &= -\int_\Omega \delta T\nabla\cdot(k\nabla T)\,d\Omega \quad \rightarrow\text{integration by parts} \\
&= \int_\Omega \nabla\delta T\cdot k\nabla T d\Omega - \int_\Omega \nabla\cdot\Big(\delta T k\nabla T\Big)d\Omega \quad \rightarrow\text{divergence theorem} \\
&= \int_\Omega \nabla\delta T\cdot k\nabla T d\Omega - \oint_\Gamma \delta T k\nabla T\cdot\vec{n}d\Gamma \\
&= \int_\Omega \nabla\delta T\cdot k\nabla T d\Omega + \int_{\Gamma_D} \cancelto{0}{\delta T}\,(-k\nabla T\cdot\vec{n})\,d\Gamma + \int_{\Gamma_N}\delta T\,(-k\nabla T\cdot\vec{n})\,d\Gamma \\
&= \int_\Omega \nabla\delta T\cdot k\nabla T d\Omega + \underbrace{\int_{\Gamma_N}\delta T q_n d\Gamma}_{\text{Heat Neumann term}}
\end{aligned} \tag{C.137}$$

The treatment of **Heat Neumann term** is in section C.21. For **Heat conduction term**, the spatial discretization in matrix form on one element reads

$$W^e_{\text{Heat conduction}} = \langle\delta T\rangle\int_{\Omega^e}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}k\nabla\langle N\rangle\{T\}d\Omega \tag{C.138}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat conduction}} = -\sum_e\int_{\Omega^e}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}k^{i-1}\nabla\langle N\rangle d\Omega \tag{C.139}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$\begin{aligned}
[K_t]_{\text{Heat conduction}} = -\sum_e\int_{\Omega^e}&\big(\nabla\langle N\rangle\big)^{\mathrm{T}}k^{i-1}\nabla\langle N\rangle \\
&+\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\left(k'\right)^{i-1}\nabla\langle N\rangle\{T\}^{i-1}\langle N\rangle d\Omega
\end{aligned} \tag{C.140}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat conduction}} = \sum_e\int_{\Omega^e}\big(\nabla\langle N\rangle\big)^{\mathrm{T}}k^{i-1}\nabla\langle N\rangle\{T\}^{i-1}d\Omega \tag{C.141}$$

C.20 Heat source term

$$-s \tag{C.142}$$

The corresponding weak formulation is

$$W_{\text{Heat source}} = -\int_{\Omega} \delta T s d\Omega \tag{C.143}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{Heat source}} = -\langle \delta T \rangle \int_{\Omega} \{N\} s d\Omega \tag{C.144}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat source}} = 0 \tag{C.145}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{Heat source}} = \sum_{e} \int_{\Omega^e} \{N\} \left(s'\right)^{i-1} \langle N \rangle d\Omega \tag{C.146}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat source}} = -\sum_{e} \int_{\Omega^e} \{N\} s^{i-1} d\Omega \tag{C.147}$$

C.21 Heat Neumann term

The weak form can be written as

$$W_{\text{Heat Neumann}} = \int_{\Gamma_N} \delta T q_n d\Gamma \tag{C.148}$$

The spatial discretization in matrix form on one element reads

$$W^{e}_{\text{Heat Neumann}} = \langle \delta T \rangle \int_{\Gamma^e_N} \{N\} q_n d\Gamma \tag{C.149}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat Neumann}} = 0 \tag{C.150}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{Heat Neumann}} = -\sum_{e} \int_{\Gamma^e_N} \{N\} \left(q'_n\right)^{i-1} \langle N \rangle d\Gamma \tag{C.151}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat Neumann}} = \sum_{e} \int_{\Gamma^e_N} \{N\} q_n^{i-1} d\Gamma \tag{C.152}$$

C.22 Heat penalty term

$$T - T_m = 0 \tag{C.153}$$

The interface constraint at the interface is imposed by penalty method. The variational formulation can be expressed as

$$\Pi = \frac{1}{2}\beta \left(T - T_m\right)^2 \tag{C.154}$$

The variational derivative of Π yields

$$\delta\Pi = \delta T \beta \left(T - T_m\right) \tag{C.155}$$

Then, the weak form can be written as

$$W_{\text{Heat penalty}} = \int_{\Gamma_I} \delta T \beta \left(T - T_m\right) d\Gamma \tag{C.156}$$

The spatial discretization in matrix form on one element reads

$$W^e_{\text{Heat penalty}} = \langle \delta T \rangle \int_{\Gamma^e_I} \{N\} \beta \Big(\langle N \rangle \{T\} - T_m \Big) d\Gamma \tag{C.157}$$

- Contribution to coefficient matrix by substitution method

$$[K]_{\text{Heat penalty}} = -\sum_e \int_{\Gamma^e_I} \{N\} \beta \langle N \rangle d\Gamma \tag{C.158}$$

- Contribution to tangential coefficient matrix by Newton-Raphson method

$$[K_t]_{\text{Heat penalty}} = -\sum_e \int_{\Gamma^e_I} \{N\} \beta \langle N \rangle d\Gamma \tag{C.159}$$

- Contribution to residual vector

$$\{R\}_{\text{Heat penalty}} = \sum_e \int_{\Gamma^e_I} \{N\} \beta \Big(\langle N \rangle \{T\}^{i-1} - T_m \Big) d\Gamma \tag{C.160}$$

C.23 LS time term

$$\frac{\partial \phi}{\partial t} \tag{C.161}$$

The weak form is

$$W_{\text{LS time}} = \int_\Omega \delta\phi \frac{\partial \phi}{\partial t} d\Omega \tag{C.162}$$

After time discretization by forward Euler method, it becomes

$$W_{\text{LS time}} = \int_\Omega \delta\phi \frac{\phi - \phi^{t-\Delta t}}{\Delta t} d\Omega \tag{C.163}$$

On one element, it can be discretized in space as

$$W^e_{\text{LS time}} = \langle \delta\phi \rangle \int_{\Omega^e} \{N\} \langle N \rangle \frac{\{\phi\} - \{\phi\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.164}$$

- Contribution to the coefficient matrix by substitution method

$$[K]_{\text{LS time}} = -\sum_e \int_{\Omega^e} \{N\}\langle N\rangle \frac{1}{\Delta t} d\Omega \tag{C.165}$$

- Contribution to the residual vector

$$\{R\}_{\text{LS time}} = \sum_e \int_{\Omega^e} \{N\}\langle N\rangle \frac{\{\phi\}^{i-1} - \{\phi\}^{t-\Delta t}}{\Delta t} d\Omega \tag{C.166}$$

C.24 LS convection term

$$F\|\nabla\phi\| \tag{C.167}$$

If forward Euler scheme is employed, the weak form can be written as

$$W_{\text{LS convection}} = \int_\Omega \delta\phi F \left\|\nabla\phi^{t-\Delta t}\right\| d\Omega \tag{C.168}$$

On one element, it can be discretized in space as

$$W^e_{\text{LS convection}} = \langle\delta\phi\rangle \int_{\Omega^e} \{N\}\langle N\rangle\{F\} \left\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\right\| d\Omega \tag{C.169}$$

- Contribution to the coefficient matrix by substitution method

$$[K]_{\text{LS convection}} = 0 \tag{C.170}$$

- Contribution to the residual vector

$$\{R\}_{\text{LS convection}} = \sum_e \int_{\Omega^e} \{N\}\langle N\rangle\{F\} \left\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\right\| d\Omega \tag{C.171}$$

C.25 LS GLS term

If forward Euler scheme is employed, the weak form can be written as

$$\sum_e \int_{\Omega^e} \Big(F\nabla\delta\phi \cdot \frac{\nabla\phi^{t-\Delta t}}{\|\nabla\phi^{t-\Delta t}\|}\Big)\tau^e \Big(\frac{\phi - \phi^{t-\Delta t}}{\Delta t} + F\left\|\nabla\phi^{t-\Delta t}\right\|\Big) d\Omega \tag{C.172}$$

On one element, it can be discretized in space as

$$\begin{aligned} W^e_{\text{LS GLS}} = \langle\delta\phi\rangle \int_{\Omega^e} \big(\langle N\rangle\{F\}\big)\left(\nabla\langle N\rangle\right)^{\mathrm{T}} \frac{\nabla\langle N\rangle\{\phi\}^{t-\Delta t}}{\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\|}\tau^e \\ \Big(\langle N\rangle\frac{\{\phi\} - \{\phi\}^{t-\Delta t}}{\Delta t} + \langle N\rangle\{F\}\left\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\right\|\Big) d\Omega \end{aligned} \tag{C.173}$$

- Contribution to the coefficient matrix by substitution method

$$[K]_{\text{LS GLS}} = -\sum_e \int_{\Omega^e} \left(\langle N\rangle\{F\}\right)\left(\nabla\langle N\rangle\right)^{\mathrm{T}} \frac{\nabla\langle N\rangle\{\phi\}^{t-\Delta t}}{\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\|}\tau^e \frac{1}{\Delta t}\langle N\rangle d\Omega \tag{C.174}$$

- Contribution to the residual vector

$$\begin{aligned}\{R\}_{\text{LS GLS}} = \sum_e \int_{\Omega^e} \left(\langle N\rangle\{F\}\right)\left(\nabla\langle N\rangle\right)^{\mathrm{T}} \frac{\nabla\langle N\rangle\{\phi\}^{t-\Delta t}}{\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\|}\tau^e \\ \Big(\langle N\rangle\frac{\{\phi\}^{i-1}-\{\phi\}^{t-\Delta t}}{\Delta t} + \langle N\rangle\{F\}\left\|\nabla\langle N\rangle\{\phi\}^{t-\Delta t}\right\|\Big) d\Omega\end{aligned} \tag{C.175}$$

C.26 Spread ortho term

$$sign(\phi)\nabla\phi\cdot\nabla F \tag{C.176}$$

The corresponding weak form is

$$W_{\text{Spread ortho}} = \int_\Omega \delta F sign(\phi)\nabla\phi\cdot\nabla F d\Omega \tag{C.177}$$

The spatial discretization on one element can be written as

$$W^e_{\text{Spread ortho}} = \langle\delta F\rangle \int_{\Omega^e} \{N\} sign(\phi)\langle\phi\rangle\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle\{F\} d\Omega \tag{C.178}$$

- Contribution to the coefficient matrix by director method

$$[K]_{\text{Spread ortho}} = \sum_e \int_{\Omega^e} \{N\} sign(\phi)\langle\phi\rangle\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle d\Omega \tag{C.179}$$

C.27 Spread GLS term

The weak form is

$$\sum_e \int_{\Omega^e} \Big(\nabla\delta F\cdot\nabla\phi\Big)\tau^e\Big(\nabla\phi\cdot\nabla F\Big) d\Omega \tag{C.180}$$

The spatial discretization on one element reads

$$W^e_{\text{Spread GLS}} = \langle\delta F\rangle \int_{\Omega^e} \big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle\{\phi\}\tau^e\langle\phi\rangle\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle\{F\} d\Omega \tag{C.181}$$

- Contribution to the coefficient matrix by direct method

$$[K]_{\text{Spread GLS}} = \sum_e \int_{\Omega^e} \big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle\{\phi\}\tau^e\langle\phi\rangle\big(\nabla\langle N\rangle\big)^{\mathrm{T}}\nabla\langle N\rangle d\Omega \tag{C.182}$$

Appendix D

1D illustration of the ill-conditioned system by XFEM

The sample governing equation with arbitrary unknown denoted by p is

$$\frac{dp}{dx} = 1 \tag{D.1}$$

To make the problem well-posed, one Dirichlet-type boundary condition is set on the left end (see Fig. D.1):

$$p_1 = 10 \tag{D.2}$$

As shown in Fig. D.1, a strong discontinuity appears within element ②. The constraint at the discontinuity can be assumed as below:

$$[\![p]\!] = 2 \tag{D.3}$$

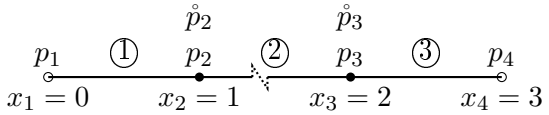


Figure D.1 – The sample mesh containing 3 elements; a strong discontinuity resides on element ② with coordinate $x_c \in (1, 2)$.

The weak formulation is

$$\int_\Omega \delta p \frac{dp}{dx} dx = \int_\Omega \delta p dx \tag{D.4}$$

The discretization form is

$$\sum_{e=①}^{③} \Big(\int_{\Omega^e} \{N\}\langle N_x\rangle dx \{p\} - \int_{\Omega^e} \{N\} dx = 0 \Big) \tag{D.5}$$

The geometric approximation (mapping) between the real element and reference element as shown in Fig. D.2 can be written as

$$x = \langle \bar{N} \rangle \{x\} \tag{D.6}$$

where $\langle \bar{N} \rangle = \langle \ \frac{1}{2}(1-\xi) \quad \frac{1}{2}(1+\xi) \ \rangle$.

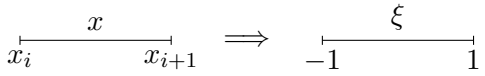


Figure D.2 – The real element (left) and reference element (right).

After mapping from real element to the reference element, Eq. D.5 becomes

$$\sum_{e=①}^{③} \Big(\int_{\Omega^e} \{\bar{N}\} \langle \bar{N}_\xi \rangle J^{-1} |J| d\xi \{p\} - \int_{\Omega^e} \{\bar{N}\} |J| d\xi = 0 \Big) \tag{D.7}$$

or in the condensed form

$$\mathbf{K}\{p\} - \mathbf{F} = 0 \tag{D.8}$$

where

$$\mathbf{K} = \sum_e \int_{\Omega^e} \{\bar{N}\} \langle \bar{N}_\xi \rangle J^{-1} |J| d\xi \tag{D.9a}$$

$$\mathbf{F} = \sum_e \int_{\Omega^e} \{\bar{N}\} |J| d\xi \tag{D.9b}$$

and J is the Jacobian matrix, it can be calculated as

$$J = \frac{dx}{d\xi} = \langle \bar{N}_\xi \rangle \{x\} = \frac{x_{i+1} - x_i}{2} \tag{D.10}$$

In the present study, it's assumed that $x_{i+1} - x_i = 1$. Then $J = \frac{1}{2}$, $|J| = \frac{1}{2}$ and $J^{-1} = 1$. In a fully enriched element, the shape (interpolation) function of unknown p can be expressed as

$$\{\bar{N}\} = \begin{Bmatrix} \bar{N}_1 \\ \bar{N}_2 \\ \bar{M}_1 \\ \bar{M}_2 \end{Bmatrix} = \begin{Bmatrix} \frac{1}{2}(1-\xi) \\ \frac{1}{2}(1+\xi) \\ \frac{1}{2}(1-\xi)\big(sign(\phi) - sign(\phi_1)\big) \\ \frac{1}{2}(1+\xi)\big(sign(\phi) - sign(\phi_2)\big) \end{Bmatrix} \tag{D.11}$$

and its derivative with respect to ξ is

$$\{\bar{N}_\xi\} = \begin{Bmatrix} \bar{N}_{1,\xi} \\ \bar{N}_{2,\xi} \\ \bar{M}_{1,\xi} \\ \bar{M}_{2,\xi} \end{Bmatrix} = \begin{Bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2}\big(sign(\phi) - sign(\phi_1)\big) \\ \frac{1}{2}\big(sign(\phi) - sign(\phi_2)\big) \end{Bmatrix} \tag{D.12}$$

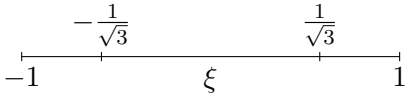


Figure D.3 – The Gauss-Legendre quadrature with two points.

Gauss-Legendre quadrature with two points as shown in Fig. D.3 is employed to do the numerical integration. It is stated as

$$\int_{-1}^{1} f(\xi)d\xi = w_1 f(\xi_1) + w_2 f(\xi_2) \tag{D.13}$$

where $w_1 = w_2 = 1$; $\xi_1 = -\frac{1}{\sqrt{3}}$ and $\xi_2 = \frac{1}{\sqrt{3}}$.

Build the coefficient matrix $\mathbf{K}^e$ and load vector $\mathbf{F}^e$ on each element:

Element ① $\langle\ p_1 \quad p_2 \quad \mathring{p}_2\ \rangle$

$$\mathbf{K}^e = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{F}^e = \begin{Bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{Bmatrix} \tag{D.14}$$

Element ② $\langle\ p_2 \quad p_3 \quad \mathring{p}_2 \quad \mathring{p}_3\ \rangle$

In the splitting element as shown in Fig. D.4, Eq. D.7 can be written as

$$\begin{aligned} 0 &= \int_{\Omega^e} \{N\}\langle N_x\rangle dx\{p\} - \int_{\Omega^e} \{N\}dx \\ &= \int_{-1}^{1} \{\bar{N}\}\langle \bar{N}_\xi\rangle J^{-1}|J|d\xi\{p\} - \int_{-1}^{1} \{\bar{N}\}|J|d\xi \\ &= \int_{-1}^{\xi_c} \{\bar{N}\}\langle \bar{N}_\xi\rangle J^{-1}|J|d\xi\{p\} - \int_{-1}^{\xi_c} \{\bar{N}\}|J|d\xi + \int_{\xi_c}^{1} \{\bar{N}\}\langle \bar{N}_\xi\rangle J^{-1}|J|d\xi\{p\} - \int_{\xi_c}^{1} \{\bar{N}\}|J|d\xi \\ &= \sum_{i=1}^{4} \left(\{\bar{N}\}\langle \bar{N}_\xi\rangle J^{-1}|J|\Big|_{\xi_{G_i}} w_{G_i}\{p\} - \{\bar{N}\}|J|\Big|_{\xi_{G_i}} w_{G_i} \right) \end{aligned} \tag{D.15}$$

The corresponding elementary coefficient matrix and load vector are

$$\mathbf{K}^e = \begin{Bmatrix} -x_c+3 \\ x_c-1 \\ 0 \\ -2x_c+2 \end{Bmatrix} \langle\ -\tfrac{1}{2} \quad \tfrac{1}{2} \quad 0 \quad -1\ \rangle (x_c-1) + \begin{Bmatrix} -x_c+2 \\ x_c \\ -2x_c+4 \\ 0 \end{Bmatrix} \langle\ -\tfrac{1}{2} \quad \tfrac{1}{2} \quad -1 \quad 0\ \rangle (2-x_c) \tag{D.16a}$$

$$\mathbf{F}^e = \frac{1}{2}(x_c-1) \begin{Bmatrix} -x_c+3 \\ x_c-1 \\ 0 \\ -2x_c+2 \end{Bmatrix} + \frac{1}{2}(2-x_c) \begin{Bmatrix} -x_c+2 \\ x_c \\ -2x_c+4 \\ 0 \end{Bmatrix} \tag{D.16b}$$

Element ③ $\langle\ p_3 \quad p_4 \quad \mathring{p}_3\ \rangle$

$$\mathbf{K}^e = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{F}^e = \begin{Bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{Bmatrix} \tag{D.17}$$

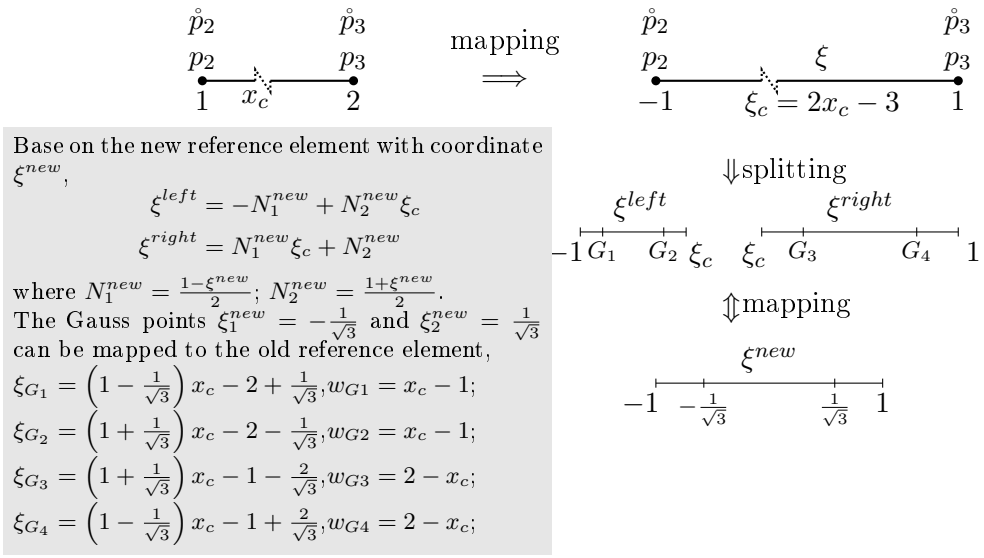


Figure D.4 – Split and mapping of element ②.

Imposing interface constraint by penalty method $\langle\ p_2 \quad p_3 \quad \mathring{p}_2 \quad \mathring{p}_3\ \rangle$

The corresponding variational form is

$$\Pi = \frac{\beta}{2}\big(\llbracket p \rrbracket - 2\big)^2 \tag{D.18}$$

The weak formulation is

$$\begin{aligned} w &= \int_{\Gamma_I} \llbracket \delta p \rrbracket \beta \big(\llbracket p \rrbracket - 2\big) d\Gamma \\ &= \langle\ \delta p_2 \quad \delta p_3 \quad \delta \mathring{p}_2 \quad \delta \mathring{p}_3\ \rangle \begin{Bmatrix} \llbracket \bar{N}_1 \rrbracket \\ \llbracket \bar{N}_2 \rrbracket \\ \llbracket \bar{M}_1 \rrbracket \\ \llbracket \bar{M}_2 \rrbracket \end{Bmatrix} \beta \left(\langle\ \llbracket \bar{N}_1 \rrbracket \quad \llbracket \bar{N}_2 \rrbracket \quad \llbracket \bar{M}_1 \rrbracket \quad \llbracket \bar{M}_2 \rrbracket\ \rangle \begin{Bmatrix} \llbracket p_2 \rrbracket \\ \llbracket p_3 \rrbracket \\ \llbracket \mathring{p}_2 \rrbracket \\ \llbracket \mathring{p}_3 \rrbracket \end{Bmatrix} - 2 \right) \end{aligned} \tag{D.19}$$

The corresponding elementary coefficient matrix and load vector are shown below:

$$\mathbf{K}^e = \beta \begin{Bmatrix} 0 \\ 0 \\ 4 - 2x_c \\ 2x_c - 2 \end{Bmatrix} \langle\ 0 \quad 0 \quad 4 - 2x_c \quad 2x_c - 2\ \rangle, \quad \mathbf{F}^e = 2\beta \begin{Bmatrix} 0 \\ 0 \\ 4 - 2x_c \\ 2x_c - 2 \end{Bmatrix} \tag{D.20}$$

Assembly

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2}(x_c-1)(3-x_c) \\ -\frac{1}{2}(2-x_c)^2 & \frac{1}{2}(x_c-1)(3-x_c) \\ +\frac{1}{2}(2-x_c)^2 & 0 & -(2-x_c)^2 & -(x_c-1)(3-x_c) \\ 0 & -\frac{1}{2}(x_c-1)^2 \\ -\frac{1}{2}(2-x_c)x_c & \frac{1}{2}(x_c-1)^2 \\ +\frac{1}{2}(2-x_c)x_c \\ -\frac{1}{2} & \frac{1}{2} & -(2-x_c)x_c & -(x_c-1)^2 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -(2-x_c)^2 & (2-x_c)^2 & 0 & -2(2-x_c)^2 \\ +\beta(4-2x_c)^2 & \beta(4-2x_c)(2x_c-2) \\ 0 & (x_c-1)^2 & -(x_c-1)^2 & 0 & \beta(4-2x_c)(2x_c-2) & 2(x_c-1)^2 \\ +\beta(2x_c-2)^2 \end{bmatrix} \begin{Bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ \mathring{p}_2 \\ \mathring{p}_3 \end{Bmatrix} = \begin{Bmatrix} \frac{1}{2}+? \\ \frac{1}{2} \\ +\frac{1}{2}(x_c-1)(3-x_c) \\ +\frac{1}{2}(2-x_c)^2 \\ \frac{1}{2}(x_c-1)^2 \\ +\frac{1}{2}(2-x_c)x_c \\ +\frac{1}{2} \\ \frac{1}{2} \\ (2-x_c)^2 \\ +2\beta(4-2x_c) \\ -(x_c-1)^2 \\ +2\beta(2x_c-2) \end{Bmatrix} \tag{D.21}$$

Imposing Dirichlet boundary condition $p_1 = 10$

$$\begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2}(x_c-1)(3-x_c) \\ -\frac{1}{2}(2-x_c)^2 & \frac{1}{2}(x_c-1)(3-x_c) \\ +\frac{1}{2}(2-x_c)^2 & 0 & -(2-x_c)^2 & -(x_c-1)(3-x_c) \\ -\frac{1}{2}(x_c-1)^2 \\ -\frac{1}{2}(2-x_c)x_c & \frac{1}{2}(x_c-1)^2 \\ +\frac{1}{2}(2-x_c)x_c \\ -\frac{1}{2} & \frac{1}{2} & -(2-x_c)x_c & -(x_c-1)^2 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ -(2-x_c)^2 & (2-x_c)^2 & 0 & -2(2-x_c)^2 \\ +\beta(4-2x_c)^2 & \beta(4-2x_c)(2x_c-2) \\ (x_c-1)^2 & -(x_c-1)^2 & 0 & \beta(4-2x_c)(2x_c-2) & 2(x_c-1)^2 \\ +\beta(2x_c-2)^2 \end{bmatrix} \begin{Bmatrix} p_2 \\ p_3 \\ p_4 \\ \mathring{p}_2 \\ \mathring{p}_3 \end{Bmatrix} = \begin{Bmatrix} 5+\frac{1}{2} \\ +\frac{1}{2}(x_c-1)(3-x_c) \\ +\frac{1}{2}(2-x_c)^2 \\ \frac{1}{2}(x_c-1)^2 \\ +\frac{1}{2}(2-x_c)x_c \\ +\frac{1}{2} \\ \frac{1}{2} \\ (2-x_c)^2 \\ +2\beta(4-2x_c) \\ -(x_c-1)^2 \\ +2\beta(2x_c-2) \end{Bmatrix} \tag{D.22}$$

The penalty number is assumed to be 1. Below gives the variation of the condition number as x_c starting off at $x = 1.5$ and approaching to $x = 1$:

As $x = 1.50000$, condition number: $\approx 2.1 \times 10^1$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$;

As $x = 1.10000$, condition number: $\approx 1.3 \times 10^2$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$;

As $x = 1.01000$, condition number: $\approx 1.3 \times 10^4$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$;

As $x = 1.00100$, condition number: $\approx 1.3 \times 10^6$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$;

As $x = 1.00010$, condition number: $\approx 1.3 \times 10^8$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$;

As $x = 1.00001$, condition number: $\approx 1.3 \times 10^{10}$; $P_2 = 11$, $P_3 = 14$, $P_4 = 15$, $\mathring{P}_2 = 1$, $\mathring{p}_3 = 1$.

If $\mathring{p}_3$ is set zero as x_c is close to 1, it turns out

As $x = 1.01000$, condition number: ≈ 13.36; $P_2 = 10.960$, $P_3 = 13.980$, $P_4 = 14.980$, $\mathring{P}_2 = 1.010$;

As $x = 1.00100$, condition number: ≈ 14.69; $P_2 = 10.996$, $P_3 = 13.998$, $P_4 = 14.998$, $\mathring{P}_2 = 1.001$;

As $x = 1.00010$, condition number: ≈ 14.72; $P_2 = 11.000$, $P_3 = 14.000$, $P_4 = 15.000$, $\mathring{P}_2 = 1.000$;

As $x = 1.00001$, condition number: ≈ 14.72; $P_2 = 11.000$, $P_3 = 14.000$, $P_4 = 15.000$, $\mathring{P}_2 = 1.000$.