ÉCOLE CENTRALE DES ARTS ET MANUFACTURES

«École Centrale Paris»

THÈSE

Présentée par

Mathieu MORLIGHEM

pour l'obtention du

GRADE DE DOCTEUR

Spécialité : **Mécanique**

Ice sheet properties inferred by combining numerical modeling and remote sensing data

Détermination de propriétés des glaciers polaires par modélisation numérique et télédétection

Soutenue le 22 Décembre 2011 devant un jury composé de :

Président de jury :	M. BEN DHIA Hachmi	École Centrale Paris
Rapporteurs :	M. PATTYN Frank	Université Libre de Bruxelles
	M. VILOTTE Jean-Pierre	IPG Paris
Directeur de thèse :	M. AUBRY Denis	École Centrale Paris
Examinateurs :	M. LAROUR Éric	Jet Propulsion Laboratory
	M. RIGNOT Éric	University of California Irvine
Invité :	M. DUREISSEIX David	INSA Lyon

Laboratoire de Mécanique des Sols, Structures et Matériaux – CNRS UMR 8579

2011 - N° 2011ECAP0062

"The test of science is its ability to predict."
Richard Feynman

"It's hard to make predictions, especially about the Future"
variously attributed to Yogi Berra, Neils Bohr or Mark Twain

Abstract

Ice sheets are amongst the main contributors to sea level rise. They are dynamic systems; they gain mass by snow accumulation, and lose it by melting at the ice-ocean interface, surface melting and iceberg calving at the margins. Observations over the last three decades have shown that the Greenland and Antarctic ice sheets have been losing more mass than they gain. How the ice sheets respond to this negative mass imbalance has become today one of the most urgent questions in understanding the implications of global climate change. The Intergovernmental Panel on Climate Change (IPCC) has indeed identified the contribution of the ice sheets as a key uncertainty in sea level rise projections. Numerical modeling is the only effective way of addressing this problem. Yet, modeling ice sheet flow at the scale of Greenland and Antarctica remains scientifically and technically very challenging.

This thesis focuses on two major aspects of improving ice sheet numerical models. The first consists of determining non-observable ice properties using inverse methods. Some parameters, such as basal friction or ice shelf hardness, are difficult to measure and must be inferred from remote sensing observations. Inversions are developed here for three ice flow models of increasing complexity: MacAyeal/Morland's shelfy-stream model, Blatter/Pattyn's higher order model and the full-Stokes model. The inferred parameters are then used to initialize large-scale ice sheet models and to determine the minimum level of complexity required to capture ice dynamics correctly.

The second aspect addressed in this work is the improvement of dataset consistency for ice sheet modeling. Available datasets are often collected at different epochs and at varying spatial resolutions, making them not readily usable for numerical simulations. We devise here an algorithm based on the conservation of mass principle and inverse methods to construct ice thicknesses that are consistent with velocity measurements. This approach therefore avoids the artificial mass redistributions that occur in existing algorithms for mapping ice thickness, hence considerably improving ice sheet model initialization.

The advances made here are important steps towards the ultimate objective of accurate characterization of ice sheets and the realistic modeling of their evolution.

Key-words: glaciology, glaciers, ice caps, ice sheet dynamics, numerical modeling, inverse methods, PDE-constrained optimization, mass conservation.

Résumé

Les calottes polaires, ou inlandsis, sont parmi les principaux contributeurs à la montée des océans. Ces systèmes dynamiques gagnent de la masse par accumulation de neige, et en perdent par fonte au contact de l'océan et à la surface, ainsi que par le vêlage d'icebergs. Depuis plus de trois décennies, les observations ont montré que les calottes polaires de l'Antarctique et du Groenland perdent plus de masse qu'ils n'en gagnent. L'évolution des glaciers suite à ce déséquilibre de masse est devenue aujourd'hui l'une des problématiques les plus importantes des implications du changement climatique. Le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) a identifié la contribution des glaciers comme l'un des facteurs clés d'incertitude de prédiction de l'élévation du niveau des mers. La modélisation numérique est le seul outil efficace pour répondre à cette question. Cependant, modéliser l'écoulement de glace à l'échelle du Groenland ou de l'Antarctique représente un défi à la fois scientifique et technique.

Deux aspects clés de l'amélioration de la modélisation des glaciers sont abordés dans cette thèse. Le premier consiste à déterminer certaines propriétés non mesurables de la glace par méthode inverse. La friction ou la rigidité des barrières de glace, sont des paramètres qui ne peuvent être mesurés directement et doivent donc être déduits à partir d'observations par télédétection. Nous appliquons ici ces inversions pour trois modèles d'écoulement de glace de complexité croissante: le modèle bidimensionnel de MacAyeal/Morland, le modèle dit d'ordre supérieur de Blatter/Pattyn et le modèle full-Stokes. Les propriétés ainsi calculées sont ensuite utilisées pour initialiser des modèles grande-échelle et pour déterminer le degré de complexité minimum nécessaire pour reproduire correctement la dynamique des glaciers.

Le second aspect abordé dans ce travail est l'amélioration de la consistance des données pour la modélisation numérique. Les données disponibles sont souvent issues de campagnes de mesures s'étalant sur plusieurs années et dont résolutions spatiales varient, ce qui rend leur utilisation pour des simulations numériques difficiles. Nous présentons ici un algorithme basé sur la conservation de la masse et les méthodes inverses pour construire des épaisseurs de glace qui sont consistantes avec les mesures de vitesse. Cette approche empêche la redistribution artificielle de masse qu'engendrent généralement les autres méthodes de cartographie de l'épaisseur de glace, ce qui améliore considérablement l'initialisation des modèles découlement de glace.

Les avancées présentées ici sont des étapes importantes afin de mieux caractériser de manière précise les glaciers et de modéliser leur évolution de manière réaliste.

Mots-clés: glaciologie, glaciers, calottes polaires, modélisation numérique, méthodes inverses, optimisation sous contraintes d'EDP, conservation de la masse.

Acknowledgements

I would like first to express my gratitude to my advisor and friend Éric Rignot for having accepted almost five years ago to take me in his team for an internship. With his enthusiasm and his inspiration, he makes *Glaciology* the most fascinating field of Earth sciences.

This thesis would not have been possible without my co-advisor Denis Aubry, I thank him for supervising this work, for his thoughtful advice and guidance along these years despite the geographical distance. My thanks also go to Hachmi Ben Dhia who has always been enthusiastic about this work.

I would like to thank my colleagues, Éric Larour, Hélène Seroussi and Ala Khazendar, for their superb partnership and excellent teamwork. Éric Larour took care of me since the beginning and has always been available and helpful during these three years. It is a pleasure to work with him. I also thank Daria Halkides and Nicole Schlegel, who greatly improved the quality of this thesis.

I am grateful to my family for their encouragement and to Hélène who chose to accompany me on this journey by working on her own Ph.D.

I am very grateful to École Centrale Paris and the MSSMat Laboratory for making this project possible.

Finally, I would like to thank the Jet Propulsion Laboratory, the California Institute of Technology, and the National Aeronautics and Space Administration for providing all the funds and facilities for the project.

Contents

List of Figures

List of Tables

Notations

Matrices, Vectors and scalars

Scalar variables are generally written in lower case letters, matrices in upper case bold letters and vectors in italic bold letters. The components of a vector or a matrix are written in lower case. For example: the components of a vector $\boldsymbol{X}$ with regards to some particular coordinate system are written x_i and those of a matrix $\mathbf{A}$ are a_{ij}.

We use $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$ for the three unit vectors that define a Cartesian coordinate system. Note that the suffix relates to the individual vectors and not the vector components.

Set symbols

$\partial\Omega$	:	Bounding surface of the open Ω
$\mathcal{L}^p(\Omega)$	:	Lebesgue space of p-power-integrable functions defined on Ω
$\mathcal{L}^2(\Omega)$	:	Lebesgue space of square-integrable functions defined on Ω
$\mathcal{H}^1(\Omega)$	:	Sobolev space $\mathcal{W}^{1,2}(\Omega)$
$\mathcal{M}_{m,n}$	:	Set of $m \times n$ matrices with entries in $\mathbb{R}$
$\mathcal{M}_n$	:	Set of square matrices with entries in $\mathbb{R}$ of dimension n

Algebraic symbols

$\cdot$	:	Euclidian space inner product
$\langle \cdot, \cdot \rangle$	:	Inner product of an inner product space $\mathcal{V}$
$\langle \cdot, \cdot \rangle_{\mathcal{V}',\mathcal{V}}$	:	Duality brackets of a space $\mathcal{V}$ and its dual $\mathcal{V}'$
$\mathbf{I}$	:	Identity matrix
$\mathbf{A}^T$	:	Transpose of a matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	:	Inverse of a regular square matrix $\mathbf{A}$
$\mathbf{A}^{-T}$	:	Transpose of the inverse of a regular square matrix $\mathbf{A}$
$\mathrm{Tr}(\mathbf{A})$	:	Trace of a square matrix $\mathbf{A}$
a_{ij}	:	Component of line i and column j of matrix $\mathbf{A}$
A_e	:	Effective value or second invariant of a square matrix $\mathbf{A}$
x_i	:	Component i of vector $\boldsymbol{X}$

Functional symbols

∇f	:	Gradient of the scalar function f
$\nabla \cdot \boldsymbol{v}$	:	Divergence of the vector function $\boldsymbol{v}$
$\nabla \cdot \boldsymbol{\sigma}$	:	Vector divergence of the tensor $\boldsymbol{\sigma}$
$f_{\vert_\Omega}$	:	Restriction of the function f to the domain Ω

Glacier geometry

b	:	Glacier's lower surface z-coordinate (m)
s	:	Glacier's upper surface z-coordinate (m)
H	:	Glacier's Thickness (m)

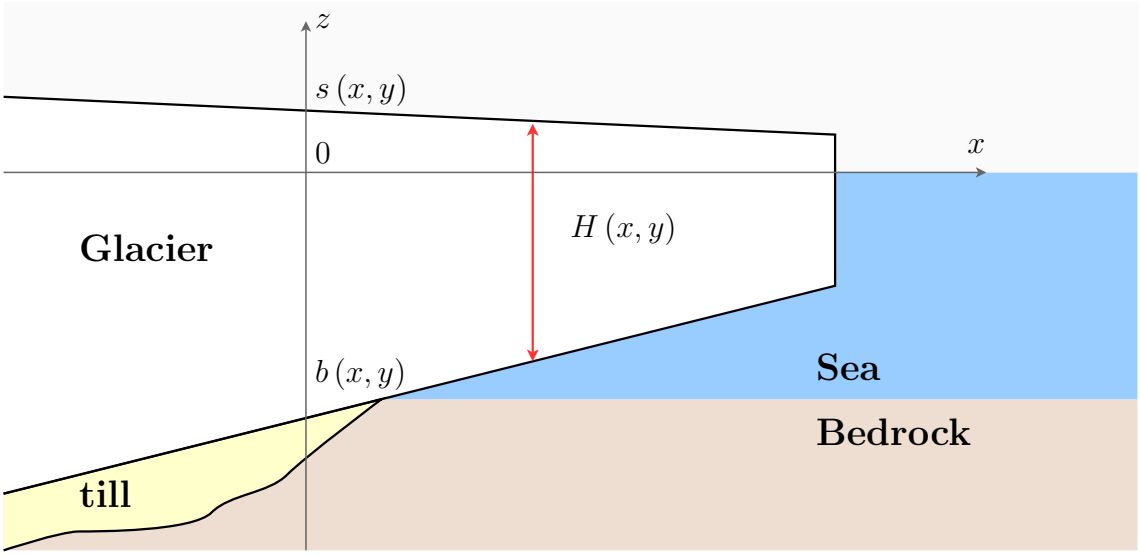


Figure 1: Schematic cross section of a glacier

Physical parameters and constants

B	:	Glen's ice hardness (Pa $a^{1/n}$)
c	:	Ice specific heat capacity (2093 J kg^{-1} K^{-1})
c_{pM}	:	Mixed layer specific heat (3974 J kg^{-1} K^{-1})
$\boldsymbol{g}$	:	Acceleration due to gravity (9.81 m s^{-1})
g	:	Acceleration due to gravity norm (9.81 m s^{-1})
G	:	Geothermal heat flux ($\sim$0.05 W m^{-2})
k	:	Basal drag parameter (Pa $^{\frac{1-r}{2}}$ $s^{\frac{s}{2}}$ $m^{-\frac{s}{2}}$)
k_{th}	:	Ice thermal conductivity (2.4 W m^{-1} K^{-1})
L	:	Ice specific latent heat of fusion (3.34 $\times$ 10^5 J kg^{-1} 3.34 $\times$ 10^5)
$\dot{M}_b$	:	Basal melting rate (m a^{-1} ice equivalent)
$\dot{M}_s$	:	Surface mass balance (m a^{-1} ice equivalent)

n	:	Glen's flow law exponent (dimensionless)
T_{pmp}	:	Pressure melting point (K)
$\boldsymbol{v}$	:	Ice velocity vector (m s^{-1})
v_x, v_y, v_z	:	velocity components (m s^{-1})
p	:	Ice pressure (Pa)
β	:	Rate of change of melting point with pressure (9.8×10^{-8} K Pa^{-1} 9.8×10^{-8})
γ	:	Thermal exchange velocity ($\sim 1.00 \times 10^{-4}$ m s^{-1})
$\dot{\boldsymbol{\varepsilon}}$	:	Strain rate tensor (s^{-1})
$\dot{\varepsilon}_e$	:	Effective strain rate tensor (s^{-1})
ρ	:	Ice density (916 kg m^{-3})
ρ_w	:	Water density (1000 kg m^{-3})
$\boldsymbol{\sigma}$	:	Cauchy stress tensor (Pa)
$\boldsymbol{\sigma}'$	:	Deviatoric stress tensor (Pa)
σ'_e	:	Effective shear stress (Pa)
$\boldsymbol{\tau}_b$	:	Friction stress (Pa)

Introduction

"The tradition of glacier studies that we inherit draws upon two great legacies of the eighteenth and nineteenth centuries: classical physics and romantic enthusiasm for Nature. This compelling mixture of the classical and romantic has drawn many of us to glaciology and brings both substance and vitality to the science. It is easy to undervalue the romantic contribution, but in glaciology it would be a mistake to do so." (Clarke, 1987)

History of ice sheet mechanics

Scientific perspectives on the nature of glaciers and their role in the Earth System have evolved dramatically over the past centuries. For example, in the middle ages, popular belief held rock crystal to be a harder form of ice (Walker and Waddington, 1988). Scientists started to give attention to glacier flow only in the eighteenth century. Johann Jacob Scheuchzer (1672 - 1733), a Swiss naturalist, proposed in 1705 that glaciers would expand due to meltwater that flows into the cracks and fissures and refreezes: the *dilation theory*. This idea was later revived by Toussaint von Charpentier in 1819. Louis Agassiz (1807 -

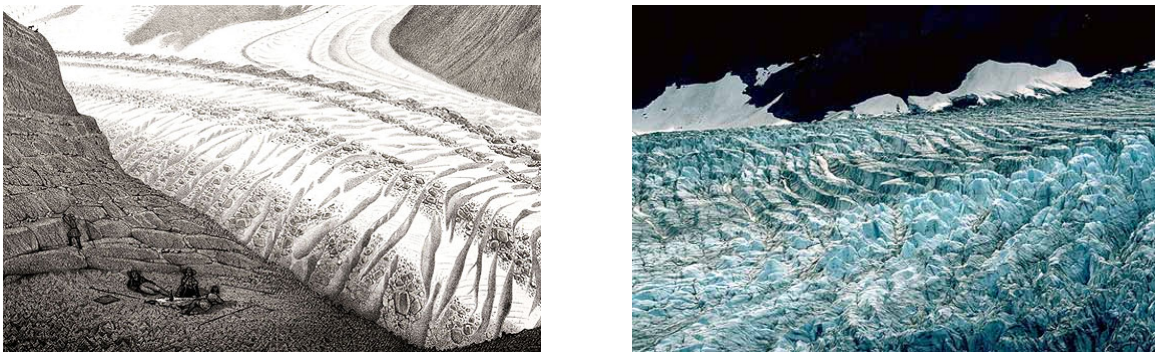


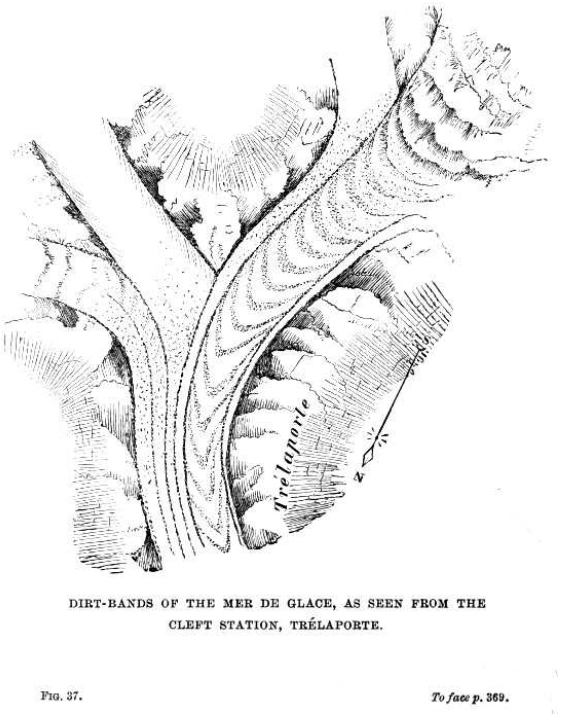
Figure 2: Fissures and cracks in the dilation theory. left: A picnic near the Zermatt Glacier, from Louis Agassiz, *Etudes sur les glaciers*, 1840. Right: Crevasses on Exit Glacier, Alaska (provided by the National Snow and Ice Data Center).

1873), a Swiss geologist, was an ardent defender of the dilation theory. In *"Etudes sur les glacier (1840)"*, he came to the following conclusions, that only careful fieldwork could prove wrong:

- The ice moves faster at the edges than at the center (greater heat from the valley walls)
- Glaciers do not move in winter (no meltwater in the cracks)

- Glaciers are reservoirs of cold (they cause meltwater to refreeze)

Another theory was developed by Altmann (1751) and Gruner (1760). They suggested that rather than advancing by dilation, a glacier slid on its bed under its own weight assisted by melting on its sides. The glacier would move forward, being stopped intermittently by obstacles, until these were overcome by melting of the ice beneath the glacier, or by the accumulative weight due to ice and snow above and behind it (Walker and Waddington, 1988).



DIRT-BANDS OF THE MER DE GLACE, AS SEEN FROM THE CLEFT STATION, TRÉLAPORTE.

Fig. 37. *To face p.* 369.

Figure 3: Ogives (arch-shaped sediment bands) or Forbes bands on Mer de glace at Chamonix. Left: schematic of Mer de glace from Tyndall (1896). Right: photo of Mer de glace (Wikipedia, 2011).

James David Forbes (1809 - 1868), a Scottish physicist, showed in 1842 that neither the sliding theory, nor the dilation theory could explain the observations. The sliding theory, for example, could not explain why glaciers could move down a shallow valley, the slope of which might be less than 3°, or, how a rigid glacier could move over an uneven bed and around corners. The dilation theory, on the other hand, was also hard to believe for Forbes, who thought the resistance to dilation must be considerable both across and along the glacier, and consequently one might expect most of the expansion on freezing to be directed upwards, towards the surface, rather than along the length of the glacier.
Several observations showed that the dilation theory was inadequate:

- The ice moves faster in the middle than on the edges (Agassiz 1841)
- Glaciers move in the winter (Forbes 1942)
- Glacial temperatures are near the melting point of water and are not reservoirs of cold (Agassiz 1840)

Forbes was the first to believe that glaciers behave like viscous bodies. But the lack of a general physical and mathematical theory of materials rheology at that time made it hard for Forbes to explain ice viscosity. It was paradoxical to allow solid ice, of brittle nature, to flow like a fluid. Thus, Forbes' theory was the source of a most bitter and acrimonious controversy (Walker and Waddington, 1988). Forbes himself gave a number of contradictory explanations for the deformation of ice. His only criterion for acceptance appears to have been that it had to be compatible with his own viscous theory. In the 1880s, it was generally admitted that ice could be slowly bent or drawn in tension without fracture, marking the threshold of present theories of glacier flow and creep of ice.

It is only in the past 50 years, by the application of modern ideas in material sciences and metallurgy that a proper understanding of the mechanism of glacier flow has been achieved (Paterson, 1994). This followed the realization that, because ice is a crystalline solid, it should deform like other crystalline solids, such as metals, at temperatures near their melting points. John W. Glen in the 1950s (Glen, 1955) derived a "flow-law" for ice from laboratory experiments. His power-law was confirmed by in-situ measurements of the closure of boreholes and tunnels (Nye, 1953). John F. Nye developed a mathematical representation of glacier flow (Nye, 1957a; Blatter et al., 2010). He extended Glen's flow law to cover multiaxial states of stress by writing Glen's flow law in tensorial form (Nye, 1957b). He became the first to apply plasticity theory to understand glacier flow: in his theory, no flow takes place until a "yield stress" is attained. The mechanisms of sliding were enunciated by Weertman (1957), who developed a sliding law for glaciers moving over a wet bed, and a flow law for floating ice spreading under gravity. These, and other glaciologists, constitute the founding generation of modern glaciology. A detailed history of modern glaciology can be found in Blatter et al. (2010).

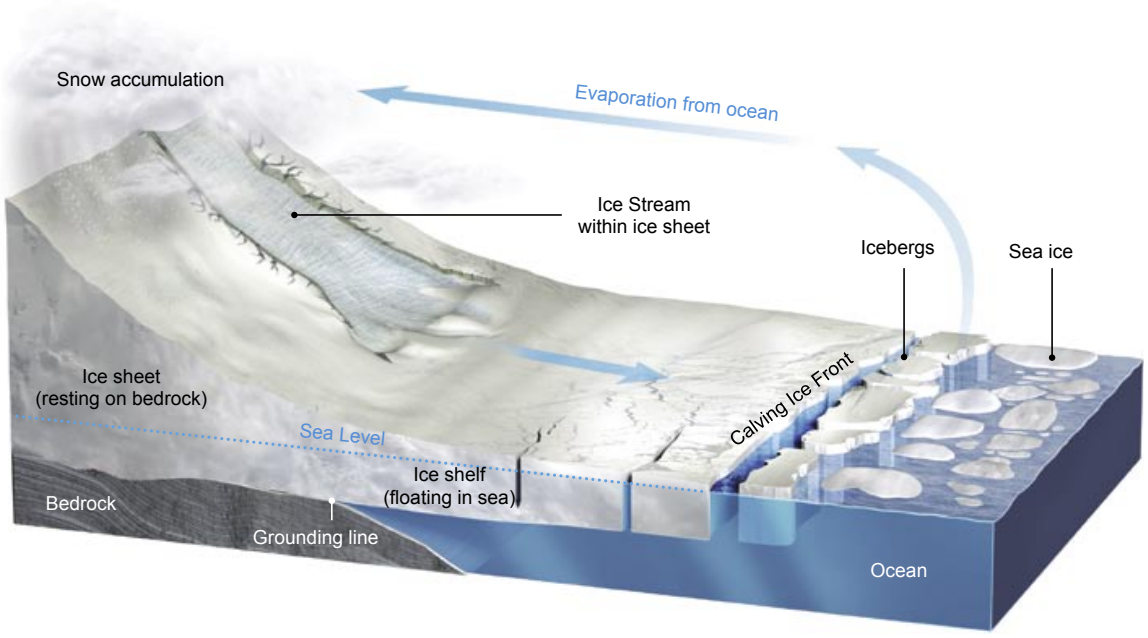


Figure 4: Schematic diagram, Illustration courtesy of Kevin Hand, Scientific American, adapted from Bell (2007)

Ice Sheet Systems

Ice sheets are masses of glacier ice that cover more than 50,000 km^2 of the Earth's surface. They are also called *continental glaciers* or *ice caps*. The only current ice sheets are confined to polar regions: the Antarctic and the Greenland ice sheets. Ice sheets are formed by

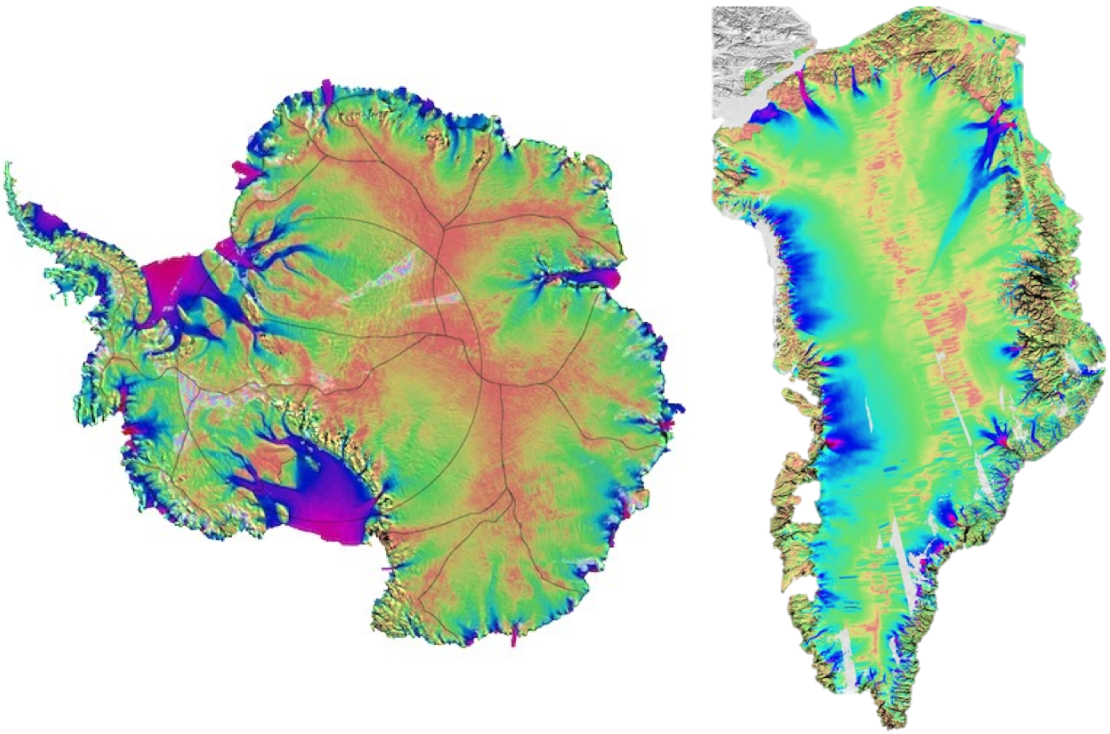


Figure 5: Antarctica and Greenland InSAR derived surface velocity, logarithmic scale (Rignot et al., 2011)

the slow transformation of snow into ice, by compaction. Solid ice then deforms under its own weight as a very viscous fluid (Fig. 4). The flow of an ice sheet is not uniform; it is organized like a hydrological system (Fig. 5), with rivers of ice channelling much of the ice flow through narrow corridors or exit gates. The glaciers typically result from the merging of many tributaries, which meander and reach far out in the interior of the ice sheet. Only a limited number of fast glaciers, called *ice streams* or *outlet glaciers*, control the ice sheet discharge into the oceans. In Greenland, a set of 30 outlet glaciers control most of the discharge. In Antarctica, 40 to 50 glaciers control most of the ice discharge. The ice may melt along low lying coastal areas and reach the ocean either to form a floating extension, ice shelf, or calve directly into the sea to form icebergs.

Ice Sheet Response To Global Warming

In the last 100 years, Earth's average surface temperature increased by about 0.8°C with about two thirds of the increase occurring over just the last three decades (IPCC-AR4, 2007). Many scientists agree that documented increases in the Earth's temperature in the past half century are a direct result of increasing concentration of manmade greenhouse gases (Hegerl

et al., 2007).

Scientific evidence supports the claim that current sea level rise is largely caused by global warming, and that sea level will continue to increase over the coming century and beyond. There are two major roughly equal contributors to sea level rise:

1. expansion of ocean water in response to higher temperatures (thermal expansion)
2. increases in the amount of meltwater and ice discharge from ice sheets.

Global warming has caused annual ice loss from the Antarctic ice sheet to surge by 75% in a decade (Rignot et al., 2008). Instead of simply adding huge volumes of meltwater to the sea, it has been shown that rising temperatures are causing glaciers, including those in Alaska, Greenland and now Antarctica, to break up and slip into the ocean at a faster rate than expected. Antarctica supports a layer of ice up to four kilometer thick and as big as the United States and Mexico combined. Antarctic ice contains 90% of the world's ice (70% of the world's fresh water).

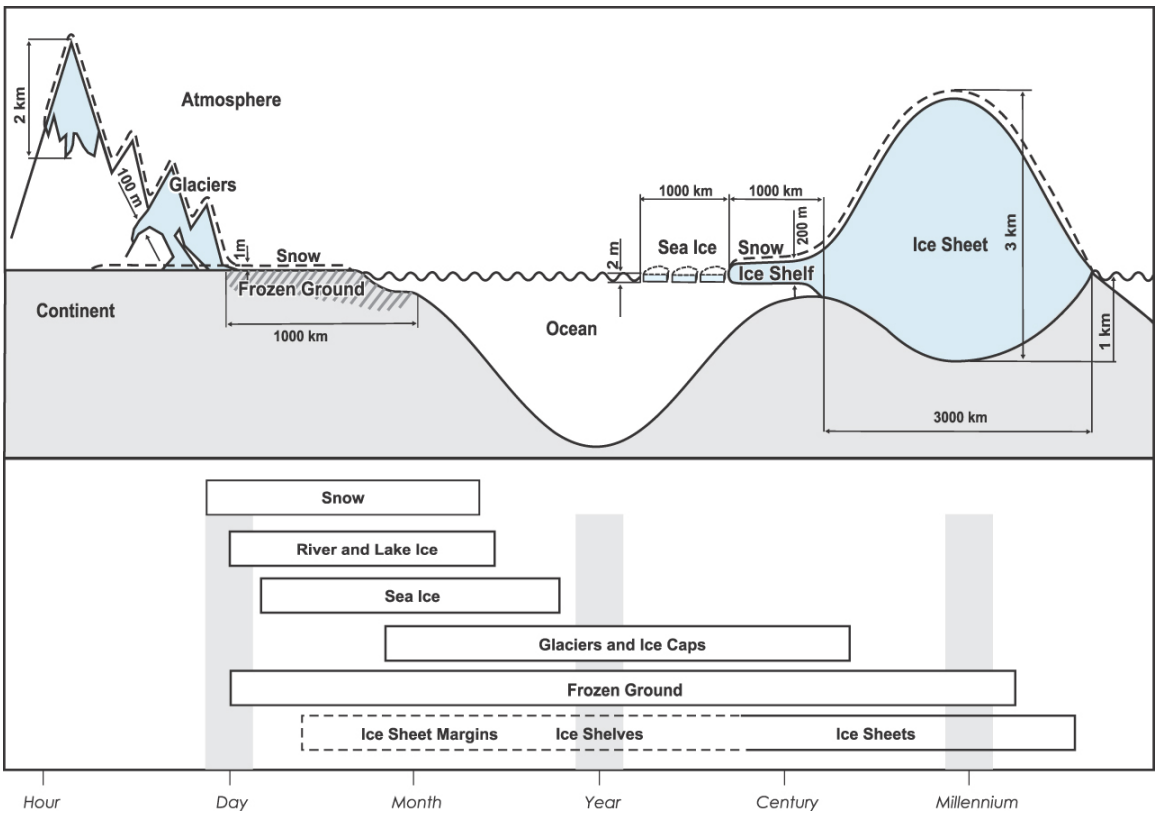


Figure 6: Components of the cryosphere and their time scales from the IPCC 4th assessment report (Lemke et al., 2007)

Fig. 6 shows all the components of the cryosphere. Even if they are all affected by global warming, the time scales at which changes can be observed are still significantly different depending on their dynamic and thermodynamic characteristics. Table 2 shows the potential of sea level rise for all components of the cryosphere.

If all the ice of Greenland and Antarctica were to melt it would cause a global sea level increase of nearly 64 meters. Although this scenario is unlikely, studies show that sea level

Ice component	**Area** (10^6 km^2)	**Ice volume** (10^6 km^3)	**Sea Level Equivalent** (m)
Snow on land	1.9 - 45.2	0.0005 - 0.005	0.001 - 0.01
Sea ice	19 - 27	0.019 - 0.025	~ 0
Glaciers and ice caps	0.51 - 0.54	0.05 - 0.13	0.15 - 0.37
Ice shelves	1.5	0.7	~ 0
Ice sheets	14.0	27.6	63.9
Greenland	1.7	2.9	7.3
Antarctica	12.3	24.7	56.6
Seasonally frozen ground	5.9 - 48.1	0.006 - 0.065	~ 0
Permafrost	22.8	0.011 - 0.037	0.03 - 0.10

Table 2: Area, volume and potential sea level equivalent of cryospheric components from the IPCC 4th assessment report (Lemke et al., 2007)

could rise between 80 cm and 2 meters by the end of the century (Pfeffer et al., 2008) threatening ecosystems, altering the flow of ocean currents and causing numerous economic, social and political severe consequences.
There are three mechanisms by which ice sheets can lose mass:

1. Direct surface melt
2. Melting at the ice-ocean interface
3. Iceberg calving

Historically, direct surface melt was thought to be the primary mechanism of ice loss. Current measurements indicate that it represents around 50% of ice loss in Greenland, and is less than 10% of the ice loss from the Antarctic ice sheet. Indeed, the surface temperatures in Antarctica are below the pressure melting point almost year round, which prevents ice from melting.

Melting at the-ice ocean interface has been shown to represent around 50% of ice loss for both continents (i.e., Rignot et al. (2008, 2010)). The remaining mechanism, iceberg discharge, represents only a small fraction of ice loss for the Greenland ice sheet, and must account for about 40% of the ice discharge from the Antarctic ice sheet.

Global warming is affecting the ice sheets in two ways: increased surface temperatures and increased ocean temperatures. The snow accumulation seems to remain unchanged (i.e. Ettema et al. (2009)).

Effect of warmer surface temperatures

An increase in air temperatures leads to enhanced surface melt. Even if the mass lost by increased melt is almost negligible, meltwater affects the ice sheet dynamics and induces mass loss by *dynamic thinning*.

Increased surface melt and drainage of supraglacial lakes increases the water content at the base of the glacier. This increased lubrication may lead to glacier speed-up, A phenomenon called *Zwally effect* (Zwally et al., 2002). The climatic significance of the effect

is controversial. Schoof (2010) for example showed that an increase in surface melt could suppress ice thinning.

Meltwater also percolates into the depths of the glaciers, possibly resulting in *cryo-hydrologic warming* (Phillips et al., 2010). This mechanism warms the entire ice column which leads to a decrease of the ice viscosity, such that the ice accelerates. This mechanism is still not well understood.

Effect of warmer ocean temperatures

An increase in ocean temperatures can have dramatic effects on the ice sheets, because the enhanced basal melting may trigger destabilizing feedbacks (Joughin and Alley, 2011). The loss of buttressing of the glacier due to ice shelf collapse, front terminus or grounding-line retreat has been observed on many glaciers (e.g. De Angelis and Skvarca (2003); Rignot et al. (2004); Nick et al. (2009)) and is responsible for most of the observed increase in ice discharge (Dupont and Alley, 2005; Gagliardini et al., 2010). It is sometimes referred to as the *Helheim effect*, after a glacier in Greenland whose ice front retreat correlates with an ice speed-up. The ice-ocean interactions play a major role in ice dynamics, since increased ocean temperatures can increase dramatically ice shelf thinning, ice front and grounding-line retreat, which results in a loss of buttressing. This loss of buttressing can lead to a significant acceleration of the tributary glaciers (Rott et al., 2002; De Angelis and Skvarca, 2003; Rignot et al., 2004). For example, Pine Island Glacier, in West Antarctica, has been thinning, accelerating and retreating since at least the 1970's (Rignot, 2008a). The ice stream has sped-up by 40 % since 1975, its grounding-line retreated inland at a rate of 1 km/yr during the 1990's, and its thickness decreased along a section extending at least 150 km upstream of the grounding-line (Bindschadler, 2002; Rignot, 2006; Thomas et al., 2011). These dramatic changes are believed to be due to warmer ocean temperatures.

Intergovernmental Panel on Climate Change

The United Nations created the Intergovernmental Panel on Climate Change (IPCC) in 1988. It is a scientific body tasked with evaluating the risk of climate change caused by human activity. The IPCC does not carry out research, nor does it monitor climate or related phenomena. A main activity of the IPCC is publishing special reports on topics relevant to the implementation of the UN Framework Convention on Climate Change (UNFCCC). The IPCC bases its assessment mainly on peer reviewed and published scientific literature. In its third Assessment Report (2001) the IPCC expected the Antarctic ice sheet to gain mass:

"The Antarctic ice sheet is likely to gain mass because of greater precipitation, while the Greenland ice sheet is likely to lose mass because the increase in runoff will exceed the precipitation increase" (IPCC-AR3, 2001).

Scientific evidence has shown however that Antarctica is losing mass, especially the West Antarctic Ice Sheet (WAIS), because of its marine-based configuration (e.g. Alley et al. (2005); Rignot and Thomas (2002); Rignot et al. (2008)). As a result, in its latest report, the IPCC stated that global climate models are fundamentally limited because they do not include changes in ice sheet dynamics:

"Future changes in the Greenland and Antarctic ice sheet mass, particularly due to changes in ice flow, are a major source of uncertainty that could increase sea level rise

projections. [...] Large-scale ocean circulation changes beyond the 21st century cannot be reliably assessed because of uncertainties in the melt-water supply from the Greenland ice sheet and model response to the warming." (IPCC-AR4, 2007).

Ice Sheet modeling at the Jet Propulsion Laboratory

The Jet Propulsion Laboratory (Pasadena, California) is a research center that builds and operates unmanned spacecraft for the National Aeronautics and Space Administration (NASA). At least 10 JPL satellites and instruments orbit Earth, returning images while monitoring atmospheric activity, and various aspects of the ocean and the cryosphere. This unmatched data collection capability is used to better understand the Earth system and the effects of climate change. Models are an important area of research at the lab, because they help define mission requirements. Models are also useful in interpreting satellite observations.

One important aspect of climate modeling is the representation of the ice sheets. The ISSM (Ice Sheet System Model) was born in 2009, as a successor to an in-house, massively parallelized, finite element software package CIELO inspired by NASTRAN (Nastran, 1972). Modeling the evolution of an ice sheet the size of a continent such as Antarctica (14,000,000 km^2) is a major challenge notably in terms of computing capability. The convergence of multiple tools (parallel solvers, dynamic memory management, object orientation, mesh anisotropy, finite element method and data assimilation), combined with the recent availability of continent-wide surface velocities (Rignot et al., 2011), are central to the ability of ISSM to perform a realistic simulation of ice flow across the entire continent of Antarctica or Greenland, at an unprecedented level of spatial detail and accuracy.

Objectives of this Thesis

Estimating current and future ice sheet contributions to sea level rise is a very active area of research. Despite recent advances in ice sheet modeling, many physical processes are still poorly understood. The aim of this work is to improve the present-day state of ice sheet models by developing calibration methods for higher-order ice sheet flow models and for reconciling datasets. Modeling Antarctica requires knowledge of several physical and geometrical parameters. Some of them can be directly measured (surface height, velocity, etc.) but not all of them (basal properties, sliding law, etc.). Inverse modeling techniques can be applied to make quantitative inferences about these characteristics. Another problem that arises is inconsistency between datasets. Low-resolution ice thickness data is usually inconsistent with high-resolution ice surface velocities (Seroussi et al., 2011), which makes ice sheet modeling difficult. This thesis addresses these two subjects: (1) the need for improved ice flow model calibration, and (2) dealing with dataset inconsistencies using optimization techniques.

This thesis is divided in six parts. The first chapter describes general concepts of continuum mechanics and inverse modeling, as a background for subsequent chapters. The second chapter derives the equations that govern ice flow from conservation laws: mass balance, momentum balance and energy balance. We present four different ice flow models: the Shallow Ice approximation, the Shelfy Stream model, a higher order model and the full-Stokes model. The third chapter deals with the numerical resolution of the models, from software architecture to mesh adaptation. This was a teamwork effort and only

the key aspects of the implementation are presented. We also show how to optimize the computational mesh using anisotropic adaptation. The fourth chapter presents how ice shelf rigidity and basal friction can be determined from InSAR-derived surface velocities. We then present an example from Pine Island Glacier, West Antarctica. We infer the basal friction for three different ice flow models corresponding to different levels of sophistication. We use the resulting patterns of basal drag to determine where each model is valid. The last chapter presents a new technique to infer ice thickness between flight track measurements using mass conservation. We also show that a straightforward calculation of the mass balance equation does not yield satisfactory results because of measurement errors.

1

Elements of continuum thermodynamics and inverse problem theory

This chapter describes general principles of continuum mechanics. We first introduce the coordinate systems that are used to describe a moving fluid: Eulerian, Lagrangian and ALE. The general conservation laws are then presented: mass conservation, balance of linear momentum, balance of angular momentum and conservation of energy. These conservation laws are very general and valid for any material body (solids and liquids). We then discuss the specific case of incompressible fluids, which is a property of ice. In a second part, we give an overview of inverse problems and present some strategies of resolution.

1.1 Coordinate systems

Two different approaches are generally adopted to describe a moving continuum. One can remain fixed in space and describe the continuum from this referential (the Eulerian approach), or one can also follow the material particles on their journey through space (the Lagrangian approach). Both approaches have their advantages and drawbacks. The Arbitrary Eulerian-Lagrangian (ALE) method (Donea and Belytschko, 1992) consists of combining the advantages of both techniques. A third referential is introduced, the reference configuration, which does not need to adhere to the material (Lagrangian) nor to be fixed in space (Eulerian).

1.1.1 Spatial, Material and Reference configurations

We introduce three different configurations corresponding to these different descriptions:

- Spatial configuration $\Omega_x \subset \mathbb{R}^3$, corresponding to the Eulerian description: each vector $\boldsymbol{x} \in \Omega_x$ represents a fixed point of the Euclidean space. Ω_x is the current domain of the continuum being described.
- Material configuration $\Omega_X \subset \mathbb{R}^3$, where each particle has a fixed material coordinate $\boldsymbol{X} \in \Omega_X$. For example, $\boldsymbol{X} \in \Omega_X$ can be taken as the coordinate of the material particle in the initial configuration, $t = 0$.
- Reference configuration $\Omega_\chi \subset \mathbb{R}^3$, where reference coordinates $\boldsymbol{\chi} \in \Omega_\chi$ are introduced to identify the grid points, which remain fixed in this configuration.

We assume that there is a one-to-one mapping between the material, the spatial and the reference configurations (Fig. 1.1). Therefore, there exist three applications:

- $\Phi : \Omega_X \times [0, T] \to \Omega_x$
- $\Psi : \Omega_\chi \times [0, T] \to \Omega_x$
- $\varphi : \Omega_X \times [0, T] \to \Omega_\chi$

that map the body from one configuration to another such that:

$$\boldsymbol{x} = \Phi(\boldsymbol{X}, t) \qquad \boldsymbol{x} = \Psi(\boldsymbol{\chi}, t) \qquad \boldsymbol{\chi} = \varphi(\boldsymbol{X}, t) \tag{1.1}$$

To ensure one-to-one mapping between these configurations, we require the determinant of the Jacobian of these transformations to be non-zero (Donea et al., 2004):

$$\det\left(\frac{\partial \Phi}{\partial \boldsymbol{X}}\right) > 0, \quad \det\left(\frac{\partial \Psi}{\partial \boldsymbol{\chi}}\right) > 0 \quad \text{and} \quad \det\left(\frac{\partial \varphi}{\partial \boldsymbol{X}}\right) > 0 \tag{1.2}$$

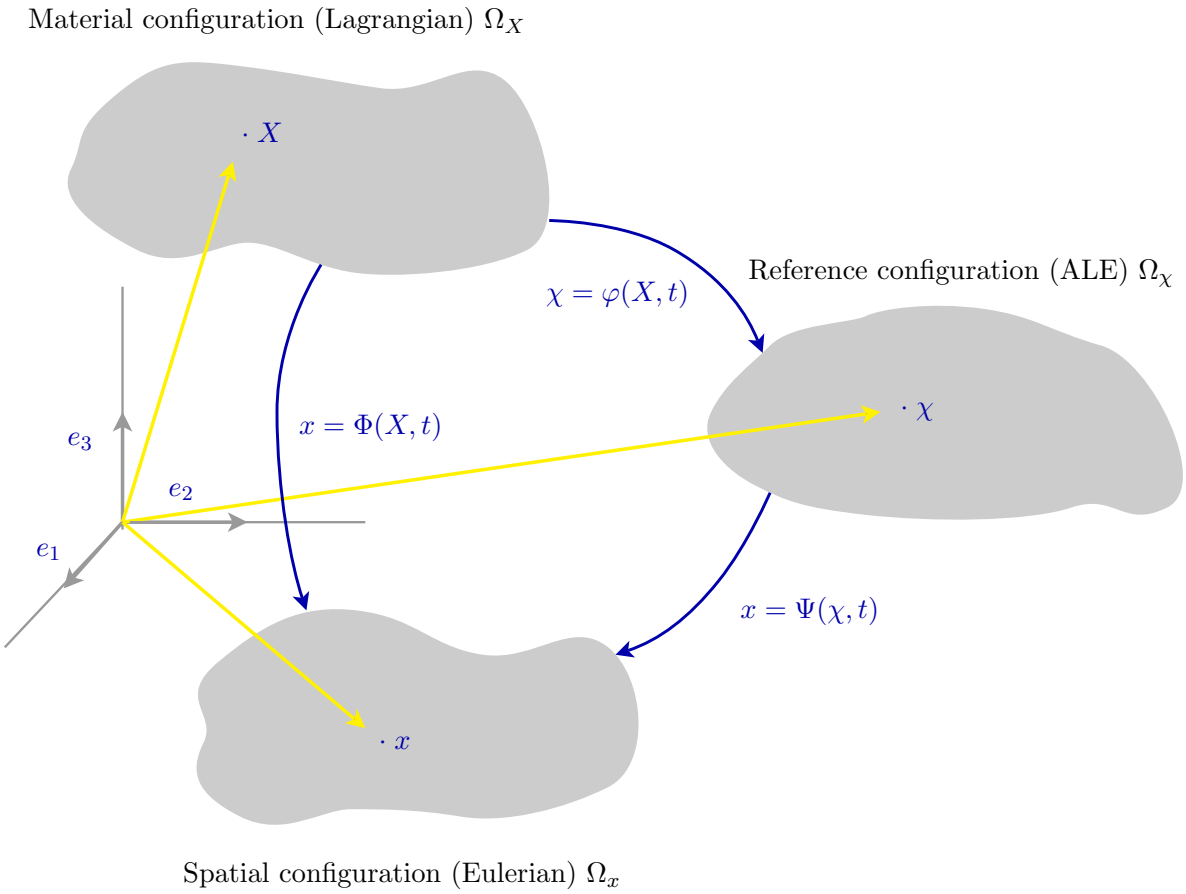


Figure 1.1: Three coordinate systems: Spatial, Material and Reference configurations

1.1.2 Material and frame velocities

One can define the material velocity, $\boldsymbol{v}$, of a particle $\boldsymbol{X}$, as the rate-of-change of the position of that particle in the spatial configuration:

Material velocity

$$\boldsymbol{v} = \frac{\partial \Phi\left(\boldsymbol{X}, t\right)}{\partial t} \tag{1.3}$$

One also defines the frame velocity, $\boldsymbol{w}$, as the velocity of any $\boldsymbol{\chi} \in \Omega_\chi$ in the spatial configuration. This velocity is sometimes referred to as the velocity of the reference coordinate or mesh velocity:

Frame velocity

$$\boldsymbol{w} = \frac{\partial \Psi\left(\boldsymbol{\chi}, t\right)}{\partial t} \tag{1.4}$$

1.1.3 Material and local time derivatives

We have to make a clear distinction between the rate-of-change of a given variable (temperature, thickness, etc.) at a specific point in space $\boldsymbol{x} \in \Omega_x$ (Eulerian approach) and the rate-of-change of that variable as we follow a particle on its journey through space (Lagrangian approach). For example, the temperature at some specific point in space may not change in time. If, however, the temperature changes from one point in space to another, the temperature of a material particle moving through that space will change with time.

A physical quantity or state variable (temperature, energy, density, etc.) is described by a function of space and time. One of three representations are used. depending on whether we are using the Eulerian, Lagrangian or ALE approach. These provide the same value at a given point in space, but require different arguments:

- $f : \Omega_x \times [0, T]$ (spatial configuration)
- $F : \Omega_X \times [0, T]$ (material configuration)
- $\tilde{f} : \Omega_\chi \times [0, T]$ (reference configuration)

The mappings between the three configurations impose:

$$f(\boldsymbol{x}, t) = F(\Phi(\boldsymbol{X}, t), t) = \tilde{f}(\Psi(\boldsymbol{\chi}, t), t) \tag{1.5}$$

The *local time derivative* is the rate-of-change at a given point in space. It is hence simply the partial derivative with respect to t:

Local time derivatives

The local time derivatives of a physical quantity in its three configurations are:

$$\begin{aligned} \frac{\partial F}{\partial t} &= \frac{\partial F(\boldsymbol{X}, t)}{\partial t} && \textit{(Lagrangian)} \\ \frac{\partial f}{\partial t} &= \frac{\partial f(\boldsymbol{x}, t)}{\partial t} && \textit{(Eulerian)} \\ \frac{\partial \tilde{f}}{\partial t} &= \frac{\partial \tilde{f}(\boldsymbol{\chi}, t)}{\partial t} && \textit{(ALE)} \end{aligned} \tag{1.6}$$

We also define the *material derivative* as the rate of change of a variable as seen by an observer following the material particle $\boldsymbol{X} \in \Omega_X$:

Material derivatives

The material derivatives of a physical quantity in its three configurations are:

$$\begin{aligned}
\frac{DF}{Dt} &= \frac{\partial F\left(\boldsymbol{X},t\right)}{\partial t} && && \textit{(Lagrangian)}\\
\frac{Df}{Dt} &= \left.\frac{\partial f\left(\Phi\left(\boldsymbol{X},t\right),t\right)}{\partial t}\right|_{\boldsymbol{X}} &= \frac{\partial f}{\partial t}+\boldsymbol{v}\cdot\nabla f && \textit{(Eulerian)}\\
\frac{D\tilde{f}}{Dt} &= \left.\frac{\partial \tilde{f}\left(\varphi\left(\boldsymbol{X},t\right),t\right)}{\partial t}\right|_{\boldsymbol{X}} &= \frac{\partial \tilde{f}}{\partial t}+\left(\boldsymbol{v}-\boldsymbol{w}\right)\cdot\nabla \tilde{f} && \textit{(ALE)}
\end{aligned} \tag{1.7}$$

Proof
See Donea and Belytschko (1992) or Donea et al. (2004). ∎

This derivative denotes the rate-of-change of the field variable as we follow the material particle through space. The material derivative is sometimes called the *total derivative* and we use a capital D symbol to distinguish it from other derivatives.

In its Eulerian form, we see that the material derivative is the sum of the local time derivative and a second term that represents the changes in f due to advection. The local time derivative accounts for all the changes in f due to other processes.

We also see that in the ALE form, if the frame velocity is zero, we recover the Eulerian case because the reference configuration remains fixed. If we take the frame velocity equal to the material velocity, we recover the Lagrangian case because the reference configuration is following the motion of the body.

1.2 Conservation laws

Certain quantities of a system are *conserved* and the conservation laws that result can be considered fundamental principles of mechanics. In mechanics, examples of conserved quantities are mass, momentum and energy. We present here the local forms of these laws, their proofs can be found in any book of continuum mechanics (e.g. Irgens (2008); Liu (2002)).

1.2.1 Mass conservation

The law of mass conservation is a fundamental concept of physics. It states that the mass of a closed system remains constant over time. A local form of mass conservation is derived:

Let $\Omega \subset \mathbb{R}^3$ (an open subset of Euclidean space) be a material body of:

- *density $\rho : \Omega \times [0,T] \to \mathbb{R}_+^*$*
- *velocity $\boldsymbol{v} : \Omega \times [0,T] \to \mathbb{R}^3$*

for a time frame $[0,T]$. The mass conservation imposes:

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T]\,, \quad \frac{D\rho}{Dt} + \rho\, \nabla \cdot \boldsymbol{v} = 0 \tag{1.8}$$

1.2.2 Balance of linear momentum

The balance of linear momentum is the generalization of Newton's second law to bodies of continuously distributed mass. The local form of linear momentum conservation requires the introduction of the stress tensor, $\boldsymbol{\sigma}$.

For the same system as above, subjected to a body force $\rho\,\boldsymbol{b} : \Omega \times [0,T] \to \mathbb{R}^3$. Let $\boldsymbol{\sigma} : \Omega \times [0,T] \to \mathcal{M}_3$ be the Cauchy stress tensor. The balance of linear momentum is:

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T]\,, \quad \rho \frac{D\boldsymbol{v}}{Dt} = \rho \boldsymbol{b} + \nabla \cdot \boldsymbol{\sigma} \tag{1.9}$$

1.2.3 Balance of angular momentum

The balance of angular momentum states that in an inertial frame the time rate-of-change of angular momentum of an arbitrary portion of a continuous body is equal to the total applied torque acting on the considered portion.

Let $\Omega \subset \mathbb{R}^3$ be a non-polar material body, and $\boldsymbol{\sigma}$ its Cauchy stress tensor. The balance of angular momentum imposes that the stress tensor be symmetric.

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T]\,, \quad \boldsymbol{\sigma} = \boldsymbol{\sigma}^T \tag{1.10}$$

1.2.4 Conservation of Energy

The last conservation law is the conservation of energy. This empirical law states that the total amount of energy in an isolated system remains constant over time. A local energy conservation equation is derived from this axiom assuming that the heat conduction follows Fourier's law.

Let $\Omega \subset \mathbb{R}^3$ be a material body of:

- *density* $\rho : \Omega \times [0, T] \to \mathbb{R}_+^*$
- *velocity* $\boldsymbol{v} : \Omega \times [0, T] \to \mathbb{R}^3$
- *temperature* $T : \Omega \times [0, T] \to \mathbb{R}_+$
- *heat capacity* $c : \Omega \times [0, T] \to \mathbb{R}_+$

If the heat conduction follows Fourier's law*, with a thermal conductivity k_{th}, and if we write Φ the local heat production, the conservation of energy imposes:*

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0, T], \quad \rho \frac{D}{Dt}(cT) = \nabla \cdot k_{th} \nabla T + \Phi \tag{1.11}$$

1.2.5 Incompressible fluids

We now examine the case of incompressible fluids. We will see in Chap. 2 that ice is treated as an incompressible material. We first need to make a distinction between incompressible materials and incompressible flow. Let's first introduce the isothermal compressibility coefficient.

Compressibility coefficient

The isothermal compressibility coefficient of a material body $\Omega_\chi \subset \mathbb{R}^3$ of:

- *density* $\rho : \Omega \times [0, T] \to \mathbb{R}_+^*$
- *pressure* $p : \Omega \times [0, T] \to \mathbb{R}$

is defined as:

$$\chi_T = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial p} \right)_T \tag{1.12}$$

χ_T is a measure of the relative volume change of a fluid or solid in response to a pressure (or mean stress) change at constant temperature. It is the inverse of the bulk modulus.

The volume of a homogeneous perfectly incompressible material does not change for any applied pressure. The compressibility coefficient is therefore zero and the bulk modulus infinite. We therefore have the following properties:

Incompressible material

For a homogeneous perfectly incompressible material in $\Omega \subset \mathbb{R}^3$, its mass density $\rho : \Omega \times [0,T] \to \mathbb{R}_+^*$ is constant throughout:

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T]\,, \quad \frac{\partial \rho}{\partial t} = 0 \quad \nabla \rho = \mathbf{0} \tag{1.13}$$

A flow is said to be incompressible if the density remain constant with motion, which is less restrictive:

Incompressible flow

The flow of a continuum is said to be incompressible if its density $\rho : \Omega \times [0,T] \to \mathbb{R}_+^*$ remains unchanged during motion. The rate-of-change of the density ρ as the motion of a material particle is followed is therefore zero:

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T]\,, \quad \frac{D\rho}{Dt} = 0 \tag{1.14}$$

The flow of an incompressible material is incompressible, but the reverse is not true: the atmosphere is generally modeled as an incompressible flow but air is obviously compressible. An important quantity to introduce is the strain rate tensor, which measures how the velocity components change in all directions.

Strain rate tensor

The strain rate tensor or stretch rate tensor, $\dot{\boldsymbol{\varepsilon}}$, associated with a velocity field $\boldsymbol{v}$ is defined as the symmetric part of the velocity gradient:

$$\dot{\varepsilon}_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{1.15}$$

where i, j are the three directions of the Cartesian space x, y and z.

The dot in $\dot{\boldsymbol{\varepsilon}}$ does not imply that the strain rate tensor is the time derivative of a quantity that is defined. It is only indicating that it is a *rate* of deformation. This tensor is sometimes referred to as $\mathbf{D}$ in the literature.

A simple condition of incompressible flow can now be derived by using the mass conservation equation (Eq. 1.8):

Condition of incompressibility

Let $\boldsymbol{v} : \Omega \times [0,T] \to \mathbb{R}^3$ *be the velocity of a material body* $\Omega \subset \mathbb{R}^3$*:*

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T], \quad \nabla \cdot \boldsymbol{v} = 0 \quad \Longleftrightarrow \quad \textit{incompressible flow} \tag{1.16}$$

This condition can be equivalently written in terms of strain rate tensor $\dot{\boldsymbol{\varepsilon}}$*:*

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0,T], \quad Tr\left(\dot{\boldsymbol{\varepsilon}}\right) = 0 \quad \Longleftrightarrow \quad \textit{incompressible flow} \tag{1.17}$$

This last equation (Eq. 1.17) shows that the strain rate tensor is only *deviatoric* for incompressible materials. This property affects the material constitutive law, which connects the stress tensor to the kinematic state of the body (e.g. strain tensor for elastic materials, strain rate tensor for viscous materials). Indeed, because one can apply any pressure to an incompressible solid without changing its shape, the stress cannot be uniquely determined from the deformation (Rajagopal and Saccomandi, 2006). To overcome this difficulty, the Cauchy stress tensor is split into a volumetric stress, σ_v, and deviatoric stress, $\boldsymbol{\sigma}'$:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' + \sigma_v \mathbf{I} \tag{1.18}$$

where $\boldsymbol{\sigma}'$ is the stress that is constitutively determined, and σ_v is the mean hydrostatic stress or the isotropic stress:

$$\sigma_v = \frac{1}{3}\mathrm{Tr}\left(\boldsymbol{\sigma}\right) \tag{1.19}$$

The isotropic stress does not work due to the incompressible nature of the material ($\mathrm{Tr}\left(\sigma_v \dot{\boldsymbol{\varepsilon}}\right) = 0$). The constitutive relation of perfectly compressible fluids only involve the deviatoric stress. The pressure, p, is introduced as a Lagrange multiplier (e.g. Hauret (2004)) to enforce the condition of incompressibility. It can be shown that this Lagrange multiplier is the mechanical pressure, which is the opposite of the isotropic stress:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' + \frac{1}{3}\mathrm{Tr}\left(\boldsymbol{\sigma}\right)\mathbf{I} = \boldsymbol{\sigma}' - p\mathbf{I} \tag{1.20}$$

Lastly, we define the effective strain rate and the effective shear stress, that will be needed for the ice behavior law. These are defined as the second invariants of their respective tensors:

$$\dot{\varepsilon}_e = \frac{1}{\sqrt{2}}\left(\sum_{i,j=1..3} \dot{\varepsilon}_{ij}^2\right)^{1/2} \qquad \sigma'_e = \frac{1}{\sqrt{2}}\left(\sum_{i,j=1..3} \boldsymbol{\sigma}_{ij}'^2\right)^{1/2} \tag{1.21}$$

All the notions introduced in this section are used in Chap. 2 to derive the system of equations that govern ice sheet dynamics. Some parameters of this system of equations cannot be measured directly and must be inferred from observations. The resolution of this inverse problem is discussed next.

1.3 Elements of inverse problem theory

Inverse problems are widely used in many fields of science to relate measured data to physically relevant model parameters. Generally speaking, inverse problems consist of recovering the input and characteristics of a system, given certain aspects of its output. We will show in the following chapters that solving inverse problems is very useful for recovering non-measurable parameters and improving dataset consistency.

1.3.1 Problem description

We present here a brief introduction to inverse problem theory. Interested readers can refer to textbooks on the topic (e.g., Tarantola (2005); Vogel (2002); Bonnet (2007); Wunsch (2006)).

Inverse problems can be defined in opposition to direct problems. In a direct problem, one computes a solution field $\boldsymbol{d} \in \mathcal{D}$ that is solution of a given set of equations completed by boundary conditions. The model equations and boundary conditions use some parameters, $\boldsymbol{p} \in \mathcal{P}$, which are known. For example, an explicit problem can be written as:

$$\boldsymbol{d} = G\left(\boldsymbol{p}\right) \tag{1.22}$$

where $G : \mathcal{P} \to \mathcal{D}$ is an operator representing the model equations.

For inverse problems, the solution is known, or partially known, by direct observations: $\boldsymbol{d} = \boldsymbol{d}^{\text{obs}}$, whereas desired model parameters, $\boldsymbol{p}$, are unknown. The inverse problem consists of finding these parameters in a control space, $\mathcal{P}$, such that the solution of the standard problem fits the observations:

$$\boldsymbol{p} = G^{-1}\left(\boldsymbol{d}^{\text{obs}}\right) \tag{1.23}$$

Unfortunately, most of the time, we do not have an explicit formulation for G^{-1}. Furthermore, these problems are generally either over-determined (a solution might not exist) or under-determined (the solution might not be unique) as $\text{Dim}(\mathcal{P}) \neq \text{Dim}(\mathcal{D})$. For example, in ice sheet modeling, we want to fit the surface velocities by changing the basal friction coefficient. For a two-dimensional mesh of N nodes, there are $2N$ observations (v_x and v_y), and only N parameters; therefore a solution to this inverse problem generally does not exist, as it is over-determined.

Because an explicit expression of G^{-1} does generally not exist, we write the inverse problem as a least square problem. We introduce a minimization problem such that the model output is as close as possible to the observations:

$$\min_{\boldsymbol{p} \in \mathcal{P}} \left\| G\left(\boldsymbol{p}\right) - \boldsymbol{d}^{\text{obs}} \right\| \tag{1.24}$$

Inverse problems are ill-posed, in opposition to a well-posed problem as defined by Hadamard (1902):

Hadamard's well-posedness

A mathematical problem is well-posed if all following conditions are satisfied:

1. *A solution exists*
2. *The solution is unique*
3. *The solution depends continuously on the data, in some reasonable topology*

The solution of an inverse problem might not exist, might not be unique and might be unstable. Hadamard considered that ill-posed problems were not physically and/or mathematically meaningful. However it turned out that the majority of the problems of mathematical geophysics are ill-posed. In the middle of the 20^{th} century Tikhonov developed the theory of ill-posed problem solutions (Tikhonov, 1943). He notably introduced regularization operators, which impose stability while recovering model parameters that do not contain more complicated features than can be justified by the data.

1.3.2 Inverse problem resolution strategy

To cover both explicit and implicit problems, we now write the model equations as a function of the model output and desired parameter:

$$G\left(\boldsymbol{d}, \boldsymbol{p}\right) = \mathbf{0} \tag{1.25}$$

We assume that this direct problem is well-posed, therefore for any model parameter $\boldsymbol{p} \in \mathcal{P}$, a solution $\boldsymbol{d}\left(\boldsymbol{p}\right) \in \mathcal{D}$ exists and is unique.

The observations are also generally only part of the model output. For example, in the context of ice sheet modeling, the model output might be the three-dimensional ice velocity. Only the two-dimensional surface velocities can be measured. A *model-equivalent* of the observation needs to be calculated to enable comparison in observation space $\mathcal{D}^{\text{obs}}$. We introduce the observation operator $H : \mathcal{D} \to \mathcal{D}^{\text{obs}}$, which provides this link to enable comparison of model output and observations. In other words, for a model output $\boldsymbol{d} \in \mathcal{D}$, the quantity $H\left(\boldsymbol{d}\right) \in \mathcal{D}^{\text{obs}}$ is the model-equivalent to the quantity being observed $\boldsymbol{d}^{\text{obs}}$. For the case we are interested in, H is a simple operator that calculates from three-dimensional velocities, the horizontal modeled surface velocities for the region where observations are available.

Solving for an inverse problem can be reduced to the resolution of a minimization problem, under the constraint that the model output must satisfy the model equations, which, in the problems that we will be interested in, are partial differential equations (PDE). We therefore need to solve a PDE-constrained optimization problem. The strategy that is adopted here is to apply a steepest-descent algorithm, where we follow the negative gradient of what is known as the cost function with respect to the parameter, $\boldsymbol{p}$.

Parametric optimization formulation

We first need to define a *cost function* or *objective function* that measures the misfit between the model output and the observations. We assume that the space of model output $\mathcal{D}$ is a

vector space equipped with a norm $\|\cdot\|$. We define a cost function as $j : \mathcal{D} \to \mathbb{R}_+$ such that:

$$j(\boldsymbol{d}) = \|H(\boldsymbol{d}) - \boldsymbol{d}^{\text{obs}}\| \tag{1.26}$$

We define the parametric optimization by seeking to minimize the cost function, j, with respect to the model parameter, $\boldsymbol{p}$, under the constraint that $\boldsymbol{d}$ satisfies the model equations. The optimization reads:

$$\min_{\boldsymbol{p}\in\mathcal{P},\, G(\boldsymbol{d},\boldsymbol{p})=\boldsymbol{0}} j(\boldsymbol{d}) \tag{1.27}$$

We introduce another cost function $\mathcal{J} : \mathcal{P} \to \mathbb{R}_+$, which is function of the model parameter, $\boldsymbol{p}$, such that:

$$\mathcal{J}(\boldsymbol{p}) = j(\boldsymbol{d}(\boldsymbol{p})) = \|H(\boldsymbol{d}(\boldsymbol{p})) - \boldsymbol{d}^{\text{obs}}\| \tag{1.28}$$

where $\boldsymbol{d}(\boldsymbol{p})$ is the solution of the model equations for the parameter $\boldsymbol{p}$. Therefore, minimizing $\mathcal{J}$ is equivalent to minimizing j under the constraint that $\boldsymbol{d}$ satisfies the model equations:

$$\min_{p\in\mathcal{P}} \mathcal{J}(\boldsymbol{p}) \Longleftrightarrow \min_{p\in\mathcal{P},\, G(\boldsymbol{d},\boldsymbol{p})=\boldsymbol{0}} j(\boldsymbol{d}) \tag{1.29}$$

Since we are trying to *minimize* $\mathcal{J}$, we define the local and global minima:

Local and global minima

$\bar{\boldsymbol{p}}$ *is a* local minimum *of* $\mathcal{J}$ *on* $\mathcal{P}$ *if and only if:*

$$\bar{\boldsymbol{p}} \in \mathcal{P}, \quad \exists \delta > 0, \quad \forall \boldsymbol{p} \in \mathcal{P}, \quad \|\boldsymbol{p} - \bar{\boldsymbol{p}}\| < \delta \Rightarrow \mathcal{J}(\boldsymbol{p}) \geq \mathcal{J}(\bar{\boldsymbol{p}}) \tag{1.30}$$

$\bar{\boldsymbol{p}}$ *is a* global minimum *of* $\mathcal{J}$ *on* $\mathcal{P}$ *if and only if:*

$$\bar{\boldsymbol{p}} \in \mathcal{P}, \quad \forall \boldsymbol{p} \in \mathcal{P}, \quad \mathcal{J}(\boldsymbol{p}) \geq \mathcal{J}(\bar{\boldsymbol{p}}) \tag{1.31}$$

Differentiability

To find the minimum of the objective function, we need to calculate the variation of the objective function, $\mathcal{J}$, with respect to changes in the model parameter, $\boldsymbol{p}$, because we want to apply a gradient-descent algorithm. Because $\mathcal{P}$ is not a finite-dimensional space *a priori*, it is essential to define the *gradient* of $\mathcal{J}$ for such spaces. Let $\mathcal{P}'$ be the dual space[1] of $\mathcal{P}$:

[1] For any vector space $\boldsymbol{V}$, the dual space $\boldsymbol{V'}$ is defined as the set of all linear functionals on $\boldsymbol{V}$ (scalar-valued linear maps on $\boldsymbol{V'}$)

Fréchet Derivative

The functional $\mathcal{J} : \mathcal{P} \to \mathbb{R}$ is Fréchet-differentiable in the neighborhood of $\boldsymbol{p} \in \mathcal{P}$ if a continuous linear functional $L : \mathcal{P} \to \mathbb{R}$ exists, such that:

$$\forall \boldsymbol{q} \in \mathcal{P} \quad \mathcal{J}(\boldsymbol{p}+\boldsymbol{q}) = \mathcal{J}(\boldsymbol{p}) + L(\boldsymbol{q}) + o(\boldsymbol{q}) \qquad \text{with} \qquad \lim_{\boldsymbol{q}\to\boldsymbol{0}} \frac{o(\boldsymbol{q})}{\|\boldsymbol{q}\|} = 0 \tag{1.32}$$

$L \in \mathcal{P}'$ is called the derivative *or the* gradient *of $\mathcal{J}$ on $\boldsymbol{p}$, and is written $\mathcal{J}'(\boldsymbol{p})$ or $\langle \mathcal{J}'(\boldsymbol{p}), \boldsymbol{q} \rangle_{\mathcal{P}',\mathcal{P}}$*

From now on, $\mathcal{P}$ is assumed to be a *Hilbert space*. We can apply the *Riesz representation theorem*, which states that for every element of $\mathcal{P}'$, dual of a Hilbert space $\mathcal{P}$, there exists a unique vector $\boldsymbol{v} \in \mathcal{P}$, such that:

$$L(\boldsymbol{q}) = \langle \boldsymbol{v}, \boldsymbol{q} \rangle \tag{1.33}$$

There are other definitions of differentiability. The Fréchet differentiability is strong because it requires a continuous linear functional that is common to every direction $\boldsymbol{q}$. The *Gâteaux* derivative is less strong because it is directional.

Gâteaux Derivative

The functional $\mathcal{J} : \mathcal{P} \to \mathbb{R}$ is Gâteaux-differentiable in the neighborhood of $\boldsymbol{p} \in \mathcal{P}$ if the following limits exist:

$$\forall \boldsymbol{q} \in \mathcal{P} \quad \lim_{\epsilon \to 0^+} \frac{\mathcal{J}(\boldsymbol{p}+\epsilon\boldsymbol{q}) - \mathcal{J}(\boldsymbol{p})}{\epsilon} = D\mathcal{J}(\boldsymbol{p}, \boldsymbol{q}) \tag{1.34}$$

$D\mathcal{J}(\boldsymbol{p}, \boldsymbol{q})$ is called the Gâteaux derivative *of $\mathcal{J}$ in the direction $\boldsymbol{q}$.*

This notion of differentiability is sometimes more convenient, because it is easier to prove Eq. (1.34) than Eq. (1.32). Besides, the two definitions are linked by the following proposition:

If $\mathcal{J} : \mathcal{P} \to \mathbb{R}$ is Fréchet-differentiable, then it is also Gâteaux-differentiable and its Fréchet and Gâteaux derivatives agree:

$$D\mathcal{J}(\boldsymbol{p}, \boldsymbol{q}) = L(\boldsymbol{q}) \tag{1.35}$$

The converse is not true.

This means that one can use Eq. (1.34) to determine the Fréchet derivative of a given function if it is Fréchet-differentiable.

Optimality conditions

Euler's Inequality: Necessary condition

If $\bar{\boldsymbol{p}} \in \mathcal{P}$, convex, is a local minimum of $\mathcal{J}$, differentiable at $\bar{\boldsymbol{p}}$, then:

$$\forall \boldsymbol{p} \in \mathcal{P} \quad \langle \mathcal{J}'(\bar{\boldsymbol{p}}), \boldsymbol{p} - \bar{\boldsymbol{p}} \rangle \geq 0 \tag{1.36}$$

Proof
For $\boldsymbol{p} \in \mathcal{P}$ and $\delta \in]0,1]$, because $\mathcal{P}$ is convex: $\bar{\boldsymbol{p}} + \delta(\boldsymbol{p} - \bar{\boldsymbol{p}}) \in \mathcal{P}$. If $\bar{\boldsymbol{p}}$ is a local minimum of $\mathcal{J}$:

$$\mathcal{J}(\bar{\boldsymbol{p}}) \leq \mathcal{J}(\bar{\boldsymbol{p}} + \delta(\boldsymbol{p} - \bar{\boldsymbol{p}})) \tag{1.37}$$

and therefore:

$$\frac{\mathcal{J}(\bar{\boldsymbol{p}} + \delta(\boldsymbol{p} - \bar{\boldsymbol{p}})) - \mathcal{J}(\bar{\boldsymbol{p}})}{\delta} \geq 0 \tag{1.38}$$

We can take the limit of the previous equation. ∎

The previous theorem only gives a *necessary* condition. In general, there is no *sufficient* condition except for the case of convex functionals. In practice, the objective function is almost never convex.

Adjoint method

We consider here an optimization under an infinite number of constraints, or, a finite number of constraints with values in a space of infinite dimensions. This is the case of PDE-constrained optimization, where a finite number of partial differential equations must be satisfied. For our purposes, this applies to the non-discretized ice flow equations, where the modeled velocity components are in an infinite-dimensional Hilbert space $\left(\mathcal{H}^1(\Omega)\right)^3$. The generic optimization problem we consider here is:

$$\min_{\boldsymbol{p} \in \mathcal{P},\, G(\boldsymbol{d},\boldsymbol{p}) = \boldsymbol{0}} j(\boldsymbol{d}) \tag{1.39}$$

where $G : \mathcal{D} \times \mathcal{P} \to \mathcal{V}$ is the operator representing the model equations with values in $\mathcal{V}$. It is convenient to introduce the Lagrangian associated with the problem:

Lagrangian

The Lagrangian *of the optimization problem (Eq. 1.39) is the functional $\mathcal{L} : \mathcal{D} \times \mathcal{V} \times \mathcal{P} \to \mathbb{R}$*

$$\mathcal{L}(\boldsymbol{d}, \boldsymbol{\lambda}, \boldsymbol{p}) = j(\boldsymbol{d}) + \langle \boldsymbol{\lambda}, G(\boldsymbol{d}, \boldsymbol{p}) \rangle \tag{1.40}$$

The new variable $\boldsymbol{\lambda} \in V$ of the system is called *Lagrange multiplier* or *adjoint state* for the constraint $G(\boldsymbol{d}, \boldsymbol{p}) = \boldsymbol{0}$. This Lagrange multiplier is introduced because it can be used

to calculate the gradient of the cost function $\mathcal{J}$. Indeed, we have the following relation:

$$\mathcal{J}(\boldsymbol{p}) = j(\boldsymbol{d}(\boldsymbol{p})) = \mathcal{L}(\boldsymbol{d}(\boldsymbol{p}), \boldsymbol{\lambda}, \boldsymbol{p}) \tag{1.41}$$

If we take the derivative of the previous equation with respect to the parameter $\boldsymbol{p}$ for a direction $\boldsymbol{q}$:

$$\begin{aligned} \langle \mathcal{J}'(\boldsymbol{p}), \boldsymbol{q} \rangle &= \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{d}}(\boldsymbol{d}(\boldsymbol{p}), \boldsymbol{\lambda}, \boldsymbol{p}), \langle \frac{\partial \boldsymbol{d}}{\partial \boldsymbol{p}}(\boldsymbol{p}), \boldsymbol{q} \rangle \rangle \\ &+ \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}}(\boldsymbol{d}(\boldsymbol{p}), \boldsymbol{\lambda}, \boldsymbol{p}), \langle \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{p}}(\boldsymbol{p}), \boldsymbol{q} \rangle \rangle \\ &+ \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{p}}(\boldsymbol{d}(\boldsymbol{p}), \boldsymbol{\lambda}, \boldsymbol{p}), \boldsymbol{q} \rangle \end{aligned} \tag{1.42}$$

The idea is to choose $\boldsymbol{\lambda}$ such that the first two derivatives are zero, and the gradient of the cost function is equal to the partial derivative of the Lagrangian with respect to the desired parameter, which is easy to compute in practice. We therefore require the Lagrangian to be *stationary* with respect to $\boldsymbol{d}$ and $\boldsymbol{\lambda}$. These two equations are sometimes referred to as the *normal equations* (Heimbach and Bugnion, 2009; Wunsch, 2006) or the Karush-Kuhn-Tucker conditions (Karush (1939); Kuhn and Tucker (1951)).

Requiring the Lagrangian to be stationary with respect to the Lagrange multiplier is equivalent to choosing $\boldsymbol{d}$ to be the solution of the model equations:

$$\forall \boldsymbol{\mu} \in \mathcal{V} \quad \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}}, \boldsymbol{\mu} \rangle = \langle \boldsymbol{\mu}, G(\boldsymbol{d}, \boldsymbol{p}) \rangle = \mathbf{0} \tag{1.43}$$

The stationarity of the Lagrangian with respect to the model output, $\boldsymbol{d}$, gives an equation on the Lagrange multiplier, $\boldsymbol{\lambda}$. They are the *Adjoint equations*:

$$\forall \boldsymbol{f} \in \mathcal{D} \quad \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{d}}, \boldsymbol{f} \rangle = \langle j'(\boldsymbol{d}), \boldsymbol{f} \rangle + \langle \boldsymbol{\lambda}, \langle \frac{\partial G}{\partial \boldsymbol{d}}, \boldsymbol{f} \rangle \rangle = \mathbf{0} \tag{1.44}$$

Adjoint and Cost function gradient

If $\boldsymbol{d}(\boldsymbol{p})$ *is solution of the model equations:*

$$G(\boldsymbol{d}, \boldsymbol{p}) = \boldsymbol{0} \tag{1.45}$$

and if $\boldsymbol{\lambda}(\boldsymbol{p})$ *is solution of the following variational formulation:*

$$\forall \boldsymbol{f} \in \mathcal{D} \quad \langle \boldsymbol{\lambda}, \langle \frac{\partial G}{\partial \boldsymbol{d}}, \boldsymbol{f} \rangle \rangle = -\langle j'(\boldsymbol{d}), \boldsymbol{f} \rangle \tag{1.46}$$

The gradient of the objective function with respect to the model parameter, a*, for a direction* b *is:*

$$\langle \mathcal{J}'(\boldsymbol{p}), \boldsymbol{q} \rangle = \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{p}}(\boldsymbol{d}(\boldsymbol{p}), \boldsymbol{\lambda}(\boldsymbol{p}), \boldsymbol{p}), \boldsymbol{q} \rangle \tag{1.47}$$

1.3.3 Steepest-descent numerical algorithms

The adjoint method allows us to compute the gradient of the cost function $\mathcal{J}$ with respect to the parameter $\boldsymbol{p}$. The minimization problem being:

$$\min_{\boldsymbol{p} \in \mathcal{P}} \mathcal{J}(\boldsymbol{p}) \tag{1.48}$$

The idea of a descent algorithm is to move from a given $\boldsymbol{p_n}$ to the next point $\boldsymbol{p_{n+1}}$ by following the line of greatest slope from an initial point $\boldsymbol{p_0} \in \mathcal{P}$:

$$\boldsymbol{p_{n+1}} = \boldsymbol{p_n} - \mu \boldsymbol{g_n} \tag{1.49}$$

where μ is a positive scalar coefficient. If μ is small enough:

$$\mathcal{J}(\boldsymbol{p_{n+1}}) = \mathcal{J}(\boldsymbol{p_n}) - \mu \langle \mathcal{J}'(\boldsymbol{p_n}), \boldsymbol{g_n} \rangle + \mathcal{O}(\mu^2) \tag{1.50}$$

The fastest way to descend is therefore to take $\boldsymbol{g_n}$ collinear to the gradient of the cost function on $\boldsymbol{p_n}$: $\mathcal{J}'(\boldsymbol{p_n})$. This method is called the gradient descent algorithm.

- **Initialization**: choose $\boldsymbol{p_0} \in V$
- **Iteration** $n \geq 0$:

$$\boldsymbol{p_{n+1}} = \boldsymbol{p_n} - \mu \frac{\mathcal{J}'(\boldsymbol{p_n})}{\|\mathcal{J}'(\boldsymbol{p_n})\|} \tag{1.51}$$

The scalar μ (the step size) must be chosen for each iteration. This can be done in several ways. The simplest is the "fixed step size algorithm" for which $\mu \in \mathbb{R}^+$ is fixed. This method is simple and easy to implement, but in most cases it is difficult to find a good step size μ.

The solution is to make it small enough to acheive a good solution, causing the model run to be computationally expensive. A better approach is the "optimized step size algorithm", where $\mu_n \in \mathbb{R}^+$ is chosen such that:

$$\mathcal{J}\left(\boldsymbol{p_{n+1}}\right) = \inf_{\mu \in \mathbb{R}^+} \mathcal{J}\left(\boldsymbol{p_n} - \mu \mathcal{J}'\left(\boldsymbol{p_n}\right)\right) \tag{1.52}$$

The minimization algorithm is presented in Fig. 1.2. We use an optimized step size projected adjoint-derived gradient descent algorithm. Each iteration involves the following steps:

1. Solve a forward problem (compute $\boldsymbol{d}$)
2. Solve the adjoint problem (compute $\boldsymbol{\lambda}$)
3. Compute the objective function gradient
4. Find a step size that gives the new distribution of the inverted parameter

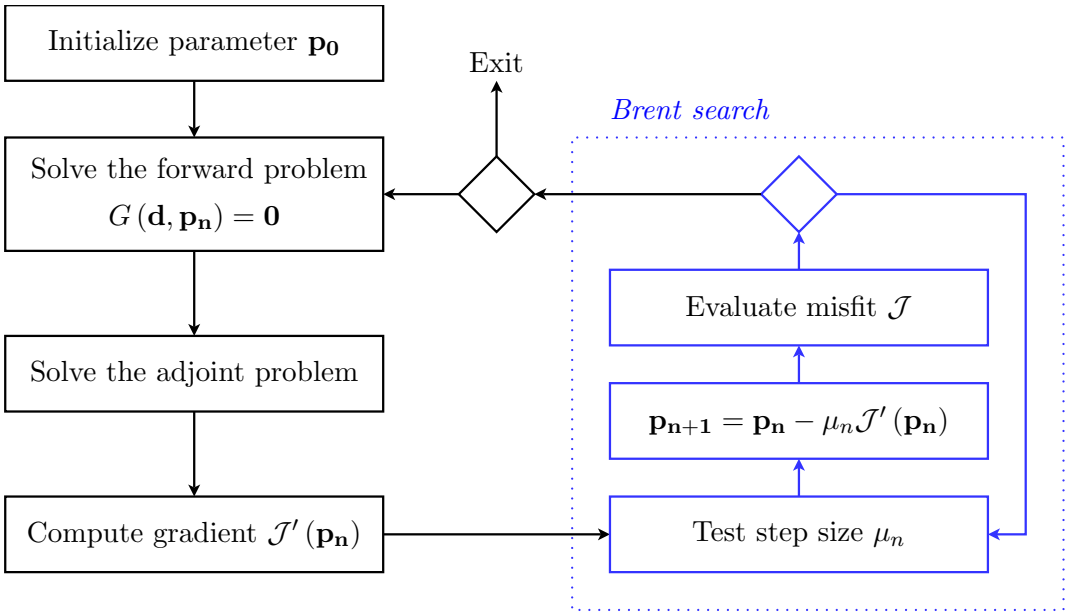


Figure 1.2: Minimization algorithm

The step size is usually estimated by a Brent search method described in App. D. We stop whenever the cost function reaches a value below a given threshold.

Very often, one wants $\boldsymbol{p} \in \mathcal{Q} \subset \mathcal{P}$, where $\mathcal{Q}$ is a non-empty bounded subset of $\mathcal{P}$. For example, we do not want the basal friction to be more than 200 kPa, which bounds the admissible space for the friction parameter. For each iteration, n, we use the operator of projection onto $\mathcal{Q}$, $P_{\mathcal{Q}}$:

$$\boldsymbol{p_{n+1}} = P_{\mathcal{Q}}\left(\boldsymbol{p_n} - \mu_n \mathcal{J}'\left(\boldsymbol{p_n}\right)\right) \tag{1.53}$$

For example, $\mathcal{P} = \mathbb{R}^M$ and we might want $\boldsymbol{p}$ to be in a bounded space $\mathcal{Q} = \prod_{i=1}^{M} [a_i, b_i]$:

$$P_{\mathcal{Q}}(\boldsymbol{p}) = \boldsymbol{p'} \quad \text{where} \quad \boldsymbol{p}'_i = \min\left(\max\left(a_i, \boldsymbol{v}_i\right), b_i\right) \tag{1.54}$$

1.3.4 Regularization

Because of its ill-posedness, the solution of the inverse problem is generally unstable: small changes in the observations, $\boldsymbol{d}^{\text{obs}}$, such as noise, can induce large changes to the solution $\boldsymbol{p}$. Tikhonov's idea was to add explicitly all *a priori* information on the model parameters to give preference to a particular solution with desirable properties (Tarantola, 2005). A regularizing operator, $\mathcal{R}$, is added to the function to be minimized:

$$\min_{\boldsymbol{p}\in\mathcal{P}} \left(\mathcal{J}(\boldsymbol{p}) + \alpha\mathcal{R}(\boldsymbol{p})\right) \tag{1.55}$$

where α is the regularization factor. For example, if $\boldsymbol{p_0}$ is an *a priori* estimate of $\boldsymbol{p}$, and we know that the solution of the inverse problem, $\boldsymbol{p}$, should not deviate too much from this initial guess, we can solve the following regularized inverse problem:

$$\min_{\boldsymbol{p}\in\mathcal{P}} \left(\mathcal{J}(\boldsymbol{p}) + \alpha\|\boldsymbol{p} - \boldsymbol{p_0}\|\right) \tag{1.56}$$

α is known as the Tikhonov parameter. A large α will lead to a solution that is close to the initial guess, but the model solution could deviate significantly from the observations, whereas a small α will give priority to fit the observations. This Tikhonov parameter must therefore be well balanced in order to find a solution that is close enough to the initial guess, while avoiding over-fitting the observations, especially when there is noise in the observations.

In some applications, one wants the solution to be smooth. One way to accomplish this is to add a Tikhonov regularization that penalizes *wiggles* in the solution:

$$\min_{\boldsymbol{p}\in\mathcal{P}} \left(\mathcal{J}(\boldsymbol{p}) + \alpha\frac{1}{2}\|\nabla\boldsymbol{p}\|^2\right) \tag{1.57}$$

Again, α needs to be adjusted so that the model fits the observations reasonably well, while forcing the parameter to be smooth enough at the same time.

Several methods have been developed to find the best value for α. One of them is Morozov's discrepancy principle (Morozov, 1984). It consists of using any α that satisfies:

$$\|H(\boldsymbol{d}) - \boldsymbol{d}^{\text{obs}}\| = \delta \tag{1.58}$$

where δ is "empirically" fixed according to the quality of the measurements and the ability of the model G to reproduce observations (Bonnet, 2007; Puel and Aubry, 2011). In this context, δ is often taken as the noise level in $\boldsymbol{d}^{\text{obs}}$. In ice sheet modeling, the noise level

in the observations is very small (a few meters per year) and δ is chosen to be orders of magnitude larger.

In practice, we employ the *L-curve analysis* (Hansen, 2000; Jay-Allemand et al., 2011) to find a Tikhonov parameter that is large enough to prevent the formation of wiggles in the solution, but small enough so that the model fits the observations well. This method is a tradeoff curve between the two quantities that both should be controlled: the cost function and the regularizing term. The L-curve analysis consists of calculating the cost function, $\mathcal{J}$, and the regularizing term, $\mathcal{R}$, for different values of α. The results are displayed on a log-log plot (Fig. 1.3).

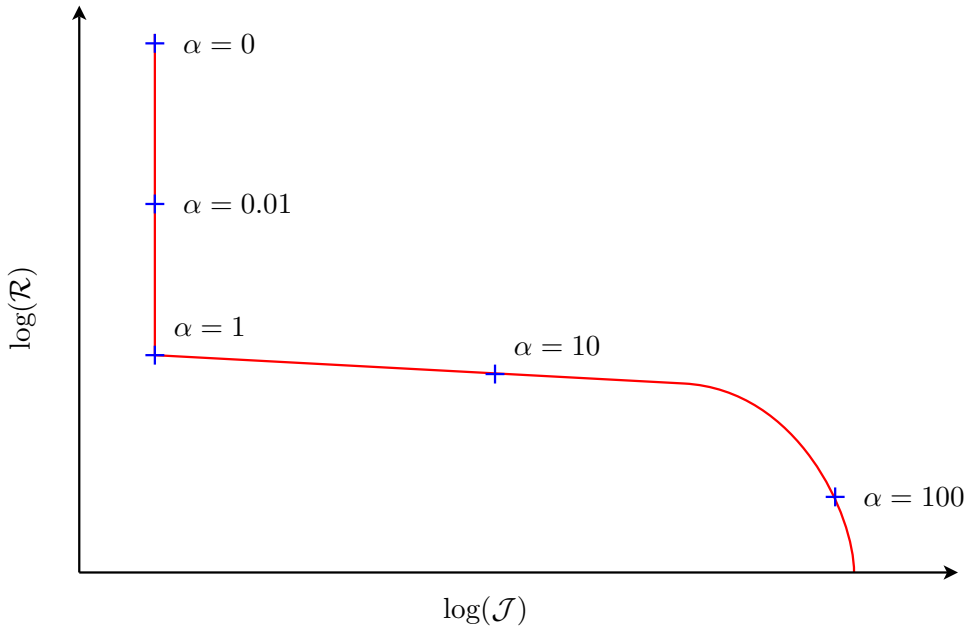


Figure 1.3: Generic L-curve analysis: log-log plot of the cost function, $\mathcal{J}$, with respect to a regularizing term, $\mathcal{R}$, for different values of Tikhonov parameter, α. The optimal value is $\alpha = 1$.

The L-curve generally presents a corner, which allows to objectively estimate the optimal regularization parameter. The corner corresponds to the maximum value of α, such that the cost function is not too much affected by the regularization. In the case of Fig. 1.3, the optimal value of α is 1.

Conclusion

In the first part of Chap. 1, we presented the different coordinate systems that are employed in the modeling of fluid flow, such as ice sheets: Eulerian, Lagrangian and ALE. We then derived all the conservation laws that govern continuum thermodynamics: the conservation of mass, angular and linear momentum, and energy. We also presented the specific case of perfectly incompressible bodies. Ice sheet thermodynamics is derived from these general principles.

In the second part this chapter, we introduced some tools to solve inverse problems using the adjoint method. Inverse problems are generally transformed into minimization problems, in which an objective function has to be optimized under constraints. These constraints are generally PDEs. We showed that the optimization of a discretized system leads to the same expression for the gradient of the cost function. Since this problem is ill-posed, we have to keep in mind that a solution might not be unique and the optimization algorithm might not converge.

2

Large-scale ice sheet thermodynamics

We now apply the general principles from Chap. 1 to ice masses. This chapter is divided into three sections. In the first section we introduce the equations of ice sheet mass balance. The second section deals with the momentum balance, the details of the approximations generally employed in ice sheet modeling, and describes four common ice sheet models of different levels of sophistication. The final section describes the thermal model from the energy balance.

2.1 Ice sheet mass balance

2.1.1 Incompressibility

Ice is treated as an incompressible material. Two effects are therefore neglected: thermal expansion (effect of temperature) and isothermal compressibility (effect of pressure). We analyze here these two effects.

Ice density is usually taken as $\rho = 917$ kg m^{-3}. This value strictly applies to temperatures near 0°C and near atmospheric pressure. This is a good assumption for mountain glaciers and the upper layers of ice sheets (Cuffey and Paterson, 2010, p.12). The linear thermal expansion of ice is approximately 5×10^{-5} °C^{-1}, so the ice density for a given temperature difference ΔT is given by:

$$\rho = \frac{917}{(1 + a\Delta T)^3} \tag{2.1}$$

In Antarctica, ice temperature can reach −40°C. Ice density can therefore increase to about 922 kg m^{-3}.

The second effect is isothermal compressibility. Indeed, the confining pressure, p, also increases ice density. The compressibility as defined in Eq. (1.12) for ice is (Cuffey and Paterson, 2010, p.12):

$$\chi_T = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial p}\right)_T \simeq 1.2 \times 10^{-10} \text{ Pa}^{-1} \tag{2.2}$$

In the center of the East Antarctic Ice Sheet, ice thickness is typically around 4 km. The pressure should thus increase the ice density from 917 to about 921 kg m^{-3}. Because ice at these depths is close to melting point, there is little temperature effect to increase the density further.

At mid-range depths in the ice sheet, where both low temperature and high pressure prevail, ice density reaches its highest value around 923 kg m^{-3} (Cuffey and Paterson, 2010, p.13). This represents a deviation of about 0.6% from a constant density of 917 kg m^{-3}. Considering ice as incompressible is therefore a good assumption for what we are interested in (Hooke (2005), p.13 and Paterson (1994)).

2.1.2 Mass balance equation

Since ice is treated as an incompressible material, its mass balance equation is reduced to the continuity equation (Eq. 1.16):

$$\forall \boldsymbol{x} \in \Omega \quad \forall t \in [0, T], \quad \nabla \cdot \boldsymbol{v} = 0 \tag{2.3}$$

2.1.3 Mass balance boundary conditions

Mass is locally added to the system and removed from system through snow fall, melting, evaporation, wind transport, or other phenomena. The *specific balance rate* represents a change of mass per unit area per unit time, and usually has units of kg m^{-2} yr^{-1}. The specific mass balance is generally subdivided into surface and basal mass balances, as englacial mass balance is negligible. We write $\dot{M}_s(x, y, t)$ the surface mass balance, in m yr^{-1} ice equivalent. Accumulation is positive and ablation negative. Similarly we write $\dot{M}_b(x, y, t)$ the melting/freezing rate, in m yr^{-1} ice equivalent. Melting is positive.

It is convenient to first define the notion of material surfaces. In the special case where the surface and basal mass balances are zero, the upper and basal surfaces of a glacier would be material surfaces:

Material surface

A material surface *is a surface which consists of the same material particles at all time.*

Material surfaces have a fundamental property, which is that the material derivative of the functional that describes their position is zero:

Material surface property

Let $S \subset \Omega$ be a surface defined as:

$$S = \{\boldsymbol{x} \in \Omega, \quad \forall t \in [0, T] \quad F(\boldsymbol{x}, t) = 0\} \tag{2.4}$$

where $F : \Omega \times [0, T] \to \mathbb{R}$. We have the following property:

$$\frac{DF}{Dt} = 0 \quad \Longleftrightarrow \quad F(\boldsymbol{x}, t) = 0 \text{ defines a material surface} \tag{2.5}$$

Proof
We use here the classical ALE notations:

(i) Consider a point on the material surface defined by (2.4) at $t = 0$. Let $X \in \Omega_X$ be the Lagrangian coordinate of the material particle at $\boldsymbol{P}$. Since this surface is a material surface, it must consist of the same particles at all times. The material particle $\boldsymbol{P}$ will then stay on the surface forever and

$$\forall t \in [0, T], \qquad F(\Phi(X, t), t) = 0 \tag{2.6}$$

The rate-of-change of F along path of $\boldsymbol{P}$ is therefore zero and this is true for any material particle on the surface. Therefore:

$$F(\boldsymbol{x}, t) = 0 \text{ defines a material surface} \quad \Longrightarrow \quad \frac{DF}{Dt} = 0 \tag{2.7}$$

(ii) Assume $DF/Dt = 0$. It follows from the definition of the material derivative that the value of $F(\Phi(X, t), t)$ does not change as some material particle $\boldsymbol{P}$ is followed. This

implies that $F(\Phi(X,t),t) = C$ where C is a constant. The set of all such material particles fulfilling $F(\boldsymbol{x},t) = C$ forms a surface for some t. The equation $F(\boldsymbol{x},t) = C$ hence defines a material surface for any value of C. Setting $C = 0$ concludes the proof. ■

Glacier surface kinematics

Let the vertical position of the glacier surface be described as a function of x and y:

$$F(x,y,z,t) = s(x,y,t) - z = 0 \tag{2.8}$$

The upper surface of a glacier *is not* a material surface as described in the previous section. It would be a material surface if there was no accumulation/ablation. If we take the surface mass balance into account, we have a kinematic boundary condition at the surface:

Kinematic boundary condition at the surface

Let $F(x,y,z,t) = s(x,y,t) - z = 0$ be the equation of the glacier's surface and $\dot{M}_s(x,y,t)$ the accumulation/ablation rate. We have the following equation:

$$\frac{\partial s}{\partial t} + v_x(s)\frac{\partial s}{\partial x} + v_y(s)\frac{\partial s}{\partial y} - v_z(s) = \dot{M}_s \tag{2.9}$$

Proof
If there was no accumulation ablation, the glacier's surface would be a material surface. Therefore we would have:

$$\frac{DF}{Dt} = \frac{\partial F}{\partial t} + (\boldsymbol{v}\cdot\nabla)F = 0 \tag{2.10}$$

This would give:

$$\frac{\partial s}{\partial t} = -v_x(s)\frac{\partial s}{\partial x} - v_y(s)\frac{\partial s}{\partial y} + v_z(s) \tag{2.11}$$

If there is accumulation and ablation, the surface rate-of-change is offset by this mass flux, and $\dot{M}_s$ must be added to $\partial s/\partial t$:

$$\frac{\partial s}{\partial t} = -v_x(s)\frac{\partial s}{\partial x} - v_y(s)\frac{\partial s}{\partial y} + v_z(s) + \dot{M}_s \tag{2.12}$$

■

where $v_x(s)$ is the x component of the velocity at the glacier surface: $v_x(s) = v_x(x,y,s(x,y))$ (ditto for y and z).

Glacier's bed kinematics

Let the vertical position of the glacier bed be described as a function of x and y:

$$F(x,y,z,t) = b(x,y,t) - z = 0 \tag{2.13}$$

The glacier's lower surface would be a material surface if there was no melting/freezing but in general this is not the case.

Kinematic boundary condition at the bed

Let $F(x,y,z,t) = b(x,y,t) - z = 0$ be the equation of the glacier's bed and $\dot{M}_b(x,y,t)$ the melting/freezing rate. We have the following equation:

$$\frac{\partial b}{\partial t} + v_x(b)\frac{\partial b}{\partial x} + v_y(b)\frac{\partial b}{\partial y} - v_z(b) = \dot{M}_b \tag{2.14}$$

Proof
Analogous argument ∎

2.2 Ice sheet momentum balance

This section presents the three ice flow models that are implemented in ISSM, from the most complex to the simplest. They all use the set of equations described in the first part of this chapter (mass balance, momentum equilibrium and constitutive relation), but have different assumptions to simplify these initial equations to different degrees. It is expected that the simpler the equations, the faster the computation time, but the less accurate the model will be in terms of evaluation of physical quantities. Other approximations of ice flow are described in Hindmarsh (2004). We here limit ourselves to three approximations: Blatter/Pattyn's Higher-order 3d Model (Blatter, 1995; Pattyn, 2003), MacAyeal/Morland's Shelfy-stream 2d model (Morland, 1987; MacAyeal, 1989) and Hutter's Shallow Ice Approximation (Hutter, 1983).

2.2.1 Quasistatic approximation (Stokes flow)

We now derive the governing equations of ice flow from the initial linear momentum balance equation (Eq. 1.9). There are two body forces acting on glaciers: the acceleration due to gravity, $\boldsymbol{g}$, and the Coriolis force due to the Earth rotation. We write $\boldsymbol{\Omega}$ the angular velocity vector of the rotating reference frame. The conservation of linear momentum becomes:

$$\rho \frac{D\boldsymbol{v}}{Dt} = \rho \left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla) \boldsymbol{v} \right) = \nabla \cdot \boldsymbol{\sigma} + \rho \, \boldsymbol{g} - 2\rho \, \boldsymbol{\Omega} \times \boldsymbol{v} \tag{2.15}$$

We aim to simplify this equation. To this end we estimate the order of magnitude of the various terms and eliminate those that are small compared to the others. We do this by scaling all the variables with their respective magnitude: for any variable a, we write $a = A_0 a^*$, where A_0 is constant and $a^* = \mathcal{O}(1)$.

$$\begin{array}{llll} \boldsymbol{v} = V_0 \, \boldsymbol{v}^* & \boldsymbol{g} = G_0 \, \boldsymbol{g}^* & \boldsymbol{r} = R_0 \, \boldsymbol{r}^* & \boldsymbol{\Omega} = \Omega_0 \, \boldsymbol{\Omega}^* \\ \boldsymbol{\sigma} = \sigma_0 \, \boldsymbol{\sigma}^* & \rho = \rho_0 \, \rho^* & t = T_0 \, t^* & \end{array} \tag{2.16}$$

where $\boldsymbol{r} = (x, y, z)$ is the position vector. Eq. (2.15) becomes:

$$\frac{\rho_0 V_0}{T_0} \, \rho^* \frac{\partial \boldsymbol{v}^*}{\partial t^*} + \frac{\rho_0 V_0^2}{R_0} \, \rho^* \, (\boldsymbol{v}^* \cdot \nabla^*) \, \boldsymbol{v}^* = \frac{\sigma_0}{R_0} \, \nabla \cdot \boldsymbol{\sigma}^* + \rho_0 G_0 \, \rho^* \boldsymbol{g}^* - 2\rho_0 \Omega_0 V_0 \, \rho^* \boldsymbol{\Omega}^* \times \boldsymbol{v}^* \tag{2.17}$$

If we multiply this equation by $\dfrac{R_0}{\rho_0 V_0^2}$

$$\frac{R_0}{T_0 V_0} \, \frac{\partial \boldsymbol{v}^*}{\partial t^*} + (\boldsymbol{v}^* \cdot \nabla^*) \, \boldsymbol{v}^* = \frac{\sigma_0}{\rho_0 V_0^2} \, \frac{1}{\rho^*} \, \nabla \cdot \boldsymbol{\sigma}^* + \frac{R_0 G_0}{V_0^2} \, \boldsymbol{g}^* - \frac{2 R_0 \Omega_0}{V_0} \, \rho^* \boldsymbol{\Omega}^* \times \boldsymbol{v}^* \tag{2.18}$$

Four dimensionless numbers can be introduced: the Strouhal number, the Reynolds number the Froude number and the Rossby number:

$$St = \frac{R_0}{T_0 V_0} \qquad Re = \frac{\rho_0 V_0^2}{\sigma_0} \qquad Fr = \frac{V_0^2}{R_0 G_0} \qquad Ro = \frac{V_0}{2\Omega_0 R_0} \tag{2.19}$$

Introducing these parameters in the last equation gives:

$$St\,\frac{\partial \boldsymbol{v}^*}{\partial t^*} + (\boldsymbol{v}^* \cdot \nabla^*)\,\boldsymbol{v}^* \;=\; \frac{1}{Re}\frac{1}{\rho^*}\,\nabla \cdot \boldsymbol{\sigma}^* + \frac{1}{Fr}\boldsymbol{g}^* - \frac{1}{Ro}\rho^* \boldsymbol{\Omega}^* \times \boldsymbol{v}^* \tag{2.20}$$

The order of magnitude of these parameters is given in table 2.1 (Reist, 2005). The Reynolds number and the Froude number are at least 10 orders of magnitude smaller than the Strouhal number for a mountain glacier, and ice sheet and an ice stream.

Variable	**Glacier**	**Ice sheet**	**Ice stream**
V_0	10^{-6}	10^{-5}	10^{-4}
G_0	10	10	10
R_0	10^4	10^6	10^5
Ω_0	10^{-4}	10^{-4}	10^{-4}
σ_0	10^5	10^5	10^5
ρ_0	10^3	10^3	10^3
T_0	R_0/V_0	R_0/V_0	R_0/V_0
St	1	1	1
Re	10^{-14}	10^{-12}	10^{-10}
Fr	10^{-17}	10^{-17}	10^{-14}
Ro	10^{-6}	10^{-7}	10^{-7}

Table 2.1: Estimated orders of magnitude of scaling parameters

The comparison of the different orders of magnitudes shows that the acceleration and inertia are negligible, even in the most extreme surges or stream flow that may occur in a glacier. This type of fluid flow is called *Stokes flow* or *creeping flow*. The Coriolis force is also negligible as it is at least seven orders of magnitude smaller than the pressure gradient. These assumptions show that the Navier-Stokes equations can be reduced to a quasi-static model:

$$\nabla \cdot \boldsymbol{\sigma}' - \nabla p + \rho\,\boldsymbol{g} = \boldsymbol{0} \tag{2.21}$$

2.2.2 Ice Constitutive Equations: Glen-Nye flow law

Ice is considered a perfectly isotropic incompressible viscous material (Hooke (2005) p.13, Paterson (1994)). The stress tensor is therefore expressed as a function of the strain rate

and the pressure:

$$\boldsymbol{\sigma} = 2\mu\dot{\boldsymbol{\varepsilon}} - p\mathbf{I} \tag{2.22}$$

where the viscosity μ is a scalar, because ice is assumed to be isotropic. It would be a fourth-order tensor for anisotropic materials.The most common flow law is Glen's flow law, based on Glen's experiments (Glen, 1955). He used the form:

$$\dot{\varepsilon}_e = \left(\frac{\sigma'_e}{B}\right)^n \tag{2.23}$$

where B is a viscosity parameter, which increases as the ice becomes stiffer ($MPa.a^{\frac{1}{n}}$), n is empirically determined (most studies have found that $n \simeq 3$ (Hooke (2005) p.15, Paterson (1994) p.86 and Cuffey and Paterson (2010) p.55) and is known as the flow law exponent. $\dot{\varepsilon}_e$ and σ'_e are the effective strain rate and the effective shear stress (Eq. 1.21).

Nye (1957b) extended this flow law to cover multiaxial states of stress by writing Glen's flow law in tensororial form. Since ice is assumed to be isotropic, the principal axes of the deviatoric stress $\boldsymbol{\sigma}'$ and the strain rate tensor $\dot{\boldsymbol{\varepsilon}}$ coincide. The ice that constitutes glaciers is neither isotropic nor perfectly incompressible, but this approximation is a convenient starting point for calculations of glacier flow. If the principal axes of deviatoric stress and strain rate coincide, the flow law becomes (Hooke, 2005, p.269):

$$\dot{\boldsymbol{\varepsilon}} = \frac{\sigma'^{n-1}_e}{B^n}\boldsymbol{\sigma}' \tag{2.24}$$

Then, combining Glen's flow law to eliminate $\boldsymbol{\sigma}_e$, we obtain:

$$\dot{\boldsymbol{\varepsilon}} = \frac{\dot{\varepsilon}_e^{\frac{n-1}{n}}}{B}\boldsymbol{\sigma}' \tag{2.25}$$

The viscosity, μ (Eq. 2.22), is therefore defined by:

$$\mu = \frac{B}{2\left(\dot{\varepsilon}_e^{1-\frac{1}{n}}\right)} \tag{2.26}$$

This flow law is widely employed in ice sheet modeling (Cuffey and Paterson, 2010, p.61). Nevertheless, ice is not an isotropic material because the viscosity of polycrystalline ice depends strongly on the c-axis fabric, which in glaciers often displays some degree of preferred orientation (Cuffey and Paterson, 2010, p.78). This is ignored by nearly all theoretical analyses of glacial flow. Some anisotropic laws are being developed (e.g. Gillet-Chaulet et al. (2005); Ma et al. (2010); Thorsteinsson (2001)), but are beyond the scope of this thesis. For the remainder of this thesis, we will always rely of Glen's flow law, which provides a good first estimate of ice behavior.

2.2.3 Mechanical boundary conditions

The mass conservation, the momentum balance and the constitutive relation of ice give a set of equations that completely describe the ice sheet system dynamics. However, one needs boundary conditions to constrain these equations. There are many boundary conditions depending on the interface: ice/atmosphere boundary (air pressure), ice/ocean boundary (water pressure) and ice till boundary (friction). These boundary conditions are described in this section.

Ice/Atmosphere boundary

Since atmospheric pressure, p_0, is negligible compared to the ice lithostatic pressure, the upper surface of the glacier is considered to be a free surface.

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = -p_0 \boldsymbol{n} \simeq \mathbf{0} \tag{2.27}$$

Ice/Water boundary

Water applies pressure on the ice front equal to $p_w = -\rho_w g z$:

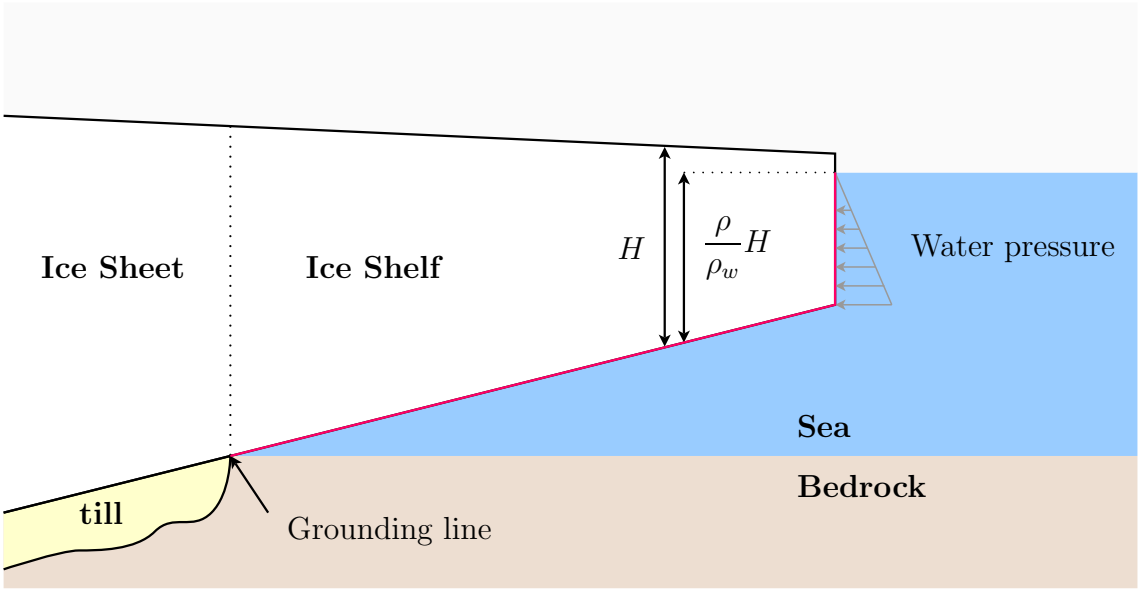


Figure 2.1: Water pressure on the front of an ice shelf

The boundary condition is thus:

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = -p_w \boldsymbol{n} \tag{2.28}$$

Ice/Till interface

The Ice/Till interface is constrained by two boundary conditions. The first one is a Neumann boundary condition describing basal friction (tangent to the interface) and the other one is

a Dirichlet boundary condition that ensures a non-interpenetrating condition between the ice and the bedrock following Eq. (2.14):

$$\boldsymbol{v} \cdot \boldsymbol{n} = \dot{M}_b \tag{2.29}$$

Basal friction is a fundamental control on ice stream dynamics that remains poorly understood. Generally, a viscous friction law is employed (Paterson (1994) p.151 and Cuffey and Paterson (2010) p.240):

$$\|\boldsymbol{v}_b\| = kN^{-q}\|\boldsymbol{\tau}_b\|^p \tag{2.30}$$

- $\boldsymbol{v}_b$ is the velocity component, tangential to the bedrock surface
- $N = g\rho H + \rho_w b$ is the effective pressure at the base, b.
- $\boldsymbol{\tau}_b$ is the friction stress component, tangential to the bedrock surface: $\boldsymbol{\tau}_b = \boldsymbol{\sigma} \cdot \boldsymbol{n} - \sigma_{nn}\boldsymbol{n}$
- k, q and p are positive constants

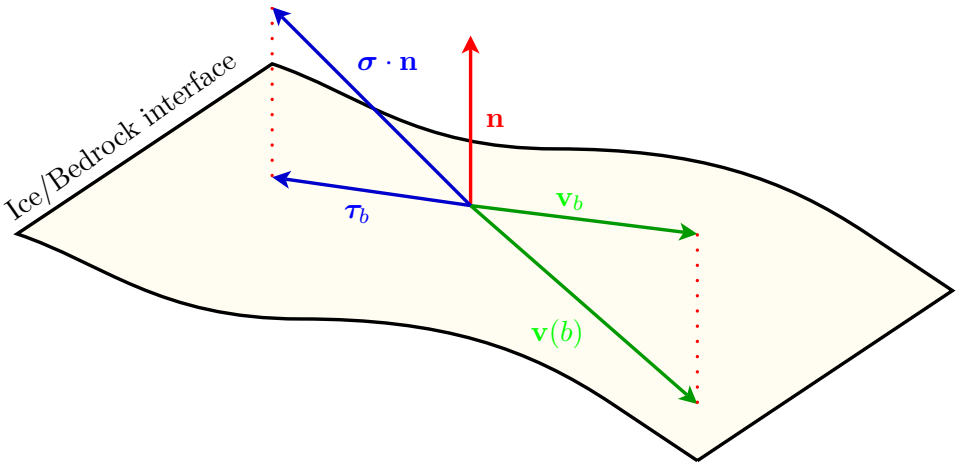


Figure 2.2: Ice/Till interface geometry

Weertman's relation is a special case of this law for which $p = 3$ and $q = 1$. This friction law may be written in vector form:

$$\boldsymbol{\tau}_b \quad = \quad -k^2 N^r \|\boldsymbol{v}\|^{s-1}\boldsymbol{v}_b \quad \stackrel{def}{=} \quad -\alpha^2 \boldsymbol{v}_b \tag{2.31}$$

with $r = q/p$ and $s = 1/p$. We use k^2 to be sure that this quantity is positive (i.e. stress opposes the motion).

2.2.4 Full Stokes Model (1845)

The full Stokes model (FS) consists in resolving Eq. (2.21) without approximation. It is a 3d incompressible ice flow model. The pressure p becomes a variable like the velocity $\boldsymbol{v} = (v_x, v_y, v_z)$.

Full Stokes Approximation

With no further approximation, Eq. (2.21) is:

$$\begin{cases} \frac{\partial}{\partial x}\left(2\mu\frac{\partial v_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_x}{\partial z} + \mu\frac{\partial v_z}{\partial x}\right) - \frac{\partial p}{\partial x} = 0 \\ \frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(2\mu\frac{\partial v_y}{\partial y}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_y}{\partial z} + \mu\frac{\partial v_z}{\partial y}\right) - \frac{\partial p}{\partial y} = 0 \\ \frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial z} + \mu\frac{\partial v_z}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial v_y}{\partial z} + \mu\frac{\partial v_z}{\partial y}\right) + \frac{\partial}{\partial z}\left(2\mu\frac{\partial v_z}{\partial z}\right) - \frac{\partial p}{\partial z} - \rho g = 0 \end{cases} \tag{2.32}$$

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0$$

2.2.5 Blatter/Pattyn's Higher-order 3d Model (1995/2003)

Blatter (1995) introduced a three-dimensional higher-order ice sheet model (HO), that has been written in terms of velocities by Pattyn (2003). It is a numerical model in which the vertical velocity is deduced from the horizontal components of velocity that are solved independently, which makes this model much less computationally intensive than solving for the Full-Stokes equations. The two main assumptions of this model are:

1. the horizontal gradients of the vertical velocity are small compared to the vertical gradient of the horizontal velocity,
2. Bridging effects (described below) are negligible.

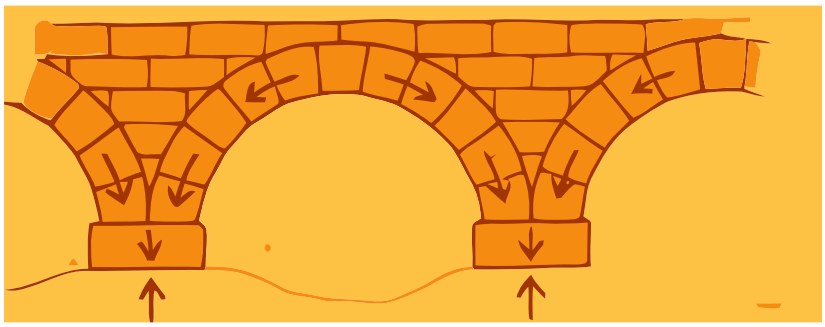


Figure 2.3: Bridging effect, arch bridges use a semicircular structure to push the weight outward along the arch to the abutments

The bridging effect is shown in Fig. 2.3. In the case of bridges, the weight of the arch is pushed outwards. The bridge piers support their own weight and the weight of the arch. The variations in vertical normal stress lead to shear-stress gradients. A paper by van der Veen and Whillans (1989) suggests that bridging effects could be important at distances that are shorter than a few ice thickness at most. At small scales, part of the glacier can act like a bridge in which the vertical stress at the underside of the bridge span is less than the weight of ice above, while that under the abutment exceeds the weight of material above.

Neglecting the bridging effect for the ice means that the vertical normal stress is purely lithostatic, or equal to the weight of ice vertically above. If the Bridging effect (van der Veen and Whillans, 1989) is negligible, the vertical equation of the momentum balance (Eq. 2.21) is simplified to:

$$\frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z} - \rho g = 0 \tag{2.33}$$

Blatter/Pattyn's Higher-order Approximation

Under the following assumptions:

- $\frac{\partial \boldsymbol{\sigma}_{xz}}{\partial x} \ll \frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z}$
- $\frac{\partial \boldsymbol{\sigma}_{yz}}{\partial y} \ll \frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z}$
- $\frac{\partial v_z}{\partial x} \ll \frac{\partial v_x}{\partial z}$
- $\frac{\partial v_z}{\partial y} \ll \frac{\partial v_y}{\partial z}$

Eq. (2.21) is:

$$\begin{cases} \frac{\partial}{\partial x}\left(4\mu\frac{\partial v_x}{\partial x} + 2\mu\frac{\partial v_y}{\partial y}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_x}{\partial z}\right) = \rho g\frac{\partial s}{\partial x} \\ \frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(4\mu\frac{\partial v_y}{\partial y} + 2\mu\frac{\partial v_x}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_y}{\partial z}\right) = \rho g\frac{\partial s}{\partial y} \\ v_z(x,y,z) = v_z(x,y,b) - \int_{b(x,y)}^{z} \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} dz' \end{cases} \tag{2.34}$$

Proof

With the first and second assumptions (bridge effect negligible), the third equation of Eq. (2.32) reduces to:

$$\frac{\partial}{\partial z}\left(2\mu\frac{\partial v_z}{\partial z}\right) - \frac{\partial p}{\partial z} - \rho g = 0 \tag{2.35}$$

If we integrate this equation from z to the surface s, we obtain:

$$p\left(x,y,s\left(x,y\right)\right)-p\left(x,y,z\right)=2\,\mu\frac{\partial v_z}{\partial z}\bigg|_s-2\,\mu\frac{\partial v_z}{\partial z}\bigg|_z-\rho g\left(s-z\right) \tag{2.36}$$

The boundary conditions at the ice/atmosphere interface Eq. (2.27) impose:

$$\boldsymbol{\sigma}_{zz}\left(x,y,s\left(x,y\right)\right)=2\,\mu\frac{\partial v_z}{\partial z}\bigg|_s-p\left(x,y,s\left(x,y\right)\right)\simeq 0 \tag{2.37}$$

Finally the pressure is:

$$p\left(x,y,z\right)=2\mu\frac{\partial v_z}{\partial z}+\rho g\left(s-z\right) \tag{2.38}$$

With this approximation, full Stokes equations (Eq. 2.32) become:

$$\begin{cases}\dfrac{\partial}{\partial x}\left(2\mu\dfrac{\partial v_x}{\partial x}\right)+\dfrac{\partial}{\partial y}\left(\mu\dfrac{\partial v_x}{\partial y}+\mu\dfrac{\partial v_y}{\partial x}\right)+\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial v_x}{\partial z}-\mu\dfrac{\partial v_z}{\partial x}\right)=\rho g\dfrac{\partial s}{\partial x}\\[2ex] \dfrac{\partial}{\partial x}\left(\mu\dfrac{\partial v_x}{\partial y}+\mu\dfrac{\partial v_y}{\partial x}\right)+\dfrac{\partial}{\partial y}\left(2\mu\dfrac{\partial v_y}{\partial y}\right)+\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial v_y}{\partial z}-\mu\dfrac{\partial v_z}{\partial y}\right)=\rho g\dfrac{\partial s}{\partial y}\\[2ex] p(z)=2\mu\dfrac{\partial v_z}{\partial z}+\rho g\left(s-z\right)\end{cases} \tag{2.39}$$

$$\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}=0$$

Lastly, the third and fourth assumptions simplify the first two equations:

$$\begin{cases}\dfrac{\partial}{\partial x}\left(2\mu\dfrac{\partial v_x}{\partial x}\right)+\dfrac{\partial}{\partial y}\left(\mu\dfrac{\partial v_x}{\partial y}+\mu\dfrac{\partial v_y}{\partial x}\right)+\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial v_x}{\partial z}\right)=\rho g\dfrac{\partial s}{\partial x}\\[2ex] \dfrac{\partial}{\partial x}\left(\mu\dfrac{\partial v_x}{\partial y}+\mu\dfrac{\partial v_y}{\partial x}\right)+\dfrac{\partial}{\partial y}\left(2\mu\dfrac{\partial v_y}{\partial y}\right)+\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial v_y}{\partial z}\right)=\rho g\dfrac{\partial s}{\partial y}\\[2ex] p(z)=2\mu\dfrac{\partial v_z}{\partial z}+\rho g\left(s-z\right)\end{cases} \tag{2.40}$$

$$\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}=0$$

The vertical component of the velocity is deduced from the incompressibility equation, which is integrated from b to z. ■

2.2.6 MacAyeal/Morland's Shelfy-stream 2d model (1989)

This model was first introduced for an unconfined ice shelf by Morland (1987), and extended by MacAyeal (1989) for ice flow over a viscous basal sediment. It is usually referred to as the *Shelfy-Stream* or *Shallow-Shelf* approximation (SSA). It is a 2d model that includes Blatter/Pattyn's assumptions, and, the horizontal components of the velocity are assumed to be constant with depth. The basal drag associated with deforming basal sediment is assumed not to induce significant vertical gradients of the horizontal velocity.

MacAyeal's Shelfy-stream Approximation

Under the following assumptions:

- $\frac{\partial \boldsymbol{\sigma}_{xz}}{\partial x} \ll \frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z}$
- $\frac{\partial \boldsymbol{\sigma}_{yz}}{\partial y} \ll \frac{\partial \boldsymbol{\sigma}_{zz}}{\partial z}$
- $\frac{\partial v_x}{\partial z} = 0$
- $\frac{\partial v_y}{\partial z} = 0$
- $\dot{\varepsilon}_{xz} = 0$
- $\dot{\varepsilon}_{yz} = 0$
- *bedrock slope negligible*

Eq. (2.21) *is:*

$$\begin{cases} \frac{\partial}{\partial x}\left(4H\bar{\mu}\frac{\partial v_x}{\partial x} + 2H\bar{\mu}\frac{\partial v_y}{\partial y}\right) + \frac{\partial}{\partial y}\left(H\bar{\mu}\frac{\partial v_x}{\partial y} + H\bar{\mu}\frac{\partial v_y}{\partial x}\right) = \rho g H\frac{\partial s}{\partial x} + \alpha^2 v_x \\ \frac{\partial}{\partial y}\left(4H\bar{\mu}\frac{\partial v_y}{\partial y} + 2H\bar{\mu}\frac{\partial v_x}{\partial x}\right) + \frac{\partial}{\partial x}\left(H\bar{\mu}\frac{\partial v_x}{\partial y} + H\bar{\mu}\frac{\partial v_y}{\partial x}\right) = \rho g H\frac{\partial s}{\partial y} + \alpha^2 v_y \end{cases} \tag{2.41}$$

Where $\bar{\mu}$ *is the depth-averaged non-linear viscosity defined as:* $H\overline{\mu} = \int_b^s \mu dz$, *and* α *the friction coefficient (Eq. 2.31).*

Proof

These assumptions are stronger than Blatter/Pattyn's. Therefore, we can use Pattyn's higher order model equations (Eq. 2.34) as a starting point. If we use the first two assumptions, the first equations of Pattyn's model become:

$$\begin{cases} \frac{\partial}{\partial x}\left(4\mu\frac{\partial v_x}{\partial x} + 2\mu\frac{\partial v_y}{\partial y}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) = \rho g\frac{\partial s}{\partial x} \\ \frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(4\mu\frac{\partial v_y}{\partial y} + 2\mu\frac{\partial v_x}{\partial x}\right) = \rho g\frac{\partial s}{\partial y} \end{cases} \tag{2.42}$$

One can integrate these equations from the bed $b(x,y)$ to the surface $s(x,y)$:

$$\begin{cases} \displaystyle\int_{b(x,y)}^{s(x,y)} \left(\frac{\partial}{\partial x}\left(4\mu\frac{\partial v_x}{\partial x} + 2\mu\frac{\partial v_y}{\partial y}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_x}{\partial z}\right)\right) dz = \rho g H \frac{\partial s}{\partial x} \\ \displaystyle\int_{b(x,y)}^{s(x,y)} \left(\frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(4\mu\frac{\partial v_y}{\partial y} + 2\mu\frac{\partial v_x}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu\frac{\partial v_y}{\partial z}\right)\right) dz = \rho g H \frac{\partial s}{\partial y} \end{cases}$$

We can use Leibniz integral rule (Eq. C.2), which gives:

$$\begin{aligned} &\frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} \left(4\mu\frac{\partial v_x}{\partial x} + 2\mu\frac{\partial v_y}{\partial y}\right) dz + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} \left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) dz \\ &\quad - \sigma'_{xx}(s)\frac{\partial s}{\partial x} - \sigma'_{xy}(s)\frac{\partial s}{\partial x} + \sigma'_{xz}(s) + \sigma'_{xx}(b)\frac{\partial b}{\partial x} + \sigma'_{xy}(b)\frac{\partial b}{\partial x} - \sigma'_{xz}(b) = \rho g H \frac{\partial s}{\partial x} \end{aligned} \tag{2.43}$$

$$\begin{aligned} &\frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} \left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) dz + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} \left(4\mu\frac{\partial v_y}{\partial y} + 2\mu\frac{\partial v_x}{\partial x}\right) dz \\ &\quad - \sigma'_{yx}(s)\frac{\partial s}{\partial x} - \sigma'_{yy}(s)\frac{\partial s}{\partial x} + \sigma'_{yz}(s) + \sigma'_{yx}(b)\frac{\partial b}{\partial x} + \sigma'_{yy}(b)\frac{\partial b}{\partial x} - \sigma'_{yz}(b) = \rho g H \frac{\partial s}{\partial y} \end{aligned} \tag{2.44}$$

The boundary conditions at the ice/atmosphere interface Eq. (2.27) imposes:

$$\boldsymbol{\sigma}'(s)\cdot\boldsymbol{n} = \mathbf{0} \tag{2.45}$$

If we assume that the bedrock is horizontal, the boundary conditions at the ice/till interface is:

$$\begin{aligned} \left(\boldsymbol{\sigma}'(b)\cdot\boldsymbol{n}\right)_x &= -\alpha^2 v_x \\ \left(\boldsymbol{\sigma}'(b)\cdot\boldsymbol{n}\right)_y &= -\alpha^2 v_y \end{aligned} \tag{2.46}$$

Eq. (2.43) and Eq. (2.44) become:

$$\begin{cases} \displaystyle\frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} \left(4\mu\frac{\partial v_x}{\partial x} + 2\mu\frac{\partial v_y}{\partial y}\right) dz + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} \left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) dz = \rho g H \frac{\partial s}{\partial x} + \alpha^2 v_x \\ \displaystyle\frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} \left(\mu\frac{\partial v_x}{\partial y} + \mu\frac{\partial v_y}{\partial x}\right) dz + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} \left(4\mu\frac{\partial v_y}{\partial y} + 2\mu\frac{\partial v_x}{\partial x}\right) dz = \rho g H \frac{\partial s}{\partial y} + \alpha^2 v_y \end{cases}$$

One of the assumptions here is that horizontal velocities (v_x and v_y) do not depend on z: there is no need to use Leibniz integral law here. ∎

2.2.7 Hutter's Shallow Ice Approximation (1983)

Hutter (1983) introduced the most widely used ice sheet flow model: the Shallow Ice Approximation (SIA). It is the most computationally efficient model, but it is also one of the simplest because of the number of assumptions that are made as only vertical shear is taken into account. SIA is also derived from Blatter/Pattyn's model, but the assumptions are opposed to the one of MacAyeal/Morland's.

Shallow Ice Approximation

Under the following assumptions:

- *The only components of the stress tensor that are not negligible are* σ'_{xz} *and* σ'_{yz}
- $\frac{\partial v_z}{\partial x} \ll \frac{\partial v_x}{\partial z}$
- $\frac{\partial v_z}{\partial y} \ll \frac{\partial v_y}{\partial z}$

Eq. (2.21) is:

$$\begin{cases} v_x = v_x(b) - 2\left(\rho g\right)^n \dfrac{\partial s}{\partial x}\left(\nabla s \cdot \nabla s\right)^{\frac{n-1}{2}} \displaystyle\int_b^z \left(\frac{s-z}{B}\right)^n dz \\ v_y = v_y(b) - 2\left(\rho g\right)^n \dfrac{\partial s}{\partial y}\left(\nabla s \cdot \nabla s\right)^{\frac{n-1}{2}} \displaystyle\int_b^z \left(\frac{s-z}{B}\right)^n dz \end{cases} \tag{2.47}$$

Proof
Under these strong assumptions, the stress balance equation (Eq. 2.21) is reduced to:

$$\begin{cases} \dfrac{\partial \sigma'_{xz}}{\partial z} - \dfrac{\partial p}{\partial x} = 0 \\ \dfrac{\partial \sigma'_{yz}}{\partial z} - \dfrac{\partial p}{\partial y} = 0 \\ -\dfrac{\partial p}{\partial z} - \rho g = 0 \end{cases} \tag{2.48}$$

The pressure is therefore simply linear with depth: $p(z) = \rho g\left(s - z\right)$. *We can reduce the system to the first two equations only:*

$$\begin{aligned} \frac{\partial \sigma'_{xz}}{\partial z} &= \rho g \frac{\partial s}{\partial x} \\ \frac{\partial \sigma'_{yz}}{\partial z} &= \rho g \frac{\partial s}{\partial y} \end{aligned} \tag{2.49}$$

The boundary conditions on the upper surface impose that $\sigma'_{xz}(s) = \sigma'_{yz}(s) = 0$, *which gives*

an analytical solution from these two stress components:

$$\begin{aligned}\sigma'_{xz} &= -\rho g \frac{\partial s}{\partial x}(s-z) \\ \sigma'_{yz} &= -\rho g \frac{\partial s}{\partial y}(s-z)\end{aligned} \tag{2.50}$$

The effective deviatoric stress is thus:

$$\sigma'^{\,2}_e = \sigma'^{\,2}_{xz} + \sigma'^{\,2}_{yz} = (\rho g)^2 (s-z)^2 (\nabla s \cdot \nabla s) \tag{2.51}$$

Glen's flow-law in its effective stress form (Eq. 2.24) imposes:

$$\begin{aligned}\dot{\varepsilon}_{xz} &= \frac{\sigma'^{\,n-1}_e}{B^n}\sigma'_{xz} \\ \dot{\varepsilon}_{yz} &= \frac{\sigma'^{\,n-1}_e}{B^n}\sigma'_{yz}\end{aligned} \tag{2.52}$$

With the second and third assumptions, this gives:

$$\begin{aligned}\frac{\partial v_x}{\partial z} &= -\frac{2}{B^n}(\rho g)^n (s-z)^n (\nabla s \cdot \nabla s)^{\frac{n-1}{2}} \frac{\partial s}{\partial x} \\ \frac{\partial v_y}{\partial z} &= -\frac{2}{B^n}(\rho g)^n (s-z)^n (\nabla s \cdot \nabla s)^{\frac{n-1}{2}} \frac{\partial s}{\partial y}\end{aligned} \tag{2.53}$$

The integral of this equation between b and z gives the final equations. ∎

2.3 Ice sheet energy balance

2.3.1 Thermal model

A thermal model is essential to predicting an ice sheet system's evolution, as several physical properties of the ice are functions of the temperature. The viscosity parameter, B, is highly temperature-dependent. Temperature also affects melting, which in turn affects the glacier geometry and sliding.

We will see in Chap. 3 that it is convenient to employ the ALE formulation (see Chap. 1). The thermal model in the ALE configuration is the following.

ALE thermal model

Assume ice to be enclosed in a volume $\Omega \subset \mathbb{R}^3$ with the following properties

- *density* $\rho : \Omega \times [0,T] \to \mathbb{R}_+^*$
- *velocity* $\boldsymbol{v} : \Omega \times [0,T] \to \mathbb{R}^3$
- *temperature* $T : \Omega \times [0,T] \to \mathbb{R}_+$
- *heat capacity* $c : \Omega \times [0,T] \to \mathbb{R}_+$
- *thermal conductivity* $k_{th} : \Omega \times [0,T] \to \mathbb{R}_+$

and a source term (here due to viscous heating) Φ*. For a frame velocity* $\boldsymbol{w}$ *the equation of conservation of energy is:*

$$\rho \frac{\partial (cT)}{\partial t} = -\rho \ (\boldsymbol{v} - \boldsymbol{w}) \cdot \nabla (cT) + \nabla \cdot k_{th} \nabla T + \Phi \tag{2.54}$$

Proof
We have shown that the conservation of energy (Eq. 1.11) was:

$$\rho \frac{D}{Dt} (cT) = \nabla \cdot k_{th} \nabla T + \Phi \tag{2.55}$$

The advection appears when one takes the Eulerian or ALE version of the material derivative In Eulerian, this equation becomes:

$$\rho \frac{\partial (cT)}{\partial t} = -\rho \, \boldsymbol{v} \cdot \nabla (cT) + \nabla \cdot k_{th} \nabla T + \Phi \tag{2.56}$$

and in ALE, the frame velocity $\boldsymbol{w}$ *must be accounted for:*

$$\rho \frac{\partial (cT)}{\partial t} = -\rho \ (\boldsymbol{v} - \boldsymbol{w}) \cdot \nabla (cT) + \nabla \cdot k_{th} \nabla T + \Phi \tag{2.57}$$

■

Several assumption can be employed for ice (Hooke, 2005, p117). These assumptions simplify the previous equations by taking c and k_{th} constant.

Simplified thermal model

Under the following assumptions:

1. *spatial and thermal dependence of the heat capacity, c, are neglected*
2. *spatial and thermal dependence of the thermal conductivity, k_{th}, are neglected*

Eq. (2.54) becomes:

$$\frac{\partial T}{\partial t} = -\left(\boldsymbol{v} - \boldsymbol{w}\right) \cdot \nabla T + \frac{k_{th}}{\rho c} \Delta T + \frac{\Phi}{\rho c} \tag{2.58}$$

The local heat transfer is thus a result of advection, conduction and internal deformation heating. In many thermal models, the horizontal advection is neglected (Paterson (1994), p.216 and Hooke (2005), p.118). Horizontal conduction is often neglected as well (Hulbe and MacAyeal (1999) and Hooke (2005), p.118). These assumptions are not employed here.

Viscous heating

The viscous heating of the ice can equivalently be written as:

$$\Phi = Tr\left(\boldsymbol{\sigma}\dot{\boldsymbol{\varepsilon}}\right) = 2\dot{\varepsilon}_e \sigma'_e = 4\mu\dot{\varepsilon}_e^2 \tag{2.59}$$

Proof
The energy of deformation is

$$\Phi = Tr\left(\boldsymbol{\sigma}\dot{\boldsymbol{\varepsilon}}\right) \tag{2.60}$$

One can use the deviatoric stress tensor and use the incompressibility:

$$\Phi = Tr\left(\boldsymbol{\sigma}'\dot{\boldsymbol{\varepsilon}} - P\dot{\boldsymbol{\varepsilon}}\right) = Tr\left(\boldsymbol{\sigma}'\dot{\boldsymbol{\varepsilon}}\right) - p\ Tr\left(\dot{\boldsymbol{\varepsilon}}\right) = Tr\left(\boldsymbol{\sigma}'\dot{\boldsymbol{\varepsilon}}\right) \tag{2.61}$$

And in terms of effective strain rate $\dot{\varepsilon}_e$ and effective deviatoric stress $\boldsymbol{\sigma}_e$:

$$\begin{aligned}\Phi^2 &= \left(Tr\left(\boldsymbol{\sigma}'\dot{\boldsymbol{\varepsilon}}\right)\right)^2 = \left(\sum_{i,j} \sigma'_{ij}\dot{\varepsilon}_{ji}\right)^2 = \left(\sum_{i,j} \sigma'_{ij}\dot{\varepsilon}_{ij}\right)^2 = \left(\sum_{i,j} \sigma'_{ij}\dot{\varepsilon}_{ij}\right)\left(\sum_{k,l} \sigma'_{kl}\dot{\varepsilon}_{kl}\right) \\ &= \left(\sum_{i,j} 2\mu\dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}\right)\left(\sum_{k,l} \sigma'_{kl}\frac{1}{2\mu}\sigma'_{kl}\right) = \sum_{i,j}\dot{\varepsilon}_{ij}^2 \sum_{kl} {\sigma'_{kl}}^2 = 2\dot{\varepsilon}_e^2\ \ 2{\sigma'_e}^2\end{aligned} \tag{2.62}$$

■

2.3.2 Thermal boundary conditions

Ice/Atmosphere interface

On the upper surface, the temperature is imposed as the mean annual air temperature at the surface, T_s. Indeed, the temperature measured at 10 or 15 m in a glacier, beneath the

zone of seasonal variation, equals the mean annual air temperature (Cuffey and Paterson, 2010, p.404). This assumption is made in the vast majority of ice sheet thermal models (i.e., Hulbe and MacAyeal (1999); Pattyn (2003); Morlighem et al. (2010)).

Ice/Water boundary

On the Ice/Water interface, the simplest boundary condition is given by Holland and Jenkins (1999). A heat flux is imposed at the boundary that is proportional to the temperature difference between the ice shelf surface T_b and the ocean $T = T_{pmp}$.

$$k_{th}\,\nabla T|_b\,.\boldsymbol{n} \simeq -k_{th}\,\left.\frac{\partial T}{\partial z}\right|_b = -\rho_w c_{pM}\gamma\,(T_b - T_{pmp}) \tag{2.63}$$

With:

- $\boldsymbol{n} = (n_x, n_y, n_z)$ normal vector pointing outward
- c_{pM} mixed layer (Holland and Jenkins, 1999, p5) specific heat capacity (J kg^{-1} K^{-1})
- γ thermal exchange velocity (m s^{-1})
- T_{pmp} pressure melting point (melting point of ice under pressure in K) (Paterson, 1994, p.212)

If the ocean temperature is above the ice shelf base temperature, the heat flux is positive with respect to the normal (heat goes from the ocean to the ice shelf), whereas when the ice shelf base is warmer than the ocean, the flux is negative (heat goes from the ice shelf to the ocean). Interested readers can refer to Holland and Jenkins (1999) for more details.

Ice/Till boundary

The basal thermal boundary condition is supplied by a geothermal flux, G, at the ice sheet bottom and the heat due to basal friction:

$$k_{th}\,\nabla T|_b \cdot \boldsymbol{n} \simeq -k_{th}\,\left.\frac{\partial T}{\partial z}\right|_b = G + \boldsymbol{\tau}_b \cdot \boldsymbol{u}_b \tag{2.64}$$

The basal temperature in the ice mass is kept at the pressure melting point $T_{pmp} = 273.15 - \beta p$ (Paterson, 1994, p212) whenever it is reached (Pattyn, 2003). In other words, the boundary condition at the ice/till interface is an imposed flux. When this imposed flux generates ice temperatures above the pressure melting point, the boundary condition is changed into an imposed temperature $T_b = T_{pmp}$.

Conclusion

In this chapter, we used the mass conservation, momentum balance and Glen's flow law to define the system of equations governing ice flow. This system of equations constitutes the full-Stokes equations. This model requires huge computational resources in terms of computer memory and execution time: it is a 3d model that solves 4 unknowns (v_x, v_y, v_z, p). Thus, we also presented two simplified models: Blatter-Pattyn's higher-order model (BP) and MacAyeal-Morland's shelfy-stream model (SSA). Both models are hierarchically derived from the full-Stokes equations. BP still uses a 3d mesh, but decorrelates the horizontal velocity components from the vertical component. There are hence only 2 unknowns (v_x, v_y). MacAyeal goes further by vertically integrating the equations, neglecting vertical shear. Finally, we presented the simple Shallow Ice Approximation (SIA) that gives analytical equations for the horizontal velocities. These four models offer four different refinement levels and their computational efficiency is also different, which is very useful for modeling a continental ice sheet, such as Antarctica, since one can use different models on different areas depending on the accuracy needed and the phenomena involved. In the last section of this chapter, we presented the thermal model and its boundary conditions, derived from the energy balance.

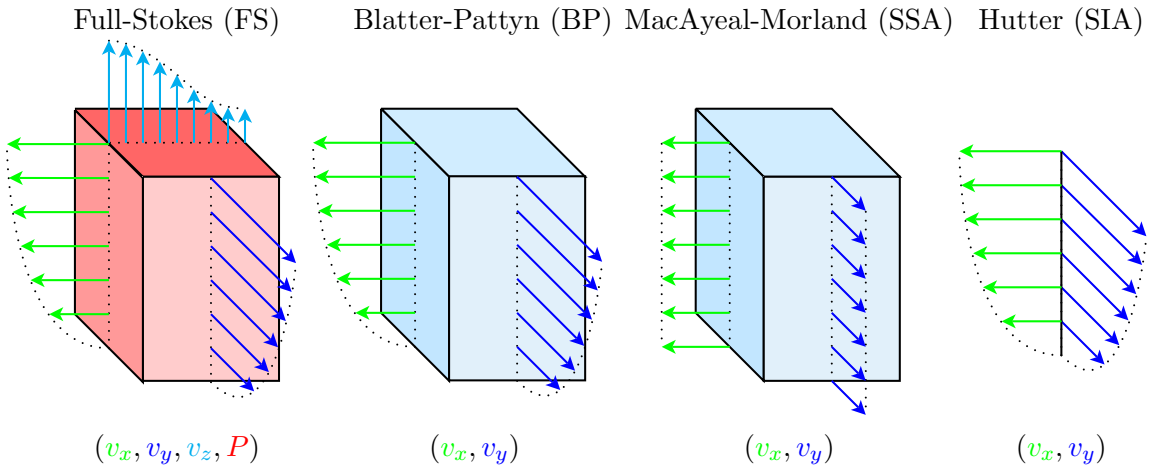


Figure 2.4: Ice flow models

3

Large-scale ice sheet numerical modeling

The equations presented in Chap. 2 have been implemented in the massively-parallelized, multi-purpose, finite element framework provided by the Ice Sheet System Model (ISSM). In this chapter, we first present a general algorithmic strategy for solving transient simulations of ice sheet flow using the finite element method and an Arbitrary Lagrangian Eulerian configuration (Chap. 1). We then discuss the software architecture designed for massively parallel runs. Only the key aspects of the implementation are presented, as this was a team effort. Finally, we present an algorithm for anisotropic mesh adaptation.

3.1 Numerical modeling strategy

3.1.1 Sequential resolution

In the previous chapter, we introduced the thermomechanical equations that govern large-scale ice sheet dynamics. These are summarized below:

$$\text{Mass balance equation} \quad \begin{cases} \nabla \cdot \boldsymbol{v} = 0 \\ \dfrac{\partial s}{\partial t} + v_x\left(s\right)\dfrac{\partial s}{\partial x} + v_y\left(s\right)\dfrac{\partial s}{\partial y} - v_z\left(s\right) = \dot{M}_s \\ \dfrac{\partial b}{\partial t} + v_x\left(b\right)\dfrac{\partial b}{\partial x} + v_y\left(b\right)\dfrac{\partial b}{\partial y} - v_z\left(b\right) = \dot{M}_b \end{cases}$$

$$\text{Momentum balance equations} \quad \nabla \cdot \boldsymbol{\sigma}' - \nabla p + \rho\, \boldsymbol{g} = \boldsymbol{0}$$

$$\text{Constitutive equation:} \quad \boldsymbol{\sigma}' = 2\mu\dot{\boldsymbol{\varepsilon}}$$

$$\text{Glen's flow law} \quad \mu = \frac{B}{2\left(\dot{\varepsilon}_e^{\,1-\frac{1}{n}}\right)}$$

$$\text{Energy balance equation} \quad \frac{\partial T}{\partial t} = -\left(\boldsymbol{v} - \boldsymbol{w}\right)\cdot\nabla T + \frac{k_{th}}{\rho c}\Delta T + \frac{\Phi}{\rho c}$$

There are 7 unkowns: the three velocity components (v_x, v_y, v_z), the pressure (p), the temperature (T) and the surface and bed elevations (s,b). For grounded ice, only the surface elevation is unknown, as the bedrock topography is assumed to remain constant. All of these variables are coupled, the geometry directly affects the velocities, which are also affected by the temperature, since the viscosity, B, is temperature dependent. The velocities in turn affect the geometry of the system and the advection in the thermal model. To reduce the numerical cost that a resolution of the full system requires, we use a sequential coupling in an explicit formulation. The steps are described in Fig. 3.1.

We start with an initial geometry, velocity (observed or taken as zero) and temperature fields. We first compute the temperature for this time step. The viscosity is then updated following its temperature dependency. We then perform a diagnostic to compute the new velocities. We go to the next time increment (following the CFL condition) and calculate the new geometry of the system. We stop whenever the final time is exceeded.

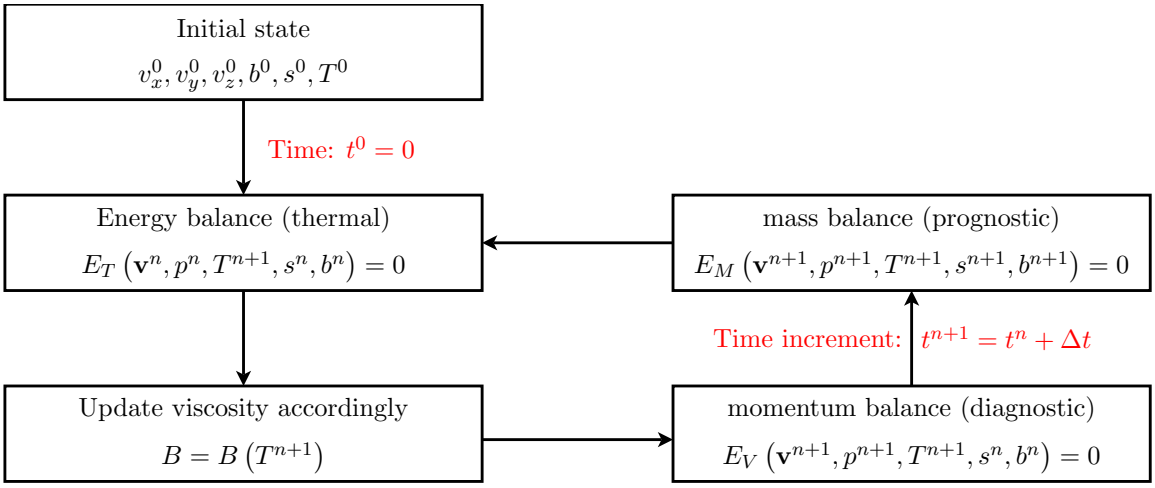


Figure 3.1: Solution sequence of a transient ice flow simulation. E_T, E_V and E_M symbolize the thermal, mechanical and prognostic equations respectively

3.1.2 Finite-Element Method

To solve the three problems (thermal, diagnostic and prognostic), ISSM relies on the Finite Element Method (FEM). This method is not described in detail here. Interested readers can refer to (Zienkiewicz and Taylor, 1989). To summarized, one begins with a weak or variational formulation of the model equations, then applies the Galerkin approximation. To illustrate, let's consider the MacAyeal-Morland equations for an ice shelf:

$$\begin{cases} \dfrac{\partial}{\partial x}\left(4H\bar{\mu}\dfrac{\partial v_x}{\partial x} + 2H\bar{\mu}\dfrac{\partial v_y}{\partial y}\right) + \dfrac{\partial}{\partial y}\left(H\bar{\mu}\dfrac{\partial v_x}{\partial y} + H\bar{\mu}\dfrac{\partial v_y}{\partial x}\right) = \rho g H\dfrac{\partial s}{\partial x} \\ \dfrac{\partial}{\partial y}\left(4H\bar{\mu}\dfrac{\partial v_y}{\partial y} + 2H\bar{\mu}\dfrac{\partial v_x}{\partial x}\right) + \dfrac{\partial}{\partial x}\left(H\bar{\mu}\dfrac{\partial v_x}{\partial y} + H\bar{\mu}\dfrac{\partial v_y}{\partial x}\right) = \rho g H\dfrac{\partial s}{\partial y} \end{cases} \tag{3.1}$$

For this two dimensional model, we consider two boundary conditions: water pressure applied on the ice front Γ_σ, and a homogeneous Dirichlet boundary condition applied on the other boundaries Γ_u:

$$\begin{cases} \boldsymbol{\sigma}' \cdot \boldsymbol{n} = \left(\dfrac{1}{2}\rho g H^2 - \dfrac{1}{2}\rho_w g b^2\right)\boldsymbol{n} & \text{on } \Gamma_\sigma \\ \boldsymbol{v} = 0 & \text{on } \Gamma_u \end{cases} \tag{3.2}$$

We define the force $\boldsymbol{f} = \left(\dfrac{1}{2}\rho g H^2 - \dfrac{1}{2}\rho_w g b^2\right)\boldsymbol{n}$.

Weak Formulation

We define the kinematically admissible velocity field as:

$$\mathcal{V} = \left\{\boldsymbol{v} \in \left(\mathcal{H}^1(\Omega)\right)^2, \quad \forall \boldsymbol{x} \in \Gamma_u \quad \boldsymbol{v} = \boldsymbol{0}\right\} \tag{3.3}$$

$\mathcal{H}^1(\Omega)$ is the standard notation for the Sobolev space, which contains all the square-integrable functions defined on Ω that have a square-integrable first derivative. It can be shown that the primal problem (Eq. (3.1) and Eq. (3.2)) is strictly equivalent to its weak formulation:

$$\forall \boldsymbol{\varphi} = (\varphi_x, \varphi_y) \in \mathcal{V} \tag{3.4}$$

$$\begin{aligned}\int_\Omega 2H\mu\left(2\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)\frac{\partial \varphi_x}{\partial x}+H\mu\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)\frac{\partial \varphi_x}{\partial y}d\Omega \\ +\int_\Omega 2H\mu\left(2\frac{\partial v_y}{\partial y}+\frac{\partial v_x}{\partial x}\right)\frac{\partial \varphi_y}{\partial x}+H\mu\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)\frac{\partial \varphi_y}{\partial y}d\Omega \\ =\int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\varphi}\,d\Gamma-\int_\Omega \rho g H\nabla s\cdot\boldsymbol{\varphi}\,d\Omega\end{aligned} \tag{3.5}$$

Galerkin approximation

Now, instead of using the entire space $\mathcal{V}$, which is an infinite-dimensional vector space, we use a sub-space $\mathcal{V}_h$, which is finite-dimensional. Let $\{\boldsymbol{\varphi}_i, \quad i = 1 \dots n\}$ be a basis of this space. We now consider all basis functions to be test functions in the weak formulation: $\forall i \in [1 \quad n]$:

$$\begin{aligned}\int_\Omega 2H\mu\left(2\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)\frac{\partial \varphi_{ix}}{\partial x}+H\mu\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)\frac{\partial \varphi_{ix}}{\partial y}d\Omega \\ +\int_\Omega 2H\mu\left(2\frac{\partial v_y}{\partial y}+\frac{\partial v_x}{\partial x}\right)\frac{\partial \varphi_{iy}}{\partial x}+H\mu\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)\frac{\partial \varphi_{iy}}{\partial y}d\Omega \\ =\int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\varphi}_i\,d\Gamma-\int_\Omega \rho g H\nabla s\cdot\boldsymbol{\varphi}_i\,d\Omega\end{aligned} \tag{3.6}$$

If we approximate the solution by decomposing $\boldsymbol{v}$ on this basis, we get:

$$\boldsymbol{v} = \sum_{j=1}^{n} v_j \boldsymbol{\varphi}_j \tag{3.7}$$

The weak formulation becomes:

$$\begin{aligned}\sum_{j=1}^{n} v_j \int_\Omega 2H\mu\left(2\frac{\partial \varphi_{jx}}{\partial x}+\frac{\partial \varphi_{jy}}{\partial y}\right)\frac{\partial \varphi_{ix}}{\partial x}+H\mu\left(\frac{\partial \varphi_{jx}}{\partial y}+\frac{\partial \varphi_{jy}}{\partial x}\right)\frac{\partial \varphi_{ix}}{\partial y}d\Omega \\ +\sum_{j=1}^{n} v_j \int_\Omega 2H\mu\left(2\frac{\partial \varphi_{jy}}{\partial y}+\frac{\partial \varphi_{jx}}{\partial x}\right)\frac{\partial \varphi_{iy}}{\partial x}+H\mu\left(\frac{\partial \varphi_{jx}}{\partial y}+\frac{\partial \varphi_{jy}}{\partial x}\right)\frac{\partial \varphi_{iy}}{\partial y}d\Omega \\ =\int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\varphi}_i\,d\Gamma-\int_\Omega \rho g H\nabla s\cdot\boldsymbol{\varphi}_i\,d\Omega\end{aligned} \tag{3.8}$$

This gives n equations (all test functions $\boldsymbol{\varphi}_i$) and n unknowns (the velocity components v_j). Let $\mathbf{K}$ be a matrix representing the left hand side, such that:

$$k_{ij} = \int_\Omega 2H\mu \left(2\frac{\partial\varphi_{jx}}{\partial x} + \frac{\partial\varphi_{jy}}{\partial y}\right)\frac{\partial\varphi_{ix}}{\partial x} + H\mu\left(\frac{\partial\varphi_{jx}}{\partial y} + \frac{\partial\varphi_{jy}}{\partial x}\right)\frac{\partial\varphi_{ix}}{\partial y}d\Omega + \int_\Omega 2H\mu \left(2\frac{\partial\varphi_{jy}}{\partial y} + \frac{\partial\varphi_{jx}}{\partial x}\right)\frac{\partial\varphi_{iy}}{\partial x} + H\mu\left(\frac{\partial\varphi_{jx}}{\partial y} + \frac{\partial\varphi_{jy}}{\partial x}\right)\frac{\partial\varphi_{iy}}{\partial y}d\Omega \quad (3.9)$$

and $\boldsymbol{F}$ a vector, such that:

$$f_i = \int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\varphi}_i\, d\Gamma - \int_\Omega \rho g H \nabla s\cdot\boldsymbol{\varphi}_i\, d\Omega \quad (3.10)$$

For a linear viscosity, the solution of the problem is equivalent to solving the following linear system:

$$\mathbf{K}\boldsymbol{U} = \boldsymbol{F} \quad (3.11)$$

In the case of Glen's flow law with $n = 3$, the viscosity depends on the velocity, and the above system becomes non-linear. In such cases, we rely on the iterative fixed-point method (or Picard's scheme). The solution of this system gives all the components of the velocity on the basis of the solution space, and the velocity vector can be recovered. The basis of the solution space is usually chosen such that, for a mesh discretization of the continuous domain, each basis function is equal to 1 on one vertex and 0 on all other vertices. The discretization of the model domain is therefore crucial, as it defines the solution space on which the solution is sought.

The finite element method is widely employed in solid mechanics, but is a relative newcomer to computational fluid dynamics. It provides advantages over both the finite difference method and the finite volume method, because it can use unstructured meshes and possesses the *best approximation* property (Céa's Lemma) when applied to elliptic and parabolic problems at relatively low Péclet numbers. In other words, the Galerkin solution $\boldsymbol{v}_h \in \mathcal{V}_h$ is as close to the exact solution $\boldsymbol{v} \in \mathcal{V}$ as any other vector in $\mathcal{V}_h$.

3.1.3 Mesh deformation with ALE

The Eulerian approach is widely used in fluid dynamics, as large distortions of the continuum motion can be easily handled. The computational mesh remains fixed. The disadvantage of this approach is that moving interfaces are difficult to follow and require specific treatment, such as the level-set method. In the context of ice sheet modeling, the surface and the ice shelf base can move dramatically and these changes must be precisely captured: the driving stress is proportional to the surface gradient. A fine description of the surface elevation is therefore required. An Eulerian approach would be difficult to apply.

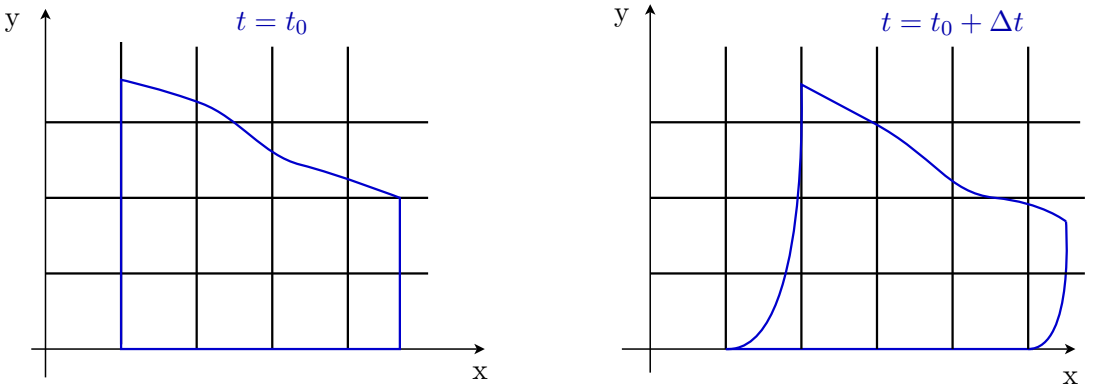


Figure 3.2: Eulerian mesh after a time increment. The two-dimensional mesh in black remains fixed, while the system represented with a blue line moves.

The Lagrangian approach is generally used in structural mechanics. Each node of the computational mesh follows a given material particle. This approach makes it easy to track free surfaces and interfaces between different materials. However, the Lagrangian approach is not well suited for large distortions, because it requires frequent remeshing to avoid an excessive distortion of the finite element mesh. This is the case for ice sheets, especially in fast moving regions where large shearing deformations can take place.

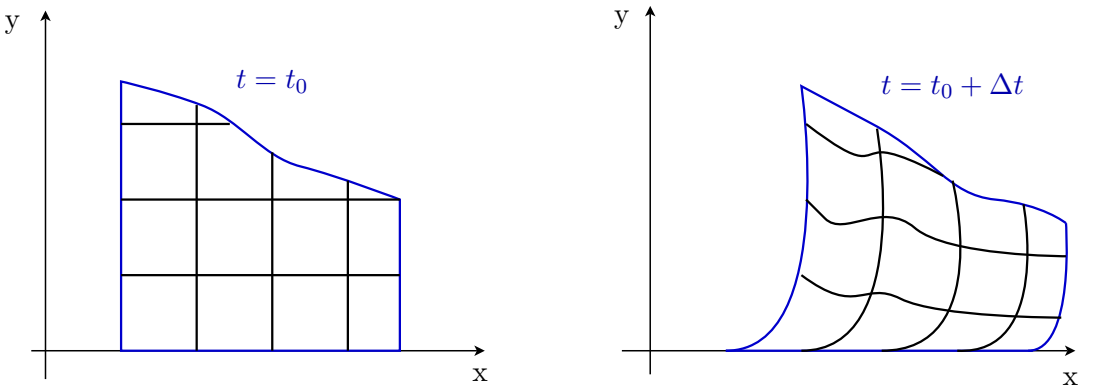


Figure 3.3: Lagrangian mesh after a time increment. The two-dimensional mesh in black follows the deformation of the system represented in blue.

The Arbitrary Eulerian-Lagrangian method (Donea and Belytschko, 1992) consists in combining the advantages of both techniques: the nodes of the mesh may move arbitrarily alleviating the drawbacks of the traditional Lagrangian and Eulerian approach. This method was first introduced for finite difference, and extended to finite elements (Hughes et al., 1981). It is now a popular tool for simulation of continuum mechanics problems with large shear deformation such as fluid flow and metal forming. Compared to pure Eulerian methods, it is also better suited for moving boundaries and large volume changes of the computational domain.

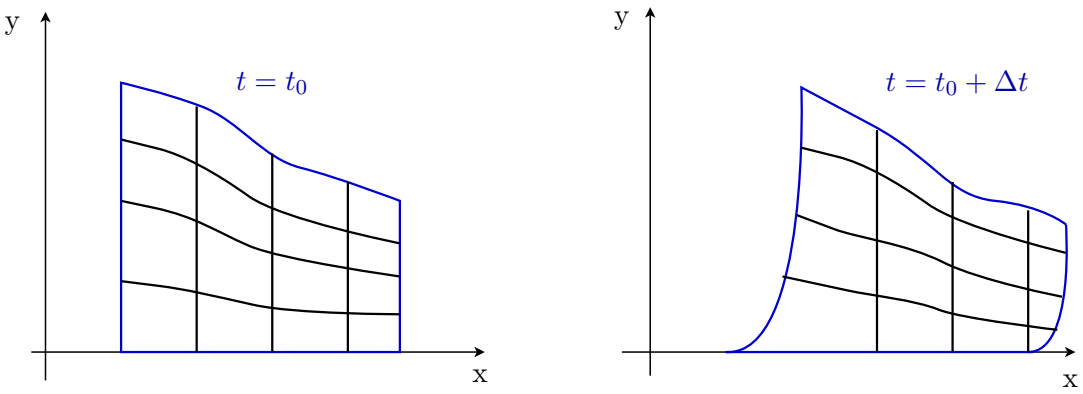


Figure 3.4: ALE mesh after a time increment. The mesh in black moves arbitrarily: here the horizontal layers follow the vertical evolution of the system surface, but the vertical edges remain fixed. The mesh follows the boundary of the system, but is not dramatically deformed.

In the context of ice sheet modeling, it is convenient to use a mesh that is horizontally fixed in space (Eulerian), and vertically equidistributed between the moving base and the surface. The mesh velocity is zero horizontally, and its vertical component is calculated after each iteration as the ratio between the node vertical displacement and the time step. This velocity is the frame velocity that needs to be incorporated in the ALE formulation of the system equations. Only the thermal model is affected by the mesh deformation, because the mechanical model is quasi-static.

3.1.4 ISSM Software Architecture

The Ice Sheet System Model (ISSM) is a massively-parallelized object-oriented framework that we have been developing since 2009. It is inspired by NASTRAN (Nastran, 1972), which is a Finite Element Analysis program that was originally developed for NASA in the late 1960s. ISSM is already capable of tackling systems with several millions of degrees of freedom within times on the order of hours, including cases where full-Stokes is used (Larour et al., 2012). We present here briefly some of the characteristics of ISSM: code languages, parallelism and solvers.

Languages

ISSM relies mainly on the C and C++ languages (Kernighan, 1988; Stroustrup, 1997), and is hosted in MATLAB (Mathworks, 2008), a common scientific platform.

MATLAB is an interpreted language (no need to compile), which is platformm independent and comes with a extensive library of predefined functions. It also comes with very powerful plotting capabilities, which makes it ideal for display and analysis of large and complex datasets. But as with any interpreted language, it is slow when dealing with loops. The C/C++ language on the other hand is a fast language for which many compilers and math kernels are well optimized. It also gives strong emphasis to the use of polymorphism, which brings object oriented methodologies to its full potential. ISSM combines the advantages of both languages: the C/C++ core is interfaced with the MATLAB environment, using the MATLAB External API (mex modules).

Parallelism

ISSM can be run in serial/sequential mode (only one CPU), within MATLAB, but its focus is mainly on massively parallel computations, run on large clusters such as the NASA Advanced Supercomputing (NAS) Pleiades cluster (which possesses 111,104 cores). ISSM is at its core a parallel architecture and uses a static domain decomposition (DD) methodology. This subdivision has two main purposes :

1. to distribute element computations to CPUs in an even manner to minimize memory requirements
2. to distribute a system of equations evenly amongst CPUs for maximum computational efficiency

Each element of the system "belongs" to only one CPU. The subdomains are connected at the interfaces by sharing degrees of freedom. Nodes on the boundary between two subdomains are assigned to only one subdomain. Other CPUs have "clones" of these same nodes, but the "master" nodes belong to one CPU. Communication between CPUs is required when the matrices are assembled so that the connections between subdomains can be made.

ISSM uses METIS (Karypis and Kumar, 1998) to partition the elements. METIS ensures that each partition has an equal number of elements, and that there are as few elements intersecting a partition boundary as possible, in order to minimize the amount of communication required between processes. Fig. 3.5 shows an example of element partitioning for 26 CPUs (one color per CPU).

Parallelism is achieved by using the Message Passing Interface (Gropp et al., 1996; Gropp and Lusk, 1996). This library is flexible enough to allow runs on distributed as well as shared memory clusters (Fig. 3.6). The difficulty that arises from the use of distributed clusters is that the memory is not shared. Each CPU has its elements and cannot access properties of elements that are not in its partition in a simple manner. This had to be taken into account in the code development.

Solver

Parallel objects and solvers rely on the Portable Extensible Toolkit for Scientific Computation package (PETSc) (Balay et al., 1997, 2008, 2009). PETSc is an object-oriented toolkit for the parallel numerical solution of PDEs. PETSc provides implementations of basic objects, such as matrices and vectors, linear solvers (primarily Krylov methods with a variety of preconditioners), and nonlinear solvers (primarily Newton-type methods).

Most of the time, ISSM relies on a direct solver. The MUltifrontal Massively Parallel Sparse direct Solver, or MUMPS (Amestoy et al., 2001, 2006). MUMPS is a software application for the solution of large sparse systems of linear algebraic equations on distributed memory parallel computers. It has been designed to solve a wide range of problems, which makes it ideal for the problems encountered in ice sheet modeling.

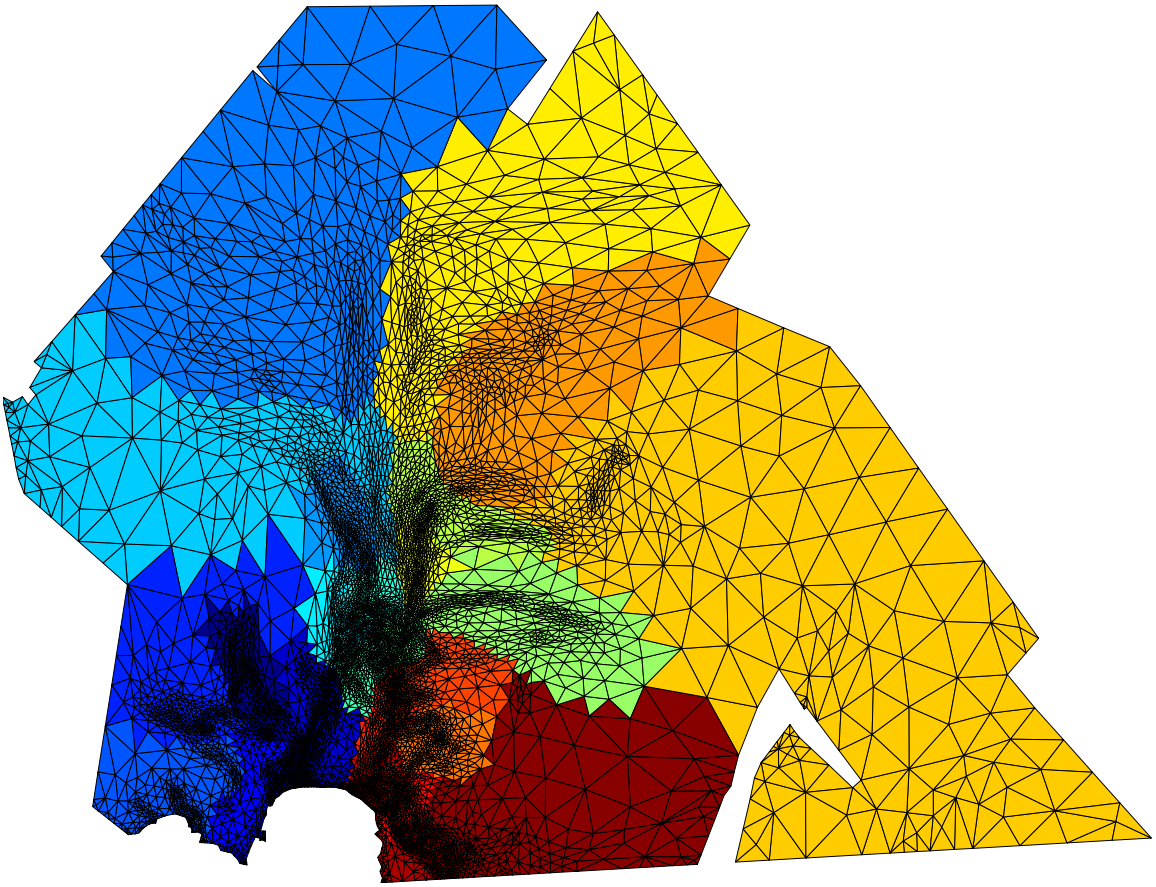


Figure 3.5: Example of Mesh partitioning, using a domain decomposition method on Pine Island Glacier (each color represents one of the 26 partitions)

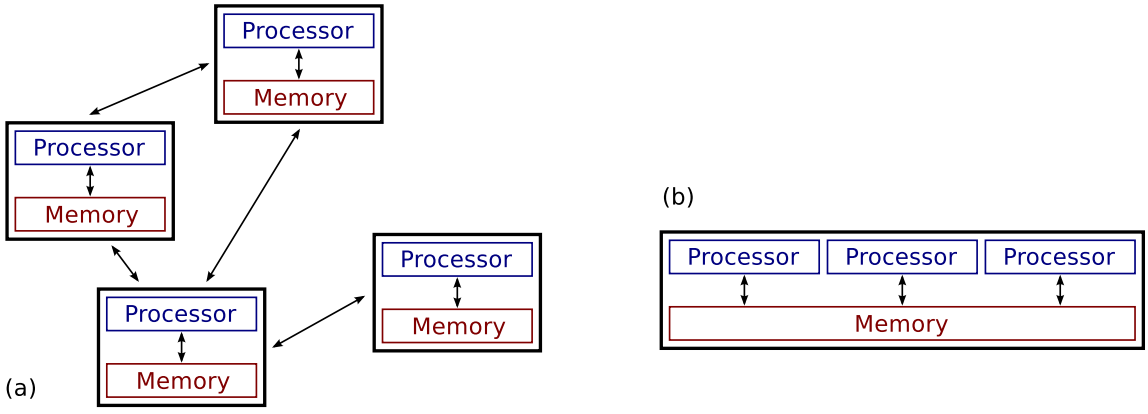


Figure 3.6: (a) Distributed-memory architecture. (b) Shared-memory architecture

3.2 Anisotropic mesh adaptation

One of the major challenges of large scale modeling is dealing with computationally intensive problems. Mesh generation is a crucial component, as one has to minimize the number of elements in the computational mesh without losing accuracy. Here we rely on static anisotropic mesh adaptation following the work of Habashi et al. (2000); Dompierre et al. (2002) and Ait-Ali-Yahia et al. (2002). The idea is to generate a mesh that minimizes interpolation error for a given field, generally the velocity. Smaller elements are introduced where the variations in velocities are high, and larger elements are used where the ice is stagnant. We detail here how this adaption is automated.

3.2.1 Interpolation Error

Let $u : \mathcal{V}^d \to \mathbb{R}$ be a scalar function defined in a space of dimension d, $\mathcal{V}^d \subset \mathbb{R}^d$, which is infinitely differentiable on its domain of definition. Let $u^h : \mathcal{V}^d \to \mathbb{R}$ be an interpolant of u, taken here as piecewise linear (See Fig.3.7). In our case, u can be the known magnitude of the observed surface velocity or the bedrock topography, and u^h the $P1$ interpolation of this field onto a triangular mesh that is to be optimized. We define the absolute error as follows:

Interpolation error

The error between a function $u : \mathcal{V}^d \to \mathbb{R}$ and its linear interpolant $u^h : \mathcal{V}^d \to \mathbb{R}$ is defined as:

$$\forall \boldsymbol{x} \in \mathcal{V}^d \quad \epsilon(\boldsymbol{x}) = \left| u(\boldsymbol{x}) - u^h(\boldsymbol{x}) \right| \tag{3.12}$$

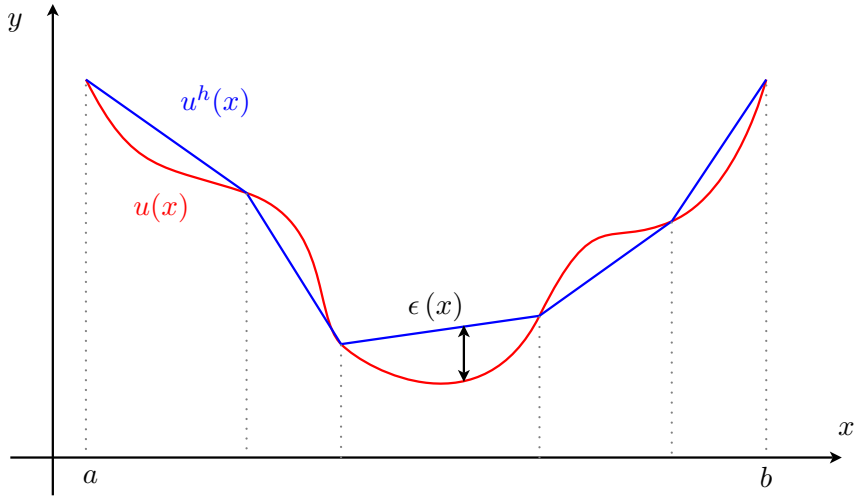


Figure 3.7: 1d interpolation error $\epsilon(x) = \left| u(x) - u^h(x) \right|$

In our case, we have a triangulation $\mathcal{T}_h$ of the domain. Fields, such as surface velocities, ice thickness or bed topography, are interpolated onto this mesh. We want to minimize the error between the exact fields and their interpolations. We therefore need an upper bound on the error for each element $K \in \mathcal{T}_h$, the idea being to decrease the size of the elements where

the error, ϵ, is high, and increase the size of the elements where this error is small. Alauzet and Frey (2003) showed that the interpolation error of $P1$ elements (piecewise linear) in a space of dimension d is as described below.

Generalized interpolation absolute error

Let u^h be a $P1$ interpolation of an infinitely differentiable field u, for a triangulation $\mathcal{T}_h$. For every element K of this triangulation we have the following error upper bound:

$$\forall K \in \mathcal{T}_h \quad \|u - u^h\|_{\infty,K} \leq c_d \sup_{\boldsymbol{x},A,B \in K} \langle \overrightarrow{AB}, |\mathbf{H_u}(\boldsymbol{x})| \overrightarrow{AB} \rangle \tag{3.13}$$

where:

- $c_d = \frac{1}{2}\left(\frac{d}{d+1}\right)^2$ *is a constant that depends only on the space dimension d*
- $\mathbf{H_u}(\boldsymbol{x})$ *is the Hessian matrix of u on $\boldsymbol{x}$, and $|\mathbf{H_u}(\boldsymbol{x})|$ the* modified *Hessian matrix for which all eigenvalues are positive.*

3.2.2 Anisotropic formulation of the interpolation error

Let E_K be the set of all edges for an element K of the triangulation. Any vector $\overrightarrow{AB} \subset K$ can be expressed as a linear combination of the edges of the triangle K. Eq. (3.13) becomes:

$$\forall K \in \mathcal{T}_h \quad \|u - u^h\|_{\infty,K} \leq c_d \max_{\boldsymbol{x} \in K} \max_{\boldsymbol{e} \in E_K} \langle \boldsymbol{e}, |\mathbf{H_u}(\boldsymbol{x})| \boldsymbol{e} \rangle \tag{3.14}$$

In practice, the right hand side is difficult to estimate, since the maximum of the Hessian matrix $|\mathbf{H_u}(\boldsymbol{x})|$ is unknown. To overcome this, we assume that we can build a metric tensor $\overline{\mathbf{M}}(K)$ that verifies:

$$\forall \boldsymbol{e} \in E_K \quad \max_{x \in K} \langle \boldsymbol{e}, |\mathbf{H_u}(\boldsymbol{x})| \boldsymbol{e} \rangle \leq \langle \boldsymbol{e}, \overline{\mathbf{M}}(K) \boldsymbol{e} \rangle \tag{3.15}$$

In the next section, we will show how the Hessian matrix is estimated and how the metric tensor, $\overline{\mathbf{M}}$, is built. This gives the following inequality:

$$\forall K \in \mathcal{T}_h \quad \|u - u^h\|_{\infty,K} \leq c_d \max_{\boldsymbol{e} \in E_K} \langle \boldsymbol{e}, \overline{\mathbf{M}}(K) \boldsymbol{e} \rangle \tag{3.16}$$

We define the error ϵ_K for any element of the triangulation as:

$$\epsilon_K = c_d \max_{\boldsymbol{e} \in E_K} \langle \boldsymbol{e}, \overline{\mathbf{M}}(K) \boldsymbol{e} \rangle \tag{3.17}$$

This means that for any element K of the mesh, the interpolation error of $P1$ elements is strictly proportional to the square of its longest edge in the metric $\overline{\mathbf{M}}(K)$. The control of the element edges allows us to control the error interpolation.

3.2.3 Mesh quality

In the mesh generation procedure, the goal is to equi-distribute the interpolation error in each direction over each element to control the global interpolation error. Let ϵ be the maximum authorized error for all elements of the mesh. For any element K of the mesh, (3.17) we impose that:

$$\forall \boldsymbol{e} \in E_K \quad \epsilon = c_d \max_{\boldsymbol{e} \in E_K} \langle \boldsymbol{e}, \overline{\mathbf{M}}(K)\, \boldsymbol{e} \rangle \tag{3.18}$$

If we define the metric tensor as:

$$\mathbf{M}(K) = \frac{c_d}{\epsilon}\, \overline{\mathbf{M}}(K) \tag{3.19}$$

the edges must satisfy:

$$\forall \boldsymbol{e} \in E_K \quad \langle \boldsymbol{e}, \mathbf{M}(K)\, \boldsymbol{e} \rangle = 1 \Longleftrightarrow \left(l_{\mathbf{M}(K)}(\boldsymbol{e}) \right)^2 = 1 \tag{3.20}$$

where $l_{\mathbf{M}(K)}(\boldsymbol{e})$ is the length of the edge $\boldsymbol{e}$ in the metric $\mathbf{M}(K)$. This means that the interpolation error is ϵ if the edge length in the metric $\mathbf{M}(K)$ is 1. We then have to modify the scalar product that is used to evaluate the distance in the mesh generator with the anisotropic metric tensor $\mathbf{M}(K)$, which replaces the standard Euclidean metric.

In 2d, the quality of the elements can be estimated as being the proportion of equilateral triangles in the metric $\mathbf{M}(K)$. For a triangle, the quality Q is the ratio of its volume, $v_{\mathbf{M}}(K)$ and the sum of its edge lengths in the metric. A coefficient is added so that the quality is 1 for equilateral triangles.

$$Q(K) = 4\sqrt{3} \frac{v_{\mathbf{M}}(K)}{\sum_{i=1}^{3} l_{\mathbf{M}}(\boldsymbol{e}_i)^2} \tag{3.21}$$

3.2.4 Metric construction

We show here how to construct the metric $\mathbf{M}$, which is defined for each node of the mesh and linearly interpolated to be continuous over the domain. We also want to control the smallest and largest sizes of element edges. We use the following notations:

- ϵ maximum interpolation error allowed
- h_{min} smallest edge length allowed
- h_{max} longest edge length allowed

We employ here the metric proposed by Alauzet and Frey (2003); Alauzet (2003) and Alauzet and Saiac (2007). For each node of the mesh, the metric is defined as:

$$\mathbf{M} = \mathbf{R}\tilde{\mathbf{\Lambda}}\mathbf{R}^{-1} \tag{3.22}$$

where $\mathbf{R}$ is the matrix of eigenvectors of the Hessian matrix $\mathbf{H_u}$. The matrix $\tilde{\mathbf{\Lambda}}$ is defined as:

$$\tilde{\mathbf{\Lambda}} = \begin{pmatrix} \tilde{\lambda_1} & 0 \\ 0 & \tilde{\lambda_2} \end{pmatrix} \tag{3.23}$$

$$\tilde{\lambda_i} = \min\left(\max\left(\frac{c_d\,|\lambda_i|}{\epsilon}, \frac{1}{h_{max}^2}\right), \frac{1}{h_{min}^2}\right) \tag{3.24}$$

where λ_i are the eigenvalues of the Hessian matrix $\mathbf{H_u}$ and c_d the constant that appears in (3.13).

This metric tensor is anisotropic, as its principal directions are given by the eigenvectors and its lengths by the modified eigenvalues of the Hessian matrix.

We have introduced in the metric definition the minimum and maximum edge lengths to avoid any "unrealistic" metric tensor. We would, for instance, have elements of infinite length where the solution is linear. The minimum length is used to avoid having a concentration of very small elements close to singularities.

If the adaptation is done with multiple fields, for example surface velocities and bedrock topography simultaneously, one needs to use two different metric tensors $\mathbf{M}_1$ and $\mathbf{M}_2$ that are based on their respective Hessian matrices. Alauzet and Frey (2003) showed that the relation

$$\mathbf{N} = {\mathbf{M}_1}^{-1}\mathbf{M}_2 \tag{3.25}$$

provides a metric tensor, $\mathbf{N}$, that satisfies both metric tensors $\mathbf{M}_1$ and $\mathbf{M}_2$. The tensor $\mathbf{M}_1$ being symmetric can be diagonalized. This operation is called simultaneously matrix reduction. This process can be repeated with all the metrics that need to be taken into account.

The final metric tensor is also smoothed to avoid large variations of element size. We limit, for example, the ratio between two neighboring elements to be at most 2. This means that the characteristic length of two connected elements varies by a maximum factor of 2.

3.2.5 Hessian matrix estimate

The upper bound of the error estimate equation (3.13) involves the Hessian matrix of the *exact* solution, u, and is hence also present in the metric construction (3.22). But in practice, the exact solution, u, is not always known: we only know a discrete solution u_h. In ice sheet modeling, for instance, we have InSAR derived velocity maps that have a spatial resolution of about 100 meters. These velocities are interpolated on the mesh node, which constitutes our discretized u_h. If InSAR data are not available, one can also use a model result, which would give the velocity estimate for each node. Since u is interpolated, it is piecewise linear and its Hessian matrix is zero over the elements and infinite over the edges. The Hessian matrix $\mathbf{H_u}$ is estimated over each element by a double $\mathcal{L}^2$ projection or a least mean square method (See Alauzet (2003); Alauzet and Saiac (2007) for more details).

3.2.6 Adaptation algorithm

In ice sheet modeling, the unknown is the velocity field $\boldsymbol{v}$. An estimate of this velocity field is given by InSAR data. This observed velocity is interpolated on a given initial mesh. The metric is computed using (3.22) for each node of the mesh. Then we use an edge-based anisotropic mesh optimization methodology inspired by Frey (2001) (YAMS) and Hecht (2006) (BAMG) to generate a new mesh. The main steps of the adaptation algorithm are as follows:

1. Generate a new mesh with only vertices on the domain outline (boundary)
2. Apply to all the vertices of the initial mesh:
   (a) Calculate the distance between the current vertex to the closest vertex of the new mesh
   (b) If this distance is greater than 1: insert the vertex in the new mesh, while keeping the triangulation Delaunay conforming
3. Apply to all the edges of the initial mesh:
   (a) Calculate the edge length in the metric provided
   (b) If the edge length is smaller than 1, go to next edge
   (c) Calculate the number of subdivisions (floor of the edge length)
   (d) Calculate the position of the additional vertices
   (e) Introduce additional vertices one by one in the new mesh so that the triangulation remains Delaunay conforming

The observed velocities are re-interpolated on this new mesh and the previous step is repeated several times to make sure that all the ice streams are captured (See Fig.3.8).

During development, we tried to base our metric on fields such as the bed topography or surface elevation, but the final mesh was always close to isotropic and uniform. We therefore base our metric on the surface velocities, but limit the maximum element size to capture the geometrical features that we would miss otherwise.

Once a satisfying mesh is reached, it is kept fixed horizontally in time. It is therefore a *static* mesh adaptation. We assume that ice streams and shear margins remain at the same locations as today for short term projections, which is supported by observations.

3.2.7 Application to Jakobshavn Isbræ

An example of static adaptive anisotropic mesh refinement is shown on Fig. 3.9 for Jakobshavn Isbræ, West Greenland. Fig. 3.9a shows a uniform mesh (with elements of approximately equal area) comprising about 1,500 elements. Fig. 3.9b shows a mesh refined using InSAR surface velocities with the same number of elements. Shear margins are well captured by the algorithm, while the interior of the ice sheet, where ice flow deformation is weak, is captured by a coarser mesh. Fig. 3.10 shows the quality of the elements for both meshes and the distribution of the adapted mesh is centered around 1, which shows that most of the elements of the adapted mesh are close to equilateral in their Euclidean metric. Comparison between the two meshes shows that for the same number of elements, static anisotropic mesh refinement captures ice flow far more efficiently, resulting in tremendous computational gains.

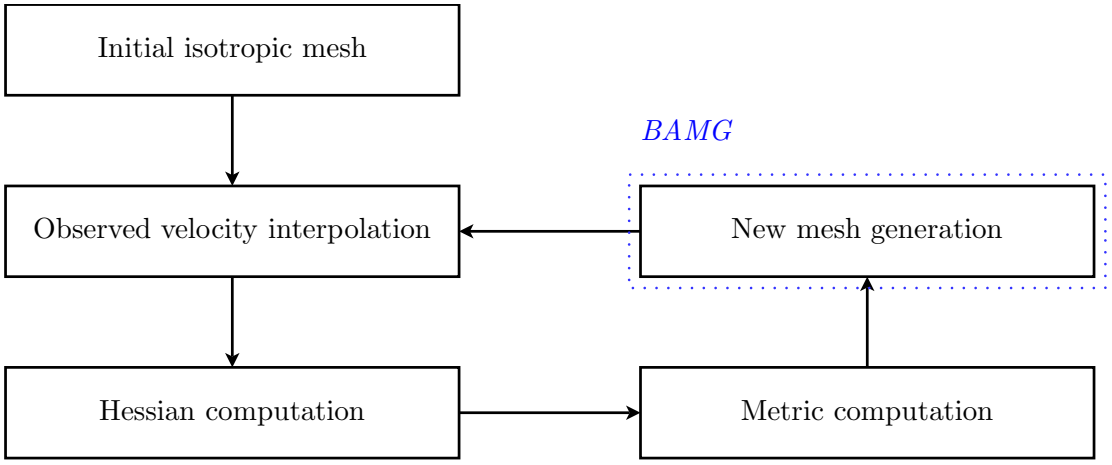


Figure 3.8: Anisotropic mesh adaptation procedure

Conclusion

In this chapter, we presented how ISSM solves the equations that govern transient ice sheet flow. These equations are solved sequentially using the finite element method. We then briefly presented key details regarding ISSM and its parallel architecture. In the second part, we showed how the computational mesh can be optimized to minimize the number of elements (which limits the computational cost) while preserving precision. We now have all the tools to model an ice sheet flow. Nevertheless, some parameters, such as basal friction, cannot be directly measured, but are required to run a simulation. To overcome this difficulty we use inverse modeling. This is the subject of the next chapter.

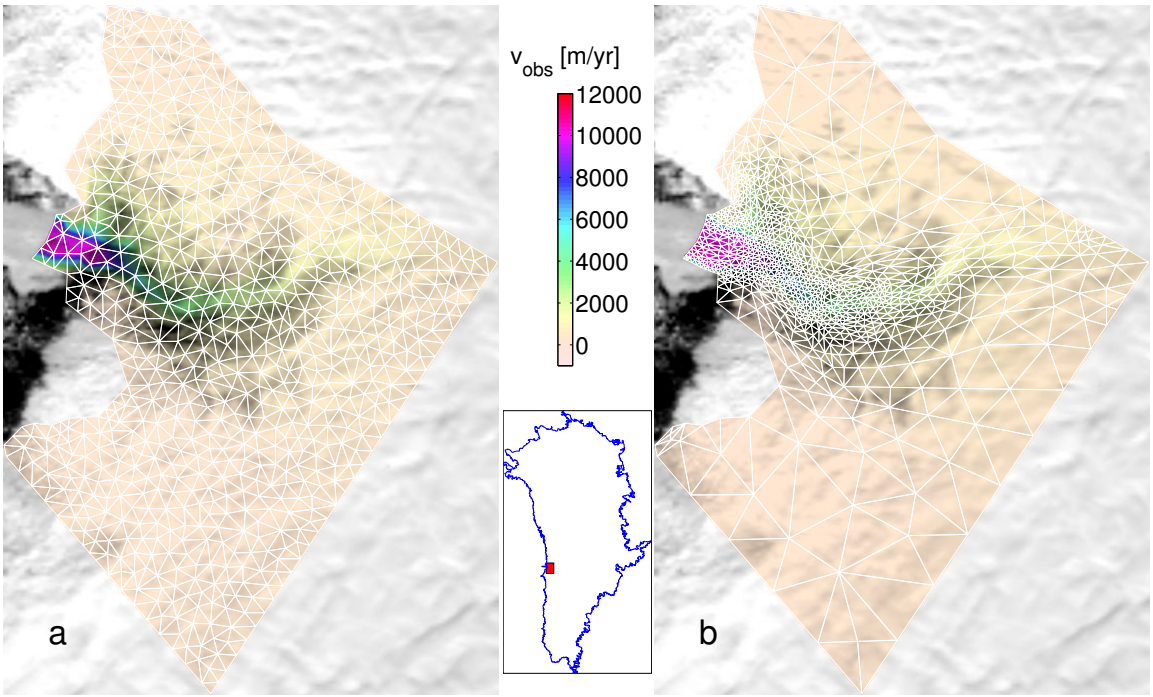


Figure 3.9: Uniform and adapted mesh of Jakobshavn Isbræ, West Greenland. (a) InSAR surface velocity from Rignot (2008b) interpolated on a uniform mesh (in white). (b) InSAR surface velocity interpolated on adapted mesh. Both meshes comprise 1,500 elements.

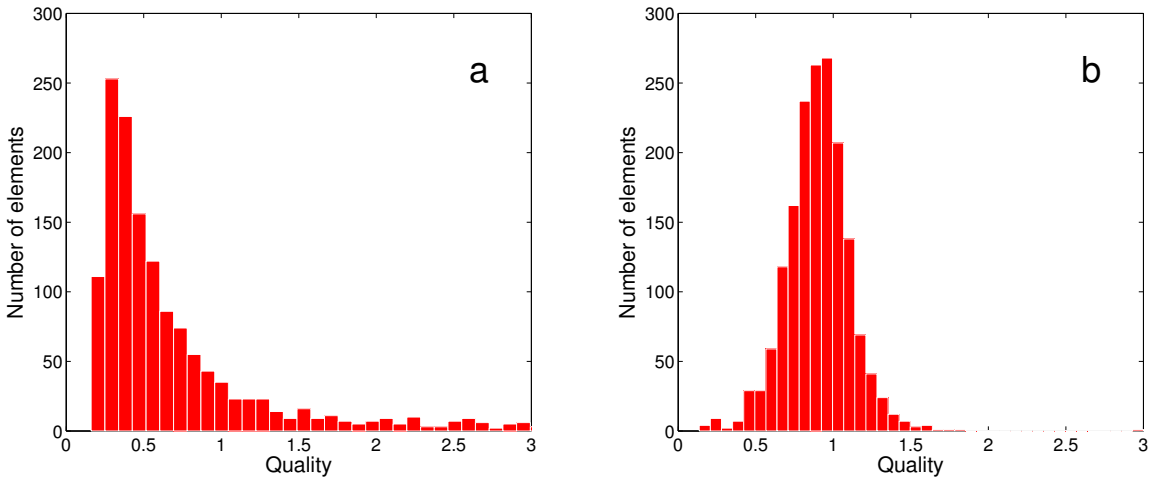


Figure 3.10: Histograms of mesh quality: (a) uniform mesh, (b) adapted mesh

4

Inferring ice sheet properties from surface observations

Many ice sheet properties can be directly measured from remote sensing or in-situ observations (surface height, velocity, etc.). Other parameters, such as basal friction or ice hardness, are poorly understood, because they cannot be measured directly. Accurate knowledge of some of these parameters is critical, because they control ice speed. MacAyeal (1992, 1993) introduced inverse problems in glaciology, sometimes referred to as "control methods". The idea is to use observed surface properties, namely ice velocity, to infer these poorly known parameters. We extend this technique to higher-order and full-Stokes models, and present some significant improvements in solving these inverse problems.

4.1 Inverse problems in ice sheet modeling

4.1.1 Poorly-known parameters

One of the central difficulties in many models of glacier and ice sheet flow lies in the prescription of boundary conditions at the bed. Often, processes that occur there dominate the evolution of the ice mass, as they control the speed at which the ice is able to slide over the bed. The basal friction expression chosen here is (See section 2.2.3):

$$\boldsymbol{\tau}_b \quad = \quad -\alpha^2 \boldsymbol{v}_b \tag{4.1}$$

The constant α depends on the thermal and mechanical properties of the ice and inversely on the bed roughness. It is extremely difficult to evaluate the spatial distribution of this parameter given the number of unknowns. In most cases, α is treated as an adjustable parameter and inferred using control methods. (MacAyeal, 1992, 1993) applied an inverse method on Ice Stream E, Antarctica, with his 2d Shelfy Stream model to infer the distribution of shear stress under ice streams using surface velocity data. Joughin et al. (2004b) modified the method to yield a direct inversion for the basal stress corresponding to a weak plastic bed.

Another parameter that is poorly known is the ice hardness of the ice shelves. The ice viscosity is defined by Eq. (2.26)

$$\mu = \frac{B}{2\dot{\varepsilon}_e^{1-\frac{1}{n}}}$$

The ice hardness is B. It is highly temperature dependent, as shown in Table 4.1.

The spatial distribution of temperature can vary significantly for a glacier of several thousands of km^2. Moreover the ice hardness also depends on the fabric, impurity, water content (Paterson, 1994) and damage (fractures) that can lead to ice shelf collapse (Khazendar et al., 2007, 2009, 2011).

Rommelaere and MacAyeal (1997) extended MacAyeal's algorithm to infer the ice hardness, B, of the Ross ice shelf, Antarctica, with the same 2d shelfy stream model, from surface velocity observations. Larour (2005) applied a similar method to infer the ice hardness of the Ronne ice shelf, Antarctica.

The problem with these two parameters is that they both control the ice velocity. Decreasing B is generally equivalent to decreasing the basal friction, k. If we invert for the two parameters at the same time and at the same locations, the solution is not unique because different distributions of B and k can have the same effect on the surface velocities.

Temperature ($^\circ C$)	B (kPa $a^{1/3}$)
0	167
-2	236
-5	271
-10	402
-15	478
-20	571
-25	696
-30	854
-35	1055
-40	1313
-45	1655
-50	1994

Table 4.1: viscosity parameter B as a function of the ice temperature (Paterson, 1994, p.97)

The solution that we propose here is to limit the inversion of ice hardness to floating ice only, and rely on thermal modeling to infer the ice hardness on grounded ice (Table 4.1). By doing so, we limit the number of solutions.

Special care must be taken when analyzing the results of the inversion, because the inferred parameters have been calculated so that the model fits the observations. The inferred basal friction, for example, might not be the *true* basal friction, but is the friction that must be applied to the model in order to match the velocity measurements. The pattern of basal friction might depend on the model assumptions and parameterization, as we will see in the application to Pine Island Glacier. If, for example, the thermal model is not accurate, the inferred basal friction will compensate for these deviations in order for the model to fit the measurements. The basal friction might not reflect the true physical distribution, but rather the errors in ice temperatures.

4.1.2 Methodology

Several approaches can be employed to solve these inverse problems. The goal is to minimize the misfit between the modeled ice velocity resulting from the solution of the stress-balance equations, and surface measurements of ice velocity derived from satellite synthetic-aperture radar interferometry (InSAR). This technique, described further in App. A, consists of measuring the surface displacement from space between two passes. We apply here the classical adjoint method introduced in Chap. 1.

Other approaches have been proposed. Arthern and Gudmundsson (2010) presents a method that does not require a separate adjoint model, by solving a Robin problem. Another completely different approach has been investigated by Raymond and Gudmundsson (2009), which consists of estimating basal properties of glaciers using a probabilistic Bayesian approach to estimate bedrock topography and basal slipperiness at the same time, from surface velocities and surface elevation. This latter technique is still under development: its application on real cases is complex as many assumptions must be taken, such as space periodicity, which are not verified in real cases.

4.2 Continuous and discrete inverse problem formulations

An important question that is not always addressed is the following: should we discretize the optimization problem first and solve a discrete optimization problem, or should we optimize the continuous problem first and obtain a set of equations to discretize? The first approach is often referred to as "Discretize-Optimize" (DO) and the second approach is known as "Optimize-Discretize" (OD).

In Chap. 1, we derived the adjoint and the gradient from a continuous formulation and adopted an OD approach, but this is not always the preferred approach. In most climate models, the DO approach is used (Wunsch, 2006; Heimbach and Bugnion, 2009), because the adjoint is generated by automatic-differentiation (AD). The gradient of the cost function is generated by a program from the implementation of the discretized model equations. An important challenge in optimization is that, in general, these two steps do not commute. As a result, the two different approaches could lead to two different solutions (Haber and Hanson, 2007).

We derive here the adjoint and gradient for a simple optimization problem of ice sheet modeling, which consists of identifying the ice viscosity, assumed to be linear for the sake of simplicity, with the full-Stokes equations from surface velocity measurements. We then compare the two approaches, OD and DO. This section has two purposes: (1) introduce the technique to derive the adjoint equations and gradient in detail for a simplified case and (2) compare the Discretize-Optimize and Optimize-Discretize approaches.

4.2.1 Simplified optimization problem

As an illustration, we consider here a simplified problem of identification. We invert for the ice viscosity assumed to be linear (μ is a scalar independent of the velocity) for a full-Stokes problem, using a quadratic cost function. The method of derivation of the adjoint and gradient presented in this section would be similar for other models, other cost functions and other parameters.

We want to find the viscosity, $\mu \in \mathcal{K}$, assumed to be linear here, that minimizes the misfit between the modeled and measured horizontal surface velocities. We introduce the local cost function, j, such that:

$$j\left(\boldsymbol{v}\right) = \frac{1}{2}\left(\left(v_x - v_x^{\text{obs}}\right)^2 + \left(v_y - v_y^{\text{obs}}\right)^2\right) \tag{4.2}$$

where $\boldsymbol{v} = (v_x, v_y)$ is the modeled velocity and $\boldsymbol{v}^{\text{obs}} = \left(v_x^{\text{obs}}, v_y^{\text{obs}}\right)$ is the observed velocity. We define the objective function as the integral of the local cost function over the glacier upper surface, S:

$$J\left(\boldsymbol{v}\right) = \int_S j\left(\boldsymbol{v}\right) dS \tag{4.3}$$

As in Chap. 1, we introduce another cost function, $\mathcal{J}$, which is a function of the ice viscosity, μ, such that:

$$\mathcal{J}\left(\mu\right) = J\left(\boldsymbol{v}\left(\mu\right)\right) \tag{4.4}$$

where $\boldsymbol{v}(\mu)$ is the solution of the full-Stokes equations for the viscosity μ. It is convenient to introduce $\mathcal{J}$ because finding the viscosity that minimizes J under the constraint that the velocity must satisfy the model equations, is equivalent to minimizing $\mathcal{J}$ wihtout constraint.

The model domain is Ω and we impose homogeneous Dirichlet constraints on Γ_u and a force $\boldsymbol{f}$ on the Neumann boundary Γ_σ. This is a PDE-constrained optimization problem:

$$\begin{aligned}
&\text{Objective function:} \quad J(\boldsymbol{v}) = \frac{1}{2}\int_S \left(v_x - v_x^{\text{obs}}\right)^2 + \left(v_y - v_y^{\text{obs}}\right)^2 dS \\
&\text{PDE constraint:} \quad \begin{cases} \nabla\cdot\boldsymbol{\sigma} + \rho\,\boldsymbol{g} &= \boldsymbol{0} & \text{in } \Omega \\ \nabla\cdot\boldsymbol{v} &= 0 & \text{in } \Omega \\ \boldsymbol{\sigma} &= 2\mu\dot{\boldsymbol{\varepsilon}} - p\mathbf{I} & \text{in } \Omega \\ \boldsymbol{v} &= \boldsymbol{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{f} & \text{on } \Gamma_\sigma \end{cases} \\
&\text{Control parameter:} \quad \mu
\end{aligned} \tag{4.5}$$

4.2.2 Continuous formulation

Let's first derive the adjoint equations and gradient using a continuous formulation. The *state variables* are $(\boldsymbol{v}, p)$ and we introduce the *adjoint state* $(\boldsymbol{\lambda_v}, \lambda_p)$. The Lagrangian of this optimization problem is:

$$\mathcal{L}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu) = J(\boldsymbol{v}) + \int_\Omega \boldsymbol{\lambda_v}\cdot(\nabla\cdot\boldsymbol{\sigma} + \rho\,\boldsymbol{g})\,d\Omega + \int_\Omega \lambda_p \nabla\cdot\boldsymbol{v}\,d\Omega \tag{4.6}$$

We introduce the following notation: $\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}$ and $\boldsymbol{\sigma}_{(\boldsymbol{v},p)}$ are respectively the strain rate tensor and the stress tensor associated to the velocity, $\boldsymbol{v}$, and pressure, p. It is convenient to express the Lagrangian in different forms, using several integrations by parts.

We have the following identities:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu) &= J(\boldsymbol{v}) + \int_\Omega \boldsymbol{\lambda_v}\cdot(\nabla\cdot\boldsymbol{\sigma} + \rho\,\boldsymbol{g})\,d\Omega + \int_\Omega \lambda_p\nabla\cdot\boldsymbol{v}\,d\Omega \\
\mathcal{L}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu) &= J(\boldsymbol{v}) + \int_\Omega \boldsymbol{\lambda_v}\cdot(\rho\,\boldsymbol{g})\,d\Omega - \int_\Omega 2\mu\, Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right)d\Omega \\
&\quad + \int_\Omega p\,\nabla\cdot\boldsymbol{\lambda_v}\,d\Omega + \int_{\Gamma_\sigma}\boldsymbol{\lambda_v}\cdot\boldsymbol{f}\,dS + \int_\Omega \lambda_p\nabla\cdot\boldsymbol{v}\,d\Omega \\
\mathcal{L}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu) &= J(\boldsymbol{v}) + \int_\Omega \boldsymbol{v}\cdot\nabla\cdot\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}d\Omega - \int_{\Gamma_\sigma}\boldsymbol{v}\cdot\left(\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}\cdot\boldsymbol{n}\right)dS \\
&\quad + \int_\Omega p\,\nabla\cdot\boldsymbol{\lambda_v}\,d\Omega + \int_{\Gamma_\sigma}\boldsymbol{\lambda_v}\cdot\boldsymbol{f}\,dS + \int_\Omega \boldsymbol{\lambda_v}\cdot(\rho\,\boldsymbol{g})\,d\Omega
\end{aligned} \tag{4.7}$$

PROOF
We integrate by part the second term of the Lagrangian, $\mathcal{L}$:

$$\begin{aligned}
\int_\Omega \boldsymbol{\lambda_v} \cdot \nabla \cdot \boldsymbol{\sigma} d\Omega &= -\int_\Omega Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} \boldsymbol{\sigma}_{(\boldsymbol{v},p)}\right) d\Omega + \int_{\partial\Omega} \boldsymbol{\lambda_v} \cdot \left(\boldsymbol{\sigma}_{(\boldsymbol{v})} \cdot \boldsymbol{n}\right) dS \\
&= -\int_\Omega Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} \boldsymbol{\sigma}_{(\boldsymbol{v},p)}\right) d\Omega + \int_{\Gamma_\sigma} \boldsymbol{\lambda_v} \cdot \boldsymbol{f} dS \\
&= -\int_\Omega Tr\left(2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right) d\Omega + \int_\Omega p \, \nabla \cdot \boldsymbol{\lambda_v} d\Omega + \int_{\Gamma_\sigma} \boldsymbol{\lambda_v} \cdot \boldsymbol{f} dS \\
&= \int_\Omega \boldsymbol{v} \cdot \nabla \cdot 2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} d\Omega - \int_{\partial\Omega} \boldsymbol{v} \cdot \left(\left(2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}\right) \cdot \boldsymbol{n}\right) dS \\
&\quad + \int_\Omega p \, \nabla \cdot \boldsymbol{\lambda_v} d\Omega + \int_{\Gamma_\sigma} \boldsymbol{\lambda_v} \cdot \boldsymbol{f} dS
\end{aligned}$$

We also integrate by part the last term of the Lagrangian:

$$\begin{aligned}
\int_\Omega \lambda_p \nabla \cdot \boldsymbol{v} d\Omega &= -\int_\Omega \boldsymbol{v} \cdot \nabla \lambda_p d\Omega + \int_{\partial\Omega} \lambda_p \boldsymbol{v} \cdot \boldsymbol{n} dS \\
&= -\int_\Omega \boldsymbol{v} \cdot \nabla \lambda_p d\Omega + \int_{\partial\Omega} \boldsymbol{v} \cdot \left(\lambda_p \mathbf{I} \cdot \boldsymbol{n}\right) dS
\end{aligned} \tag{4.8}$$

If we combine these two equations, we have the following Lagrangian:

$$\begin{aligned}
\mathcal{L}\left(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu\right) &= J\left(\boldsymbol{v}\right) + \int_\Omega \boldsymbol{v} \cdot \nabla \cdot 2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} d\Omega - \int_\Omega \boldsymbol{v} \cdot \nabla \lambda_p d\Omega \\
&\quad - \int_{\partial\Omega} \boldsymbol{v} \cdot \left(\left(2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}\right) \cdot \boldsymbol{n}\right) dS + \int_{\partial\Omega} \boldsymbol{v} \cdot \left(\lambda_p \mathbf{I} \cdot \boldsymbol{n}\right) dS \\
&\quad + \int_\Omega p \, \nabla \cdot \boldsymbol{\lambda_v} d\Omega + \int_{\Gamma_\sigma} \boldsymbol{\lambda_v} \cdot \boldsymbol{f} dS + \int_\Omega \boldsymbol{\lambda_v} \cdot \left(\rho \, \boldsymbol{g}\right) d\Omega
\end{aligned} \tag{4.9}$$

And finally:

$$\begin{aligned}
\mathcal{L}\left(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu\right) &= J\left(\boldsymbol{v}\right) + \int_\Omega \boldsymbol{v} \cdot \nabla \cdot \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} d\Omega - \int_{\Gamma_\sigma} \boldsymbol{v} \cdot \left(\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} \cdot \boldsymbol{n}\right) dS \\
&\quad + \int_\Omega p \, \nabla \cdot \boldsymbol{\lambda_v} d\Omega + \int_{\Gamma_\sigma} \boldsymbol{\lambda_v} \cdot \boldsymbol{f} dS + \int_\Omega \boldsymbol{\lambda_v} \cdot \left(\rho \, \boldsymbol{g}\right) d\Omega
\end{aligned} \tag{4.10}$$

■

Adjoint and gradient

If $(\boldsymbol{v}, p)$ is the solution of the forward model:

$$\left\{\begin{array}{rcll} \nabla\cdot\boldsymbol{\sigma}+\rho\,\boldsymbol{g} &=& 0 & \text{in } \Omega \\ \nabla\cdot\boldsymbol{v} &=& 0 & \text{in } \Omega \\ \boldsymbol{\sigma} &=& 2\mu\dot{\boldsymbol{\varepsilon}}-p\mathbf{I} & \text{in } \Omega \\ \boldsymbol{v} &=& \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}\cdot\boldsymbol{n} &=& \boldsymbol{f} & \text{on } \Gamma_\sigma \end{array}\right. \tag{4.11}$$

If μ and $\boldsymbol{f}$ do not depend on $\boldsymbol{v}$ or p and if $(\boldsymbol{\lambda_v}, \lambda_p)$ is the solution of the adjoint problem:

$$\left\{\begin{array}{rcll} \nabla\cdot\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} &=& 0 & \text{in } \Omega \\ \nabla\cdot\boldsymbol{\lambda_v} &=& 0 & \text{in } \Omega \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} &=& 2\mu\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}-\lambda_p\mathbf{I} & \text{in } \Omega \\ \boldsymbol{\lambda_v} &=& \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}\cdot\boldsymbol{n} &=& \mathbf{0} & \text{on } \Gamma_\sigma\backslash S \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}\cdot\boldsymbol{n} &=& j'(\boldsymbol{v}) & \text{on } S \end{array}\right. \tag{4.12}$$

Then the gradient of the objective function with respect to the viscosity, μ, for a direction ν is:

$$\langle \mathcal{J}'(\mu), \nu\rangle = \langle \frac{\partial \mathcal{L}}{\partial \mu}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu), \nu\rangle \tag{4.13}$$

Proof

We write $\mathcal{V}_v$ the space of kinematically admissible velocities, $\mathcal{V}_p$ the space of admissible pressures. Similarly, $\mathcal{W}_v$ and $\mathcal{W}_p$ are the spaces of admissible adjoint velocities and pressures. If $(\boldsymbol{v}, p)$ is solution of the forward problem, then:

$$\mathcal{J}(\mu) = \mathcal{L}(\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p, \mu) \tag{4.14}$$

if we take the derivative of this equation with respect to the viscosity and apply the chain rule:

$$\begin{aligned} \langle \mathcal{J}'(\mu), \mu'\rangle &= \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}}(\boldsymbol{v}, p), \langle \frac{\partial \boldsymbol{v}}{\partial \mu}, \mu'\rangle\rangle + \langle \frac{\partial \mathcal{L}}{\partial p}(\boldsymbol{v}, p), \langle \frac{\partial p}{\partial \mu}, \mu'\rangle\rangle \\ &+ \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda_v}}(\boldsymbol{v}, p), \langle \frac{\partial \boldsymbol{\lambda_v}}{\partial \mu}, \mu'\rangle\rangle + \langle \frac{\partial \mathcal{L}}{\partial \lambda_p}(\boldsymbol{v}, p), \langle \frac{\partial \lambda_p}{\partial \mu}, \mu'\rangle\rangle \\ &+ \langle \frac{\partial \mathcal{L}}{\partial \mu}, \mu'\rangle \end{aligned} \tag{4.15}$$

The first form of the Lagrangian in Eq. (4.7) gives:

$$\forall \boldsymbol{\mu_v} \in \mathcal{W}_v \quad \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda_v}}, \boldsymbol{\mu_v}\rangle = \int_\Omega \boldsymbol{\mu_v}\cdot(\nabla\cdot\boldsymbol{\sigma}+\rho\,\boldsymbol{g})\,d\Omega \tag{4.16}$$

$$\forall \mu_p \in \mathcal{W}_p \quad \langle \frac{\partial \mathcal{L}}{\partial \lambda_p}, \mu_p \rangle = \int_\Omega \mu_p \nabla \cdot \boldsymbol{v} d\Omega \tag{4.17}$$

And because $(\boldsymbol{v}, p)$ is solution of the forward model, these derivatives vanish:

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda_v}} (\boldsymbol{v}, p) = \mathbf{0} \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \lambda_p} (\boldsymbol{v}, p) = \mathbf{0} \tag{4.18}$$

The third form of the Lagrangian in Eq. (4.7) gives:

$$\forall \boldsymbol{w} \in \mathcal{V}_v \quad \langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}}, \boldsymbol{w} \rangle = \langle J'(\boldsymbol{v}), \boldsymbol{w} \rangle + \int_\Omega \boldsymbol{w} \cdot \nabla \cdot \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} d\Omega - \int_{\Gamma_\sigma} \boldsymbol{w} \cdot \left(\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} \cdot \boldsymbol{n} \right) dS \tag{4.19}$$

$$\forall q \in \mathcal{V}_p \quad \langle \frac{\partial \mathcal{L}}{\partial p}, q \rangle = \int_\Omega q \, \nabla \cdot \boldsymbol{\lambda_v} d\Omega \tag{4.20}$$

And because $(\boldsymbol{\lambda_v}, \lambda_p)$ is solution of the adjoint problem, these derivatives are zero:

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}} (\boldsymbol{\lambda_v}, \lambda_p) = \frac{\partial \mathcal{L}}{\partial p} (\boldsymbol{\lambda_v}, \lambda_p) = \mathbf{0} \tag{4.21}$$

We therefore have for $(\boldsymbol{v}, p)$ and $(\boldsymbol{\lambda_v}, \lambda_p)$ the following equation:

$$\langle \mathcal{J}'(\mu), \nu \rangle = \langle \frac{\partial \mathcal{L}}{\partial \mu} (\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p), \nu \rangle \tag{4.22}$$

■

Now that there is a simple relationship between the gradient of the objective function and the Lagrangian, we can deduce an analytical formulation of this gradient:

Expression of the gradient

Under the same assumptions, the gradient of $\mathcal{J}$ with respect to the viscosity, μ, for a direction ν is:

$$\mathcal{J}'(\mu) = -2 Tr \left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})} \right) \tag{4.23}$$

Proof

The second form of the Lagrangian in Eq. (4.7) gives:

$$\langle \frac{\partial \mathcal{L}}{\partial \mu} (\boldsymbol{v}, p, \boldsymbol{\lambda_v}, \lambda_p), \nu \rangle = - \int_\Omega 2\nu Tr \left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})} \right) d\Omega \tag{4.24}$$

■

Summary

- Forward equations:

$$\left\{\begin{array}{rcll} \nabla\cdot\boldsymbol{\sigma}+\rho\,\boldsymbol{g} &=& 0 & \text{in } \Omega \\ \nabla\cdot\boldsymbol{v} &=& 0 & \text{in } \Omega \\ \boldsymbol{\sigma} &=& 2\mu\dot{\boldsymbol{\varepsilon}}-p\mathbf{I} & \text{in } \Omega \\ \boldsymbol{v} &=& \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}\cdot\boldsymbol{n} &=& \boldsymbol{f} & \text{on } \Gamma_\sigma \end{array}\right. \tag{4.25}$$

- Adjoint equations:

$$\left\{\begin{array}{rcll} \nabla\cdot\boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} &=& 0 & \text{in } \Omega \\ \nabla\cdot\boldsymbol{\lambda_v} &=& 0 & \text{in } \Omega \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} &=& 2\mu\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}-\lambda_p\mathbf{I} & \text{in } \Omega \\ \boldsymbol{\lambda_v} &=& \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}\cdot\boldsymbol{n} &=& \mathbf{0} & \text{on } \Gamma_\sigma\backslash S \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)}\cdot\boldsymbol{n} &=& j'(\boldsymbol{v}) & \text{on } S \end{array}\right. \tag{4.26}$$

- Cost function gradient:

$$\mathcal{J}'(\mu) = -2\mathrm{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right) \tag{4.27}$$

4.2.3 Discrete formulation

We now derive the gradient of the objective function from the discrete finite element equations. The model equations once discretized can be written in matrix form:

$$\mathbf{K}\boldsymbol{U} = \boldsymbol{F} \tag{4.28}$$

where $\mathbf{K}$ is the stiffness matrix, $\boldsymbol{F}$ is the load vector and $\boldsymbol{U}$ is a vector that holds the unknown nodal values of the velocity (See Chap. 3). There are $N \in \mathbb{N}$ model equations:

$$\forall i \in [\![1\ N]\!] \quad f_i - \sum_{j=1}^{N} k_{ij}u_j = 0 \tag{4.29}$$

Now, we want to derive the gradient of the cost function, $\mathcal{J}$, with respect to a discretized parameter. The ice viscosity, μ, is discretized on the mesh as follows:

$$\mu = \sum_{k=1}^{M} \mu_k \phi_k(\boldsymbol{x}) \tag{4.30}$$

where $\phi_k : \mathbb{R}^3 \to \mathbb{R}$ are the nodal functions. The gradient of the cost function is here a vector that measures the variations of the cost function with respect to each component, μ_k:

$$\frac{\partial\mathcal{J}}{\partial\boldsymbol{\mu}} = \left[\frac{\partial\mathcal{J}}{\partial\mu_1} \quad \cdots \quad \frac{\partial\mathcal{J}}{\partial\mu_k} \quad \cdots \quad \frac{\partial\mathcal{J}}{\partial\mu_M}\right]^T \tag{4.31}$$

The Lagrangian associated to the minimization of the objective function, under the constraints of the N model equations is[1]:

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{\Lambda}) = J(\boldsymbol{U}) + \sum_{i=1}^{N} \lambda_i \left(f_i - \sum_{j=1}^{N} k_{ij} u_j \right) = J(\boldsymbol{U}) + \boldsymbol{\Lambda}^T (\boldsymbol{F} - \mathbf{K}\boldsymbol{U}) \tag{4.32}$$

We can here express the gradient of the objective function with respect to every μ_k.

Adjoint and gradient

If $\boldsymbol{U}$ is solution of the forward model:

$$\mathbf{K}\boldsymbol{U} = \boldsymbol{F} \tag{4.33}$$

If the viscosity μ and the load vector $\boldsymbol{f}$ do not depend on the discretized velocity, $\boldsymbol{U}$, and if $\boldsymbol{\Lambda}$ is solution of the adjoint problem:

$$\mathbf{K}^T \boldsymbol{\Lambda} = \frac{\partial J}{\partial \boldsymbol{U}} \tag{4.34}$$

where $\dfrac{\partial J}{\partial \boldsymbol{U}} = \left(\dfrac{\partial J}{\partial \boldsymbol{u}_1}, \ldots, \dfrac{\partial J}{\partial \boldsymbol{u}_N} \right)^T$, then the gradient of the objective function with respect to the nodal viscosity, μ_k, is:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = \frac{\partial \mathcal{L}}{\partial \mu_k}(\boldsymbol{U}, \boldsymbol{\Lambda}) \tag{4.35}$$

Proof

If $\boldsymbol{U}$ is solution of the forward model, we have:

$$\mathcal{J}(\mu_k) = \mathcal{L}(\boldsymbol{U}, \boldsymbol{\Lambda}, \mu_k) \tag{4.36}$$

If we take the derivative of this equation with respect to μ_k

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial u_i}(\boldsymbol{U}, \boldsymbol{\Lambda}, \mu_k) \frac{\partial u_i}{\partial \mu_k} + \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \lambda_i}(\boldsymbol{U}, \boldsymbol{\Lambda}, \mu_k) \frac{\partial \lambda_i}{\partial \mu_k} + \frac{\partial \mathcal{L}}{\partial \mu_k} \tag{4.37}$$

If we admit that neither the stiffness matrix $\mathbf{K}$ nor the load vector $\boldsymbol{F}$ depend on the velocity (in practice the stiffness matrix depends on the velocity but its effect is neglected), we have the following derivatives:

$$\frac{\partial \mathcal{L}}{\partial u_i} = \frac{\partial J}{\partial u_i} - \sum_{j=1}^{N} \lambda_j k_{ji} \tag{4.38}$$

[1] We could equivalently write $\mathcal{L}(\boldsymbol{U}, \boldsymbol{\Lambda}) = J(\boldsymbol{U}) + \boldsymbol{\Lambda}^T (\mathbf{K}\boldsymbol{U} - \boldsymbol{F})$, and this would only change the sign of the Lagrange multipliers. The gradient would then be of opposite sign.

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = f_i - \sum_{j=1}^{N} k_{ij} u_j \tag{4.39}$$

Because $\boldsymbol{U}$ is solution of the forward model:

$$\forall i \in [\![1\ N]\!] \quad \sum_{j=1}^{N} k_{ij} u_j - f_i = 0 \tag{4.40}$$

And if $\boldsymbol{\Lambda}$ is solution of the adjoint problem:

$$\forall i \in [\![1\ N]\!] \quad \sum_{j=1}^{N} \lambda_j k_{ji} = \frac{\partial J}{\partial u_i} \tag{4.41}$$

Therefore, for $\boldsymbol{U}$ solution of the forward problem and $\boldsymbol{\Lambda}$ solution of the adjoint problem:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = \frac{\partial \mathcal{L}}{\partial \mu_k}(\boldsymbol{U}, \boldsymbol{\Lambda}, \mu_k) \tag{4.42}$$

■

Expression of the gradient

Under the same assumptions as before, the gradient of $\mathcal{J}$ with respect to the viscosity at the node k, μ_k, is:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = -\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial k_{ij}}{\partial \mu_k} u_j \lambda_i = -\int_{\Omega} 2\phi_k \, Tr\left(\dot{\boldsymbol{\epsilon}}_{(\boldsymbol{\Lambda})} \dot{\boldsymbol{\epsilon}}_{(\boldsymbol{U})}\right) d\Omega \tag{4.43}$$

Proof

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = \frac{\partial \mathcal{L}}{\partial \mu_k}(\boldsymbol{U}, \boldsymbol{\Lambda}, \mu_k) = \frac{\partial}{\partial \mu_k}\left(J(\boldsymbol{U}) + \sum_{i=1}^{N} \lambda_i \left(f_i - \sum_{j=1}^{N} k_{ij} u_j\right)\right) \tag{4.44}$$

and only the stiffness matrix depends on the viscosity μ_k, which gives the first equality. It can be shown that the finite element formulation gives the second equality. ■

Summary

- Forward equations:

$$\mathbf{K}\boldsymbol{U} = \boldsymbol{F} \tag{4.45}$$

- Adjoint equations:

$$\mathbf{K}^T \mathbf{\Lambda} = \frac{\partial J}{\partial \boldsymbol{U}} \tag{4.46}$$

- Cost function gradient:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = -\int_\Omega 2\phi_k \mathrm{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\mathbf{\Lambda})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{U})}\right) d\Omega \tag{4.47}$$

4.2.4 Are these two approaches equivalent?

Adjoint equations

It can be shown that the stiffness matrix associated to the full-Stokes equations is symmetric positive definite. We therefore have:

$$\mathbf{K}^T \mathbf{\Lambda} = \mathbf{K} \mathbf{\Lambda} = \frac{\partial J}{\partial \boldsymbol{U}} \tag{4.48}$$

This means that the adjoint uses the exact same stiffness matrix as the full-Stokes solution but where the forcing is only the derivative of the objective function with respect to each velocity nodal value. This objective function is evaluated at the surface, since only surface velocities are available. The discretized adjoint is therefore equivalent to the following continuous formulation:

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} &= 0 & \text{in } \Omega \\ \nabla \cdot \boldsymbol{\lambda_v} &= 0 & \text{in } \Omega \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} &= 2\mu \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})} - \lambda_p \mathbf{I} & \text{in } \Omega \\ \boldsymbol{\lambda_v} &= \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} \cdot \boldsymbol{n} &= \mathbf{0} & \text{on } \Gamma_\sigma \backslash S \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v}, \lambda_p)} \cdot \boldsymbol{n} &= j'(\boldsymbol{v}) & \text{on } S \end{cases} \tag{4.49}$$

This shows that the discretized adjoint is strictly equivalent to the discretization of the continuous adjoint. The advantage of continuous adjoints is that there is no need to know how the forward model is solved. One can even use different methods to solve the forward problem and the adjoint equations (for example finite element and finite difference).

Gradients

The discretized form of the continuous gradient (OD approach) is:

$$\mathcal{J}'(\mu) = -2\mathrm{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\mathbf{\Lambda})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{U})}\right) \tag{4.50}$$

and the k^{th} component of the gradient of the discrete problem (OD approach) is:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = -\int_\Omega 2\phi_k \mathrm{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\mathbf{\Lambda})} \dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{U})}\right) d\Omega \tag{4.51}$$

These two gradients are not equal, as one is obviously a vector, and the other a scalar field. The continuous gradient was defined such that:

$$\mathcal{J}(\mu + \varepsilon\nu) = \mathcal{J}(\mu) + \varepsilon\langle \mathcal{J}'(\mu), \nu\rangle + \mathcal{O}(\varepsilon^2) \tag{4.52}$$

and in the discrete case, we have:

$$\mathcal{J}(\boldsymbol{\mu} + \varepsilon\boldsymbol{\nu}) = \mathcal{J}(\boldsymbol{\mu}) + \varepsilon \left[\frac{\partial \mathcal{J}}{\partial \boldsymbol{\mu}}\right]^T \boldsymbol{\nu} + \mathcal{O}(\varepsilon^2) \tag{4.53}$$

We see here a fundamental difference between the two gradients: the continuous gradient is not directly sensitive to the discretization because it has been defined for any direction. The discrete gradient, on the other hand, is designed to evaluate the sensitivity of the cost function with respect to each *nodal* value of the discretized viscosity. How this difference affects the inversion is discussed in the next section.

We can recover the discrete gradient from the continuous gradient, since the k^{th} component of the discrete gradient is the sensitivity of the cost function in the direction of the nodal function ϕ_k:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = \langle \mathcal{J}'(\mu), \phi_k\rangle \tag{4.54}$$

Because the discretized continuous gradient is $\mathcal{J}'(\mu) = -2\text{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\Lambda})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{U})}\right)$ and by definition of the scalar product in this space:

$$\frac{\partial \mathcal{J}}{\partial \mu_k} = -\int_\Omega 2\phi_k \text{Tr}\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\Lambda})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{U})}\right) d\Omega \tag{4.55}$$

4.2.5 Comparison of the two gradients

To illustrate the differences between the two approaches, we focus here on the identification of basal friction. The identification of linear viscosity was chosen for its pedagogical value as the equations were easy to derive. We move here to the identification of basal friction that is further described in the next section. We consider here a square ice sheet and a sinusoid basal drag (Fig. 4.1a). We first calculate the velocity for this pattern of basal friction (Fig. 4.1b).

We take the exact same mesh and now take the model results (Fig. 4.1b) as observations. We set the basal drag to a constant value (Fig. 4.2a), which corresponds to an almost constant velocity (Fig. 4.2b). From this initial state, we try to match the surface velocities calculated with the sinusoidal basal drag. Because it is the exact same model, we could theoretically find the exact solution of the inverse problem and find the sinusoidal pattern.

The inferred basal friction with a continuous gradient (OD approach) is shown in Fig. 4.3a. It is smooth and looks like the exact basal drag. On the other hand, the basal friction inferred using the discrete gradient (DO approach) is extremely noisy and the length of the wiggles is about the same as the mesh size (Fig. 4.4a). The modeled velocity of the approaches matches reasonably well the observed velocities (Fig. 4.1b).

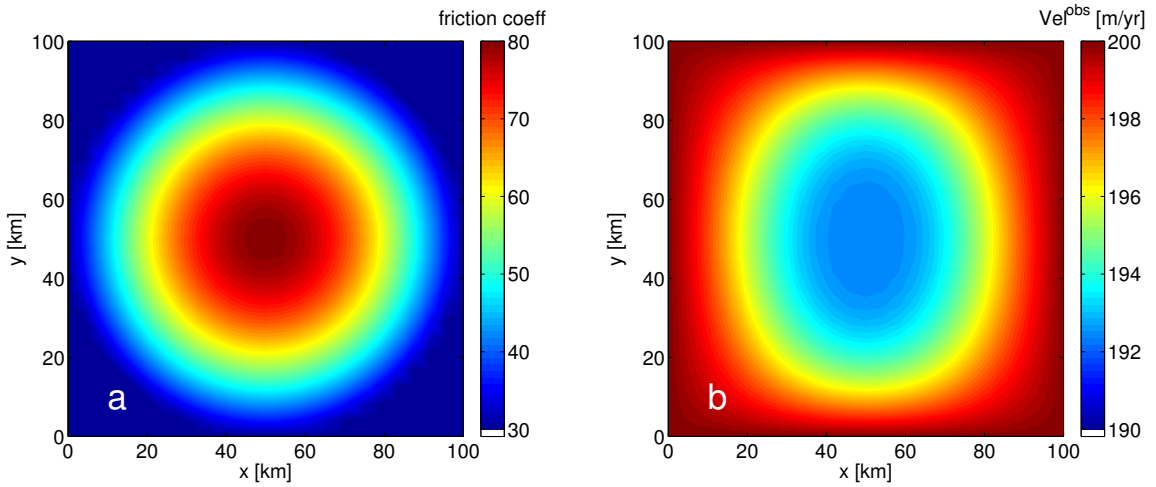


Figure 4.1: Exact basal drag and observed velocities (a) "true" basal drag (b) observed velocities (forward model for this basal drag)

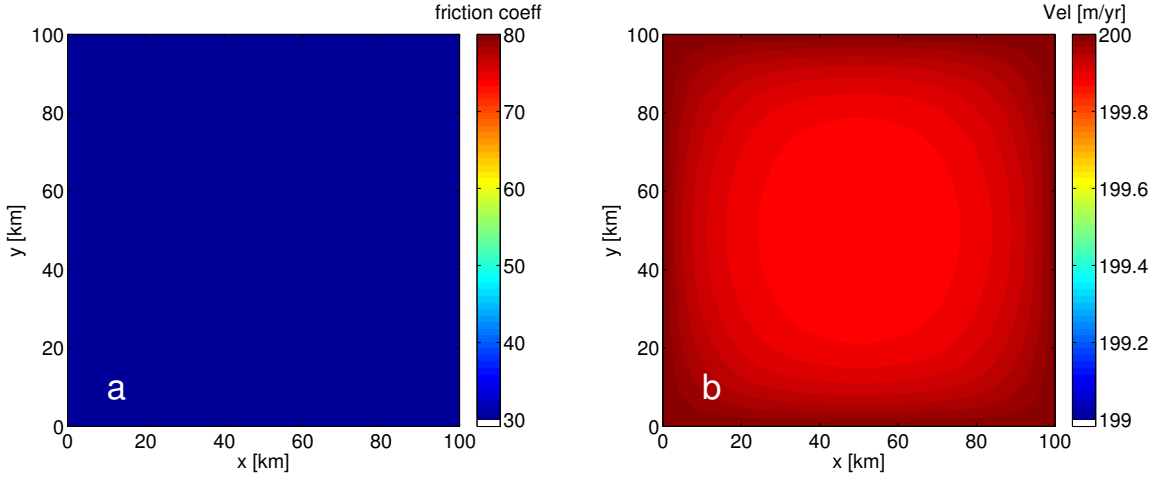


Figure 4.2: (a) Initial basal drag (b) modeled velocities for this basal drag

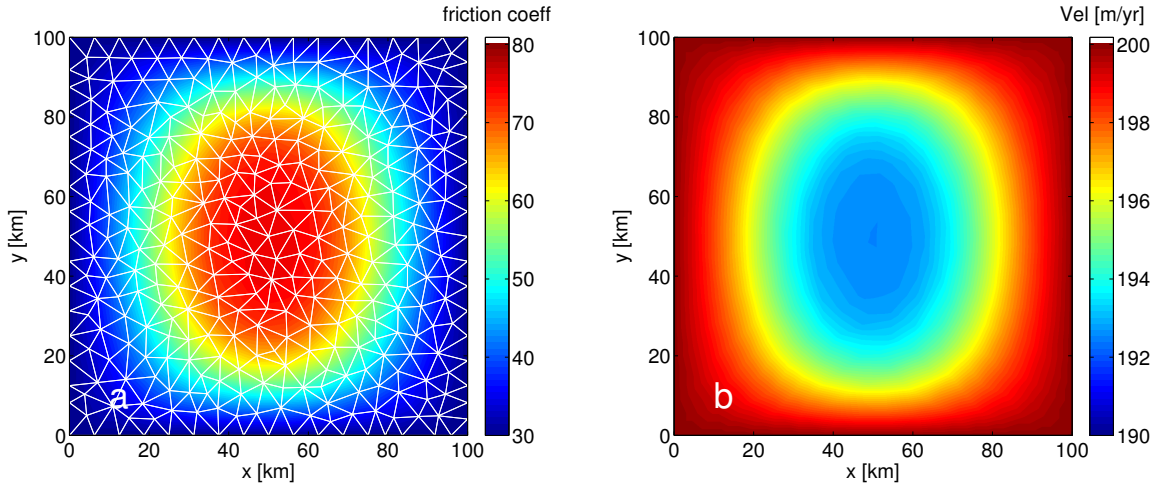


Figure 4.3: Inferred basal drag with continuous gradient (a) inferred basal drag (b) modeled velocities

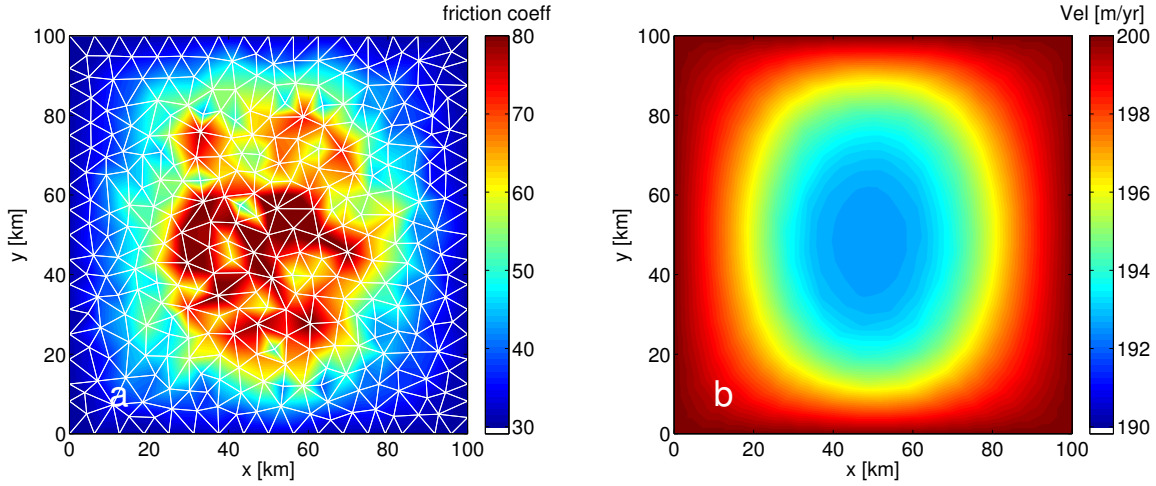


Figure 4.4: Inferred basal drag with discrete gradient (a) inferred basal drag (b) modeled velocities

The wiggles that arise from the DO approach are due to the nature of the discrete gradient, which gives priority to nodes that are connected to the largest number of elements. Indeed, the gradient for the node i measures the sensitivity of the cost function for the i^{th} nodal value of the basal drag. If the elements connected to node i are large, the gradient on i will be large, whereas if the surface of the elements connected to node i is small, the cost function will be less sensitive to the basal friction on node i because it affects a smaller region of the model. The gradient being irregular, the inferred basal friction becomes noisy. The continuous gradient on the other hand is very smooth because it generally involves smooth fields.

There are two ways one can remedy the problem of regularity of the DO approach. The first, the most common method, is to add a regularizing term, R, that penalizes wiggles in the cost function (see Tikhonov regularization in Chap. 1):

$$J(\boldsymbol{v}) = \frac{1}{2}\|\boldsymbol{v} - \boldsymbol{v}^{\text{obs}}\|^2 + R(a) \tag{4.56}$$

with:

$$R(a) = \frac{1}{2}\alpha\|\nabla a\|^2 \tag{4.57}$$

where α is a parameter used to adjust the influence of the regularization with respect to the initial cost function. A large α will result in a smoother basal drag but surface velocities deviate more from the observations whereas a small α will result in a good fit with observations but the basal drag may vary strongly. Fig. 4.5a shows the basal drag for the unstructured mesh with the addition of regularization. The basal drag is smoother and the surface velocities (Fig. 4.5b) well captured. The addition of regularization is therefore crucial so that the inversion is less mesh-dependent for the DO approach.

The second method is to use uniform structured meshes. Nodes of structured meshes have the same weight because the total area associated with each node is the same for every node (except on the boundaries). This is why the gradient of the cost function with respect to the nodal values of the basal drag is naturally smoother, which makes the inferred basal friction also smoother (Fig. 4.5). Most global climate model are based on finite difference schemes and therefore use structured grids. This explains why most of them are not affected by these problems of regularity.

In this thesis we always use the Optimize-Discretize approach. We are in the process of implementing automatic differentiation in ISSM and the problem of regularization will need to be be dealt with. One solution is to precondition the gradient by dividing each component by the integral of its associated nodal function. This process will reduce the gradient's dependency on the mesh's structure.

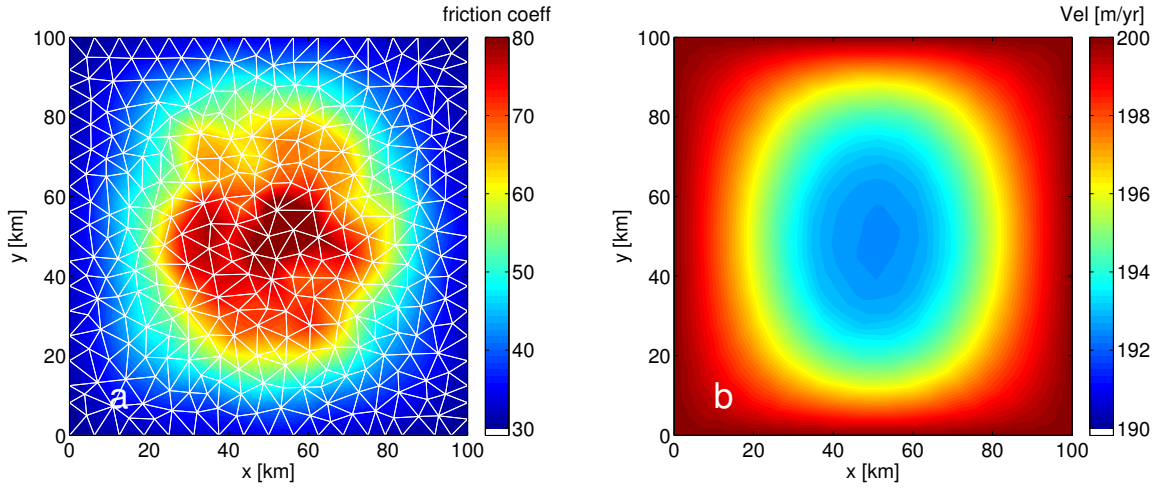


Figure 4.5: Inferred basal drag for the unstructured mesh with Tikhonov regularization (a) inferred basal drag (b) modeled velocities

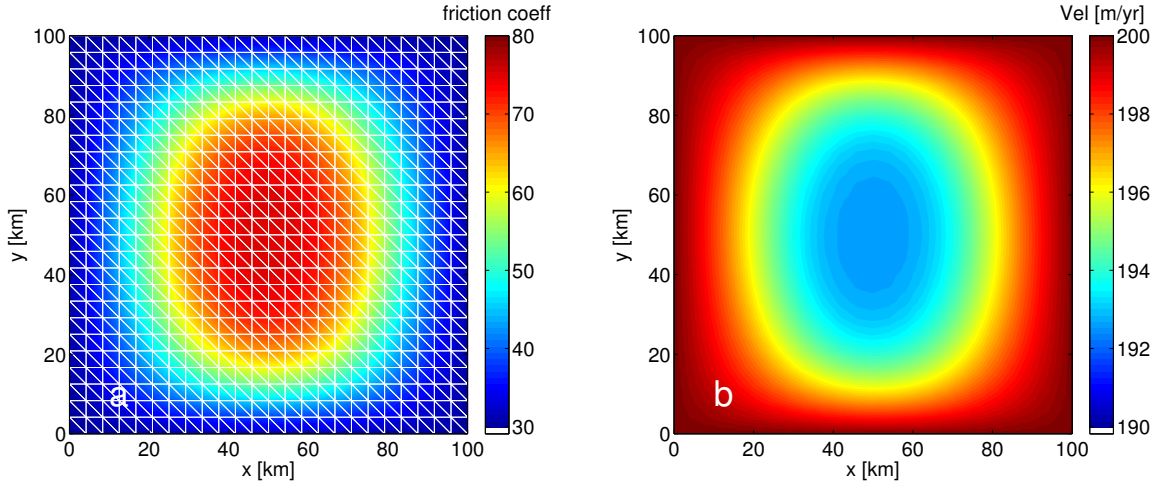


Figure 4.6: Inferred basal drag for the structured mesh (a) inferred basal drag (b) modeled velocities

4.3 Adjoints and gradients for viscosity and basal drag inversions

Here we present the adjoint equations and the gradients used to infer ice hardness and basal friction from surface velocities with three different ice flow models: Morland/MacAyeal's Shelfy-Stream approximation (SSA), Blatter/Pattyn's Higher Order model (HO) and the full-Stokes model (FS). We do not detail the derivation of these adjoints, since the same methodology as the one presented in the previous section (Section 4.2.2) is employed.

4.3.1 Adjoint problems for the three models

We detail here the adjoints for the three ice flow models. These adjoints are all incomplete in the sense that the viscosity is assumed to be linear. This assumption is not correct as the viscosity, μ, depends on the strain rate, but it allows an easier calculation of the adjoint state for the three ice flow models and this approximation is widely employed (MacAyeal, 1993). The viscosity is only considered linear in the adjoint equations, but the forward model includes the non-linear viscosity.

Full-Stokes adjoint

For a linear viscosity, the adjoint state $(\boldsymbol{\lambda_v}, \lambda_p)$ *of the Full-Stokes equations is the solution of the following problem:*

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} &= 0 & \text{in } \Omega \\ \nabla \cdot \boldsymbol{\lambda_v} &= 0 & \text{in } \Omega \\ \boldsymbol{\lambda_v} &= \mathbf{0} & \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} \cdot \boldsymbol{n} &= \mathbf{0} & \text{on } \Gamma_\sigma \backslash S \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v},\lambda_p)} \cdot \boldsymbol{n} &= j'(\boldsymbol{v}) & \text{on } S \end{cases} \tag{4.58}$$

With the following derivatives for the cost function derivative:

$$j'(\boldsymbol{v}) = \begin{bmatrix} v_x - v_x^{obs} \\ v_y - v_y^{obs} \\ 0 \end{bmatrix} \tag{4.59}$$

Proof
See previous section 4.2.2 ∎

The equations governing the adjoint of the full-Stokes equations are very similar to the FS equations. The only difference is that here, we apply forcing only on the ice surface, imposed as the derivative of the cost function with respect to the velocity. The third component of this forcing is always zero because the only available measurements are in the horizontal direction. The adjoint equations of HO and SSA are very similar:

Higher-Order Adjoint

For a linear viscosity, the adjoint state $\boldsymbol{\lambda_v}$ of the Higher-Order equations is solution of the following problem:

$$\begin{cases} \frac{\partial}{\partial x}\left(4\mu\frac{\partial v_x}{\partial x}+2\mu\frac{\partial v_y}{\partial y}\right)+\frac{\partial}{\partial y}\left(\mu\frac{\partial v_x}{\partial y}+\mu\frac{\partial v_y}{\partial x}\right)+\frac{\partial}{\partial z}\left(\mu\frac{\partial v_x}{\partial z}\right)=0 \quad \text{in } \Omega \\ \frac{\partial}{\partial x}\left(\mu\frac{\partial v_x}{\partial y}+\mu\frac{\partial v_y}{\partial x}\right)+\frac{\partial}{\partial y}\left(4\mu\frac{\partial v_y}{\partial y}+2\mu\frac{\partial v_x}{\partial x}\right)+\frac{\partial}{\partial z}\left(\mu\frac{\partial v_y}{\partial z}\right)=0 \quad \text{in } \Omega \\ \boldsymbol{\lambda_v} = \mathbf{0} \qquad \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v})}\cdot\boldsymbol{n} = \mathbf{0} \qquad \text{on } \Gamma_\sigma\backslash S \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v})}\cdot\boldsymbol{n} = j'(\boldsymbol{v}) \qquad \text{on } S \end{cases} \tag{4.60}$$

PROOF
Analogous proof ∎

Finally, the adjoint problem for SSA is similar except that the forcing does not appear as a boundary condition because it is a two-dimensional model.

Shelfy-Stream Adjoint

For a linear viscosity, the adjoint state $\boldsymbol{\lambda_v}$ of the Higher-Order equations is solution of the following problem:

$$\begin{cases} \frac{\partial}{\partial x}\left(4H\bar{\mu}\frac{\partial v_x}{\partial x}+2H\bar{\mu}\frac{\partial v_y}{\partial y}\right)+\frac{\partial}{\partial y}\left(H\bar{\mu}\frac{\partial v_x}{\partial y}+H\bar{\mu}\frac{\partial v_y}{\partial x}\right)=j'(v_x) \quad \text{in } \Omega \\ \frac{\partial}{\partial y}\left(4H\bar{\mu}\frac{\partial v_y}{\partial y}+2H\bar{\mu}\frac{\partial v_x}{\partial x}\right)+\frac{\partial}{\partial x}\left(H\bar{\mu}\frac{\partial v_x}{\partial y}+H\bar{\mu}\frac{\partial v_y}{\partial x}\right)=j'(v_y) \quad \text{in } \Omega \\ \boldsymbol{\lambda_v} = \mathbf{0} \qquad \text{on } \Gamma_u \\ \boldsymbol{\sigma}_{(\boldsymbol{\lambda_v})}\cdot\boldsymbol{n} = \mathbf{0} \qquad \text{on } \Gamma_\sigma \end{cases} \tag{4.61}$$

PROOF
Analogous proof ∎

4.3.2 Basal drag gradients for the three models

The gradient of the cost function with respect to the basal drag coefficient is derived from the partial derivative of the Lagrangian with respect to this parameter. Again, we do not prove these results, but the methodology is the same as the one in Section 4.2.2. It involves both the model result, $\boldsymbol{v}$, and the adjoint state, $\boldsymbol{\lambda}$:

Full-Stokes basal drag gradient

For a cost function that does not depend on basal drag (no regularization), its derivative with respect to basal friction is:

$$\begin{cases} \mathcal{J}'(\alpha) &= 0 & \text{in } \Omega \\ \mathcal{J}'(\alpha) &= -2\alpha\, \boldsymbol{v}_b \cdot \boldsymbol{\lambda}_b & \text{on } \Gamma_b \end{cases} \tag{4.62}$$

PROOF
Analogous proof is an Section 4.2.2 ∎

The gradient is zero except on the ice bedrock interface because friction is only applied at the ice sheet base. For HO, the gradient is similar but only involves the horizontal components:

Higher-Order basal drag gradient

For a cost function that does not depend on basal drag (no regularization), its derivative with respect to basal friction is:

$$\begin{cases} \mathcal{J}'(\alpha) &= 0 & \text{in } \Omega \\ \mathcal{J}'(\alpha) &= -2\alpha\ (v_x\lambda_x + v_y\lambda_y) & \text{on } \Gamma_b \end{cases} \tag{4.63}$$

Finally, for SSA, basal friction appears as a "body force" and not as a boundary condition.

Shelfy-Stream basal drag gradient

For a cost function that does not depend on basal drag (no regularization), its derivative with respect to basal friction is:

$$\mathcal{J}'(\alpha) = -2\alpha\ (v_x\lambda_x + v_y\lambda_y) \qquad \text{in } \Omega \tag{4.64}$$

4.3.3 Ice viscosity gradients for the three models

We consider here the inversion of the ice hardness B. The gradient of the cost function with respect to this parameter for all three models are:

Full-Stokes viscosity gradient

For a cost function that does not depend on the viscosity (no regularization), its derivative with respect to μ is:

$$\mathcal{J}'(\mu) = -\frac{Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda_v})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right)}{\dot{\varepsilon}_e^{1-\frac{1}{n}}} \tag{4.65}$$

PROOF
See previous section 4.2.2 ∎

Higher-Order viscosity gradient

For a cost function that does not depend on the viscosity (no regularization), its derivative with respect to μ is:

$$\mathcal{J}'(\mu) = -\frac{Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right)}{\dot{\varepsilon}_e^{1-\frac{1}{n}}} \tag{4.66}$$

Shelfy-Stream viscosity gradient

For a cost function that does not depend on the viscosity (no regularization), its derivative with respect to μ is:

$$\begin{aligned}
\mathcal{J}'(\mu) &= -H\frac{Tr\left(\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{\lambda})}\dot{\boldsymbol{\varepsilon}}_{(\boldsymbol{v})}\right)}{\dot{\varepsilon}_e^{1-\frac{1}{n}}} \\
&= -\frac{H}{2\dot{\varepsilon}_e^{1-\frac{1}{n}}}\left(\left(2\frac{\partial\lambda_x}{\partial x}\left(2\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\right)\right)+\left(2\frac{\partial\lambda_y}{\partial y}\left(2\frac{\partial v_y}{\partial y}+\frac{\partial v_x}{\partial x}\right)\right)\right. \\
&\quad \left.\left(\frac{\partial\lambda_x}{\partial y}+\frac{\partial\lambda_y}{\partial x}\right)\left(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}\right)\right)
\end{aligned} \tag{4.67}$$

4.4 Application to Pine Island Glacier

We apply here the methods developed in this chapter to a real glacier, West Antarctica. The key aspects of the experiments presented in this section have been published in Morlighem et al. (2010).

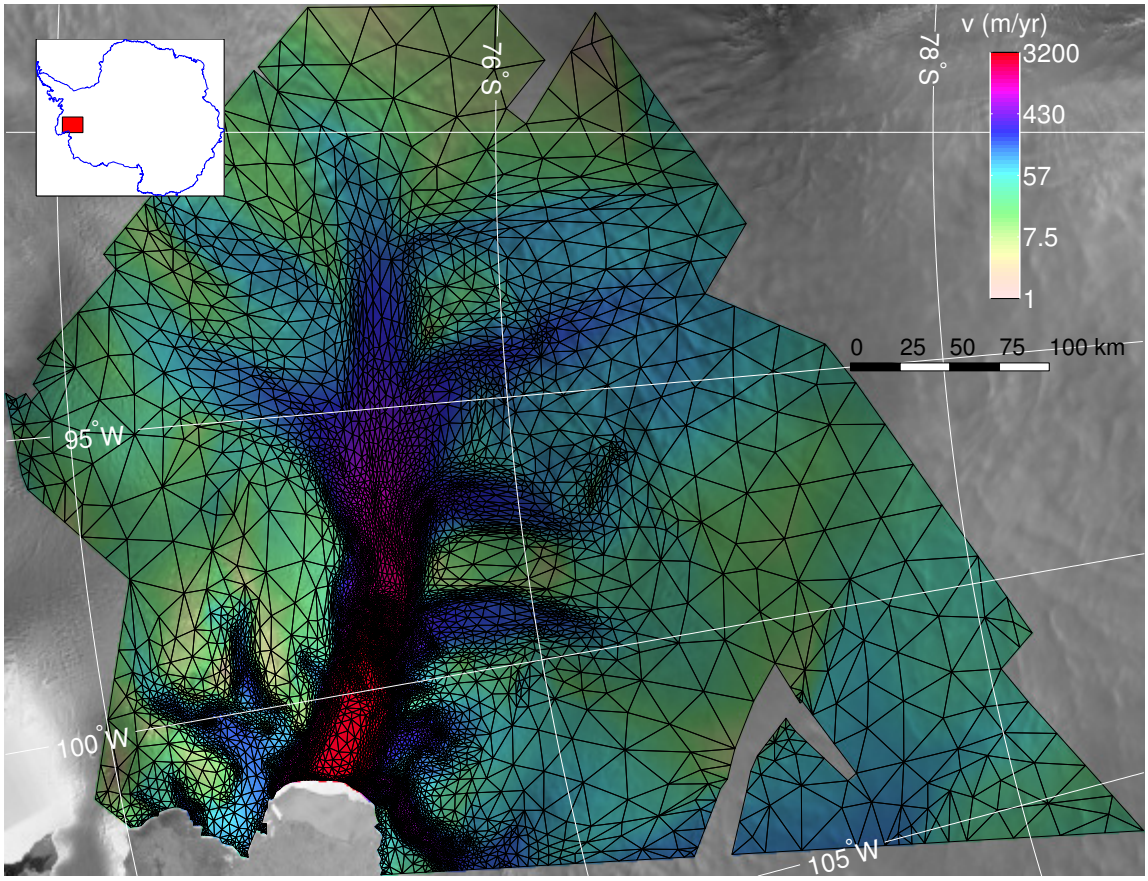


Figure 4.7: Pine Island Glacier surface velocity field in 1996 from Rignot et al. (2002) and 2D anisotropic mesh of 15,000 elements, which is vertically extruded to generate a 3D mesh.

Pine Island Glacier has been retreating, thinning and accelerating steadily since at least the 1970s (Rignot, 2008a). The glacier acceleration at present is several times larger than that estimated for the 1970-1980 and is increasing every year (Rignot, 2008a). The changes taking place on Pine Island Glacier cannot be explained using simple ice flow models such as the Shallow Ice Approximation (SIA) (Hutter, 1983) where stresses are determined locally and all components other than vertical shear are neglected. This limitation has raised the issue of the degree of sophistication needed from numerical models to reproduce the observed ice dynamics. We compare here the patterns of basal friction obtained using three different ice sheet models: full-Stokes, higher-order and shelfy stream. We will also show that the classical quadratic cost function is not optimal for ice sheet systems and will introduce an alternate solution.

4.4.1 Data and model setup

We generated an anisotropic mesh (Fig. 4.7) based on the observed surface velocities following the method presented in Chap. 3. The surface velocities are from Rignot et al. (2002).

Surface topography is from a digital elevation model of Antarctica from Bamber et al. (2009), a firn depth correction from van den Broeke (2008) and ice thickness is from Vaughan et al. (2006). We use a thermal model to calculate the ice hardness over grounded ice, assuming that the ice is in thermal steady-state. The surface temperature is the mean annual air temperature from Giovinetto et al. (1990). The geothermal heat from was provided by Maule et al. (2005).

4.4.2 Choice of objective functions

To infer the ice properties from surface observations, we must choose a cost function. We present here three different functions that are compared after runs of ten steps (we compute ten times a new gradient and follow the steepest descent). The classical cost function is the mean square error:

$$J(\boldsymbol{v}) = \iint_{\Omega} \frac{1}{2}\left(v_x - v_x^{\text{obs}}\right)^2 + \frac{1}{2}\left(v_y - v_y^{\text{obs}}\right)^2 d\Omega \tag{4.68}$$

We have noticed that this cost function is very efficient where the velocities are high, but less efficient in slow moving regions. Fig. 4.8a shows the initial misfit between the modeled and observed surface velocities. This misfit is very high on the ice shelf and on the ice stream where the velocity is highest. Fig. 4.8b shows the final velocity (after optimization) and Fig. 4.8c the final basal friction coefficient. We see that we are missing the tributary glaciers, the main ice stream is well captured but the streams that feed the trunk are missing, which also shows on the basal friction, which is pretty uniform.

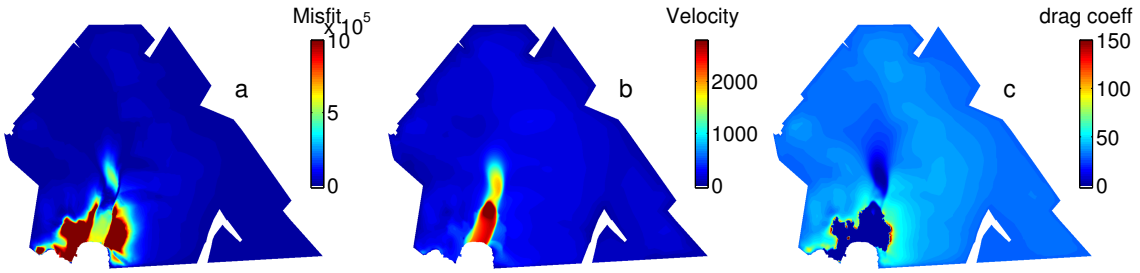


Figure 4.8: "Absolute" cost function (a) Initial misfit distribution, (b) final velocity after convergence, (c) identified basal friction

To capture slower velocity, we tried a relative misfit, defined by:

$$J(\boldsymbol{v}) = \iint_{\Omega} \left(\frac{v_x^{\text{obs}} - v_x}{v_x^{\text{obs}} + \varepsilon}\right)^2 + \left(\frac{v_y^{\text{obs}} - v_y}{v_y^{\text{obs}} + \varepsilon}\right)^2 d\Omega \tag{4.69}$$

where ε is a minimum velocity used to avoid the observed velocity being equal to zero. But this form of the objective function is not adequate because too much weight is given to

the slow moving region. Even if the absolute misfit is small, the relative misfit can reach extremely large values and the optimization procedure concentrates on these regions as shown in Fig. 4.9. The final velocity does not match the observations well, and the basal friction shows peaks, where the algorithm concentrates.

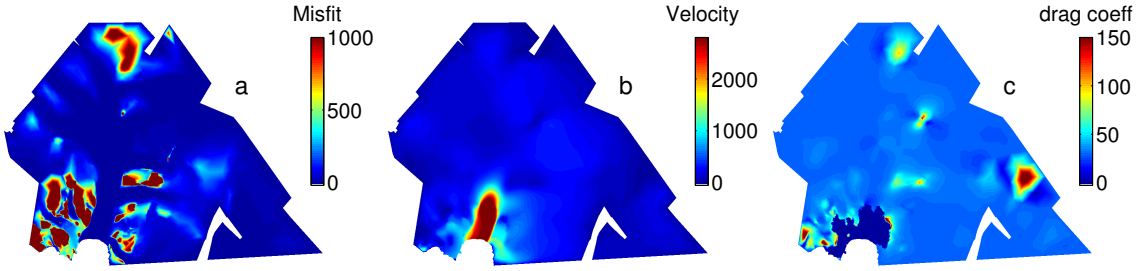


Figure 4.9: "Relative" cost function (a) Initial misfit distribution, (b) final velocity after convergence, (c) identified basal friction

The objective function that gave the best results measured the logarithmic magnitude misfit:

$$J(\boldsymbol{v}) = \iint_{\Omega} \ln\left(\frac{\sqrt{v_x^2 + v_y^2} + \varepsilon}{\sqrt{{v_x^{\text{obs}}}^2 + {v_y^{\text{obs}}}^2} + \varepsilon}\right)^2 d\Omega \tag{4.70}$$

It is *not* a norm because only the magnitude of the velocity is calculated. Nevertheless, the direction of ice flow is controled by the driving stress, which is defined by the surface slopes. The modeled flow direction is therefore naturally close to the observed direction but not the magnitude of the velocity. Fig. 4.10a shows the initial misfit. We see all the tributary

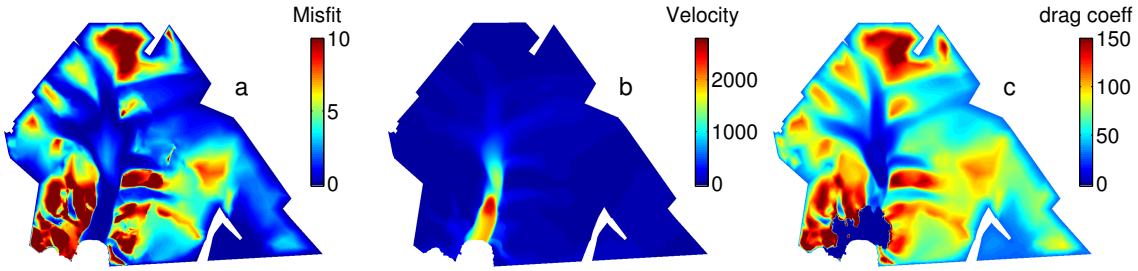


Figure 4.10: "Logarithmic" cost function (a) Initial misfit distribution, (b) final velocity after convergence, (c) identified basal friction

glaciers very clearly. The optimization after only ten steps is very satisfying because the basal friction is identified over the entire domain, not only over the regions of high velocity, and the velocity reproduces the observations very accurately. This cost function is the one that we use here to compare the patterns of basal friction between the three ice flow models.

4.4.3 Algorithm convergence

The surface velocities are shown in Fig. 4.11.

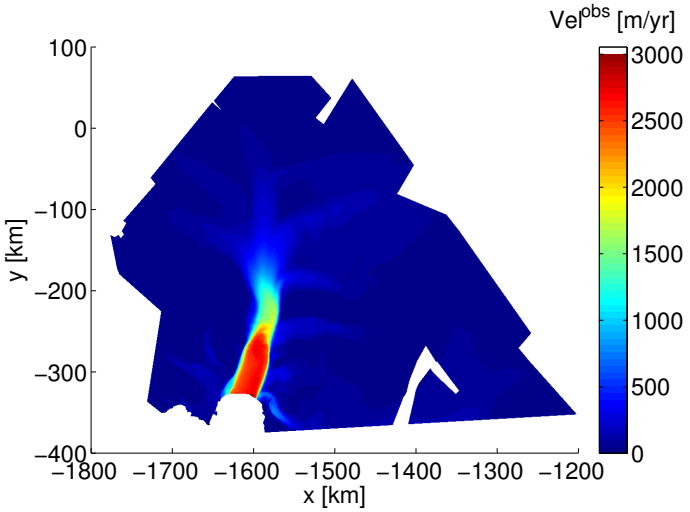


Figure 4.11: Observed surface velocity

We then initialize the basal drag coefficient with a uniform value (Fig. 4.12a). We compute the gradient of the cost function and show the results after 1, 5 and 10 iterations of the steepest descent algorithm (See Fig. 1.2 in Chap. 1). The final velocity (Fig. 4.12h) is very close to the observed velocity.

4.4.4 Comparison of basal friction patterns

We run the same experiment for the three ice flow models using the same mesh and boundary conditions. The inferred patterns of basal drag and the velocity misfits are shown in Fig. 4.13. The optimization scheme converges well for the three models and the modeled velocities reproduce the observed velocities with an excellent accuracy, even in slow-moving regions where InSAR observations are less accurate. The average misfits for the entire domain are: $\bar{M}_{SSA}$=27 m/yr for SSA, $\bar{M}_{BP}$=11.1 m/yr for BP and $\bar{M}_{FS}$=10.4 m/yr for FS. The largest errors are found on fast flow areas. On the ice stream proper, we have $\bar{M}_{SSA}$=62 m/yr, which represents 5% of the average speed in this area, $\bar{M}_{BP}$=22.9 m/yr (1.8%) and $\bar{M}_{FS}$=19.5 m/yr (1.6%).

The spatial patterns of basal drag inferred from the three models (Fig. 4.13) are similar to those inferred from simpler models (Joughin et al., 2009; Vieli and Payne, 2003). The basal drag from FS is closer to BP, as expected since SSA is the most simplified solution. In most areas, the difference in basal drag between solutions is minimal and the agreement between observed and modeled velocity remains excellent. Near the grounding-line, however, SSA and BP exhibit a high basal drag (80 kPa), while the basal drag inferred from FS is less than 10 kPa Fig. 4.14.

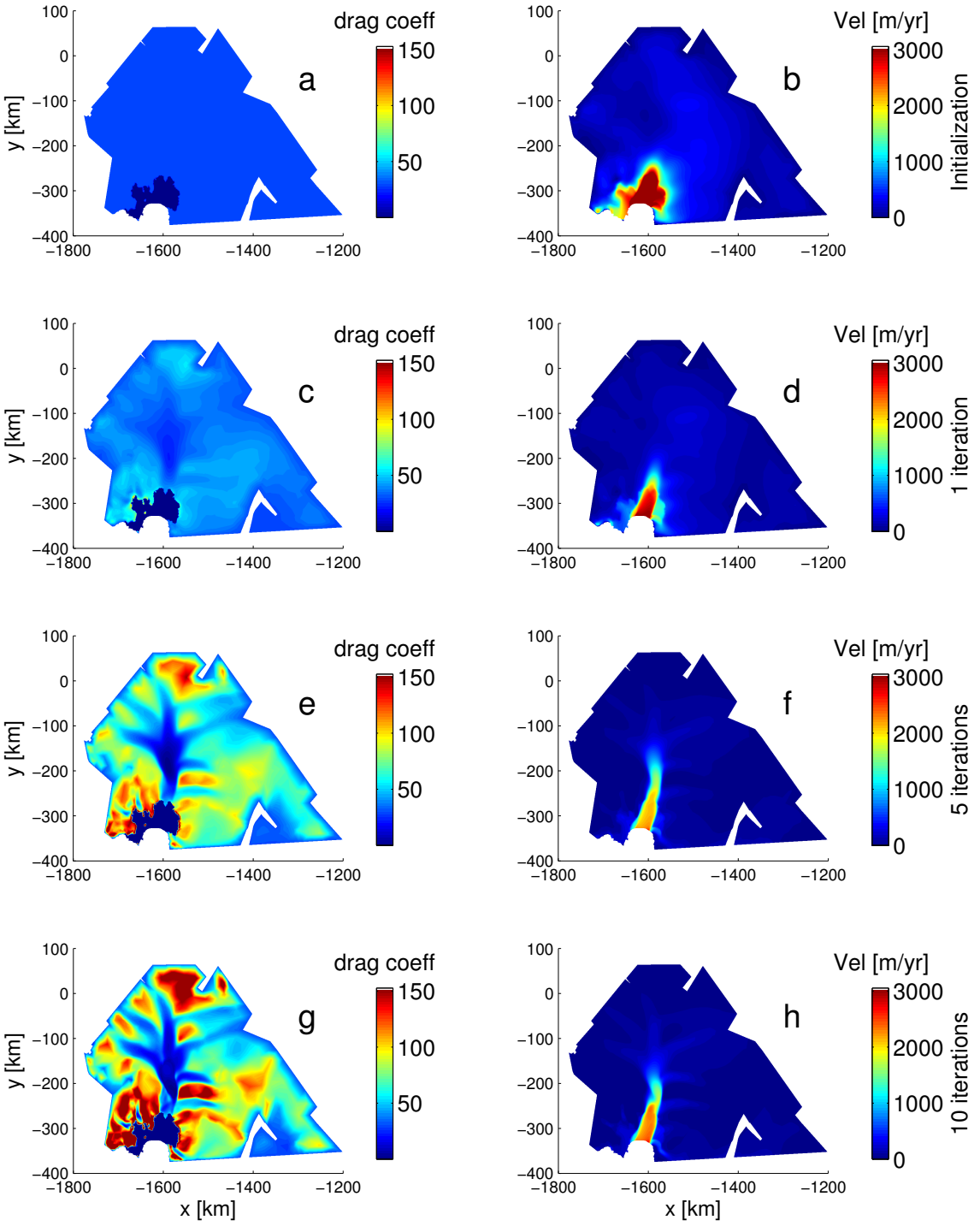


Figure 4.12: Inferred basal friction (left column) and modeled velocity (right column) after 0 (a,b), 1 (c,d), 5 (e,f) and 10 (g,h) iterations

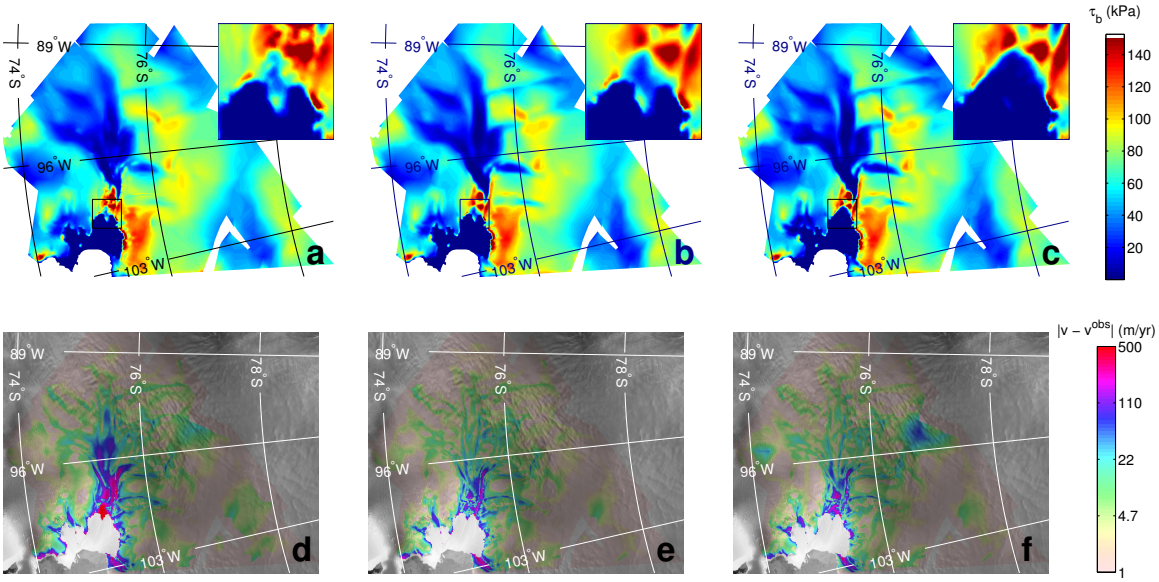


Figure 4.13: Magnitude of the basal drag, $\tau_{\mathbf{b}}$, in kPa, inferred from observations using (a) SSA (b) BP (c) FS, and velocity misfits between observed and modeled velocities in m/yr, using (d) SSA (e) BP and (f) FS.

4.4.5 Discussion

The bed elevation of Pine Island Glacier rises steeply downstream toward the grounding-line: in 20 km the bed rises from -1200m to -550m, i.e. a slope of +3%. It is in this region that the difference in basal drag patterns between the three models is the largest.

BP and SSA both neglect the bridging effect, which simplifies the vertical equation of the momentum balance. Applying the stress-free boundary condition on the upper surface gives:

$$\boldsymbol{\sigma}_{zz} = \rho g\,(z - s) \tag{4.71}$$

At the base of the glacier, we therefore have $\boldsymbol{\sigma}_{zz} = -\rho g H$, where H is the ice thickness. Although this approximation is generally true almost everywhere, our calculation of $\boldsymbol{\sigma}_{zz}$ at the base of the glacier using FS shows that $\boldsymbol{\sigma}_{zz}$ and $\rho g H$ differ by up to 2% in the grounding-line region. The vertical stress applied by the bedrock is indeed slightly larger than the ice column weight, $\rho g H$, as the rising bedrock pushes the ice upward. In FS, the rising bed reduces the ice velocity without additional basal drag. In SSA and BP, the bridging stress is neglected, and the models can only fit the data by increasing basal drag. This increase is not physical but is instead due to an incomplete physics in the two simplified models.

Interestingly, the low basal drag inferred from FS corresponds to a region where the grounding-line probably retreated between 1996 and 2007 (Rignot, 2008a). In this area, the ice surface was only 15 to 40 meters above hydrostatic equilibrium in 2002 (Thomas et al., 2004). Basal drag should not be high in this region since it is proportional to the overburden pressure, which is expected to be small. The FS solution is therefore more consistent with the ice physics near the grounding-line.

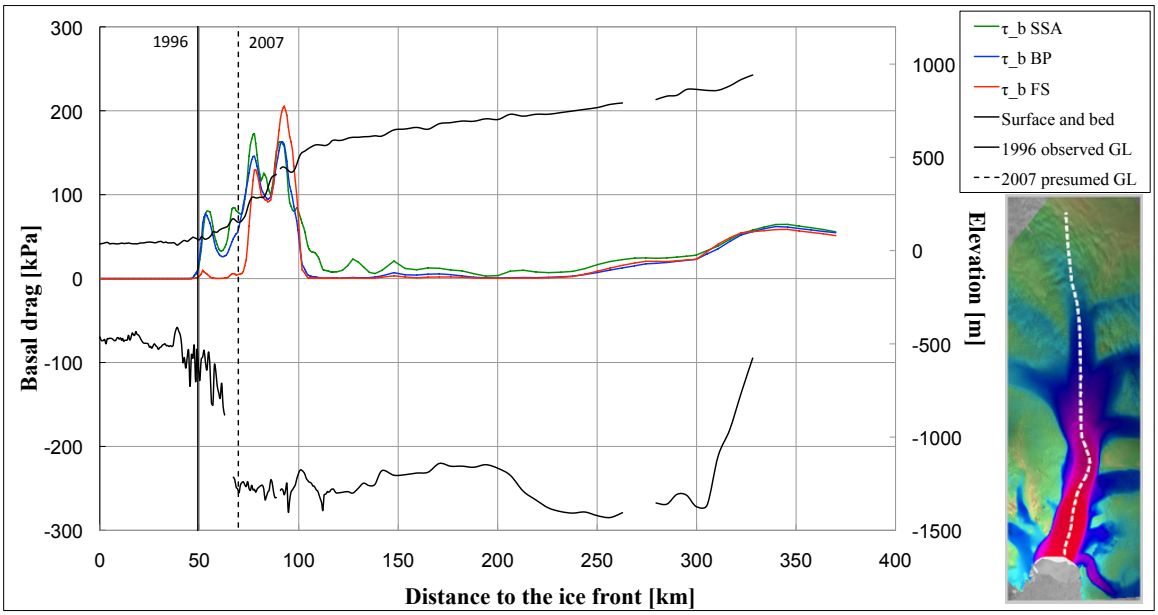


Figure 4.14: Comparison of inferred basal drag, τ_b, in kPa, SSA (green line), BP (blue line) and FS (red line) along a flow-line. Bed and surface elevations are the black lines (Thomas et al., 2004). Vertical lines are the 1996 observed (black solid) and 2007 presumed (black dashed) grounding-line positions (Rignot, 2008a).

To model dynamic glacier changes, e.g. grounding-line retreat, it would seem essential to employ a model that fully represents the ice flow dynamics. As we discussed above, SSA and BP are not adequate near the grounding-line because of simplified physics. If the FS solution is not used, the model would be initiated using values of basal drag that are too high and this would impact the simulation of grounding-line retreat.

Using a FS solution near the grounding-line is also mandated by theoretical studies to be essential to the analysis of grounding line stability (Nowicki and Wingham, 2008) and migration in dynamic models (Durand et al., 2009).

However, because FS is computationally intensive, it is prohibitive for large-scale modeling. We therefore recommend hybrid models that use a simple two-dimensional model on ice shelves, a 3D BP on grounded ice but FS near the grounding-line.

The conclusion of our experiment has a broader character than just the case of Pine Island Glacier. We would expect similar issues with other fast-moving glaciers that have a steeply rising bed near the grounding-line and high stresses of all orders in that region. This clearly suggests that near the grounding-line of ice streams, treating ice flow with the complete physics of FS is essential

Conclusion

In this chapter we detailed a very powerful method for identifying poorly-known parameters: ice hardness and basal friction, given velocity measured at the ice surface. We showed that the Optimize-Discretize approach was better suited for our problems. We then applied this technique to compare the patterns of basal friction for three ice flow models and saw that parameter identification could be used to analyze the validity of ice sheet models. We showed that FS was required near the grounding-line, but that simpler models gave good results over the rest of the model domain.

5

A mass conservation approach for mapping ice thickness

In the previous chapter, we developed a technique to optimize unknown parameters to match surface velocity measurements. We show here that even if the modeled velocities fit perfectly the observations, the inconsistencies between datasets can lead to unrealistic model behavior. The traditional method for interpolating ice thickness data from airborne radar sounding surveys onto regular grids is to employ geostatistical algorithms, e.g., kriging (Deutsch and Joumel, 1997). While this approach provides continuous maps of ice thickness, it generates products that are not consistent with ice flow dynamics and are impractical for high-resolution ice flow simulations as shown by Seroussi et al. (2011). Here, we present a novel approach that combines sparse ice thickness data collected by airborne radar sounding profilers with high-resolution swath mapping of ice velocity derived from satellite synthetic-aperture interferometry to obtain a high-resolution map of ice thickness that conserves mass and minimizes the departure from observations.

5.1 Motivations: Flux divergence anomalies

We show in this section that the velocity and thickness measurements are not consistent in terms of mass conservation and must be reconciled. We first derive a depth-averaged version of the mass balance equation, and then show that for fast moving glaciers at a high-resolution, this depth-averaged mass conservation equation is violated.

5.1.1 Depth-integrated mass conservation

Depth-averaged velocity

The depth-averaged velocity $\overline{\boldsymbol{v}} = (\overline{v}_x, \overline{v}_y)^T$ *is defined as:*

$$\begin{aligned} H\overline{v}_x &= \int_b^s v_x dz \\ H\overline{v}_y &= \int_b^s v_y dz \end{aligned} \tag{5.1}$$

2d Mass conservation

For a given point $\boldsymbol{x} \in \Omega$ *in the horizontal plane, the mass conservation (Eq. 1.16) imposes:*

$$\frac{\partial H}{\partial t} + \nabla \cdot H\overline{\boldsymbol{v}} = \dot{M}_s - \dot{M}_b \tag{5.2}$$

Proof
we integrate the equation of incompressibility Eq. (1.16) from the bed, b*, to the ice surface,* s*:*

$$\int_{b(x,y)}^{s(x,y)} \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right) dz = 0 \tag{5.3}$$

The last term is integrated:

$$v_z\left(x,y,s(x,y)\right) - v_z\left(x,y,b(x,y)\right) + \int_{b(x,y)}^{s(x,y)} \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right) dz = 0 \tag{5.4}$$

We then use the Leibniz integral rule (Eq.C.2) for the two terms under the integral which gives:

$$\begin{aligned} v_z\left(x,y,s(x,y)\right) - v_z\left(x,y,b(x,y)\right) + \\ \frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} v_x dz + v_y\left(x,y,b(x,y)\right)\frac{\partial b}{\partial y} - v_x\left(x,y,s(x,y)\right)\frac{\partial s}{\partial x} \\ + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} v_y dz + v_x\left(x,y,b(x,y)\right)\frac{\partial b}{\partial x} - v_y\left(x,y,s(x,y)\right)\frac{\partial s}{\partial y} = 0 \end{aligned} \tag{5.5}$$

So we have from the incompressibility:

$$\begin{aligned} v_z\left(s\right) - v_z\left(b\right) + \frac{\partial}{\partial x}\int_b^s v_x dz + \frac{\partial}{\partial y}\int_b^s v_y dz \\ - v_x(s)\frac{\partial s}{\partial x} - v_y(s)\frac{\partial s}{\partial y} + v_x(b)\frac{\partial b}{\partial x} + v_y(b)\frac{\partial b}{\partial y} = 0 \end{aligned} \tag{5.6}$$

and with the boundary conditions provided by the bed and surface kinematics (Eq.2.14 and Eq.2.9):

$$\frac{\partial}{\partial x}\int_{b(x,y)}^{s(x,y)} v_x dz + \frac{\partial}{\partial y}\int_{b(x,y)}^{s(x,y)} v_y dz + \frac{\partial s}{\partial t} - \dot{M}_s - \frac{\partial b}{\partial t} - \dot{M}_b = 0 \tag{5.7}$$

and $H = s - b$. ∎

5.1.2 Velocity and thickness measurements

Important ice sheet characteristics such as ice thickness, surface elevation or surface velocity are most efficiently derived from airborne and satellite platforms carrying instruments operating at different spatial resolutions and deployed at different epochs. As a consequence, datasets are not always consistent with one another, and this complicates their combined use in numerical ice sheet models. This is especially true for surface velocities and ice thicknesses.

Surface ice velocity are generally derived from satellite synthetic-aperture radar interferometry (InSAR). This technique, extensively described in App. A, consists in measuring the surface displacement from space between two passes. The resulting displacement provides a very high resolution product (30-300 m), with low error margins (a few m/yr). The surface velocities of the entire Greenland and Antarctic ice sheets are shown in Fig. 5.1.

Ice thicknesses on the other hand are difficult to measure. Airborne Ground Penetrating Radar (GPR) is generally employed. This technique provides high-resolution images of the dielectric properties of the top few hundreds of meters of the ice. When the ice is dry (as it is the case in Antarctica), the GPR can penetrate the entire ice column (several

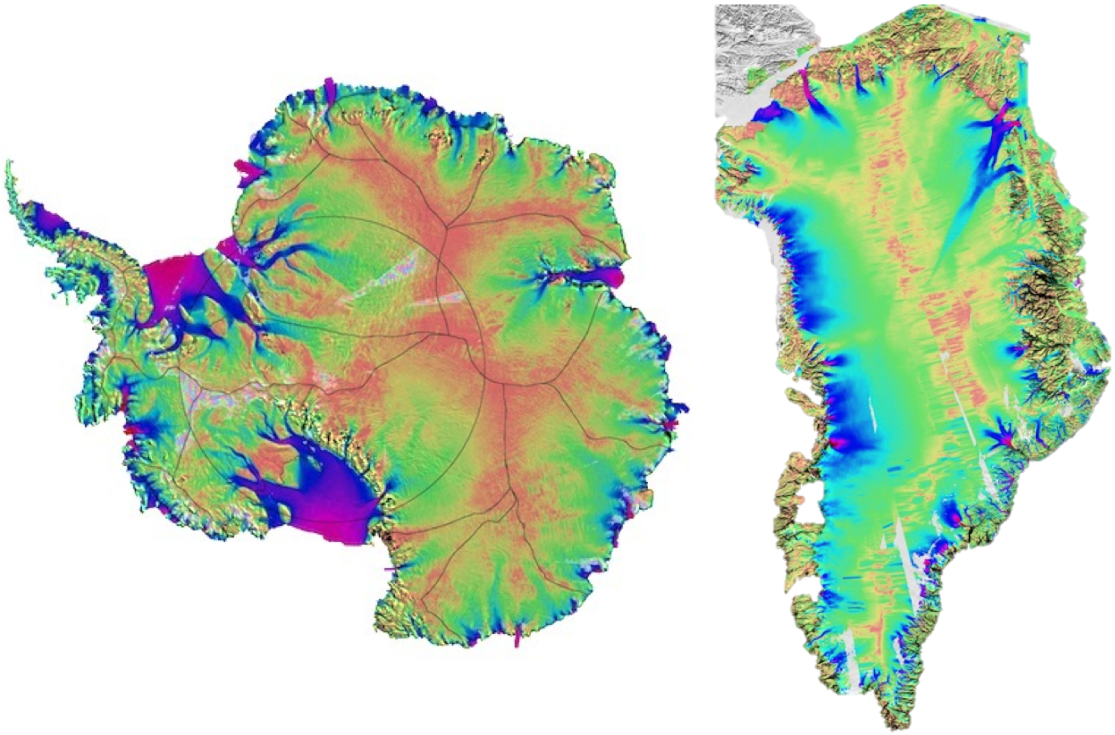


Figure 5.1: Antarctica and Greenland InSAR derived surface velocity, logarithmic scale (Rignot et al., in prep.)

thousands of meters), whereas in temperate glacier (outlet glaciers of Greenland), the pockets of water block the transmission of the electromagnetic field after a few tens of meters. When the signal reaches the bedrock, the transition between two media of different dielectric properties, ice and rock, can be detected and a measure of the travel time gives the bedrock elevation (See the CReSIS echogram Fig. 5.2). The downside of this technique is that the ice thickness is measured only on flight-tracks (Nadir-only viewing). Fig. 5.3 shows an example of NASA's Operation IceBridge (OIB) and the Center for Remote Sensing of Ice Sheet (CReSIS) Greenland and Antarctic ice thickness measurements campaigns. These linear measurements are then extrapolated to generate a complete map of ice thickness for the whole continent, using geostatistical algorithms, e.g. kriging (Deutsch and Joumel, 1997). The grid size selected for ice thickness mapping, typically 5 km in Greenland, is often much smaller than the actual average spacing between tracks (e.g. Bamber et al. (2001)), which means that the true spatial resolution of the data is much larger than the grid spacing.

5.1.3 Flux divergence of Nioghalvfjerdsfjorden

We examine first the case of Nioghalvfjerdsfjorden, also called 79North glacier, in north Greenland, because this glacier has been mapped extensively by N. Reeh et al. in the late 1990s and constitutes one of the most surveyed glaciers in Greenland (Seroussi et al., 2011).

The depth-integrated mass conservation equation (Eq. 5.2) states that the flux divergence is entirely balanced by the accumulation and melting rates and the thickness rate of change. For Nioghalvfjerdsfjorden, observations show that the sum of these terms should not exceed

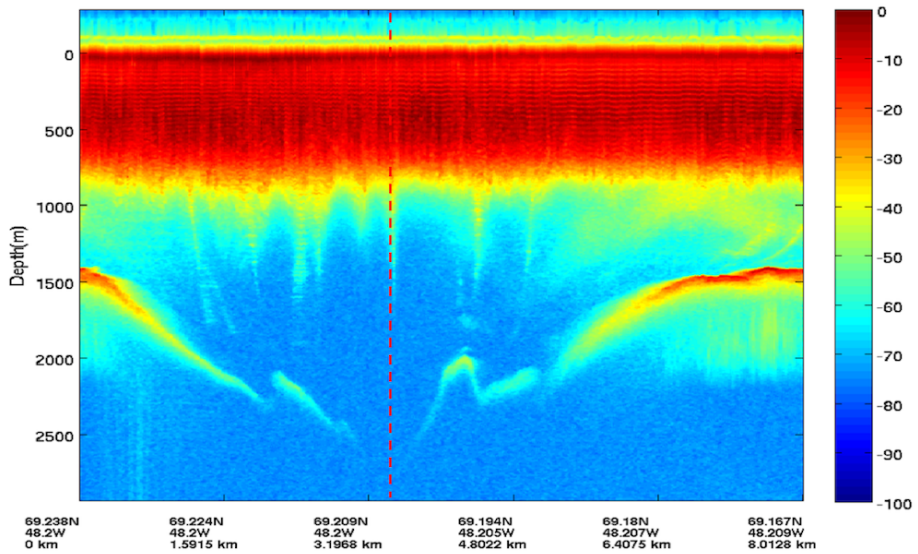


Figure 5.2: Echogram of a GPR along a flight track above Jakobshavn Glacier (Jilu Li, PhD Dissertation)

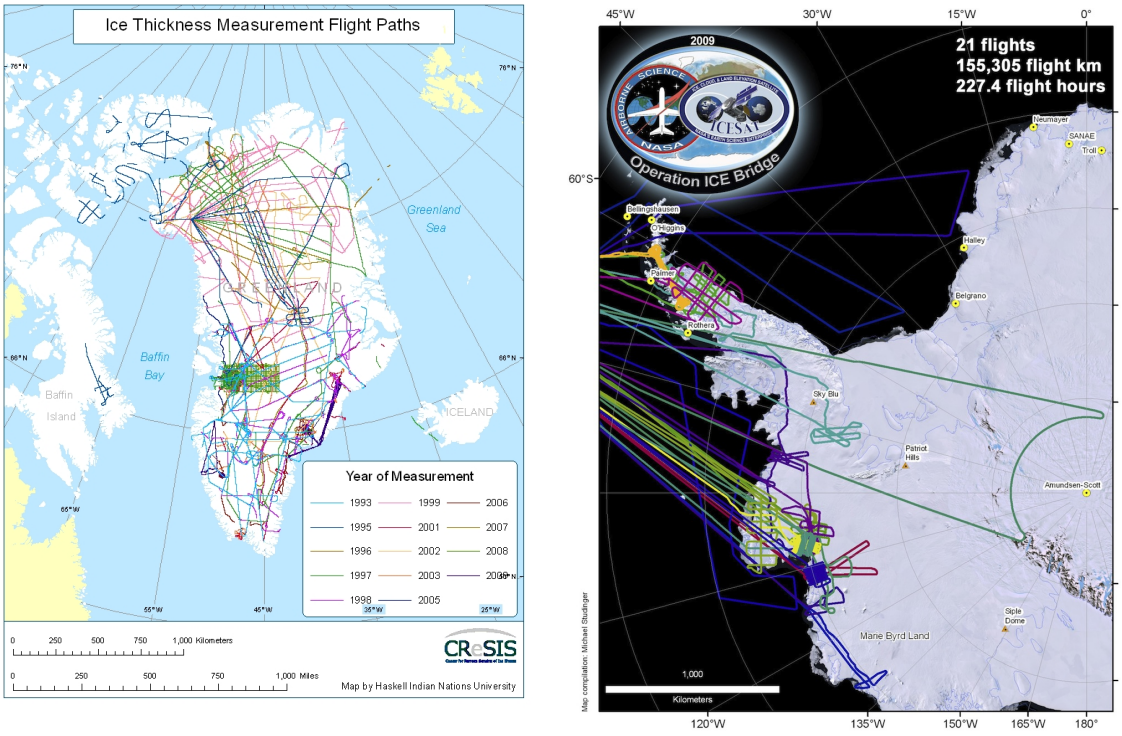


Figure 5.3: Flight lines of CReSIS (left) used to generate a thickness map of the Greenland ice sheet https://www.cresis.ku.edu/data/greenland, and flight lines of Operation IceBridge (NASA)

a few meters per year. When we combine the gridded thickness data from Reeh with InSAR velocities, we obtain a a flux divergence that exhibits large spatial variations over grounded ice, exceeding ±50 m/yr, that are physically untenable (Fig. 5.4).

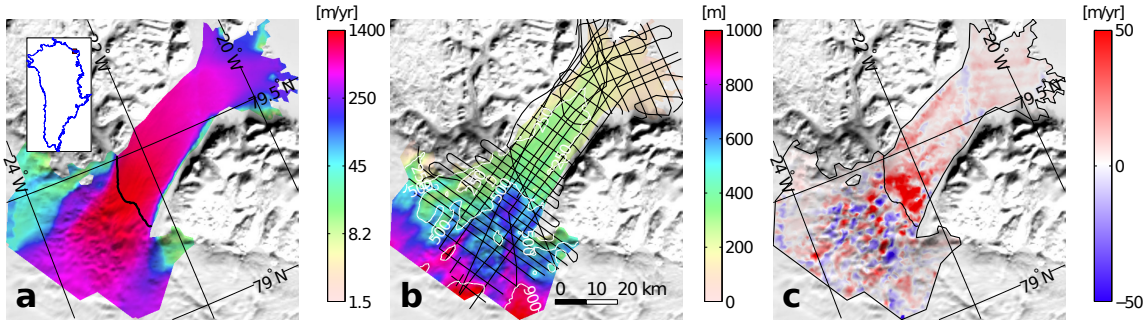


Figure 5.4: (a) Ice velocity (m/yr) of 79north Glacier, NE Greenland, measured from ERS-1/2 radar interferometry in 1996, (b) ice thickness from N. Reeh with flight tracks from 1998 indicated as black lines, (c) ice flux divergence combining ice velocity and ice thickness reveals large rates of thinning/thickening on grounded ice that are not physical.

5.1.4 Flux divergence of Jakobshavn Isbræ

We find these anomalies not only for 79North glacier but for all fast-moving glaciers that control the ice discharge from Greenland. Another example is Jakobshavn Isbræ in west Greenland (Fig. 5.5).

If we combine the the 1km SeaRISE datasets: the CReSIS thickness and InSAR derived surface velocities (Joughin et al., 2010), we get anomalous patterns of flux divergence (Fig. 5.6), that largely exceed ±200 m/yr (600 m/yr in some regions!), which makes transient ice sheet models impossible: after a one year run, the model would have deep hollows and bumps that significantly change the driving stress resulting in unphysical behavior. The square that arise from this combination are due to the bilinear interpolation from the 5 km grid to the mesh.

These variations are not caused by uncertainties in surface mass balance, errors in long-term thickness change or errors in ice velocity, but by the interpolation of ice thickness data from individual tracks to a regular grid without insuring that the results obey mass conservation. These anomalies are not minor disturbances in a model set up; they are very large and must be dealt with in order to make numerical models work. Indeed, the mass conservation equation is used by numerical models to compute the thickness rate of change. If the flux divergence is anomalous, the ice thickness is directly affected and the model immediately starts building new bumps and hollows in ice thickness from these anomalies, which yields a thickness map significantly different from the original one, regardless of the state of the mass balance of the observed area.

These anomalies were not a problem in the past for two reasons. Ice sheet models used to run at very low-resolution (20 to 50 km), and relied on long spin-ups from last interglacial maximum (Huybrechts et al., 2003; Pollard and DeConto, 2009) for initialization. These anomalies were therefore not critical, because the glacier is close to steady-state after these runs, and the flux divergence is close to zero. The computational costs and these long

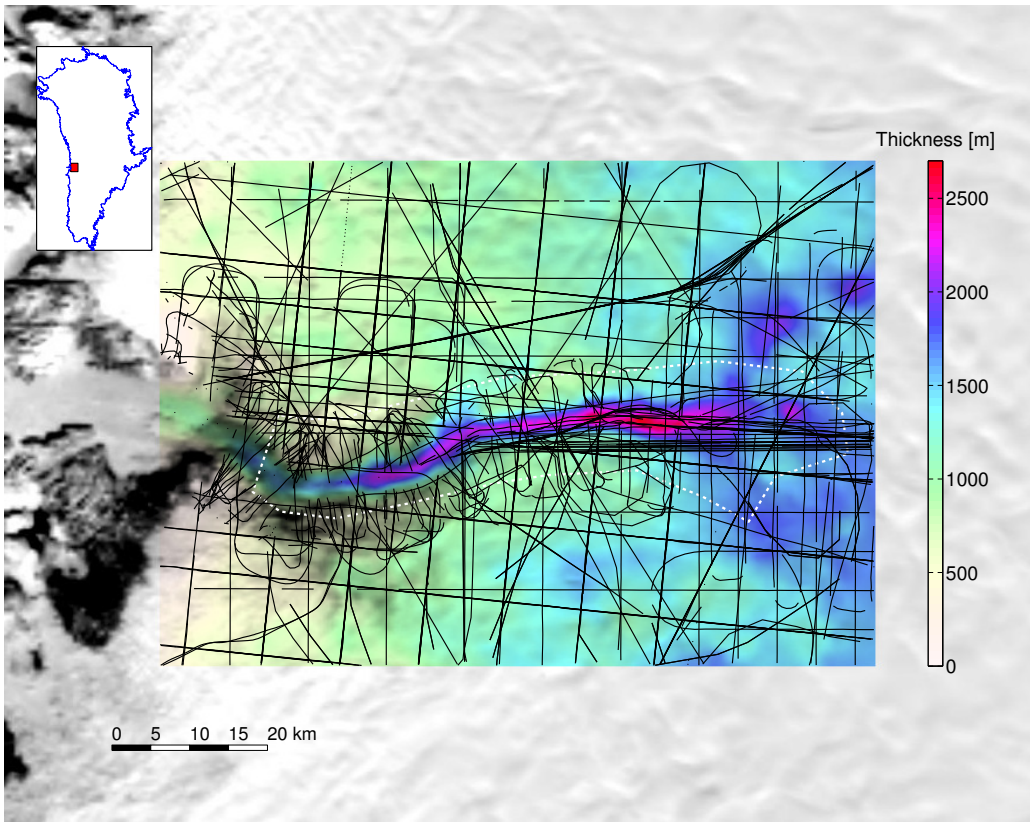


Figure 5.5: Location map of Jakobshavn Isbræ. The CReSIS flight tracks from 1993 to 2009 are shown in black and the white dotted line is the model domain

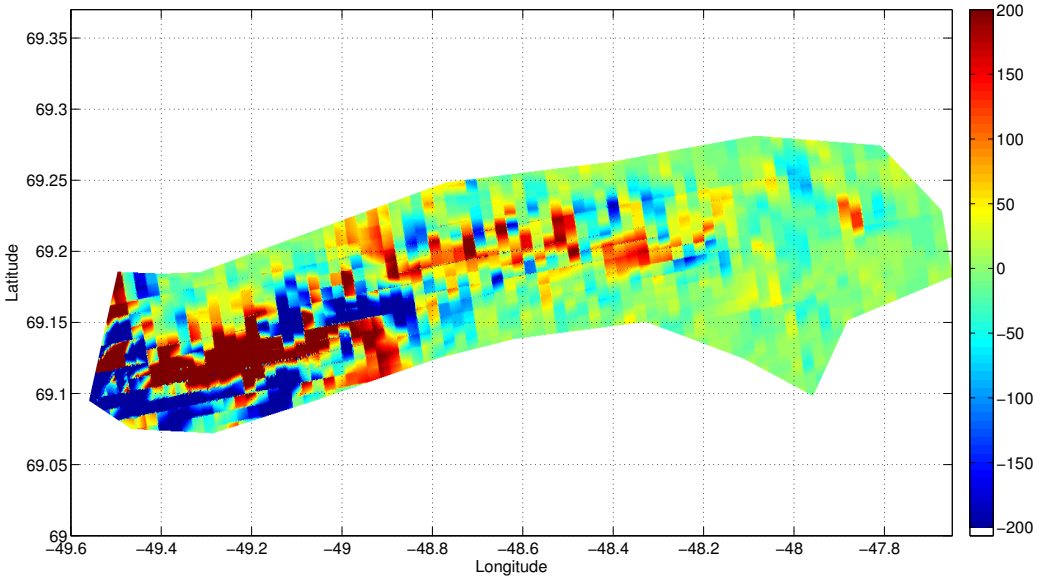


Figure 5.6: Jakobshavn Isbræ flux divergence using the SeaRISE datasets

simulations is prohibitive for ice sheet numerical models that include a higher-order or full-Stokes stress balances (Chap. 2). We therefore need to make these datasets consistents. These observations highlight the importance of consistency between datasets and the need for thicknesses that conserve mass for a given velocity field.

If we have measurements of elevation changes and a surface mass balance map, we know what the flux divergence should be. We can compute the ice thickness so that the mass conservation equation is verified as shown in the following section.

5.1.5 Balance thickness

To reconcile velocity and thickness measurements, we calculate the ice thickness that satisfies the mass conservation equation. The idea of using ice velocity to infer a spatially consistent map of ice thickness is not new. Rasmussen (1988) used this approach on Columbia Glacier using a finite difference scheme. Fastook et al. (1995) used a polynomial solution of the fourth degree derived from the Shallow Ice Approximation (SIA) to calculate the ice thickness of Jakobshavn Isbræ, in West Greenland. Warner and Budd (2000) employed the SIA to calculate the ice thickness over the Antarctic Ice Sheet using mass flux conservation (Budd and Warner, 1996). Farinotti et al. (2009) employed a method, derived from the same principle, to determine the ice volume of Swiss alpine glaciers. All these studies, however, suffer from significant deviations from the original thickness data, i.e., by hundreds of meters in the case Fastook et al. (1995) and Warner and Budd (2000), and an average 25% in Farinotti et al. (2009). We will see that the optimization sequence of the method proposed here reduces these deviations.

Let Ω be the two-dimensional ice domain and $\partial\Omega$ its boundary. We define the inflow and outflow boundaries $\Gamma_- \cup \Gamma_+ = \partial\Omega$ as follows:

$$\begin{aligned} \Gamma_- &= \{\mathbf{x} \in \partial\Omega \quad \bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \leq 0\} \\ \Gamma_+ &= \{\mathbf{x} \in \partial\Omega \quad \bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) > 0\} \end{aligned} \tag{5.8}$$

with $\mathbf{n}$ the outward-pointing unit normal vector. We also define $T \in \Omega$ as the flight tracks where data are collected within the model domain. The balance ice thickness is calculated by solving:

$$\begin{cases} \nabla \cdot H\bar{\mathbf{v}} &= \dot{a} & \text{in } \Omega \\ H &= H_{obs} & \text{on } \Gamma_- \end{cases} \tag{5.9}$$

where $\dot{a} = \dot{M}_s - \dot{M}_b - \partial H/\partial t$, is the apparent mass balance following (Farinotti et al., 2009) and H_{obs} is an observed thickness. Equation 5.9 is a special case of transport equation, with a purely convective transport. It constitutes a steady hyperbolic partial differential equation of first order. Such equations are difficult to solve numerically (e.g. Donea (1984)). Information is transported at finite speeds along the streamlines of the velocity field $\bar{\boldsymbol{v}}$. The nature of hyperbolic problems requires that boundary conditions be specified only on the inflow part Γ_-. It would be inappropriate and incorrect to prescribe any boundary conditions elsewhere.

Note that even though we call this solution "balance thickness", it shall not be confused with the steady-state glacier thickness because it incorporates the rate of thickness change, $\partial H/\partial t$. We discuss in App. B different numerical strategies to solve this equation as the

standard Continuous Galerkin approximation produces spurious oscillations, also known as *wiggles*. We implemented all these strategies and the Streamline Upwinding (SU) method proved to be the most robust method for our problem. The artifacts due to the addition of artificial diffusivity are minor and the solutions are very similar to the ones obtained with a Discontinuous Galerkin (DG) formulation. The latter is very sensitive to errors in velocities and in some cases does not converge because there is no artificial diffusion.

5.2 Balance thickness optimization

We present here a method that can be used to minimize the deviation between calculated and measured ice thicknesses. We formulate an optimization problem and allow input data to be changed according to their error margins in order for the calculated thickness to fit the measurements along track.

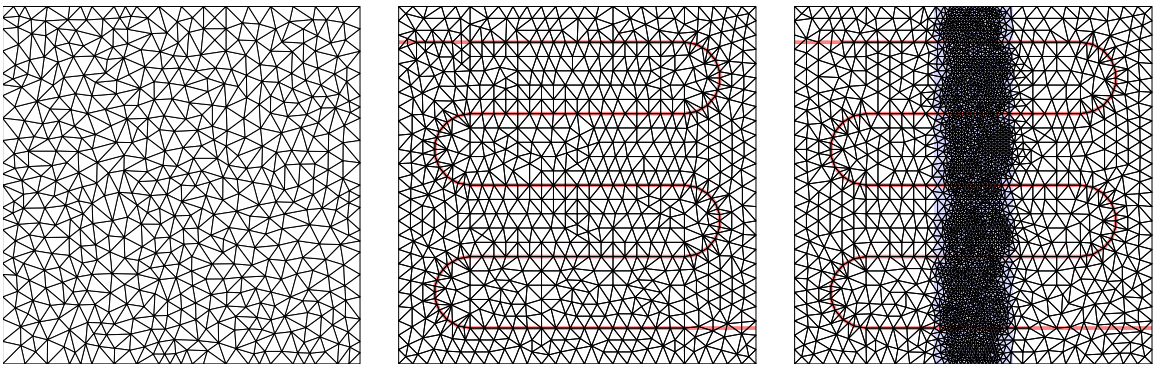
5.2.1 Inverse problem

The depth-averaged velocity is usually taken as the measured surface velocity, the accumulation is given by reliable atmospheric models, the thickness rate of change is usually small but can be derived from surface elevation change. The melting rate is usually small under grounded ice (a few cm per year), and is usually neglected.

Now, the modeled thickness usually deviates significantly from measurements on flight tracks. We can use these measurements on flight tracks to optimize all parameters with their error margins, so that the misfit between the modeled and the measured thickness is small. We formulate the following optimization problem:

$$\begin{aligned}
&\text{Objective function:} \quad J(H) = \int_T \frac{1}{2}(H - H_{obs})^2 \, d\Omega \\
&\text{PDE constraint:} \quad \begin{cases} \nabla \cdot H\bar{\boldsymbol{v}} = \dot{a} & \text{in } \Omega \\ H = H_{obs} & \text{on } \Gamma_- \end{cases} \\
&\text{Control parameters:} \quad \bar{\boldsymbol{v}}_x, \bar{\boldsymbol{v}}_y, \dot{a}
\end{aligned} \tag{5.10}$$

5.2.2 Mesh management



a) Unconstrained b) Flight track c) Flight track + Trough

Figure 5.7: Example 2d mesh management in ISSM. (a) Unconstrained unstructured regular mesh, (b) mesh constrained to follow a flight track (red), and (c) mesh constrained to follow a flight track (red) and local refinement on a central trough (blue area)

To solve the balance thickness equation, we use the finite element method. We use a 2 dimensional (2D) unstructured triangular mesh. To minimize the misfit between the calculated thickness and along-track thickness measurements, we constrain the mesh so

that there are vertices along flight tracks. To minimize interpolation errors between the measurements and the mesh, we constrain some vertices to coincide with flight tracks so we use the original data with no interpolation. Moreover, we need to refine the mesh in some regions, such as deep troughs, so that the algorithm correctly captures the ice thickness variations (Fig. 5). We implemented a mesh generator in ISSM, which uses an edge-based anisotropic mesh optimization methodology (See Fig. 5.7).

5.2.3 Adjoint and gradients

The optimization (Eq. 5.10) is PDE-constrained and we must first introduce the Lagrangian:

Lagrangian

The Lagrangian associated to Eq. (5.10) is:

$$\mathcal{L}\left(H,\lambda,\overline{\boldsymbol{v}},\dot{a}\right)=J\left(H\right)+\int_{\Omega}\lambda\left(\nabla\cdot H\overline{\boldsymbol{v}}-\dot{a}\right)d\Omega \tag{5.11}$$

where $\lambda:\Omega\rightarrow\mathbb{R}$ is the Lagrangian multiplier associated to the PDE constraint.

It is convenient to integrate by part the first term under the integral:

$$\mathcal{L}=J\left(H\right)-\int_{\Omega}H\overline{\boldsymbol{v}}\cdot\nabla\lambda d\Omega-\int_{\Omega}\lambda\dot{a}\;d\Omega+\int_{\Gamma_{-}}H_{obs}\lambda\overline{\boldsymbol{v}}\cdot\boldsymbol{n}dS+\int_{\Gamma_{+}}H\lambda\overline{\boldsymbol{v}}\cdot\boldsymbol{n}dS \tag{5.12}$$

Adjoint and gradient

If H is solution of the forward model:

$$\left\{\begin{array}{rcll}\nabla\cdot H\overline{\boldsymbol{v}} & = & \dot{a} & \text{in }\Omega\\ H & = & H_{obs} & \text{on }\Gamma_{-}\end{array}\right. \tag{5.13}$$

and if λ is solution of the adjoint problem, defined by the following variational formulation:

$$\int_{T}\varphi\left(H-H_{obs}\right)dx-\int_{\Omega}\varphi\boldsymbol{v}\cdot\nabla\lambda d\Omega+\int_{\Gamma_{+}}\varphi\lambda\boldsymbol{v}\cdot\boldsymbol{n}dS=0 \tag{5.14}$$

Then the gradients of the objective function with respect to the $\overline{\boldsymbol{v}}_x$, $\overline{\boldsymbol{v}}_y$ and $\dot{a}$ for a direction φ_{vx}, φ_{vy} and φ are:

$$\left\langle\mathcal{J}'\left(\overline{\boldsymbol{v}}_x\right),\varphi_{vx}\right\rangle=\left\langle\frac{\partial\mathcal{L}}{\partial\overline{\boldsymbol{v}}_x}\left(H,\lambda,\overline{\boldsymbol{v}},\dot{a}\right),\varphi_{vx}\right\rangle \tag{5.15}$$

$$\begin{aligned} \langle \mathcal{J}'(\bar{\boldsymbol{v}}_y), \varphi_{vy} \rangle &= \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_y}(H, \lambda, \bar{\boldsymbol{v}}, \dot{a}), \varphi_{vy} \rangle \\ \langle \mathcal{J}'(\dot{a}), \varphi \rangle &= \langle \frac{\partial \mathcal{L}}{\partial \dot{a}}(H, \lambda, \bar{\boldsymbol{v}}, \dot{a}), \varphi \rangle \end{aligned} \tag{5.16}$$

Proof
If H is solution of the forward problem, then:

$$\mathcal{J}(\bar{\boldsymbol{v}}, \dot{a}) = \mathcal{L}(H, \lambda, \bar{\boldsymbol{v}}, \dot{a}) \tag{5.17}$$

And if we take the derivative of this equation with respect to $\bar{\boldsymbol{v}}_x$ and apply the chain rule:

$$\langle \mathcal{J}'(\bar{\boldsymbol{v}}_x), \varphi_{vx} \rangle = \langle \frac{\partial \mathcal{L}}{\partial H}(H), \langle \frac{\partial H}{\partial \bar{\boldsymbol{v}}_x}, \varphi \rangle \rangle + \langle \frac{\partial \mathcal{L}}{\partial \lambda}(H), \langle \frac{\partial \lambda}{\partial \bar{\boldsymbol{v}}_x}, \varphi \rangle \rangle + \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_x}(H), \varphi \rangle \tag{5.18}$$

The first form of the Lagrangian (Eq. 5.11) gives:

$$\forall \mu \in V_\lambda \quad \langle \frac{\partial \mathcal{L}}{\partial \lambda}, \mu \rangle = \int_\Omega \mu (\nabla \cdot H\bar{\boldsymbol{v}} - \dot{a}) \, d\Omega \tag{5.19}$$

And because H is solution of the forward model, this derivative is equal to zero. The second form of the Lagrangian (Eq. 5.12) gives:

$$\forall \varphi \in V_H \quad \langle \frac{\partial \mathcal{L}}{\partial H}, \varphi \rangle = \int_T \varphi (H - H_{obs}) \, dx - \int_\Omega \varphi \bar{\boldsymbol{v}} \cdot \nabla \lambda d\Omega + \int_{\Gamma_+} \varphi \lambda \bar{\boldsymbol{v}} \cdot \boldsymbol{n} dS \tag{5.20}$$

and if λ is solution of the variational formulation that defines the adjoint (Eq. 5.14), this derivative is also zero. We showed that:

$$\langle \mathcal{J}'(\bar{\boldsymbol{v}}_x), \varphi_{vx} \rangle = \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_x}(H, \lambda), \varphi \rangle \tag{5.21}$$

We can apply a similar analysis for $\bar{\boldsymbol{v}}_y$ and $\dot{a}$. ∎

Expression of the Gradients

We have the following expressions for the gradients:

$$\begin{aligned} \mathcal{J}'(\bar{\boldsymbol{v}}_x) &\simeq -H \frac{\partial \lambda}{\partial x} \\ \mathcal{J}'(\bar{\boldsymbol{v}}_y) &\simeq -H \frac{\partial \lambda}{\partial y} \\ \mathcal{J}'(\dot{a}) &= -\lambda \end{aligned} \tag{5.22}$$

Proof
The second form of the Lagrangian (Eq. 5.12) provides the first two gradients:

$$\langle \mathcal{J}'\left(\bar{\boldsymbol{v}}_x\right), \varphi_{vx} \rangle = \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_x}\left(H, \lambda\right), \varphi_{vx} \rangle = -\int_\Omega H \varphi_{vx} \frac{\partial \lambda}{\partial x} d\Omega + \int_{\Gamma_+} H \lambda \varphi_{vx} n_x dS \tag{5.23}$$

$$\langle \mathcal{J}'\left(\bar{\boldsymbol{v}}_y\right), \varphi_{vy} \rangle = \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_y}\left(H, \lambda\right), \varphi_{vy} \rangle = -\int_\Omega H \varphi_{vy} \frac{\partial \lambda}{\partial y} d\Omega + \int_{\Gamma_+} H \lambda \varphi_{vy} n_y dS \tag{5.24}$$

and we approximate the gradients by neglecting the boundary term. The gradient with respect to the apparent mass balance is recovered likewise:

$$\langle \mathcal{J}'\left(\dot{a}\right), \varphi \rangle = \langle \frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{v}}_y}\left(H, \lambda\right), \varphi \rangle = -\int_\Omega \lambda \varphi \, d\Omega \tag{5.25}$$

■

5.2.4 Error Analysis

We want to evaluate the error δH made on the balance thickness given errors in both velocity $\delta \bar{\boldsymbol{v}}$ and apparent mass balance $\delta \dot{a}$. The calculated thickness $H + \delta H$ is solution of the mass balance equation by definition:

$$\nabla \cdot \left(H + \delta H\right)\left(\bar{\boldsymbol{v}} + \delta \bar{\boldsymbol{v}}\right) = \dot{a} + \delta \dot{a} \tag{5.26}$$

If we neglect second order terms and assume that the thickness real H is solution of the mass balance equation, the thickness error is solution of:

$$\nabla \cdot \delta H \; \bar{\boldsymbol{v}} = \delta \dot{a} - \nabla \cdot H \delta \bar{\boldsymbol{v}} \tag{5.27}$$

The error is known on the flight tracks T and δH is solution of:

$$\begin{cases} \nabla \cdot \left(\delta H \; \bar{\boldsymbol{v}}\right) &= \delta \dot{a} + \nabla \cdot H \delta \bar{\boldsymbol{v}} & \text{in } \Omega \\ \delta H &= H - H_{obs} & \text{on } \Gamma_- \cup T \end{cases} \tag{5.28}$$

This expression does not take into account the fact that the flux is well constrained upstream of each flight track. The simplest way to take that into account is to solve two problems:

$$\begin{cases} \nabla \cdot \delta H_1\left(+\bar{\boldsymbol{v}}\right) &= \delta \dot{a} + \nabla \cdot H \delta \bar{\boldsymbol{v}} & \text{in } \Omega \\ \delta H &= H - H_{obs} & \text{on } \Gamma_- \cup T \end{cases} \tag{5.29}$$

$$\begin{cases} \nabla \cdot \delta H_2\left(-\bar{\boldsymbol{v}}\right) &= \delta \dot{a} + \nabla \cdot H \delta \bar{\boldsymbol{v}} & \text{in } \Omega \\ \delta H &= H - H_{obs} & \text{on } \Gamma_- \cup T \end{cases} \tag{5.30}$$

So that the error propagates once downstream (δH_1) and once upstream (δH_2). The actual error would be the minimum of the two:

$$\delta H = \min\left(\delta H_1, \delta H_2\right) \tag{5.31}$$

To evaluate the right hand side, we take $\delta \dot{a} = 0.5$ m/yr, the error in velocities is 2 m/yr and the error in strain rate is estimated as 3 10^{-4} yr^{-1} based on observation errors [*Rignot et al, unpublished*].

5.3 Application to Nioghalvfjerdsfjorden

To illustrate, we focus on Nioghalvfjerdsfjorden and evaluate our algorithm. This experiment part of a paper recently published (Morlighem et al., 2011).

5.3.1 Data and Method

In this section, we compare three methodologies to infer the ice thickness of 79North Glacier. The goal is to illustrate that the proposed optimization scheme achieves the best results. In experiment 1, we calculate the balance thickness by directly solving Eq.5.9, with ice thickness only constrained on Γ_-; the flux divergence is imposed as equal to the apparent mass balance. In experiment 2, we impose ice thickness to be strictly equal to the measurements along the original tracks: the ice thickness is therefore constrained on both inflow boundary and flight tracks, T. Finally in experiment 3, we apply an optimization sequence that includes uncertainties in velocity and apparent mass balance to best fit the observations without strongly constraining the thickness on T.

The study area is the fast flowing portion of 79North Glacier, a large discharger of ice in north Greenland (Fig.5.8). This glacier has been extensively surveyed in the late 1990s, is not accelerating or changing in ice thickness at a significant level, hence providing a reliable, reference glaciological setting to test our algorithm. We use ice thickness from (Thomsen et al., 1997; Christensen et al., 2000) with a track spacing of 5 km on the upper part of the glacier, and 2.5 km near the grounding zone (Fig.5.8). A gridded thickness map at 1-km posting was generated by N. Reeh (unpublished, Fig.5.9a) using block kriging, here referred to as Reeh's thickness map.

Surface velocity is measured using ascending and descending tracks of the Earth Remote Sensing Satellite (ERS) 1 and 2 in year 1996 (Rignot et al., 1997). The ice stream velocity exceeds 1200 m/yr at the grounding-line. The model domain is set by the geographic limits of Reeh's thickness map. This ensures that the input ice flux, which is an important control on the solution, is well constrained. Along parts of the inflow boundary where no measurement is available, we use Reeh's thickness map values to constrain the solution by default. We exclude floating parts of the glacier from the model domain because the basal melt rates on floating ice are not known well enough. Surface mass balance is from Ettema et al. (2009) and ranges from -120 cm/yr to -50 cm/yr averaged over the time period 1961-1996. Temporal changes in ice thickness are less than 1 m/yr (Thomas et al., 2006) and are thus neglected. Basal melting on grounded ice is less than a few 10 cm per year and is neglected as well (Fahnestock et al., 2001). The apparent mass balance is therefore taken here as equal to the surface mass balance.

5.3.2 Results

When Reeh's thickness map (Fig.5.9a) is combined with the observed surface velocities, we find a flux divergence on grounded ice that exceeds $\pm$100 m/yr (Fig.5.9e), which is not physically acceptable. As mentioned earlier, if an ice flow model is initiated in this manner, deep hollows and high bumps will rapidly appear, yielding a new glacier configuration significantly different from the original one.

In experiment 1, we use an unstructured triangular mesh with a resolution of 400 m. The balance thickness shown in Fig.5.9b compares well with Reeh's thickness map, but deviates

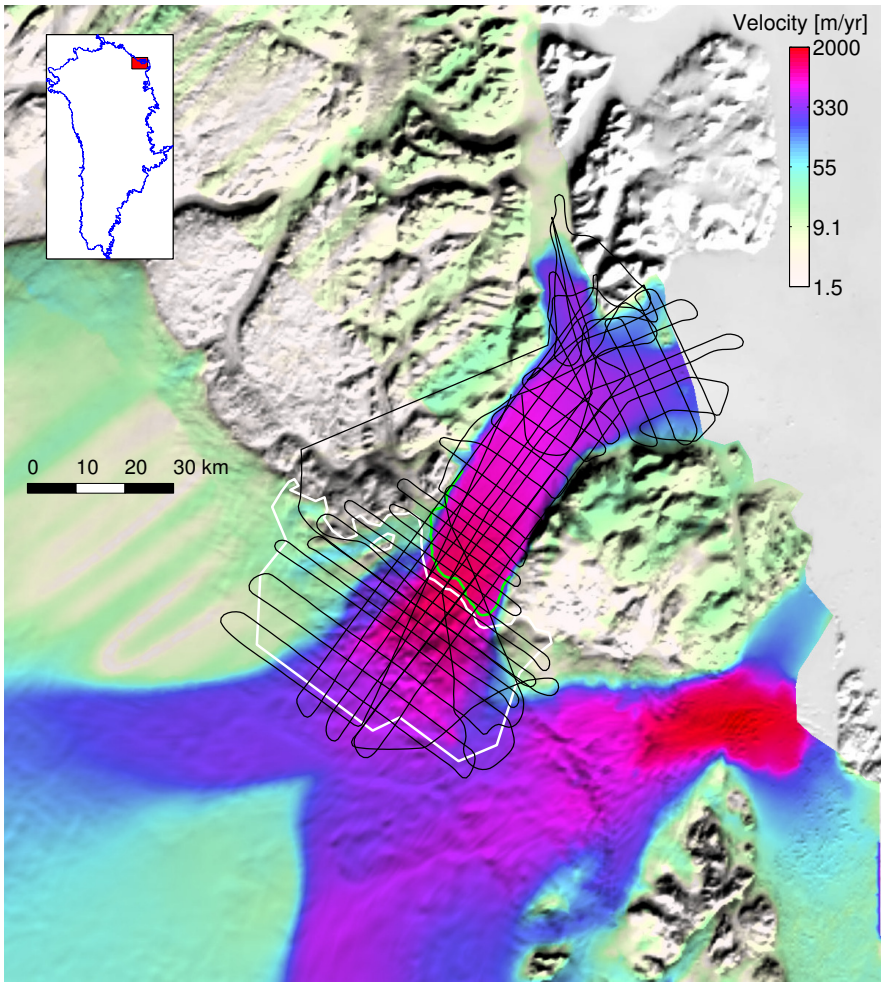


Figure 5.8: InSAR-derived surface velocity of 79North Glacier, in Greenland, color coded on a logarithmic scale and overlaid on a MODIS (Moderate-resolution Imaging Spectroradiometer) mosaic of Greenland. The flight tracks used in Reeh's thickness map are shown as black lines; the model domain is a white line; and the grounding-line is the green line.

significantly from the measurements, by several hundreds of meters in some areas, which is not satisfactory.

In experiment 2, we use the same spatial resolution but constrain the mesh to follow the flight tracks so that measurements can be imposed at their locations. Fig.5.9c shows that the results fit the original data better but deviate rapidly from the measurements in between tracks and exhibit large jumps when crossing tracks, hence resulting in a noisy appearance of both ice thickness and flux divergence (Fig.5.9g). Indeed, the balance thickness hyperbolic equation requires the thickness to be constrained only once; it is incorrect to prescribe the thickness elsewhere. We therefore need to relax the constraints along the flight tracks by formulating an optimization problem.

In experiment 3, we solve for Eq.5.9 without additional constraint (same as experiment 1) but optimize the ice velocity and apparent mass balance within tolerance levels for the calculated thickness to best fit the thickness measurements along tracks. The resulting

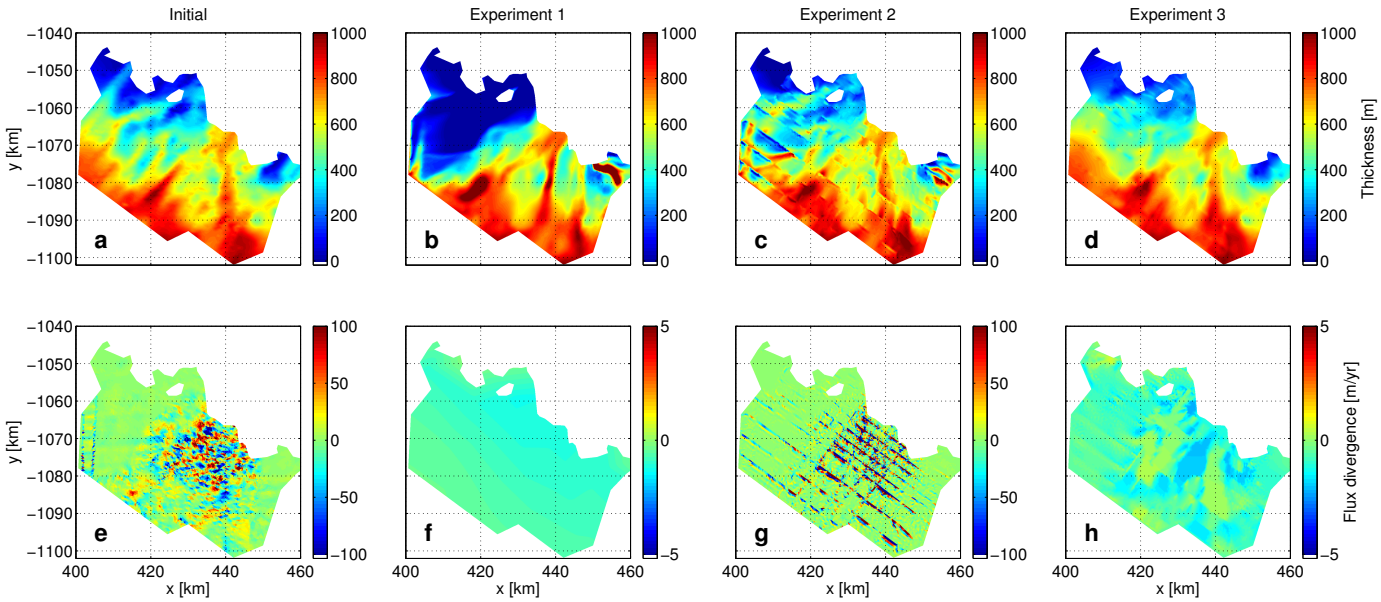


Figure 5.9: Thickness (m) of 79North Glacier, Greenland, from (a) N. Reeh, (b) experiment 1, (c) experiment 2, (d) experiment 3 (optimized); and flux divergence (m/yr) combining InSAR velocities and the thickness map from (e) N. Reeh, (f) experiment 1 (apparent mass balance), (g) experiment 2, (h) experiment 3 (optimized). Color bars associated with the flux divergence in (e) and (g) have broader ranges to maintain visibility.

thickness map (Fig.5.9d) is close to Reeh's thickness, fits the original data well, and yields anomalies in flux divergence that are two orders of magnitude lower than the flux divergence associated to Reeh's thickness map (Fig.5.9h).

Fig.5.10a compares Reeh's thickness data and the optimized balance thickness data in experiment 3 with observations. The balance thickness method results in deviations from observations that are a factor of three less than Reeh's thickness map, i.e the error decreases from 80 to 30 m.

Fig.5.10c shows error estimates for the balance thickness. The error is low along flight tracks and increases along the flow direction. Maximum errors are reached in between flight tracks, and the error increases when the track spacing increases. Errors are less than 30-50 m over the model domain. Along the sides of the domain, where ice thickness is weakly constrained by observations and ice velocity is low, errors exceed 120 m.

Comparing our results with independent measurements of ice thickness from the Center for Remote Sensing of Ice Sheet (CReSIS) and Operation IceBridge (OIB) of years 1997, 1999, 2004 and 2010 in Fig.5.10b, we find an average difference of 43 m with the balance thickness and 73 m with Reeh's thickness map. This comparison therefore illustrates that our balance thickness is reliable and more accurate than the data obtained from kriging.

5.3.3 Discussion

Our mass conservation approach provides seamless ice thickness data that conserves mass by combining ice velocity and ice thickness data in an optimum fashion, i.e., with no artifacts in ice flow divergence. Kriging provides interpolated data that do not deviate largely from

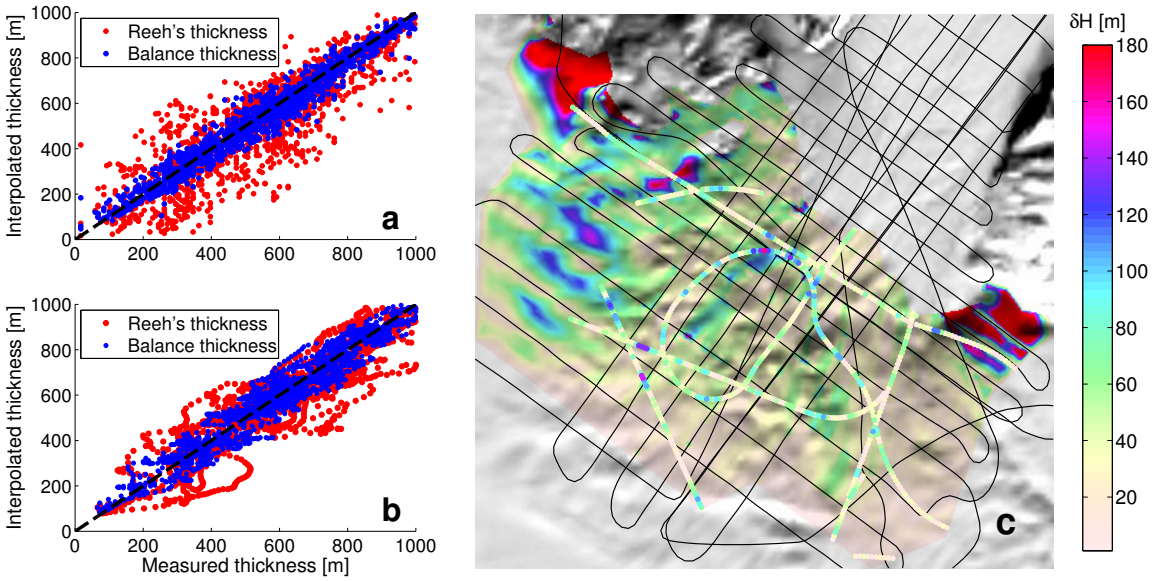


Figure 5.10: (a) Reeh's thickness (red dots) and balance thickness (blue dots) vs measurements used to generate both maps. (b) Reeh's thickness (red dots) and balance thickness (blue dots) vs CReSIS/OIB independent measurements. (c) Estimated maximum error in balance thickness (m) overlaid by the error between balance thickness and CReSIS/OIB measurements along flight lines. Flight tracks used in Reeh's thickness map are shown as black lines.

the measurements but that do not conserve mass between flight tracks. Our error analysis shows that the maximum error in ice thickness is inversely proportional to ice velocity. Our method therefore yields optimal results in fast flowing regions, which is also region highly relevant to ice sheet flow models because fast flowing regions have a major impact on ice sheet mass balance. The approach is less reliable in slow moving areas, but this is also a part of the ice sheet where anomalies in ice flux divergence are quite small (Fig.5.9e). The algorithm is therefore optimum where it matters most.

As demonstrated in this study, the solution to Eq.5.9 is very sensitive to small errors in the input data. The hyperbolic nature of this equation leads to unstable solutions, especially in slow moving areas. In order to use this method most effectively it is therefore essential to have a good understanding of the error budget of the variables in Eq.5.9.

Our error analysis provides useful guidelines for ice thickness mapping in areas where ice velocity data are available. Errors in balance thickness are proportional to the distance to the closest track, so dense tracks are required, as expected. The preferable configuration of tracks is one that crosses as many flow lines as possible in order to constrain the flux of the largest possible region of the domain. In contrast, a flight track that follows a flow line cannot constrain the flux of other flow lines. We therefore recommend use of glacier surveys with preferential tracks perpendicular to the flow direction; tracks along the flow direction will remain useful for flow line modeling and for estimating cross-track errors, but not as useful for dense interpolation of ice thickness data using a mass conservation approach.

Our method is applicable to other glaciers provided that surface velocity is known, as well as the apparent mass balance. On 79North Glacier, the apparent mass balance is

small and almost negligible. On a fast changing glacier like Jakobshavn Isbræ, the rates of thickness change are large (15-30 m/yr), the apparent mass balance is not negligible and measurements of $\partial H/\partial t$ will be needed to reduce uncertainties in balance thickness.

On floating ice shelves, our method is difficult to apply because the high rates of basal melting from the ocean are not well known, yet the anomalies in ice flux divergence are also much smaller (Rignot and Steffen, 2008). But more importantly, ice-shelf thickness may be more effectively deduced from digital maps of surface elevation assuming that ice is in hydrostatic equilibrium, so the mass conservation approach is less critical.

5.4 Application to Jakobshavn Isbræ

We apply the same experiments to the fast moving region of Jakobshavn Isbræ, West Greenland, to illustrate that this method can be applied to other glaciers. Jakobshavn Isbræ drains 6.5% of the Greenland ice sheet (Joughin et al., 2004a) and flows through a deeply eroded bedrock trough, which extends about 80 km inland of the calving front. This glacier has been extensively surveyed by the Center for Remote Sensing of Ice Sheets (CReSIS) and Operation IceBridge (OIB), who produced a high-resolution thickness map.

Here, we use the 1 km SeaRISE datasets. The CReSIS observed thickness, is combined with InSAR derived surface velocities (Joughin et al., 2010). We use a surface mass balance from the Regional Atmospheric Climate Model Ettema et al. (2009) and $\partial H/\partial t$ from the Airborne Topographic Mapper (ATM) Thomas et al. (2009). The basal melting is also assumed to be negligible (Fahnestock et al., 2001) on grounded ice.

First, we calculate ice thickness by solving the balance thickness equation, i.e., imposing ice thickness only at the inflow boundary. Fig. 5.11c shows that the algorithm creates a trough, which has approximately the same shape as that observed by radar. A kriging algorithm could not generate this trough. This illustrates that the use of surface velocities improves our knowledge of bed topography, though the calculated trough is much deeper than the one observed by CReSIS (5000 m vs. 2800 m). Hence, the modeled thickness deviates again too significantly from measurements. In the second experiment, we constrain the calculated thickness to be equal to the measurement along flight tracks. As with 79North, the resulting ice thickness (Fig. 5.11e) is noisy and exhibits jumps when crossing a flight track. The same jumps are visible in the map of ice flux divergence (Fig. 5.11f).

To apply the optimization, we first need to define the tolerance interval for the depth-averaged velocity, $\bar{\mathbf{v}}$, and the forcing, $\dot{a}$. The surface velocities in this fast moving portion of the glacier are assumed to be very close to the depth-averaged velocities, as basal shear stress is generally small under ice streams. Errors in surface velocity measurements and interpolation errors are on the order of 50 m/yr. The admissible space for $\bar{\mathbf{v}}$ is therefore defined as:

$$\forall \mathbf{x} \in \Omega \quad \bar{\mathbf{v}}\left(\mathbf{x}\right) \in \left[\mathbf{v}_{obs}\left(\mathbf{x}\right) - 50 \quad \mathbf{v}_{obs}\left(\mathbf{x}\right) + 50\right] \tag{5.32}$$

For the forcing, $\dot{a}$, the error is equal to the sum of the errors in surface mass balance, basal melting and $\partial H/\partial t$. Although the exact error is not known, we estimate that it is on the order of 3 m/yr. The admissible space for $\dot{a}$ is:

$$\forall \mathbf{x} \in \Omega \quad \dot{a}\left(\mathbf{x}\right) \in \left[\dot{a}_{obs}\left(\mathbf{x}\right) - 3 \quad \dot{a}_{obs}\left(\mathbf{x}\right) + 3\right] \tag{5.33}$$

The resulting thickness is shown in Fig. 5.11g. It is very close to the observed thickness in Fig. 5.11a, except that we find a modeled thickness that is slightly deeper in the lower reaches of the glacier than the CReSIS dataset. This could be due to a model error or to a lack of data in this region (Fig. 5.5), but might also indicate that the trough is deeper than expected in this region.

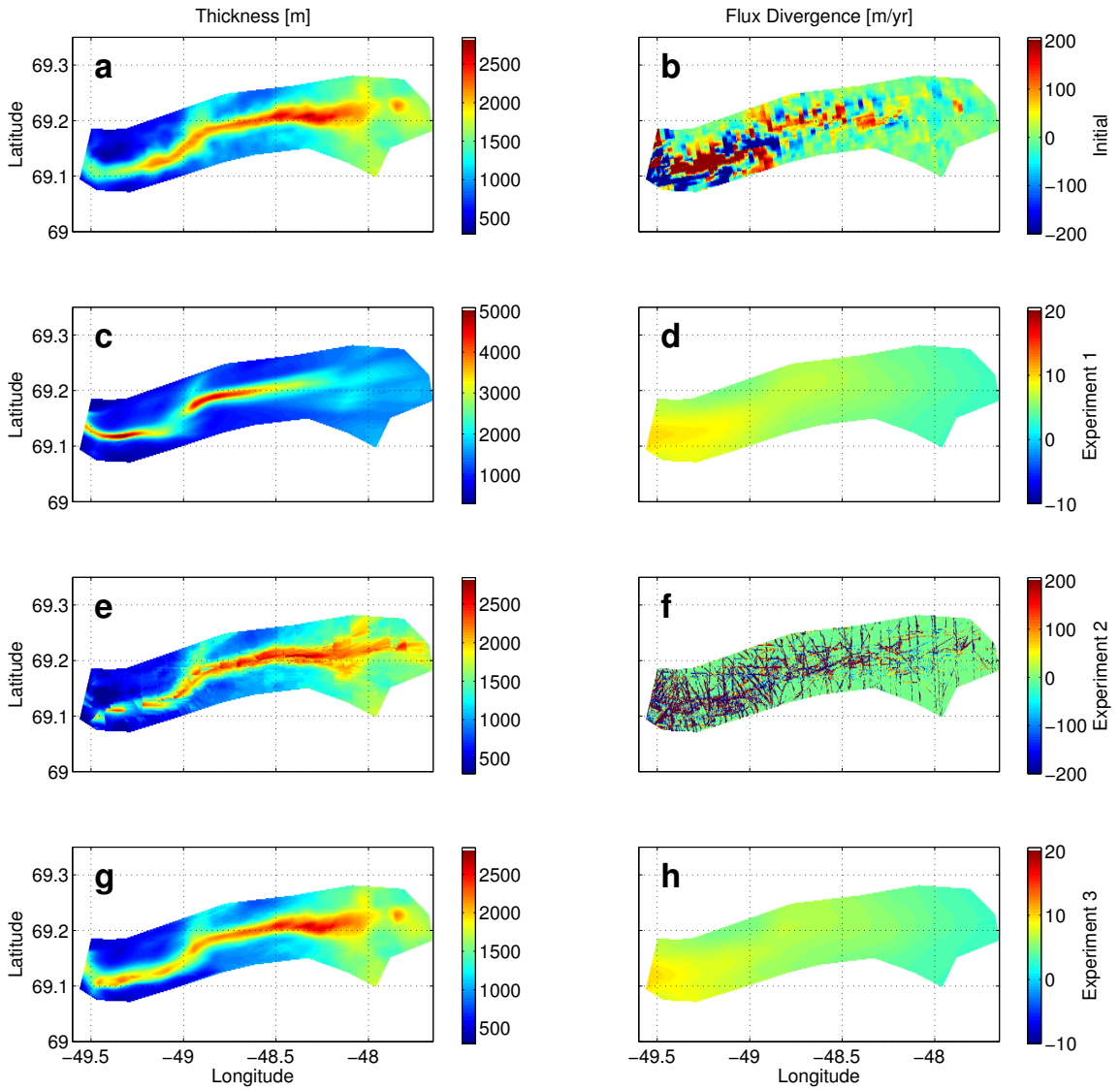


Figure 5.11: Thicknesses (m) and flux divergences (m/yr) of Jakobshavn Isbræ. (a) observed CReSIS thickness, (b) flux divergence from InSAR velocities (Joughin et al., 2010) and CReSIS thickness, (c) balanced ice thickness, (d) observed flux divergence ($\dot{M}_s - \dot{M}_b - \partial H/\partial t$), (e) constrained balanced ice thickness, (f) flux divergence from InSAR velocities and constrained balance thickness, (g) optimized balance thickness, (h) optimized flux divergence.

Conclusion

We presented in this chapter an alternate solution to the traditional mapping of ice thickness with kriging, which has shown serious limitations for ice sheet modeling applications. Our technique is applicable to any glaciated terrain for which information on apparent mass balance data is available, specifically for those areas where there are dense measurements of ice velocity. Our method is most effective in fast flowing regions, where ice sheet models need it most. It provides datasets that can readily be used in ice sheet models, fit the observations well and generate no anomalies in ice flux divergence, thereby increasing confidence in the model results. In one example application, our approach reduces errors in ice thickness by a factor 2, to less than 40 m, or 5% of the total thickness.

Conclusion

Main results

In this thesis, we developed tools to infer ice sheet properties using a combination of remote sensing and ice sheet modeling via inverse methods. We first introduced the general equations governing ice sheet dynamics and thermo-dynamics and presented three different models of varying complexity: MacAyeal/Morland's shelfy-stream model, Blatter/Pattyn's higher order model and the full-Stokes model. Inverse methods to infer basal friction and ice shelf hardness have been developed for all three models. We investigated two approches of optimization: Discretize-Optimize and Optimize-Discretize and showed that the latter was preferable for unstructured meshes. We also discussed different cost functions and found that a logarithmic misfit gave very good results. All these improvements are believed to be an original contribution developed in this thesis.

We applied these inversions to Pine Island Glacier, West Antarctica, using the same mesh and boundary conditions for the three models. We saw that the inferred patterns of basal drag were very similar, which suggest that full-Stokes is not required everywhere to model ice sheet or ice shelf flow. Nevertheless, in the grounding-line region, bridging effects cannot be neglected and only the full-Stokes model is able to capture what is believed to be the exact basal friction. Because both MacAyeal/Morland and Blatter/Pattyn neglect bridging effects, the basal friction was artificially high in order to match the observed velocity. These high values are not real, but rather reflect the fact that these two models are not valid in this region because of the steep bed and high velocities.

In the last chapter, we introduced a novel approach that combines sparse ice thickness data collected by airborne radar sounding profilers with high-resolution swath mapping of ice velocities derived from satellite synthetic-aperture interferometry, in order to obtain a high-resolution map of ice thickness that conserves mass and minimizes the departure from observations. We applied this approach to the case of Nioghalvfjerdsfjorden (79North) Glacier, a major outlet glacier in northeast Greenland that has been relatively stable in recent decades. The results show that our mass conserving method removes the anomalies in mass flux divergence, yields interpolated data that are within about 5% of the original data, and produces thickness maps that are directly usable in high-spatial resolution, high-order ice flow models. We discussed the application of this method to the broad and detailed radar surveys of ice sheet and glacier thickness.

This work was done in parallel with the development of the massively parallelized Ice Sheet System Model (ISSM). Developing the theoretical tools and then implementing them was very complementary.

Perspectives

Many processes need to be better understood and better represented in climate models. There are many capabilities that need to be implemented in ISSM. The first one is a calving-law. Even for short term projections, the ice front can move dramatically following ice shelf collapse or enhanced melting at the front. This requires the ice front to move in the model. This aspect presents several challenges because the processes involved in iceberg calving are still debated, and calving laws are difficult to derive. Implementing moving boundaries in a parallel architecture is also a technical challenge because the ice front might move from one processor to another and communication between CPUs needs to be limited.

Another area of possible improvement is basal hydrology. Basal friction inferred by control methods needs to be related to water content at the base of the ice in order to better constrain how the basal drag evolves with time. A hydrological model is already implemented in ISSM, but the relationship between water content and inferred basal drag remains difficult to derive.

Additionally, data assimilation needs to be extended to transient models. We now have observations at different times and we need to take advantage of these datasets to better constrain the models. This requires the implemention of a temporal adjoint. We are in the process of implementing an automatic differentiation tool, but a "manual" adjoint could also be developed.

Finally, grounding-line processes are one of the key aspects of ice sheet dynamics that we need to better model. We currently use a simple hydrostatic criterion, but it is well known that the ice is not in hydrostatic equilibrium in the first few kilometers downstream of the grounding-line. Several schemes based on contact mechanics have been developed and should be implemented in ISSM.

A

Synthetic aperture radar interferometry

One often says that no recent development in the field of polar remote sensing has engendered more excitement and interest than Synthetic Aperture Radar (SAR) *interferometry*, or *InSAR*. This technique uses Earth-orbiting spacecraft to measure two ice sheet parameters: surface elevation and motion. Considering that none of the currently deployed imaging radar satellite platforms was designed with interferometric applications in mind, the high quality and vast quantity of exciting results from this technique are astounding. Interferometry on glaciers started to develop in the 1990s. The goal of this section is to give the reader an overview of InSAR, since all the data we used during this internship came from InSAR. A number of reviews on theory and applications of InSAR have been written, this section is inspired by Burgmann et al. (2000), Zebker et al. (1994), Rosen et al. (2000) and Goldstein et al. (1993). This report do not deal with Radar techniques, interested readers are urged to follow up on Elachi (1988).

A.1 Overview of this techniques

A.1.1 SAR, How it works

The conventional radar (RAdio Detection And Ranging) imaging is a technique in which the target is illuminated by an electromagnetic beam of microwaves. The signal is reflected by the target and recorded by the radar. The processing of this reflected signal gives information about the target. For instance, the round-trip travel time from multiple targets on the ground are used to determine the distance to the targets. The amplitude of the signal generates a rough two-dimensional image of the target area. Bright regions in a radar image represent high amplitude of the returned wave energy, which depends on the surface slope and roughness and dielectric characteristics of the surface material. The resolution of such a real-aperture radar in space would be of about 5 to 10 km, limited by the power and size of the footprint of the radar beam. Whereas Synthetic aperture radar (SAR) combines signal processing techniques with satellite orbit information to produce a much-higher-resolution (tens of meters) radar image.

By focusing the raw radar echoes, SAR processing significantly improves the resolution of point targets in both the cross-track (range) and along-track (azimuth) (See Fig.A.1). Fine resolution in the cross-track direction is achieved by using a radar signal of high bandwidth, which improves the differentiation of radar echoes from closely spaced targets in the range direction. Focusing in the along-track direction relies on the synthetic aperture provided by a moving antenna. The radar return from a particular point target on the ground is present in successively recorded radar echoes (note that the radar footprints of successive pulses overlap). As the moving antenna passes by the image point, the Doppler frequency shift of the return signal from the point and the round trip time of the signal can be used together to differentiate image points in the azimuth direction. Because the Doppler frequency shift depends on the location of the target fore or aft of the radar beam's center, the frequency of the return from the target sweeps out a band as the target passes through the radar's footprint. With this information, a signal of high bandwidth is constructed in the azimuth direction using successive pulses of the radar. The echoes are combined to synthesize a larger antenna aperture and thus achieve much improved resolution. A typical image point (pixel) spacing in space-based SAR images is 20-100 m within a 100-km-wide swath.

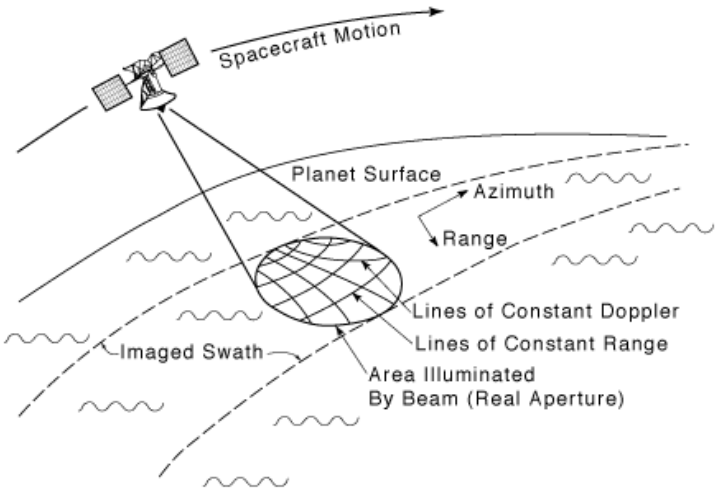


Figure A.1: Method of data acquisition of a synthetic aperture radar (SAR) system (from Engineering & Science, Spring 1991)

A.1.2 InSAR, How it works

Two SAR images can be combined to produce a radar interferogram, which can reveal information about the third dimension (elevation) of the target area or measure minute changes in the range distance between two image acquisitions. After SAR focusing, the radar image is a two-dimensional record of both the amplitudes and the phases of the backscattered signals from targets within the imaging area. The amplitude is a measure of target reflectivity, whereas the phase encodes changes at the surface.

InSAR uses the phase information in two SAR images to determine the phase difference between each pair of corresponding image points, thus producing an interferogram. Thus, if two SAR images of a target area are available and can be precisely aligned to a fraction of a pixel width, the relative phase of each image point can be determined. Relative phase data from two images taken from slightly different viewing angles provide information about changes in range to targets on the ground, and thus the surface topography at the resolution of the SAR image can be recovered with knowledge of the imaging geometry. Interestingly, the first applications of InSAR were in Earth-based studies of topography of the Moon and Venus (Zisk, 1972; Rumsey et al., 1974).

The phase difference of two images taken from the same viewpoint, but at different times, can precisely measure any shifts of the returned phase. Thus, if the glacier's surface moved toward or away from the radar between the two imaging passes, phase changes would result that can be measured with a precision corresponding to millimeter-level displacements. The first demonstration of this technique used Seasat satellite data to detect vertical motions caused by soil swelling of irrigated fields in the Imperial Valley, California (Gabriel et al 1989). In the following section we give a more detailed explanation of how InSAR can be used to measure Ice sheet's velocity and Tidal deformation.

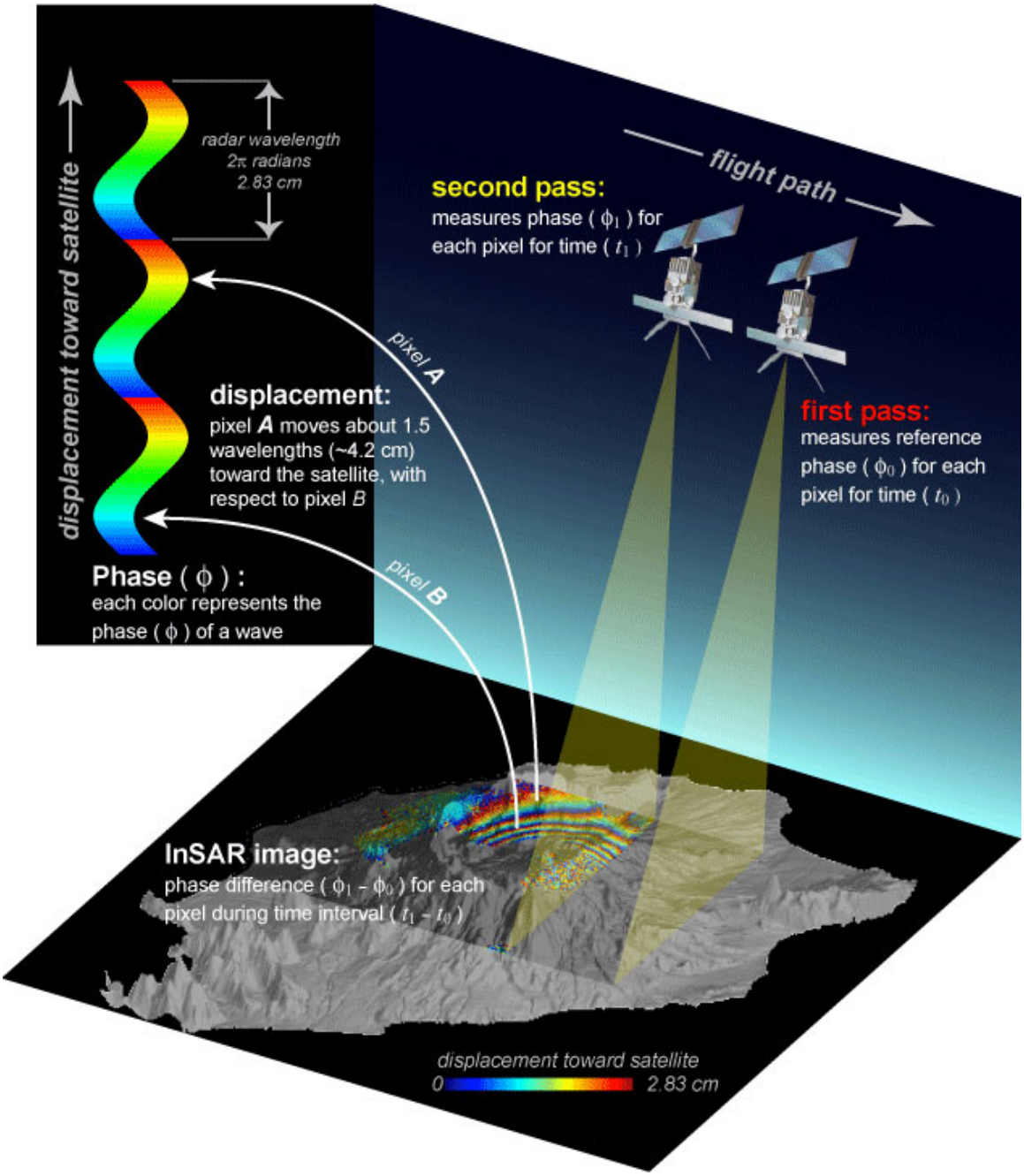


Figure A.2: Synthetic aperture radar interferometry process
(U.S. Geological Survey)

A.2 Processing of InSAR data

We have seen that interferogram images were built from two SAR images of the same ground surface and received by two antennas spatially offset so that the phase of the backscattered signal is not the same for each antenna. The two antennas can be on the same spacecraft but Jet Propulsion Laboratory often uses the same satellite that passes several times above the same area, or even two different satellites.

The interferogram is constructed by multiplying the signal of the first image with the complex conjugate of the other. The phase of the interferogram is a fringe pattern that represents the difference in propagation phase of the two SAR images. This phase difference can only be determined modulo 2π. The interferogram's amplitude (brightness) is the product of the image backscatter amplitudes. Interferograms are often displayed overlaying the brightness rendered in gray-scale, and phase as color, where each cycle of color, or fringe, represents a phase change of 2π.

The precision of InSAR depends on many factors: the atmosphere, the thermal noise, the ionosphere,... but usually, the detection limit is about 1.5 millimeters for vertical motions and about 4 millimeters for horizontal motions in the radar beam direction. The grounding-line (The point at which a tidewater glacier, or ice stream feeding an ice shelf, floats free of its bed), detected by tidal motions where the ice goes afloat, can be mapped at a resolution of -0.5 kilometer (Goldstein et al., 1993).

A.2.1 Motionless surface

In this first step, we assume that the surface is motionless, or that the two images are taken in the same time.

The phase difference measured in the case of two antennas A1 and A2 (see Fig.A.3) receiving in ping-pong mode (each antenna sends and receives its own pulse) is:

$$\Delta\Phi = \frac{4\pi}{\lambda}\left(\rho_1 - \rho_2\right) = \frac{4\pi}{\lambda}\Delta\rho \tag{A.1}$$

Using Al-Kashi's theorem, we have:

$$\rho_2^2 = \rho_1^2 + B^2 - 2\rho_1 B\cos\left(\frac{\pi}{2} - \theta + \alpha\right) = \rho_1^2 + B^2 - 2\rho_1 B\sin\left(\theta - \alpha\right) \tag{A.2}$$

Then

$$\begin{aligned} \rho_1 - \rho_2 = &\quad \frac{1}{\rho_1 + \rho_2}\left(2\rho_1 B\sin\left(\theta - \alpha\right) - B^2\right) \\ \delta\rho = &\quad \frac{1}{2\rho_1 - \delta\rho}\left(2\rho_1 B\sin\left(\theta - \alpha\right) - B^2\right) \end{aligned} \tag{A.3}$$

to the first order in $\delta\rho$:

$$\delta\rho \approx B\sin\left(\theta - \alpha\right) - \frac{B^2}{2\rho_1} \tag{A.4}$$

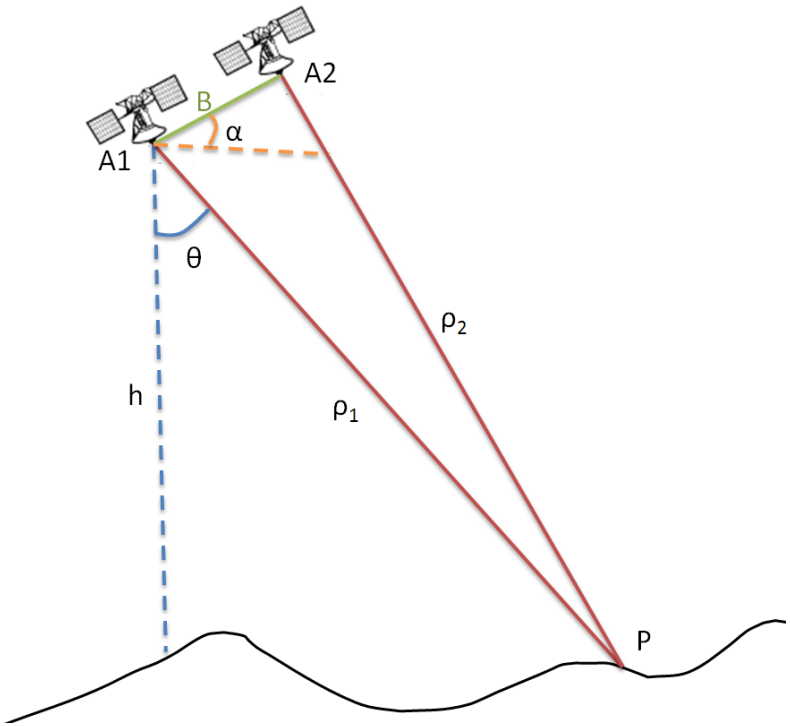


Figure A.3: SAR geometry

and finally:

$$\Delta\Phi = \frac{4\pi}{\lambda}\left[B\sin(\theta-\alpha) - \frac{B^2}{2\rho_1}\right] \tag{A.5}$$

Now we must take into account the fact that the Earth is curved, and there is a local topography. We want to decompose the previous equation in order to have an expression of this term. Let's assume that θ_0 is the look angle to each point in the image assuming zero local height. With $\delta\theta = \theta - \theta_0$:

$$\begin{aligned}\Delta\Phi &= \frac{4\pi}{\lambda}\left[B\sin(\delta\theta+\theta_0-\alpha) - \frac{B^2}{2\rho_1}\right] \\ &= \frac{4\pi}{\lambda}\left[B\sin(\delta\theta)\cos(\theta_0-\alpha) + B\cos(\delta\theta)\sin(\theta_0-\alpha) - \frac{B^2}{2\rho_1}\right]\end{aligned} \tag{A.6}$$

with $B_{\|} = Bsin(\theta_0-\alpha)$ and $B_{\perp} = Bcos(\theta_0-\alpha)$ which are in turn the parallel component of the baseline B (parallel to the look direction) and the perpendicular component (with respect to the across track direction),

$$\Delta\Phi = \frac{4\pi}{\lambda}\left[B_{\perp}\sin(\delta\theta) + B_{\|}\cos(\delta\theta) - \frac{B^2}{2\rho_1}\right] \tag{A.7}$$

A.2.2 Mobile surface

For a glacier or an Ice sheet, the target may have moved between the times t_1 and t_2 of the signal acquisition. We call V_x, V_y and V_z the three components of the velocity vector of the glacier, where x is the across track direction on the ground, y is the track direction and z the vertical direction. The target is on P1 at t_1 and P2 at t_2. The target can move because of the ice flow, but can also move vertically because of the tides if it is an ice shelf. The phase difference is a little bit trickier, we need to use the distance ρ_{int} from Fig.A.4:

$$\Phi_{12} = \Phi_1 - \Phi_2 = \frac{4\pi}{\lambda}(\rho_1 - \rho_2) = \frac{4\pi}{\lambda}(\rho_1 - \rho_{int}) + \frac{4\pi}{\lambda}(\rho_{int} - \rho_2) \tag{A.8}$$

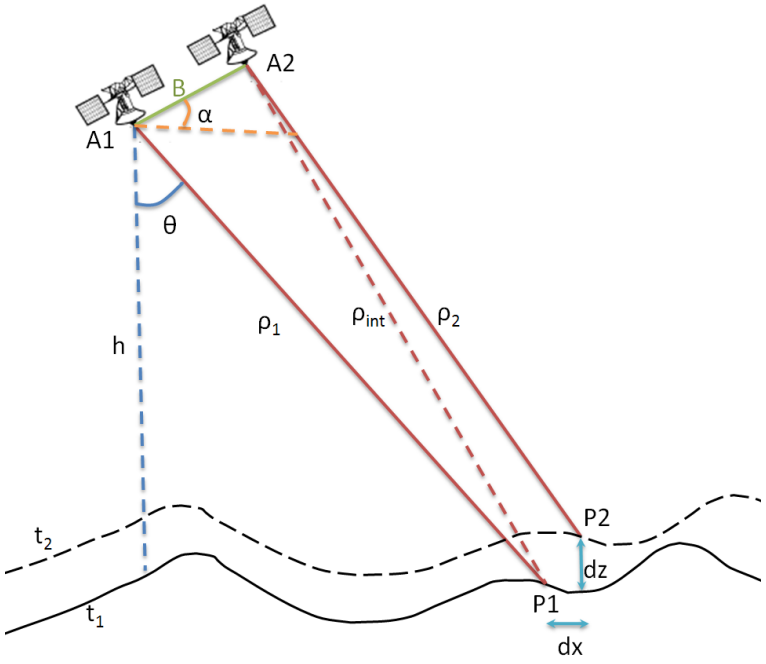


Figure A.4: InSAR geometry

For the first term, we can use the previous section, the geometry is exactly the same.

For the second term, we have the following equations:

$$\begin{aligned}
\cos(i) &= \frac{h - z_{P2} + B_{12}\sin(\alpha_{12})}{\rho_2} \\
\sin(i) &= \frac{x_{P2} - B_{12}\cos(\alpha_{12})}{\rho_2} \\
\cos(i-\delta i) &= \frac{h - z_{P1} + B_{12}\sin(\alpha_{12})}{\rho_{int}} \\
\sin(i-\delta i) &= \frac{h - z_{P1} - B_{12}\cos(\alpha_{12})}{\rho_{int}}
\end{aligned} \tag{A.9}$$

And we also have the following equations:

$$\begin{aligned}
dx &= x_{P2} - x_{P1} &&= V_x(t_2 - t_1) \\
dz &= z_{P2} - z_{P1} &&= V_z(t_2 - t_1) + (Z_2 - Z_1)_{Tide}
\end{aligned} \tag{A.10}$$

Therefore:

$$\begin{aligned}
\rho_{int}\cos(i-\delta i) - \rho_2\cos(i) &= V_z(t_2 - t_1) + (Z_2 - Z_1)_{Tide} \\
\rho_{int}\sin(i-\delta i) - \rho_2\sin(i) &= -V_x(t_2 - t_1)
\end{aligned} \tag{A.11}$$

By multiplying the first of the previous equations by $\cos(i)$ and the second one by $\sin(i)$ and then sum them, we have:

$$\begin{aligned}
\rho_{int}\cos(i-\delta i)\cos(i) + \rho_{int}\sin(i-\delta i)\sin(i) - \rho_2 &= \rho_{int}\cos\left((i-\delta i) - i\right) - \rho_2 \\
&= \rho_{int}\left(1 + \mathcal{O}\left(\delta i^2\right)\right) - \rho_2
\end{aligned} \tag{A.12}$$

So to the second order in δi:

$$\rho_{int} - \rho_2 = \left(-V_x\sin(i) + V_z\cos(i)\right)(t_2 - t_1) + (Z_2 - Z_1)_{Tide}\cos(i) \tag{A.13}$$

Finally, the phase difference can be written as follows:

$$\begin{aligned}
\Phi_{12} &= \frac{4\pi}{\lambda}\left[B_\perp\sin(\delta\theta) + B_\parallel\cos(\delta\theta) - \frac{B^2}{2\rho_1}\right] \\
&+ \frac{4\pi}{\lambda}\left(-V_x\sin(i) + V_z\cos(i)\right)(t_2 - t_1) \\
&+ \frac{4\pi}{\lambda}(Z_2 - Z_1)_{Tide}\cos(i) + \Phi_{12}^{noise}
\end{aligned} \tag{A.14}$$

In the previous equation, the first term depends only on the Earth's curvature and the topography of the surface, the baseline and the wavelength of the radar. The second term is due to the motion of the glacier, this term highlights that the interferometer measures only the projection of the displacement vector in the radar line-of-sight direction. To construct the velocity vector, observations must be made from different aspect angles. The third term depends on the change in elevation due to the tide. Fig.A.5 shows an amplitude image and the phase of Pine Island Glacier, West Antarctica.

By a very sophisticated post processing involving filters and spatial averages, it is possible to reduce the noise so that it is negligible. We also can get rid of the first term by knowing precisely the topography of the target and the exact position of the satellite. The remaining signal, which is a combination of tidal deformation and glacial motion, is called the *flattened* interferogram. In the following section, we will show how we can separate these two signals.

A.2.3 Tidal deformation

The technique used to separating the tidal deformation signal from the ice motion signal is called double difference interferometry. The double difference interferograms are constructed using three different SAR images of the same target. This allows us to generate two interferograms using Φ_{12} and Φ_{23}. To the first order, we have:

$$\begin{aligned}\Phi_{23} &= \frac{4\pi}{\lambda}\left[B_{23\perp}\sin\left(\delta\theta^{(23)}\right) + B_{23\parallel}\cos\left(\delta\theta^{(23)}\right) - \frac{B^2_{(23)}}{2\rho_1}\right] \\ &+ \frac{4\pi}{\lambda}\left(-V_x\sin(i) + V_z\cos(i)\right)(t_3 - t_2) \\ &+ \frac{4\pi}{\lambda}\left(Z_3 - Z_2\right)_{Tide}\cos(i) + \Phi^{noise}_{23}\end{aligned} \tag{A.15}$$

The two interferograms are processed to lower the noise, and the signal is flattened so that the first term is eliminated. If the ground motion is stable (which is generally correct for a glacier), the velocity vector (V_x, V_y, V_z) remains the same. If the time intervals are also similar $t_2 - t_1 = t_3 - t_2$, we can eliminate the term due to the motion by subtracting the phase differences of the two interferograms:

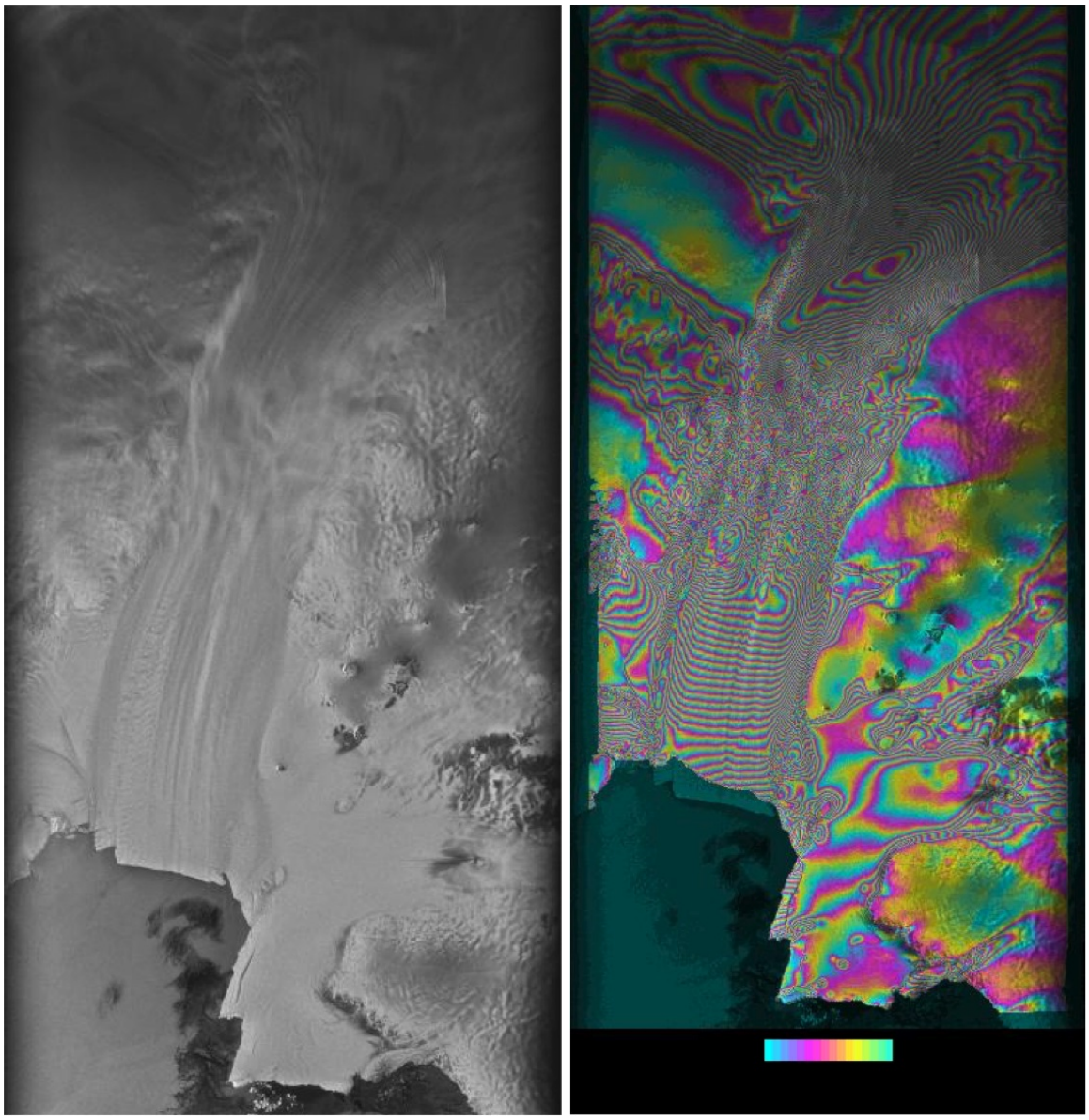
$$\Phi^{flat}_{12} - \Phi^{flat}_{23} = \frac{4\pi}{\lambda}(2Z_2 - Z_3 - Z_1)_{Tide}\cos(i) \tag{A.16}$$

This is called the *flattened tidal interferogram*. Fig.A.6 shows the tidal deformation term, this is very useful to determine the position of the grounding line, which is needed to set up the boundary conditions.

A.2.4 Unwrapping

Once we have the interferogram, it is not that easy to deduce the velocity or the displacements. The measurements of each phase are known only modulo 2π. The total phase difference between two image points in an interferogram, which may measure many multiples of 2π, needs to be determined by counting the number of phase cycles, or fringes, between the

points. This crucial step in the interferogram determination is called phase unwrapping, which has been the focus of a variety of approaches and algorithm development. The easiest way to determine the absolute phase was developed by Goldstein et al. (1988). The algorithm unwraps the phase to within a common integer multiple of 2π. Then, this constant is determined by using one point of known elevation or displacement in the image. Once the absolute phase is determined, it is relatively easy to build a field of displacement or velocity for the whole interferogram (See Fig.A.7).



(a) Amplitude of the signal (b) Phase of the signal

Figure A.5: InSAR image of Pine Island Glacier (E. Rignot)

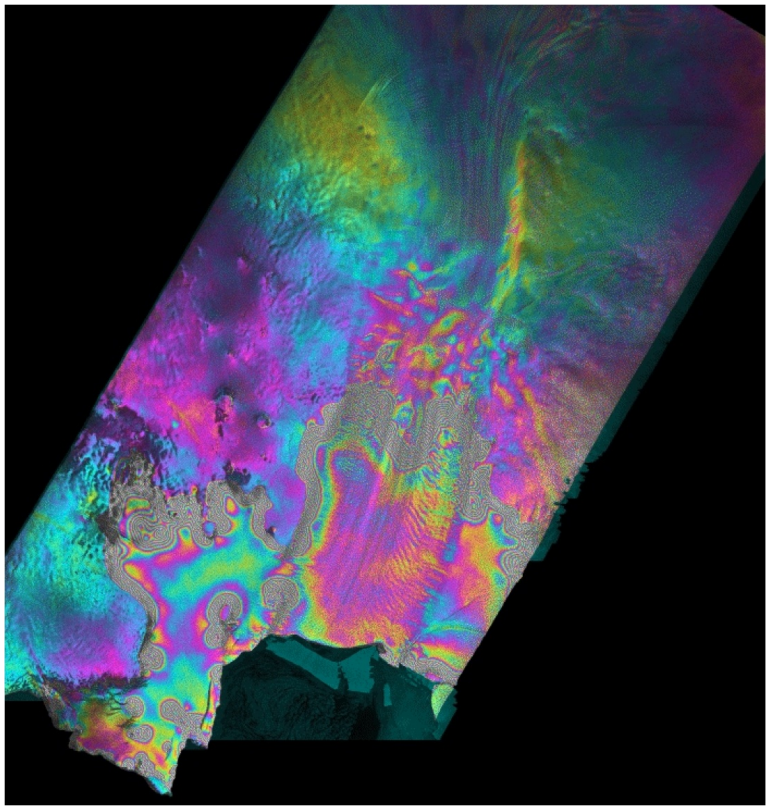


Figure A.6: Tidal deformation of Pine Island Glacier (E. Rignot)

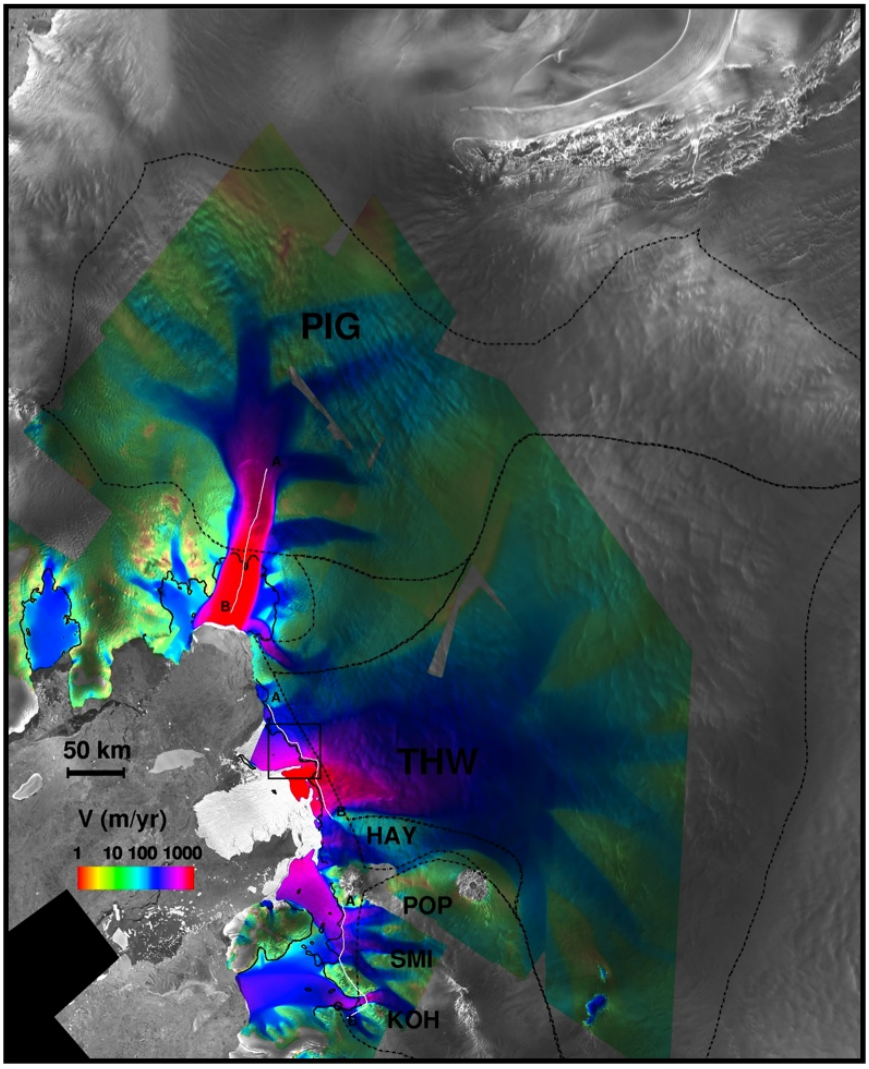


Figure A.7: InSAR derived velocity of Pine Island Glacier (E. Rignot)

B

Numerical approximations of hyperbolic problems

There are different ways to solve hyperbolic equations with the finite element method. We present here four possibilities: the artificial diffusion method, streamline upwinding, the Petrov-Galerkin method and the Discontinuous Galerkin method. To illustrate these methods, we focus on solving the balance thickness equation:

$$\begin{cases} \nabla \cdot H\bar{\mathbf{v}} &= \dot{a} & \text{in } \Omega \\ H &= H_{obs} & \text{on } \Gamma_- \end{cases} \tag{B.1}$$

B.1 Weak formulation

We derive here the variational formulation associated to the balance thickness problem (Eq. B.1). We multiply and integrate Eq. (B.1) by a weighing function φ over the domain, Ω:

$$\forall \varphi \in \mathcal{H}^1(\Omega) \quad \int_\Omega \varphi \, \nabla \cdot H\bar{\boldsymbol{v}} d\Omega = \int_\Omega \varphi \dot{a} d\Omega \tag{B.2}$$

By integrating by part the first term and using Gauss formula (Eq. C.4):

$$\forall \varphi \in \mathcal{H}^1(\Omega) \quad -\int_\Omega H\bar{\boldsymbol{v}} \cdot \nabla\varphi d\Omega + \int_{\Gamma_+} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} dS + \int_{\Gamma_-} \varphi H_{obs}\bar{\boldsymbol{v}} \cdot \boldsymbol{n} dS = \int_\Omega \varphi \dot{a} d\Omega \tag{B.3}$$

In this formulation, we apply the Dirichlet condition in a weak sense. This weak formulation can be employed but we generally apply the Dirichlet condition in the strong sense by requiring both the solution, H, and the weighing function, φ to be in the admissible, $\mathcal{V}$, defined as:

$$\mathcal{V} = \left\{ H \in \mathcal{H}^1(\Omega), \quad \forall \boldsymbol{x} \in \Gamma_- \quad H(\boldsymbol{x}) = 0 \right\} \tag{B.4}$$

which gives:

$$\forall \varphi \in \mathcal{V} \quad -\int_\Omega H\bar{\boldsymbol{v}} \cdot \nabla\varphi d\Omega + \int_{\Gamma_+} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} dS = \int_\Omega \varphi \dot{a} d\Omega \tag{B.5}$$

We define the bilinear forms $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ and $b : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, as well as the linear form $f : \mathcal{V} \to \mathbb{R}$ as:

$$\begin{aligned} a(H, \varphi) &= \int_\Omega -H\bar{\boldsymbol{v}} \cdot \nabla\varphi d\Omega \\ b(H, \varphi) &= \int_{\Gamma_+} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} dS \\ \langle f, \varphi \rangle &= \int_\Omega \varphi \, \dot{a} d\Omega \end{aligned} \tag{B.6}$$

The weak formulation is now: find $H \in \mathcal{V}$ such that

$$\forall \varphi \in \mathcal{V} \quad a(H, \varphi) + b(H, \varphi) = \langle f, \varphi \rangle \tag{B.7}$$

The bilinear forms, a and b, are non-symmetric and exhibits very unfavorable properties. Although the standard Bubnov-Galerkin approximation produces accurate solutions to elliptic and parabolic problems (for small a Péclet number), the absence of diffusion in this transport equation deprives it from *best approximation property*[1] which it is known to possess in the case of self-adjoint (symmetric) operators. The solution produces spurious oscillations, also known as *wiggles*. Due to the lack of diffusive effects, hyperbolic conservation laws admit discontinuous and, possibly, nonunique weak solutions. Such problems are particularly difficult to solve numerically, although a lot of information about the properties of exact solutions is available. The lack of robustness can be rectified by adding some artificial diffusion, using modified test functions to construct an upwind-biased finite element scheme or employing a Discontinuous Galerkin method.

B.2 Artificial diffusion

To overcome this difficulty, the first possible approximation consists of stabilizing the Galerkin operator, a, by adding a linear stabilization term:

$$a'(H,\varphi) = a(H,\varphi) + d(H,\varphi) \tag{B.8}$$

where $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is a stabilizing bilinear form. Ideally, this term should vanish if H the exact solution of the continuous problem. For practical purposes, it is sufficient to make sure that $d(H,\varphi) \to 0$ as the mesh size h goes to zero. In a classical artificial diffusion method (Kuzmin, 2010), the stabilization operator is defined as:

$$d(H,\varphi) = \int_\Omega -\nabla\varphi \cdot (\mathscr{D}\nabla H)\, d\Omega \tag{B.9}$$

where $\mathscr{D}$ is a diffusivity tensor. Solving for the new problem is equivalent to solving the following equation:

$$\nabla \cdot H\bar{\boldsymbol{v}} + \nabla \cdot (\mathscr{D}\nabla H) = \dot{a} \tag{B.10}$$

Typically, the amount of artificial diffusion depends on the local mesh size, h, and on the magnitude of the velocity vector $\bar{\boldsymbol{v}}$. The simplest way to offset the intrinsic negative diffusion of the Galerkin scheme (Brooks and Hughes, 1982; Donea, 1984) is to apply isotropic balancing dissipation of the form $\mathscr{D} = \delta\mathbf{I}$, where

$$\delta = \alpha\frac{\|\bar{\boldsymbol{v}}\|h}{2} \tag{B.11}$$

is a scalar diffusion coefficient and $\mathbf{I}$ the identity tensor. For convection-diffusion problem, α is based on the local Péclet number:

$$\alpha = \coth\left(\frac{\mathrm{Pe}_h}{2}\right) - \frac{2}{\mathrm{Pe}_h} \simeq \min\left\{1, \frac{\mathrm{Pe}_h}{6}\right\} \qquad \mathrm{Pe}_h = \frac{\|\bar{\boldsymbol{v}}\|h}{\epsilon} \tag{B.12}$$

where ϵ is the diffusivity. In this case, $\epsilon = 0$ and we therefore take $\alpha = 1$.

[1] the Galerkin solution is as close to the exact solution as any other vector in its discretized space $\mathcal{V}$

B.3 Streamline Upwinding

In multidimensional problems, the velocity field determines the direction and speed of convective transport. The results of simple artificial diffusivity results are frequently polluted by numerical crosswind diffusion that could be removed without making the scheme unstable. This led to the introduction of the *anisotropic balancing dissipation* (Donea, 1984; Kuzmin, 2010) that acts along the streamlines of the velocity but not transversely:

$$\mathscr{D} = \tau \bar{\boldsymbol{v}} \otimes \bar{\boldsymbol{v}} \qquad \tau = \frac{\delta}{\|\bar{\boldsymbol{v}}\|^2} = \frac{\alpha h}{2\|\bar{\boldsymbol{v}}\|} \tag{B.13}$$

if $\alpha = 1$, it can be shown that this anisotropic diffusion is equivalent to a first-order upwinding along the streamlines. The lack of crosswind diffusion results in a smaller deviation from the isotropic artificial diffusion.

The stabilizing bilinear form $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is:

$$\begin{aligned} d(H, \varphi) &= \int_\Omega -\nabla\varphi \cdot (\tau \bar{\boldsymbol{v}} \otimes \bar{\boldsymbol{v}} \nabla H)\, d\Omega \\ &= \int_\Omega -\tau\, (\bar{\boldsymbol{v}} \cdot \nabla\varphi)\, (\bar{\boldsymbol{v}} \cdot \nabla H)\, d\Omega \end{aligned} \tag{B.14}$$

B.4 Petrov-Galerkin Method

The Galerkin method is a method of weighted residuals: the equation is verified only for a finite number of weighing functions:

$$\forall \varphi \in \boldsymbol{V}_h \quad a(H, \varphi) + b(H, \varphi) = \langle f, \varphi \rangle \tag{B.15}$$

The idea behind Petrov-Galerkin is to use weighing functions different than the one used to decompose the unknown, H. We modify the weighing functions as follows:

$$\tilde{\varphi} = \varphi + \tau \bar{\boldsymbol{v}} \cdot \nabla\varphi \tag{B.16}$$

Fig. B.1 shows the difference between a standard Bubnov-Galerkin weighing function and a Petrov-Galerkin nonconforming weighing function in 1d for linear elements. We see that more weight is given upwind, which stabilizes the convection term.

The weak formulation becomes:

$$\forall \varphi \in \mathcal{V} \quad a(H, \varphi + \tau \bar{\boldsymbol{v}} \cdot \nabla\varphi) + b(H, \varphi + \tau \bar{\boldsymbol{v}} \cdot \nabla\varphi) = \langle f, \varphi + \tau \bar{\boldsymbol{v}} \cdot \nabla\varphi \rangle \tag{B.17}$$

With the linearity of the operators, it gives:

$$a(H, \varphi) + b(H, \varphi) + S(H, \varphi) = \langle f, \varphi \rangle \tag{B.18}$$

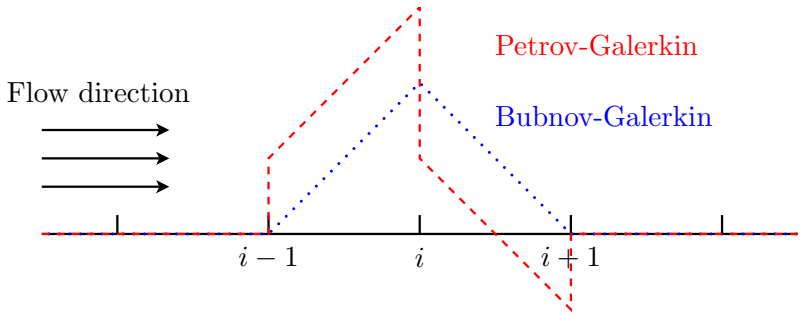


Figure B.1: One dimensional comparision between Bubnov-Galerkin and Petrov-Galerkin test functions

So we just added the operator S such that:

$$S\left(H,\varphi\right)=a\left(H,\tau\bar{\boldsymbol{v}}\cdot\nabla\varphi\right)+b\left(H,\tau\bar{\boldsymbol{v}}\cdot\nabla\varphi\right)-\left\langle f,\tau\bar{\boldsymbol{v}}\cdot\nabla\varphi\right\rangle \tag{B.19}$$

Franca et al. (2006) gives the expression of this operator for different problems.

B.5 Discontinuous Galerkin

The Discontinuous Galerkin (DG) method is becoming more and more popular in computational mechanics. This method combines features of the finite element and the finite volume framework. It was introduced in the 1970s (Reed and Hill, 1973) and has been devised to solve nonlinear hyperbolic systems and convection-dominated problem. This method is therefore well-suited for our balance thickness problem.

The idea behind the Discontinuous Galerkin method is to allow the test functions to be discontinuous across element interfaces. This leads to undefined inter-element terms or numerical fluxes, which are chosen to enforce the consistency of the numerical scheme.

B.5.1 Jump and average definitions

Let $\mathcal{T}_h$ be a regular family of decomposition of Ω into triangles K. Let $\mathcal{E}_h$ be the set of all edges of $\mathcal{T}_h$. This set is decomposed in interior edges, $\mathcal{E}_h^o$, and boundary edges, $\mathcal{E}_h^\partial$. We have of course $\mathcal{E}_h=\mathcal{E}_h^o\cup\mathcal{E}_h^\partial$.

Let $e\in\mathcal{E}_h^o$ be an interior edge shared by two triangles K_1 and K_2. The unit normal vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ on e are pointing outward toward K_1 and K_2 respectively (Fig. B.2).

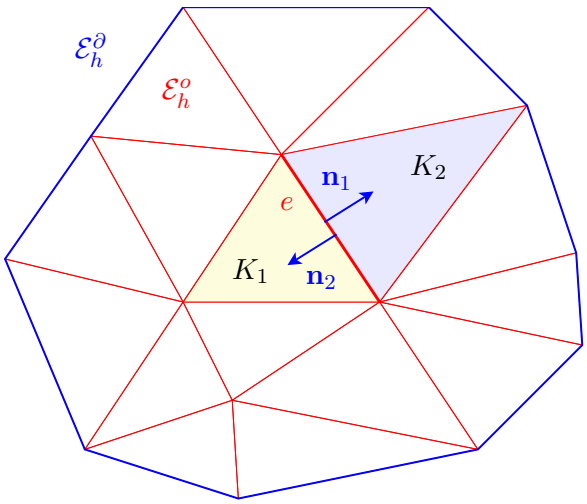


Figure B.2: Triangulation, $\mathcal{T}_h$, of the domain, Ω. Interior edges, $\mathcal{E}_h^o$, are shown in red and boundary edges, $\mathcal{E}_h^\partial$, are in blue. The edge e is shared by two triangles: K_1 and K_2.

Jump and Average operators

For any scalar function γ and any vector-valued function $\boldsymbol{v}$ piecewise smooth on $\mathcal{T}_h$, with $\gamma_i = \gamma_{|_{K_i}}$, we define the average operator accross e as:

$$\begin{aligned} &\forall e \in \mathcal{E}_h^o \qquad \{\gamma\} = \frac{1}{2}(\gamma_1 + \gamma_2) \qquad \{\boldsymbol{v}\} = \frac{1}{2}(\boldsymbol{v}_1 + \boldsymbol{v}_2) \\ &\forall e \in \mathcal{E}_h^\partial \qquad\qquad\qquad\quad \{\boldsymbol{v}\} = \boldsymbol{v} \end{aligned} \tag{B.20}$$

The jump operator is defined as follows:

$$\begin{aligned} &\forall e \in \mathcal{E}_h^o \qquad [\![\gamma]\!] = \gamma_1 \boldsymbol{n}_1 + \gamma_2 \boldsymbol{n}_2 \qquad [\![\boldsymbol{v}]\!] = \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 \\ &\forall e \in \mathcal{E}_h^\partial \qquad [\![\gamma]\!] = \gamma\, \boldsymbol{n} \end{aligned} \tag{B.21}$$

We leave the definition of $[\![\boldsymbol{v}]\!]$ and $\{\gamma\}$ undefined as they won't be used. Notice that the jump of a scalar function γ accross $e \in \mathcal{E}_h^o$ is a vector parallel to the normal of e ($\boldsymbol{n}_2 = -\boldsymbol{n}_1$), and the jump of a vector-valued function $\boldsymbol{v}$ is a scalar quantity.

B.5.2 Unstabilize discrete problem

We define the finite element space of discontinuous piecewise linear polunomial functions as:

$$\mathcal{V}_h = \left\{ \varphi \in \mathcal{H}^1(\Omega) \quad \forall K \in \mathcal{T}_h,\ \varphi_{|_K} \in P^1(K) \right\} \tag{B.22}$$

where P^1 is the usual space of polynomials of degree 1 or less. The integration by parts is done over individual elements. Eq. 5.9 is multiplied by a function $\varphi \in \mathcal{V}_h$ and integrated

over the domain Ω:

$$\forall \varphi \in \mathcal{V}_h \quad \sum_{K \in \mathcal{T}_h} \left(\int_K \varphi \, \nabla \cdot H\bar{\boldsymbol{v}} \, dK \right) = \int_\Omega \dot{a} \, \varphi \, ds \tag{B.23}$$

We integrate by part the first term of the previous equation:

$$\forall \varphi \in \mathcal{V}_h \quad \sum_{K \in \mathcal{T}_h} \left(\int_K -H\bar{\boldsymbol{v}} \cdot \nabla \varphi d\Omega + \int_{\partial K} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} \, dx \right) = \int_\Omega \dot{a} \, \varphi \, ds \tag{B.24}$$

For any vector $\boldsymbol{v}$ and scalar γ piecewise smooth on $\mathcal{T}_h$, we have the following identity[2]:

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_{\partial K} \gamma \left(\boldsymbol{v} \cdot \boldsymbol{n} \right) d\Omega &= \sum_{e \in \mathcal{E}_h^\partial} \int_e \gamma \boldsymbol{v} \cdot \boldsymbol{n} \, dx + \sum_{e \in \mathcal{E}_h^o} \int_e \gamma_1 \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \gamma_2 \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 \, dx \\ &= \sum_{e \in \mathcal{E}_h^\partial} \int_e \{\boldsymbol{v}\} \cdot [\![\gamma]\!] dx + \sum_{e \in \mathcal{E}_h^o} \int_e [\![\boldsymbol{v}]\!] \{\gamma\} + \{\boldsymbol{v}\} \cdot [\![\gamma]\!] dx \\ &= \sum_{e \in \mathcal{E}_h} \int_e \{\boldsymbol{v}\} \cdot [\![\gamma]\!] dx + \sum_{e \in \mathcal{E}_h^o} \int_e [\![\boldsymbol{v}]\!] \{\gamma\} \, dx \end{aligned} \tag{B.26}$$

This gives here:

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} \, d\Omega = \sum_{e \in \mathcal{E}_h} \int_e \{H\bar{\boldsymbol{v}}\} \cdot [\![\varphi]\!] dx + \sum_{e \in \mathcal{E}_h^o} \int_e [\![H\bar{\boldsymbol{v}}]\!] \{\varphi\} \, dx \tag{B.27}$$

The flux continuity imposes that for every edge e shared by T_1 and T_2 we have:

$$[\![H\bar{\boldsymbol{v}}]\!] = H_1\bar{\boldsymbol{v}}_1 \cdot \boldsymbol{n}_1 + H_2\bar{\boldsymbol{v}}_2 \cdot \boldsymbol{n}_2 = (H_1\bar{\boldsymbol{v}}_1 - H_2\bar{\boldsymbol{v}}_2) \cdot \boldsymbol{n}_1 = \boldsymbol{0} \tag{B.28}$$

The last term of Eq. B.27 is hence zero. We rewrite the equation by decomposing the integral over the edges belonging to Γ_- and the others:

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_{\partial K} \varphi H\bar{\boldsymbol{v}} \cdot \boldsymbol{n} \, d\Omega &= \sum_{e \in \mathcal{E}_h} \int_e \{H\bar{\boldsymbol{v}}\} \cdot [\![\varphi]\!] dx \\ &= \sum_{e \not\subset \Gamma_-} \int_e \{H\bar{\boldsymbol{v}}\} \cdot [\![\varphi]\!] dx + \sum_{e \subset \Gamma_-} \int_e H_{obs} \varphi \bar{\boldsymbol{v}} \cdot \boldsymbol{n} \, dx \end{aligned} \tag{B.29}$$

[2] For one edge $e \in \mathcal{E}_h^o$, we have the following relations:

$$\begin{aligned} [\![\boldsymbol{v}]\!] \{\gamma\} &= (\boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{v}_2 \cdot \boldsymbol{n}_2) \left(\frac{1}{2}\gamma_1 + \frac{1}{2}\gamma_2 \right) = \frac{1}{2} \left(\gamma_1 \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \gamma_2 \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \gamma_1 \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 + \gamma_2 \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 \right) \\ \{\boldsymbol{v}\} \cdot [\![\gamma]\!] &= \left(\frac{1}{2}\boldsymbol{v}_1 + \frac{1}{2}\boldsymbol{v}_2 \right) \cdot (\gamma_1 \boldsymbol{n}_1 + \gamma_2 \boldsymbol{n}_2) = \frac{1}{2} \left(\gamma_1 \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \gamma_2 \boldsymbol{v}_1 \cdot \boldsymbol{n}_2 + \gamma_1 \boldsymbol{v}_2 \cdot \boldsymbol{n}_1 + \gamma_2 \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 \right) \end{aligned}$$

And since we have $\boldsymbol{n}_1 = -\boldsymbol{n}_2$, the sum is reduced to:

$$[\![\boldsymbol{v}]\!] \{\gamma\} + \{\boldsymbol{v}\} \cdot [\![\gamma]\!] = \gamma_1 \boldsymbol{v}_1 \cdot \boldsymbol{n}_1 + \gamma_2 \boldsymbol{v}_2 \cdot \boldsymbol{n}_2 \tag{B.25}$$

This gives finally:

$$\forall \varphi \in \mathcal{V}_h \qquad \sum_{K\in\mathcal{T}_h} \int_K -H\bar{\boldsymbol{v}}\cdot\nabla\varphi d\Omega + \sum_{e\not\subset\Gamma_-}\int_e \{H\bar{\boldsymbol{v}}\}\cdot[\![\varphi]\!]dx = \int_\Omega \dot{a}\ \varphi\ dx - \sum_{e\subset\Gamma_-}\int_e H_{obs}\varphi\bar{\boldsymbol{v}}\cdot\boldsymbol{n}\ dx \quad \text{(B.30)}$$

Unstabilized discrete problem

With the following operators:

$$\begin{aligned} a_h(H,\varphi) &= \sum_{K\in\mathcal{T}_h}\int_T -H\bar{\boldsymbol{v}}\cdot\nabla\varphi\ d\Omega \\ b_h(H,\varphi) &= \sum_{e\not\subset\Gamma_-}\int_e \{H\bar{\boldsymbol{v}}\}\cdot[\![\varphi]\!]\ dx \\ (\dot{a},\varphi) &= \int_\Omega \dot{a}\ \varphi\ dx \\ \langle g,\varphi\rangle &= \sum_{e\subset\Gamma_-} -\int_e H_{obs}\varphi\bar{\boldsymbol{v}}\cdot\boldsymbol{n}\ dx \end{aligned} \quad \text{(B.31)}$$

the unstabilized discrete problem is defined as:

$$\begin{cases} \text{find } H\in\mathcal{V}_h \\ a_h(H,\varphi)+b_h(H,\varphi) = (\dot{a},\varphi)+\langle g,\varphi\rangle \end{cases} \quad \text{(B.32)}$$

B.5.3 Upwind-value numerical flux

According to Brezzi et al. (2004), the formulation Eq.B.32 is stable but only with respect to the $\mathcal{L}^2$-norm, which leads to discontinuities in the boundary data and may trigger large, nonphysical oscillations in the numerical solution.

In order to design a solution that is stable in a stronger norm, the average $\{H\bar{\boldsymbol{v}}\}$, which appears in the bilinear form b_h of Eq.B.32 is modified following the upwind scheme:

$$\{H\bar{\boldsymbol{v}}\}_u = \begin{cases} \bar{\boldsymbol{v}}H_1 & \text{if} \quad \bar{\boldsymbol{v}}\cdot\boldsymbol{n}_1 > 0 \\ \bar{\boldsymbol{v}}H_2 & \text{if} \quad \bar{\boldsymbol{v}}\cdot\boldsymbol{n}_1 < 0 \\ \bar{\boldsymbol{v}}\{H\} & \text{if} \quad \bar{\boldsymbol{v}}\cdot\boldsymbol{n}_1 = 0 \end{cases} \quad \text{(B.33)}$$

Only the normal component of the term is involved as:

$$b_h(H,\varphi) = \sum_{e \not\subset \Gamma_-} \int_e \{H\bar{\boldsymbol{v}}\} \cdot [\![\varphi]\!] \, dx \tag{B.34}$$

and $[\![\varphi]\!]$ is along the normal of the edge. We have the following identity:

$$\begin{aligned} \{H\bar{\boldsymbol{v}}\}_u \cdot \boldsymbol{n} &= \frac{1}{2}(H_{h_1} + H_{h_2})(\bar{\boldsymbol{v}} \cdot \boldsymbol{n}) + \frac{1}{2}(H_{h_1} - H_{h_2})|\bar{\boldsymbol{v}} \cdot \boldsymbol{n}| \\ &= \left(\{H\bar{\boldsymbol{v}}\} + \frac{|\bar{\boldsymbol{v}} \cdot \boldsymbol{n}|}{2}[\![\varphi]\!]\right) \cdot \boldsymbol{n} \end{aligned} \tag{B.35}$$

B.5.4 Jump-penalty stabilization

It can be shown (Brezzi et al., 2004) that by choosing the following numerical flux:

$$\{H\bar{\boldsymbol{v}}\}_u = (\{H\bar{\boldsymbol{v}}\} + c_e[\![H]\!]) \tag{B.36}$$

we achieve a stability in a norm stronger that $\mathcal{L}^2$, provided that c_e is a nonnegative function chosen on each e so that:

$$c_e \leq \theta_0 |\bar{\boldsymbol{v}} \cdot \boldsymbol{n}| \tag{B.37}$$

where θ_0 is a positive constant independent of the mesh size. We therefore replace b_h by its *stabilized* version:

$$b_h^s(H,\varphi) = b_h(H,\varphi) + \sum_{e \in \mathcal{E}_h^o} \int_e c_e [\![H]\!] \cdot [\![\varphi]\!] \, dx \tag{B.38}$$

In ISSM we chose $c_e = |\bar{\boldsymbol{v}} \cdot \boldsymbol{n}|/2$. The jump-penalty stabilization is therefore equivalent to the upwinding.

C

Useful integrals formula

Several integral formulas are used throughout this document. These formulas are summarized here with or without proof as some of them are very common in the litterature. For all these formulas:

- $V(t)$ is a volume that encloses a body
- $S(t)$ is its outer surface assumed to be a material surface
- $\varphi(x,y,z,t)$ is a scalar variable (temperature, density,...)
- $\boldsymbol{F}(x,y,z,t)$ is a vector variable (velocity,...)

C.1 Leibnitz Theorem

The *Leibnitz integral formula* also known as *Leibniz's rule* is very useful for integrals over a moving domain.

3d case

$$\frac{D}{Dt}\int_{V(t)}\varphi(x,y,z,t)\,dV=\int_{V(t)}\frac{\partial\varphi}{\partial t}dV+\int_{S(t)}\varphi(x,y,z,t)\,\boldsymbol{v}\cdot\boldsymbol{n}\,dS \tag{C.1}$$

1d case

$$\frac{D}{Dt}\int_{a(t)}^{b(t)}f(x,t)\,dx=\frac{\partial b(t)}{\partial t}f(b,t)-\frac{\partial a(t)}{\partial t}f(a,t)+\int_{a(t)}^{b(t)}\frac{\partial f}{\partial t}dx \tag{C.2}$$

C.2 Reynold's transport formula

The well-known *Reynold's transport formula* can be derived from (C.1) by taking $\varphi=\varphi(\boldsymbol{r},t)$:

$$\frac{D}{Dt}\int_{V(t)}\varphi\,dV=\int_{V(t)}\frac{\partial\varphi}{\partial t}dV+\int_{S(t)}\varphi\,\boldsymbol{v}\cdot\boldsymbol{n}\,dS \tag{C.3}$$

C.3 Integral formula of Gauss

The integral formula of Gauss is also known as the *divergence theorem*.

$$\int_{V(t)}\nabla\cdot\boldsymbol{F}dV=\int_{S(t)}\boldsymbol{F}\cdot\boldsymbol{n}\,dS \tag{C.4}$$

C.4 Another useful integral formula

The mass conservation equation (Eq. 1.8)) can be used with Reynolds transport formula (Eq. C.3) for $\boldsymbol{F} = \varphi \boldsymbol{v}$:

$$\int_{V(t)} \nabla \cdot \varphi \boldsymbol{v} dV = \int_{S(t)} \varphi \boldsymbol{v} \cdot \boldsymbol{n} \, dS \tag{C.5}$$

The Reynold's transport formula can now be rewritten:

$$\frac{D}{Dt} \int_{V(t)} \varphi \, dV = \int_{V(t)} \frac{\partial \varphi}{\partial t} dV + \int_{V(t)} \nabla \cdot \varphi \boldsymbol{v} dV \tag{C.6}$$

We can use the following identity:

$$\frac{\partial \varphi}{\partial t} + \nabla \cdot \varphi \boldsymbol{v} = \frac{D\varphi}{Dt} + \varphi \nabla \cdot \boldsymbol{v} \tag{C.7}$$

which transforms the previous equation:

$$\frac{D}{Dt} \int_{V(t)} \varphi \, dV = \int_{V(t)} \frac{D\varphi}{Dt} + \varphi \nabla \cdot \boldsymbol{v} \tag{C.8}$$

For $\varphi = \rho\theta$ where ρ is the densityand θ some arbitrary variable, we have:

$$\begin{aligned} \frac{D}{Dt} \int_{V(t)} \rho\theta \, dV &= \int_{V(t)} \frac{D}{Dt} (\rho\theta) + \rho\theta \nabla \cdot \boldsymbol{v} \\ &= \int_{V(t)} \rho \frac{D\theta}{Dt} + \theta \left(\frac{D\rho}{Dt} + \rho \nabla \cdot \boldsymbol{v} \right) \end{aligned} \tag{C.9}$$

And finally, using the incompressibility, we have:

$$\frac{D}{Dt} \int_{V(t)} \rho\theta \, dV = \int_{V(t)} \rho \frac{D\theta}{Dt} dV \tag{C.10}$$

D

Brent fitting

D.1 Golden section search

D.1.1 Introduction

We are trying to find the minimum of a single-function $f(x)$ on a segment $[a, b]$ assuming that this minimum is unique. The golden section search is similar to the bisection in the sense that one evaluates the function in the interval $[a, b]$ and obtains a new, smaller bracketing interval.

But instead of evaluating the function at one point only, the Golden section search scheme uses two interior points c and d. At each iteration, one has a triplet of points $a < c < b$, such that $f(c)$ is less than or equal to both $f(a)$ and $f(b)$, only $f(d)$ is to be computed. We continue the process of bracketing until the distance between the two outer points of the triplet is sufficiently small.

D.1.2 Strategy

Suppose that c is closer[1] to a and is a fraction R of the way between a and b, i.e.

$$\frac{c-a}{b-a} = R > 0 \tag{D.1}$$

We choose the second guess d, such that (See Fig.D.1):

$$\frac{d-c}{b-a} = R' > 0 \tag{D.2}$$

If $f(c) > f(d)$, the interval is reduced to $[c, b]$ and if $f(c) < f(d)$, the interval is reduced to $[a, d]$. The length of the new interval will hence be either $(1-R)(b-a)$ or $(R+R')(b-a)$. If we want to minimize the worst case possibility, then we will choose R' to make them equal, namely:

$$R' = 1 - 2R \tag{D.3}$$

We see that R' is positive only if $R < 0.5$. This implies that the new point is the symmetric point to c in the interval:

$$\begin{aligned} |a-b| &= |b-d| + |d-c| + |c-a| \\ &= |b-d| + (1-2R)|b-a| + R|b-a| \\ |b-d| &= R|a-b| \\ &= |c-a| \end{aligned} \tag{D.4}$$

This implies that the point d lies in the larger of the two segments $[a, c]$, $[c, b]$ (As shown on

[1] would the point c be closer to b, we would have had $\frac{b-c}{b-a} = R > 0$

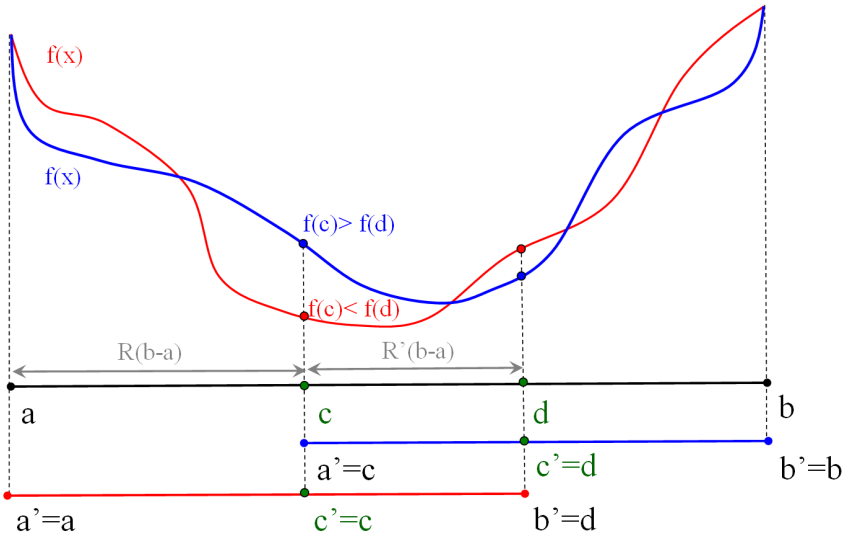


Figure D.1: Golden search scheme

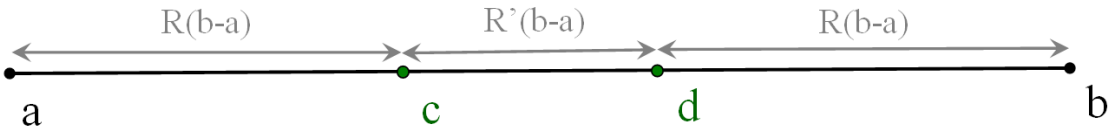


Figure D.2: Repartition of c and d

Fig.D.2).
Depending on the value of $f(d)$, the next triplet is $(a', c', b') = (a, c, d)$ or $(a', c', b') = (c, d, b)$. We want to keep the same ratio between the $\min(c' - a, b - c')$ and $b - a$ in order to apply the same procedure. This implies:

$$R = \frac{R'(b-a)}{b-a-R(b-a)} = \frac{R'}{1-R} = \frac{1-2R}{1-R} \tag{D.5}$$

Which finally yields the quadratic equation:

$$R^2 - 3R + 1 = 0 \tag{D.6}$$

Since $0 < R < 0.5$, there is only one acceptable root which is:

$$R = \frac{3-\sqrt{5}}{2} \approx 0.38197\ldots \tag{D.7}$$

D.1.3 Iterations

The previous scheme is repeated at each iteration. We have a triplet of point. The inner point can be either closer to b (it is c) or to a (it is d). The fourth point is computed as a

fraction R of the current interval $[a, b]$ located between the largest segment. The function is evaluated on this new point and compared to its value on the other inner point. We obtain a smaller bracketing interval as shown in Fig.D.3.

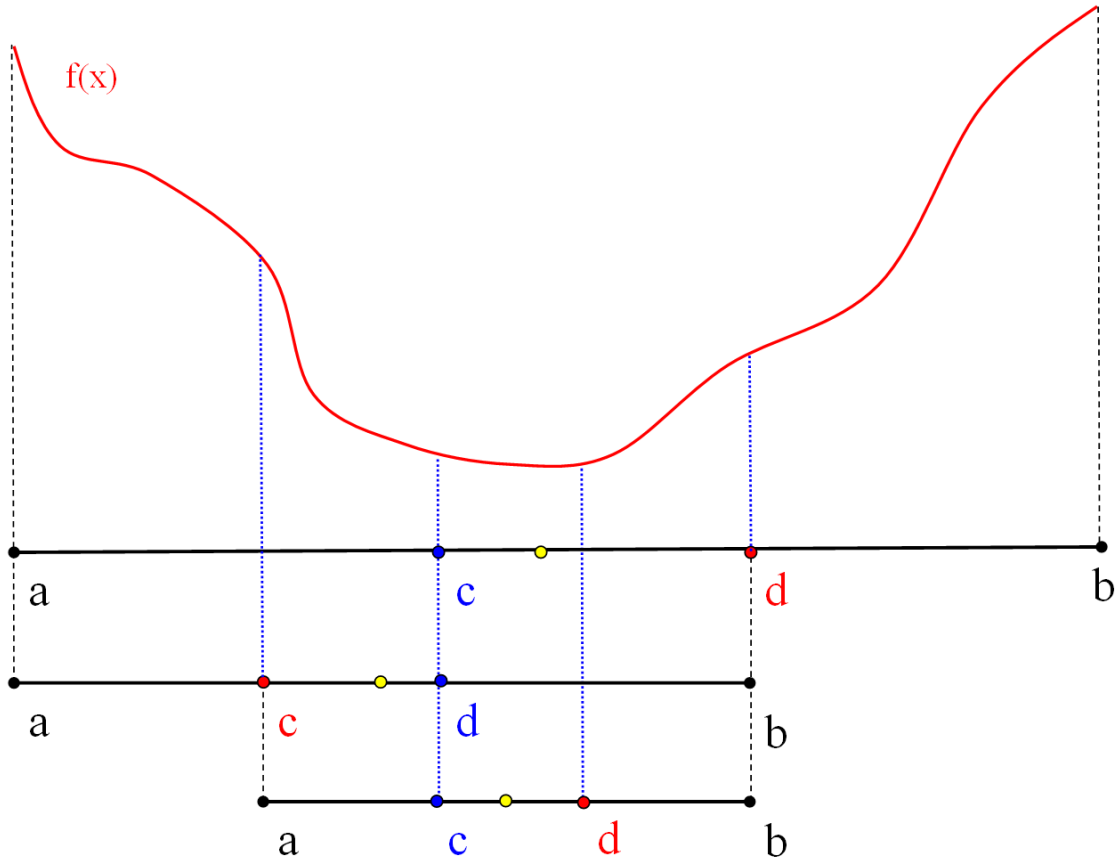


Figure D.3: Golden Search Scheme

D.2 Parabolic interpolation

D.2.1 Introduction

Exactly as in the Golden Search Scheme, we assume to have a triplet of points $a < c < b$, such that $f(c)$ is less than both $f(a)$ and $f(b)$. At each iteration, we find the quadratic function that goes through the three points $(a, f(a))$, $(b, f(b))$ and $(c, f(c))$. The minimum of this parabola is easy to compute by searching the root of its derivative. This gives the point d. The function is evaluated on this point and we obtain a smaller bracketing interval by applying the same rule as in the GSS.

D.2.2 Parabola fitting

The equation of the parabola through $f(a)$, $f(b)$ and $f(c)$ is:

$$P(x) = f(a)\frac{(x-b)(x-c)}{(a-b)(a-c)} + f(b)\frac{(x-c)(x-a)}{(b-c)(b-a)} + f(c)\frac{(x-a)(x-b)}{(c-a)(c-b)} \tag{D.8}$$

The idea is that a quadratic function is uniquely defined by three points and the previous parabola goes through the three points of the triplet, it is hence the parabola one is looking for. Its derivative is easily computed:

$$\frac{dP}{dx} = f(a)\frac{(x-b)-(x-c)}{(a-b)(a-c)} + f(b)\frac{(x-c)-(x-a)}{(b-c)(b-a)} + f(c)\frac{(x-a)-(x-b)}{(c-a)(c-b)} \tag{D.9}$$

The minimum of the parabola is d where is derivative is equal to zero:

$$0 = f(a)\frac{(d-b)-(d-c)}{(a-b)(a-c)} + f(b)\frac{(d-c)-(d-a)}{(b-c)(b-a)} + f(c)\frac{(d-a)-(d-b)}{(c-a)(c-b)} \tag{D.10}$$

We multiply with the denominator $(a-b)(b-c)(c-a)$ and get:

$$0 = f(a)(c-b)(2d-b-c) + f(b)(a-c)(2d-c-a) + f(c)(b-a)(2d-a-b) \tag{D.11}$$

This gives:

$$\begin{aligned} d &= \frac{1}{2}\frac{f(a)\left(c^2-b^2\right) + f(b)\left(a^2-c^2\right) + f(c)\left(b^2-a^2\right)}{f(a)(c-b) + f(b)(a-c) + f(c)(b-a)} \\ &= b + \frac{1}{2}\frac{f(a)(c-b)(c+b-2b) + f(b)(a-c)(a+c-2b) + f(c)(b-a)(b+a-2b)}{f(a)(c-b) + f(b)(a-c) + f(c)(b-a)} \\ &= b + \frac{1}{2}\frac{(f(a)-f(b))(c-b)^2 + (f(b)-f(c))(b-a)^2}{(f(a)-f(b))(c-b) + (f(b)-f(c))(b-a)} \end{aligned} \tag{D.12}$$

This formula fails only if three points are collinear, in which case the denominator is zero (minimum of the parabola is infinitely far away).

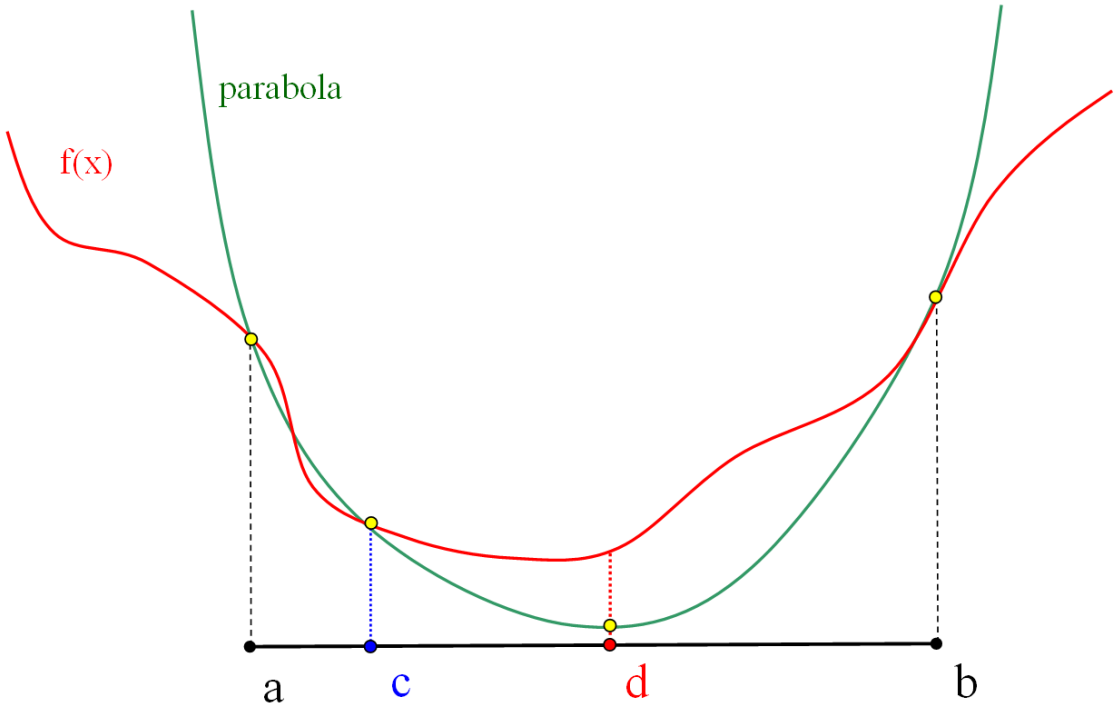


Figure D.4: Parabola fitting

Bibliography

Ait-Ali-Yahia, D., G. Baruzzi, W. Habashi, M. Fortin, J. Dompierre and M. Vallet. *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent cfd. part ii. structured grids.* Int. J. Numer. Meth. Fluids, volume 39, no. 8; pages 657–673, JUL 20 2002. ISSN 0271-2091. doi:10.1002/fld.356. (page 65).

Alauzet, F. *Adaptation de maillage anisotrope en trois dimensions. Application aux simulations instationnaires en Mecanique des fluides.* Ph.D. thesis, Universite Montpellier II, 2003. (pages 67, 68).

Alauzet, F. and P. Frey. *Estimateur d'erreur géométrique et métriques anisotropes pour l'adaptation de maillage. partie i : aspects théoriques.* Technical report, INRIA, 2003. (pages 66, 67, 68).

Alauzet, P. F., Frederic and J.-H. Saiac. *Simulation numerique en geometries complexes : apport des techniques d'adaptation de maillages. Applications aux calculs de mecanique des fuides.*, volume cours d'application 4, deuxieme annee (Ecole Centrale Paris, 2007), 182 pages. (pages 67, 68).

Alley, R., P. Clark, P. Huybrechts and I. Joughin. *Ice-sheet and sea-level changes.* Science, volume 310, no. 5747; pages 456–460, OCT 21 2005. ISSN 0036-8075. doi:10.1126/science.1114613. (page 7).

Amestoy, P. R., I. S. Duff, J. Koster and J.-Y. L'Excellent. *A Fully Asynchronous Multifrontal Solver Using Distributed Dynamic Scheduling.* SIAM J. Matrix Anal. Appl., volume 23, no. 1; pages 15–41, 2001. (page 63).

Amestoy, P. R., A. Guermouche, J.-Y. L'Excellent and S. Pralet. *Hybrid scheduling for the parallel solution of linear systems.* Parallel Computing, volume 32, no. 2; pages 136–156, 2006. (page 63).

Arthern, R. J. and G. H. Gudmundsson. *Initialization of ice-sheet forecasts viewed as an inverse robin problem.* J. Glaciol., volume 56, no. 197; pages 527–533, 2010. ISSN 0022-1430. (page 76).

Balay, S., K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith and H. Zhang. *Petsc users manual.* Technical Report ANL-95/11 - Revision 3.0.0, Argonne National Laboratory, 2008. (page 63).

Balay, S., K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith and H. Zhang. *PETSc Website*, 2009. Http://www.mcs.anl.gov/petsc. (page 63).

Balay, S., W. D. Gropp, L. C. McInnes and B. F. Smith. *Efficient Management of Parallelism in Object Oriented Numerical Software Libraries*. In E. Arge, A. M. Bruaset and H. P. Langtangen, editors, *Modern Software Tools in Scientific Computing*, (pages 163–202) (Birkhäuser Press, 1997). (page 63).

Bamber, J., R. Layberry and S. Gogineni. *A new ice thickness and bed data set for the Greenland ice sheet 1. Measurement, data reduction, and errors*. J. Geophys. Res., volume 106, no. D24; pages 33 773–33 780, DEC 27 2001. ISSN 0747-7309. (page 107).

Bamber, J. L., J. L. Gomez-Dans and J. A. Griggs. *A new 1 km digital elevation model of the Antarctic derived from combined satellite radar and laser data - Part 1: Data and methods*. The Cryosphere, volume 3, no. 1; pages 101–111, 2009. ISSN 1994-0416. URL http://www.the-cryosphere.net/3/101/2009/ (page 96).

Bell, R. E. *The unquiet ice*. Scientific American, volume February; pages 60–67, 2007. (pages xv, 3).

Bindschadler, R. *History of lower Pine Island Glacier, West Antarctica, from Landsat imagery*. J. Glaciol., volume 48, no. 163; pages 536–544, 2002. ISSN 0022-1430. doi: 10.3189/172756502781831052. (page 7).

Blatter, H. *Velocity and stress-fields in grounded glaciers: A simple algorithm for including deviatoric stress gradients*. J. Glaciol., volume 41, no. 138; pages 333–344, 1995. ISSN 0022-1430. (pages 39, 44).

Blatter, H., G. Greve and A. Abe-ouchi. *A short history of the thermomechanical theory and modelling of glaciers and ice sheets*. Annals of Glaciology, volume 56 (200); pages 187–194, 2010. (page 3, 3).

Bonnet, M. *Problemes inverses*, volume Master recherche- Ecole Centrale Paris Mention Matiere, Structures, Fluides, Rayonnement Specialite Dynamique des Structures et Systemes Couples (Ecole Centrale Paris, 2007). (pages 21, 29).

Brezzi, F., L. Marini and E. Suli. *Discontinuous Galerkin methods for first-order hyperbolic problems*. Math. Models Methods Appl. Sci., volume 14, no. 12; pages 1893–1903, DEC 2004. ISSN 0218-2025. (pages 151, 152).

van den Broeke, M. *Depth and density of the Antarctic firn layer*. Arctic Ant. Alp. Res., volume 40, no. 2; pages 432–438, MAY 2008. ISSN 1523-0430. (page 96).

Brooks, A. and T. Hughes. *Streamline upwind Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*. Comput. Methods Appl. Mech. Engrg., volume 32, no. 1-3; pages 199–259, 1982. ISSN 0045-7825. (page 146).

Budd, W. F. and R. C. Warner. *A computer scheme for rapid calculations of balance-flux distributions*. Ann. Glaciol., volume 23; pages 21–27, 1996. (page 111).

Burgmann, R., P. Rosen and E. Fielding. *Synthetic aperture radar interferometry to measure earth's surface topography and its deformation*. Annu. Rev. Earth Planet. Sci., volume 28; pages 169–209, 2000. ISSN 0084-6597. (page 131).

Christensen, E. L., N. Reeh, R. Forsberg, J. H. Jørgensen, N. Skou and K. Woelders. *A low-cost glacier-mapping system.* J. Glaciol., volume 46, no. 154; pages 531–537, 2000. ISSN 0022-1430. (page 118).

Clarke, G. *A short history of scientific investigations on glaciers.* J. Glaciol., volume Special Issue; pages 4–24, 1987. (page 1).

Cuffey, K. and W. S. B. Paterson. *The Physics of Glaciers, 4th Edition* (Elsevier, 2010). (pages 35, 35, 35, 41, 41, 41, 43, 53).

De Angelis, H. and P. Skvarca. *Glacier surge after ice shelf collapse.* Science, volume 299, no. 5612; pages 1560–1562, MAR 7 2003. ISSN 0036-8075. doi:10.1126/science.1077987. (page 7, 7).

Deutsch, C. and A. Joumel. *GSLIB. Geostatistical Software Library and User's Guide, 2nd ed.* (Oxford Univ. Press, New York, 1997). (pages 105, 107).

Dompierre, J., M. Vallet, Y. Bourgault, M. Fortin and W. Habashi. *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent cfd. part iii. unstructured meshes.* Int. J. Numer. Meth. Fluids, volume 39, no. 8; pages 675–702, JUL 20 2002. ISSN 0271-2091. doi:10.1002/fld.357. (page 65).

Donea, J. *Recent advances in computational methods for steady and transient transport problems.* Nucl Eng Des, volume 80, no. 2; pages 141–162, 1984. ISSN 0029-5493. doi:{10.1016/0029-5493(84)90163-8}. (pages 111, 146, 147).

Donea, J. and T. Belytschko. *Advances in computational mechanics.* Nucl Eng Des, volume 134, no. 1; pages 1–22, MAY 1992. ISSN 0029-5493. (pages 13, 16, 61).

Donea, J., A. Huerta, J.-P. Ponthot and A. Rodrìguez-Ferran. *Arbitrary Lagrangian-Eulerian methods*, chapter 14, (pages 413–437) (Encyclopedia of Computational Mechanics, 2004). (pages 13, 16).

Dupont, T. and R. Alley. *Assessment of the importance of ice-shelf buttressing to ice-sheet flow.* Geophys. Res. Lett., volume 32, no. 4; pages 1–4, FEB 25 2005. ISSN 0094-8276. doi:10.1029/2004GL022024. (page 7).

Durand, G., O. Gagliardini, T. Zwinger, E. Le Meur and R. Hindmarsh. *Full Stokes modeling of marine ice sheets: influence of the grid size.* Ann. Glaciol., volume 50(52); pages 109–114, 2009. (page 101).

Elachi, C. *Spaceborne radar remote sensing: Applications and techniques* (IEEE Press, 1988), 285 pages. (page 131).

Ettema, J., M. R. van den Broeke, E. van Meijgaard, W. J. van de Berg, J. L. Bamber, J. E. Box and R. C. Bales. *Higher surface mass balance of the Greenland ice sheet revealed by high-resolution climate modeling.* Geophys. Res. Lett., volume 36; pages 1–5, JUN 16 2009. ISSN 0094-8276. doi:10.1029/2009GL038110. (pages 6, 118, 123).

Fahnestock, M., W. Abdalati, I. Joughin, J. Brozena and P. Gogineni. *High geothermal heat row, basal melt, and the origin of rapid ice flow in central Greenland.* Science, volume 294, no. 5550; pages 2338–2342, DEC 14 2001. ISSN 0036-8075. (pages 118, 123).

Farinotti, D., M. Huss, A. Bauder, M. Funk and M. Truffer. *A method to estimate the ice volume and ice-thickness distribution of alpine glaciers.* J. Glaciol., volume 55, no. 191; pages 422–430, 2009. ISSN 0022-1430. (page 111, 111, 111).

Fastook, J., H. Brecher and T. Hughes. *Derived bedrock elevations, strain rates and stresses from measured surface elevations and velocities: Jakobshavns Isbrae, Greenland.* J. Glaciol., volume 41, no. 137; pages 161–173, 1995. ISSN 0022-1430. (page 111, 111).

Franca, L., G. Hauke and A. Masud. *Revisiting stabilized finite element methods for the advective-diffusive equation.* Comput. Methods Appl. Mech. Engrg., volume 195, no. 13-16; pages 1560–1572, 2006. ISSN 0045-7825. doi:10.1016/j.cma.2005.05.028. (page 148).

Frey, P. J. *Yams, a fully automatic adaptive isotropic surface remeshing procedure.* Technical Report RT-0252, INRIA, Rocquencourt, 11 2001.
URL `http://hal.inria.fr/inria-00069922/en/` (page 69).

Gagliardini, O., G. Durand, T. Zwinger, R. C. A. Hindmarsh and E. Le Meur. *Coupling of ice-shelf melting and buttressing is a key process in ice-sheets dynamics.* Geophys. Res. Lett., volume 37; pages 1–5, JUL 24 2010. ISSN 0094-8276. doi:10.1029/2010GL043334. (page 7).

Gillet-Chaulet, F., O. Gagliardini, J. Meyssonnier, M. Montagnat and O. Castelnau. *A user-friendly anisotropic flow law for ice-sheet modelling.* J. Glaciol., volume 51, no. 172; pages 3–14, 2005. ISSN 0022-1430. doi:{10.3189/172756505781829584}. (page 41).

Giovinetto, M., N. Waters and C. Bentley. *Dependence of Antarctic surface mass balance on temperature, elevation, and distance to open ocean.* J. Geophys. Res., volume 95, no. D4; pages 3517–3531, MAR 20 1990. ISSN 0148-0227. (page 96).

Glen, J. *The creep of polycrystalline ice.* Proc. R. Soc. A, volume 228, no. 1175; pages 519–538, 1955. (pages 3, 41).

Goldstein, R., H. Engelhardt, B. Kamb and R. Frolich. *Satellite radar interferometry for monitoring ice-sheet motion: Application to an antarctic ice stream.* Science, volume 262, no. 5139; pages 1525–1530, DEC 3 1993. ISSN 0036-8075. (pages 131, 134).

Goldstein, R., H. Zebker and C. Werner. *Satellite radar interferometry - two-dimensional phase unwrapping.* Radio Science, volume 23, no. 4; pages 713–720, JUL-AUG 1988. ISSN 0048-6604. (page 139).

Gropp, W., E. Lusk, N. Doss and A. Skjellum. *A high-performance, portable implementation of the MPI message passing interface standard.* Parallel Computing, volume 22, no. 6; pages 789–828, September 1996. (page 63).

Gropp, W. D. and E. Lusk. *User's Guide for* `mpich`*, a Portable Implementation of MPI.* Mathematics and Computer Science Division, Argonne National Laboratory, 1996. ANL-96/6. (page 63).

Habashi, W., J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia, M. Fortin and M. Vallet. *Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part I: general principles.* Int. J. Numer. Meth. Fluids, volume 32, no. 6; pages 725–744, MAR 30 2000. ISSN 0271-2091. (page 65).

Haber, E. and L. Hanson. *Model problems in PDE-constrained optimizatio*. Technical Report TR-2007-009, Emory Technical Report, 2007.
URL http://www.mathcs.emory.edu/~haber/pubs/ParamC.pdf (page 77).

Hadamard, J. *Sur les problèmes aux dérivées partielles et leur signification physique*. Princeton Univ. Bull., volume 13; pages 49–52, 1902. (page 21).

Hansen, P. C. *The L-curve and its use in the numerical treatment of inverse problems*. In *in Computational Inverse Problems in Electrocardiology, ed. P. Johnston, Advances in Computational Bioengineering*, (pages 119–142) (WIT Press, 2000). (page 30).

Hauret, P. *Méthodes Numériques pour la Dynamique des Structures Non-Linéaires Incompressibles à Deux Échelles*. Ph.D. thesis, Ecole Polytechnique, 2004.
URL http://www.imprimerie.polytechnique.fr/Theses/Files/Hauret.pdf (page 20).

Hecht, F. *BAMG: Bi-dimensional anisotropic mesh generator*. Technical report, FreeFem++, 2006.
URL http://www-rock-inria/gamma/cdrom/www/bamg/eng.htm (page 69).

Hegerl, G. C., F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. M. Orsini, N. Nicholls, J. E. Penner and P. A. Stott. *Understanding and attributing climate change. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* , chapter 9, (pages 665–745) (IPCC, 2007).
URL http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter9.pdf (page 4).

Heimbach, P. and V. Bugnion. *Greenland ice-sheet volume sensitivity to basal, surface and initial conditions derived from an adjoint model*. Ann. Glaciol., volume 50, no. 52; pages 67–80, 2009. ISSN 0260-3055. (pages 26, 77).

Hindmarsh, R. *A numerical comparison of approximations to the Stokes equations used in ice sheet and glacier modeling*. J. Geophys. Res., volume 109, no. F1; pages 1–15, MAR 10 2004. ISSN 0148-0227. doi:10.1029/2003JF000065. (page 39).

Holland, D. and A. Jenkins. *Modeling thermodynamic ice-ocean interactions at the base of an ice shelf*. J. Phys. Oceanogr., volume 29, no. 8, Part 1; pages 1787–1800, AUG 1999. ISSN 0022-3670. (page 53, 53, 53).

Hooke, R. L. *Principles of Glacier Mechanics* (Cambridge, 2005), 2nd edition, 763 pages. (pages 35, 40, 41, 41, 51, 52, 52).

Hughes, T., W. Liu and T. Zimmermann. *Lagrangian-eulerian finite-element formulation for incompressible viscous flows*. Comput. Methods Appl. Mech. Engrg., volume 29, no. 3; pages 329–349, 1981. ISSN 0045-7825. (page 61).

Hulbe, C. and D. MacAyeal. *A new numerical model of coupled inland ice sheet, ice stream, and ice shelf flow and its application to the West Antarctic Ice Sheet*. J. Geophys. Res., volume 104, no. B11; pages 25 349–25 366, NOV 10 1999. ISSN 0148-0227. (pages 52, 53).

Hutter, K. *Theoretical glaciology: material science of ice and the mechanics of glaciers and ice sheets* (D. Reidel Publishing Co, Dordrecht, The Netherlands, 1983), 150 pages. (pages 39, 49, 95).

Huybrechts, P., J. Gregory, I. Janssens and M. Wild. *Modelling Antarctic and Greenland volume changes during the 20th and 21st centuries forced by GCM time slice integrations.* Glob. Planet. Change, volume 42, no. 1-4; pages 83–105, JUL 2003. ISSN 0921-8181. doi:10.1016/j.gloplacha.2003.11.011. (page 109).

IPCC-AR3. *Fourth Assessment Report: Climate Change 2001: The AR3 Synthesis Report*, volume Working Group I, Summary for Policymakers (Geneva: IPCC, 2001).
URL http://www.grida.no/publications/other/ipcc_tar/ (page 7).

IPCC-AR4. *Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report* (Geneva: IPCC, 2007).
URL http://www.ipcc.ch/ipccreports/ar4-wg1.htm (pages 4, 8).

Irgens, F. *Continuum Mechanics* (Springer Berlin Heidelberg, 2008). (page 17).

Jay-Allemand, M., F. Gillet-Chaulet, O. Gagliardini and M. Nodet. *Investigating changes in basal conditions of Variegated Glacier prior to and during its 1982-1983 surge.* The Cryosphere, volume 5; pages 659–672, 2011. doi:10.5194/tc-5-659-2011. (page 30).

Joughin, I., W. Abdalati and M. Fahnestock. *Large fluctuations in speed on Greenland's Jakobshavn Isbrae glacier.* Nature, volume 432, no. 7017; pages 608–610, DEC 2 2004a. ISSN 0028-0836. doi:{10.1038/nature03130}. (page 123).

Joughin, I. and R. B. Alley. *Stability of the West Antarctic ice sheet in a warming world.* Nature Geosci., volume 4, no. 8; pages 506–513, AUG 2011. ISSN 1752-0894. doi:10.1038/NGEO1194. (page 7).

Joughin, I., D. MacAyeal and S. Tulaczyk. *Basal shear stress of the Ross ice streams from control method inversions.* J. Geophys. Res., volume 109, no. B9; pages 1–62, SEP 24 2004b. ISSN 0148-0227. doi:10.1029/2003JB002960. (page 75).

Joughin, I., B. Smith, I. Howat, T. Scambos and T. Moon. *Greenland flow variability from ice-sheet-wide velocity mapping.* J. Glaciol., volume 56; pages 416–430, 2010. (pages xvii, 109, 123, 124).

Joughin, I., S. Tulaczyk, J. Bamber, D. Blankenship, J. Holt, T. Scambos and D. Vaughan. *Basal conditions for Pine Island and Twaites Glaciers, West Antarctica, determined using satellite and airborne data.* J. Glaciol., volume 55, no. 190; pages 245–257, MAR 2009. ISSN 0148-0227. (page 98).

Karush, W. *Minima of Functions of Several Variables with Inequalities as Side Constraints.* Master's thesis, Department of Mathematics, University of Chicago, 1939. (page 26).

Karypis, G. and V. Kumar. *A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices* (University of Minnesota, 1998). (page 63).

Kernighan, B. W. *The C Programming Language* (Prentice Hall Professional Technical Reference, 1988), 2nd edition. ISBN 0131103709. (page 62).

Khazendar, A., E. Rignot and E. Larour. *Larsen B Ice Shelf rheology preceding its disintegration inferred by a control method.* Geophys. Res. Lett., volume 34, no. 19; pages 1–6, OCT 4 2007. ISSN 0094-8276. doi:10.1029/2007GL030980. (page 75).

Khazendar, A., E. Rignot and E. Larour. *Roles of marine ice, rheology, and fracture in the flow and stability of the Brunt/Stancomb-Wills Ice Shelf.* J. Geophys. Res., volume 114; pages 1–9, OCT 27 2009. ISSN 0148-0227. doi:{10.1029/2008JF001124}. (page 75).

Khazendar, A., E. Rignot and E. Larour. *Acceleration and spatial rheology of Larsen C Ice Shelf, Antarctic Peninsula.* Geophys. Res. Lett., volume 38, L09502; pages 1–6, MAY 13 2011. ISSN 0094-8276. doi:10.1029/2011GL046775. (page 75).

Kuhn, H. W. and A. W. Tucker. *Nonlinear programming.* In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, 1950*, (pages 481–492) (University of California Press, Berkeley, California, 1951). Reprinted in: *Readings in Mathematical Economics, Vol 1, Value Theory*, (P. Newman, ed.), The Johns Hopkins University Press, Baltimore, 1968, pp. 3–14. (page 26).

Kuzmin, D. *A guide to numerical methods for transport equations*, 2010. Univ. Nürnberg. URL http://www.mathematik.uni-dortmund.de/~kuzmin/cfdbook.html (pages 146, 147).

Larour, E. *Modélisation numérique du comportement des banquises flottantes, validée par imagerie satellitaire.* Ph.D. thesis, Ecole Centrale Paris, JUN 2005. (page 75).

Larour, E., H. Seroussi, M. Morlighem and E. Rignot. *Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM).* J. Geophys. Res., volume in press, 2012. doi:10.1029/2011JF002140. (page 62).

Lemke, P., J. Ren, R. Alley, I. Allison, J. Carrasco, G. Flato, Y. Fuji, G. Kaser, P. Mote, R. Thomas and T. Zhan. *Observations: Changes in Snow, Ice and Frozen Ground*, chapter 4 in IPCC 2007: Climate Change 2007: The Physical Science Basis, (page 48) (Cambridge U. Press, 2007).
URL http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter4.pdf (pages xv, xix, 5, 6).

Liu, I.-S. *Continuum Mechanics* (Springer- Verlag, 2002). (page 17).

Ma, Y., O. Gagliardini, C. Ritz, F. Gillet-Chaulet, G. Durand and M. Montagnat. *Enhancement factors for grounded ice and ice shelves inferred from an anisotropic ice-flow model.* J. Glaciol., volume 56; pages 805–812, 2010. (page 41).

MacAyeal, D. *Large-scale ice flow over a viscous basal sediment - Theory and application to Ice Stream-B, Antarctica.* J. Geophys. Res., volume 94, no. B4; pages 4071–4087, APR 10 1989. ISSN 0148-0227. (pages 39, 47).

MacAyeal, D. *The basal stress distribution of Ice Stream E, Antarctica, Inferred by Control Methods.* J. Geophys. Res., volume 97, no. B1; pages 595–603, JAN 10 1992. ISSN 0148-0227. (page 75, 75).

MacAyeal, D. *A tutorial on the use of control methods in ice-sheet modeling.* J. Glaciol., volume 39, no. 131; pages 91–98, 1993. ISSN 0022-1430. (pages 75, 75, 91).

Mathworks, T. *MATLAB (2008a, The Mathworks, Natick, MA, UNITED STATES)*, 2008. (page 62).

Maule, C. F., M. E. Purucker, N. Olsen and K. Mosegaard. *Heat flux anomalies in Antarctica revealed by satellite magnetic data.* Science, volume 309, no. 5733; pages 464–467, 2005. (page 96).

Morland, L. *Unconfined ice shelf flow.* Proceedings of Workshop on the Dynamics of the West Antarctic Ice Sheet, University of Utrecht, May 1985. Published by Reidel, volume ed. C.J. van der Veen and J. Oerlemans; pages 99–116, 1987. (pages 39, 47).

Morlighem, M., E. Rignot, H. Seroussi, E. Larour, H. Ben Dhia and D. Aubry. *Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica.* Geophys. Res. Lett., volume 37, L14502; pages 1–6, JUL 2010. doi:10.1029/2010GL043853. (pages 53, 95).

Morlighem, M., E. Rignot, H. Seroussi, E. Larour, H. Ben Dhia and D. Aubry. *A mass conservation approach for mapping glacier ice thickness.* Geophys. Res. Lett., volume 38, L19503; pages 1–6, 2011. doi:10.1029/2011GL048659. (page 118).

Morozov, V. A. *Methods for Solving Incorrectly Posed Problems* (Springer-Verlag, New York, 1984). (page 29).

Nastran, M. *Nastran theoretical volume.* Technical report, MSC, 1972. (pages 8, 62).

Nick, F. M., A. Vieli, I. M. Howat and I. Joughin. *Large-scale changes in greenland outlet glacier dynamics triggered at the terminus.* Nature Geosci., volume 2, no. 2; pages 110–114, FEB 2009. ISSN 1752-0894. doi:10.1038/NGEO394. (page 7).

Nowicki, S. M. J. and D. J. Wingham. *Conditions for a steady ice sheet-ice shelf junction.* Earth Planet. Sci. Lett., volume 265, no. 1-2; pages 246–255, JAN 15 2008. ISSN 0012-821X. (page 101).

Nye, J. *The distribution of stress and velocity in glaciers and ice-sheets.* Proc. R. Soc. A, volume 239, no. 1216; pages 113–133, 1957a. (page 3).

Nye, J. *Physical properties of crystals: Their representation by tensor and matrices* (Oxford University Press, 1957b), first edition. ISBN 9780198511656. (pages 3, 41).

Nye, J. F. *The flow law of ice from measurements in glacier tunnels, laboratory experiments and the Jungfraufirn borehole experiment.* Proc. R. Soc. A, volume 219, no. 1193; pages 477–489, 1953.
Comparison of Glen 1952 compression tests, Gerrard 1952 vertical borehole on Jungfraufirn, and behavior of cylindrical tunnel closure. Equations for 3D slow flow of ice mass: general flow law. (page 3).

Paterson, W. *The Physics of Glaciers* (Pergamon Press, Oxford, London, New York, 1994), 3rd edition. (pages xix, 3, 35, 40, 41, 43, 52, 53, 53, 75, 76).

Pattyn, F. *A new three-dimensional higher-order thermomechanical ice sheet model: Basic sensitivity, ice stream development, and ice flow across subglacial lakes.* J. Geophys. Res., volume 108, no. B8; pages 1–15, AUG 16 2003. ISSN 0148-0227. doi:10.1029/2002JB002329. (pages 39, 44, 53, 53).

Pfeffer, W. T., J. T. Harper and S. O'Neel. *Kinematic constraints on glacier contributions to 21st-century sea-level rise.* Science, volume 321, no. 5894; pages 1340–1343, SEP 5 2008. ISSN 0036-8075. doi:{10.1126/science.1159099}. (page 6).

Phillips, T., H. Rajaram and K. Steffen. *Cryo-hydrologic warming: A potential mechanism for rapid thermal response of ice sheets.* Geophys. Res. Lett., volume 7; pages 1–5, OCT 23 2010. ISSN 0094-8276. doi:10.1029/2010GL044397. (page 7).

Pollard, D. and R. DeConto. *Modelling West Antarctica ice sheet growth and collapse through the past five million years.* Nature, volume Letters 458; pages 329–332, 2009. URL http://www.nature.com/nature/journal/v458/n7236/abs/nature07809.html (page 109).

Puel, G. and D. Aubry. *Using mesh adaption for the identification of a spatial field of material properties.* International Journal for Numerical Methods in Engineering, volume Online library; page 23, 2011. ISSN 1097-0207. doi:10.1002/nme.3170. URL http://dx.doi.org/10.1002/nme.3170 (page 29).

Rajagopal, K. R. and G. Saccomandi. *On internal constraints in continuum mechanics.* Differ Equat Nonlinear Mech, volume 2006; 18572, 2006. doi:10.1155/DENM/2006/18572. URL http://www.hindawi.com/journals/ijde/2006/018572/abs/ (page 20).

Rasmussen, L. *Bed topography and mass-balance distribution of Columbia Glacier, Alaska, USA, determined from dequential serial-photography.* J. Glaciol., volume 34, no. 117; pages 208–216, 1988. ISSN 0022-1430. (page 111).

Raymond, M. J. and G. H. Gudmundsson. *Estimating basal properties of glaciers from surface measurements: a non-linear bayesian inversion approach.* The Cryosphere Discussions, volume 3, no. 1; pages 181–222, 2009. ISSN 1994-0432. URL http://www.the-cryosphere-discuss.net/3/181/2009/ (page 76).

Reed, W. H. and T. R. Hill. *Triangular mesh methods for the neutron transport equation.* Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, USA, 1973. (page 148).

Reist, A. *Mathematical analysis and numerical simulation of the motion of a glacier.* Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, 2005. (page 40).

Rignot, E. *Changes in ice dynamics and mass balance of the Antarctic ice sheet.* Phil. Trans R. Soc. A, volume 364, no. 1844; pages 1637–1655, JUL 15 2006. ISSN 1364-503X. doi:10.1098/rsta.2006.1793. Royal-Society Discussion Meeting on Evolution of the antarctic Ice Sheet, London, England, Oct 17-18, 2005. (page 7).

Rignot, E. *Changes in West Antarctic ice stream dynamics observed with ALOS PALSAR data.* Geophys. Res. Lett., volume 35, no. 12; pages 1–5, JUN 28 2008a. ISSN 0094-8276. doi:10.1029/2008GL033365. (pages xvi, 7, 95, 95, 100, 101).

Rignot, E. *PALSAR studies of ice sheet motion in Antarctica.* In *ALOS PI Symposium, Nov. 3-7 2008* (2008b). (pages xvi, 71).

Rignot, E., J. Bamber, M. van den Broeke, C. Davis, Y. Li, W. van de Berg and E. van Meijgaard. *Recent Antarctic ice mass loss from radar interferometry and regional climate modelling.* Nature Geosci., volume 1(2); pages 106–110, 2008. (pages 5, 6, 7).

Rignot, E., G. Casassa, P. Gogineni, W. Krabill, A. Rivera and R. Thomas. *Accelerated ice discharge from the Antarctic Peninsula following the collapse of Larsen B ice shelf.* Geophys. Res. Lett., volume 31, no. 18; pages 1–4, SEP 22 2004. ISSN 0094-8276. doi:10.1029/2004GL020697. (page 7, 7).

Rignot, E., S. Gogineni, W. Krabill and S. Ekholm. *North and northeast Greenland ice discharge from satellite radar interferometry.* Science, volume 276, no. 5314; pages 934–937, MAY 9 1997. ISSN 0036-8075. (page 118).

Rignot, E., M. Koppes and I. Velicogna. *Rapid submarine melting of the calving faces of west greenland glaciers.* Nature Geosci., volume 3, no. 3; pages 187–191, MAR 2010. ISSN 1752-0894. doi:10.1038/NGEO765. (page 6).

Rignot, E., J. Mouginot and B. Scheuchl. *Ice Flow of the Antarctic Ice Sheet.* Science, volume 333, no. 6048; pages 1427–1430, SEP 9 2011. ISSN 0036-8075. doi:10.1126/science.1208336. (pages xv, 4, 8).

Rignot, E. and K. Steffen. *Channelized bottom melting and stability of floating ice shelves.* Geophys. Res. Lett., volume 35, no. 2; pages 1–5, Jan 17 2008. ISSN 0094-8276. doi: 10.1029/2007GL031765. (page 122).

Rignot, E. and R. Thomas. *Mass balance of polar ice sheets.* Science, volume 297, no. 5586; pages 1502–1506, AUG 30 2002. ISSN 0036-8075. (page 7).

Rignot, E., D. Vaughan, M. Schmeltz, T. Dupont and D. MacAyeal. *Acceleration of Pine Island and Thwaites Glaciers, West Antarctica.* Ann. Glaciol., volume 34, no. Winther, JG and Solberg, R; pages 189–194, 2002. (pages xvi, 95, 96).

Rommelaere, V. and D. MacAyeal. *Large-scale rheology of the Ross Ice Shelf, Antarctica, computed by a control method.* Ann. Glaciol., volume 24; pages 43–48, 1997. (page 75).

Rosen, P. A., S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, S. Member, E. Rodríguez and R. M. Goldstein. *Synthetic aperture radar interferometry.* In *Proceedings of the IEEE*, (pages 333–382) (2000). (page 131).

Rott, H., W. Rack, P. Skvarca and H. De Angelis. *Northern Larsen Ice Shelf, Antarctica: further retreat after collapse.* In *Ann. Glaciol*, volume 34 of *Ann. Glaciol.*, (pages 277–282) (2002). ISBN 0-946417-29-6. ISSN 0260-3055. doi:10.3189/172756402781817716. (page 7).

Rumsey, H., G. Morris, R. Green and R. Goldstein. *Radar brightness and altitude image of a portion of venus.* Icarus, volume 23, no. 1; pages 1–7, 1974. ISSN 0019-1035. (page 132).

Schoof, C. *Ice-sheet acceleration driven by melt supply variability.* Nature, volume 468, no. 7325; pages 803–806, DEC 9 2010. ISSN 0028-0836. doi:10.1038/nature09618. (page 7).

Seroussi, H., M. Morlighem, E. Rignot, E. Larour, D. Aubry, H. Ben Dhia and S. S. Kristensen. *Ice flux divergence anomalies on 79north Glacier, Greenland*. Geophys. Res. Lett., volume 38, L09501; doi:10.1029/2011GL047338, 2011. (pages 8, 105, 107).

Stroustrup, B. *The C++ Programming Language* (Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997), 3rd edition. ISBN 0201889544. (page 62).

Tarantola, A. *Inverse Problem Theory* (SIAM, 2005). (pages 21, 29).

Thomas, R., E. Frederick, W. Krabill, S. Manizade and C. Martin. *Progressive increase in ice loss from Greenland*. Geophys. Res. Lett., volume 33, no. 10; pages 1–4, MAY 27 2006. ISSN 0094-8276. doi:{10.1029/2006GL026075}. (page 118).

Thomas, R., E. Frederick, W. Krabill, S. Manizade and C. Martin. *Recent changes on Greenland outlet glaciers*. J. Glaciol., volume 55, no. 189; pages 147–162, 2009. ISSN 0022-1430. (page 123).

Thomas, R., E. Frederick, J. Li, W. Krabill, S. Manizade, J. Paden, J. Sonntag, R. Swift and J. Yungel. *Accelerating ice loss from the fastest Greenland and Antarctic glaciers*. Geophys. Res. Lett., volume 38, L10502; pages 1–6, 2011. (page 7).

Thomas, R., E. Rignot, P. Kanagaratnam, W. Krabill and G. Casassa. *Force-perturbation analysis of Pine Island Glacier, Antarctica, suggests cause for recent acceleration*. In J. Jacka, editor, *Ann. Glaciol., Vol 39, 2005*, volume 39, (pages 133–138) (Int Glaciological Soc, 2004). ISBN 0-946417-34-2. (pages xvi, 100, 101).

Thomsen, H., N. Reeh, O. Olesen, C. Boggild, W. Starzer, A. Weidick and A. Higgins. *The Nioghalvfjerdsfjorden glacier project, north-east Greenland: a study of ice sheet response to climatic change*. Geol. Surv. Greenland Bull., volume 179; pages 95–103, 1997. (page 118).

Thorsteinsson, T. *An analytical approach to deformation of anisotropic ice-crystal aggregates*. J. Glaciol., volume 47, no. 158; pages 507–516, 2001. ISSN 0022-1430. (page 41).

Tikhonov, A. N. *On the stability of inverse problems*. C. R. (Doklady) Acad. Sci. SSSR (N.S.), volume 39; pages 176–179, 1943. (page 22).

Vaughan, D., H. Corr, F. Ferraccioli, N. Frearson, A. O'Hare, D. Mach, J. Holt, D. Blankenship, D. Morse and D. Young. *New boundary conditions for the West Antarctic Ice Sheet: Subglacial topography beneath Pine Island Glacier*. Geophys. Res. Lett., volume 33, no. 9; pages 1–4, MAY 4 2006. ISSN 0094-8276. doi:10.1029/2005GL025588. (page 96).

van der Veen, C. J. and I. M. Whillans. *Force budget: I. theory and numerical methods*. J. Glaciol., volume 35; pages 53–60, 1989. (page 45, 45).

Vieli, A. and A. Payne. *Application of control methods for modelling the flow of Pine Island Glacier, West Antarctica*. Ann. Glaciol., volume 36; pages 197–204, 2003. ISSN 0260-3055. (page 98).

Vogel, C. R. *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002). ISBN 0898715075. (page 21).

Walker, J. and E. Waddington. *Early discoverers XXXV Descent of glaciers: some early speculations on glacier flow and ice physics.* J. Glaciol., volume 34, no. 118; pages 342–348, 1988. (pages 1, 2, 3).

Warner, R. and W. Budd. *Derivation of ice thickness and bedrock topography in data-gap regions over Antarctica.* Ann. Glaciol., volume 31; pages 191–197, 2000. (page 111, 111).

Weertman, J. *On the sliding of glaciers.* J. Glaciol., volume 3; pages 33–38, March 1957. (page 3).

Wikipedia. *the free encyclopedia*, 2011.
URL http://en.wikipedia.org/ (pages xv, 2).

Wunsch, C. *Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications* (Cambridge University Press, 2006), 384 pages. ISBN 0521854245. (pages 21, 26, 77).

Zebker, H., P. Rosen, R. Goldstein, A. Gabriel and C. Werner. *On the derivation of coseismic displacement fields usind differential radar interferometry: The landers earthquake.* J. Geophys. Res. -Solid Earth, volume 99, no. B10; pages 19 617–19 634, OCT 10 1994. ISSN 0148-0227. (page 131).

Zienkiewicz, O. C. and R. L. Taylor. *The finite element method*, volume 1 (New York, London, 1989), fourth edition. (page 58).

Zisk, S. *Lunar topography - first radar-interferometer measurements of alphonsus-ptolemaeus-arzachel region.* Science, volume 178, no. 4064; page 977, 1972. ISSN 0036-8075. (page 132).

Zwally, H. J., W. Abdalati, T. Herring, K. Larson, J. Saba and K. Steffen. *Surface melt-induced acceleration of Greenland ice-sheet flow.* Science, volume 297, no. 5579; pages 218–222, 2002. doi:10.1126/science.1072708.
URL http://www.sciencemag.org/content/297/5579/218.abstract (page 6).