From local to global magnetospheric hybrid-Vlasov simulations

Yann Pfau-Kempf

Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

Academic dissertation in theoretical physics

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in auditorium E204 at Physicum on Kumpula campus (Gustaf Hällströmin katu 2a) on December 1st, 2016 at noon (12 o’clock).

Finnish Meteorological Institute
Helsinki, 2016
Supervising professor
Professor Hannu E. J. Koskinen, University of Helsinki, Finland

Thesis advisers
Research Professor Minna Palmroth, Finnish Meteorological Institute, Finland
Dr. Dimitry Pokhotelov, now at Mullard Space Science Laboratory, UK
Dr. Arto Sandroos, now at Cadence Design Systems, San Jose, CA 95134, USA
Dr. Sebastian von Alftan, now at CSC – IT Center for Science, Finland

Pre-examiners
Professor Mats André, Uppsala University, Sweden
Professor James F. Drake, University of Maryland, USA

Opponent
Dr. Benoît Lavraud, Institut de Recherche en Astrophysique et Planétologie, Toulouse, France

ISBN 978-952-336-000-6 (paperback)
ISSN 0782-6117

Erweko
Helsinki, 2016
In memory of Daniel Oberländer and Anita Pfau.
Vlasiator – From local to global magnetospheric hybrid-Vlasov simulations

Abstract
The Sun is the source of the solar wind, a continuous stream of electrically charged particles and magnetic fields pervading the Solar system. Its interaction with the magnetic field of the Earth, in and around the region called the magnetosphere, controls the flow of matter and energy in near-Earth space. A fundamental understanding of the physical processes at play is crucial for the building of forecasting and warning systems, as the influence of the solar wind during space storms can harm life and technology in space and on the ground. These effects, collectively known as space weather, are one of the biggest albeit least understood natural threats to society. The research effort needed includes the development of observational methods as well as theories and models, to first describe and later predict the mechanisms and consequences of space weather.

This doctoral thesis, comprising an introduction and four peer-reviewed articles, presents the hybrid-Vlasov model Vlasiator developed at the Finnish Meteorological Institute. Based on a detailed description of proton physics in space plasmas, Vlasiator allows to simulate both local contexts and the Earth’s magnetosphere on global scales. This unprecedented capability is only accessible by harnessing the power of modern supercomputers. The aim of this work is threefold. The current version of Vlasiator is documented considering physical and computational aspects, the correctness of the simulations is demonstrated by comparing to analytical theories and spacecraft observations, and new scientific results gained with this model are presented.

Vlasiator – à l’approche de l’échelle globale dans la modélisation hybride-Vlasov de la magnétosphère

Le Soleil est la source du vent solaire, un flux continu de particules chargées électriquement et de champs magnétiques se propageant dans le Système solaire. Son interaction avec le champ magnétique de la Terre, dans et autour de la région de la magnétosphère, contrôle les courants de matière et d’énergie dans l’espce proche de la Terre. Une compréhension fondamentale des processus physiques en jeu est cruciale pour la construction de systèmes de prédiction et d’alerte, étant donné que l’influence du vent solaire lors de tempêtes spatiales peut nuire à la vie et aux technologies dans l’espace et au sol. Ces effets, collectivement définis comme la météorologie spatiale, sont une des plus grandes menaces naturelles pour la société tout en étant l’une des moins bien comprises. L’effort de recherche nécessaire inclut le développement de méthodes d’observation ainsi que de théories et de modèles pour permettre de comprendre puis de prédire les mécanismes et les conséquences de la météorologie spatiale.

Cette thèse doctorale, qui se compose d’une introduction et de quatre articles évalués par des pairs, présente le modèle hybride-Vlasov Vlasiator développé à l’Institut Météorologique de Finlande. Fondé sur une description détaillée de la physique des protons dans les plasmas spatiaux, Vlasiator permet de simuler à la fois des contextes locaux et la magnétosphère de la Terre à l’échelle globale. Cette possibilité sans précédent est uniquement accessible en exploitant la puissance des supercalculateurs modernes. L’objectif de ce travail est triple. La version actuelle de Vlasiator est documentée en considérant les aspects physiques et numériques, la justesse des simulations est démontrée par la comparaison avec des théories analytiques et des observations spatiales, et de nouveaux résultats scientifiques obtenus avec ce modèle sont présentés.
Vlasiator – von lokalen zu globalen hybrid-Vlasov Simulationen der Magnetosphäre

Titel

Kurz zusammenfassung


## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstracts</td>
<td>v</td>
</tr>
<tr>
<td>Preface</td>
<td>xi</td>
</tr>
<tr>
<td>Publications</td>
<td>xiii</td>
</tr>
<tr>
<td>Variables, symbols and acronyms</td>
<td>xv</td>
</tr>
</tbody>
</table>

1 Introduction

1.1 What is plasma? .......................... 1
1.2 The solar wind, the Earth’s magnetosphere and why they matter 2
1.3 Parallel computing ...................... 6
1.4 The role of simulation in space physics ...................... 8

2 Plasma phenomena .......................... 11

2.1 Characteristic length and time scales ...................... 11
2.2 Waves and instabilities ..................... 13
2.3 Shocks ..................................... 16
2.4 Magnetic reconnection ..................... 18

3 Space plasma simulation ................... 21

3.1 Particles in cells, a rough picture ...................... 22
3.2 Particle distributions in higher dimensions ...................... 24
3.3 Magnetohydrodynamics ...................... 26
3.4 Hybrids and other combinations ...................... 29

4 Vlasiator .................................. 33

4.1 The hybrid-Vlasov model ...................... 33
4.2 Discretising a six-dimensional space ...................... 35
4.3 The field solver ...................... 38
4.4 From six to seven dimensions: time stepping ...................... 40
4.5 Making kinetic global magnetospheric simulations possible ...................... 41
5 Major code contributions
  5.1 Second-order field solver time stepping .................................. 45
  5.2 The Hall term in Ohm’s law ......................................................... 46
  5.3 System boundary conditions ....................................................... 48

6 Verification, validation and new scientific results 51
  6.1 Wave dispersion .......................................................................... 51
  6.2 Ion beam instabilities ................................................................. 52
  6.3 Ion foreshock populations ......................................................... 53
  6.4 Local foreshocks ....................................................................... 54

7 Conclusions and outlook 59

Bibliography 63

Paper I Wave dispersion in the hybrid-Vlasov model: Verification of Vlasiator 77

Paper II Vlasiator: First global hybrid-Vlasov simulations of Earth’s foreshock and magnetosheath 85

Paper III Ion distributions in the Earth’s foreshock: Hybrid-Vlasov simulation and THEMIS observations 99

Paper IV Evidence for transient, local ion foreshocks caused by dayside magnetopause reconnection 119
Preface

Someone: ‘So, what do you do?’
Me (Option 1): ‘I am a theoretical physicist.’ *End of the discussion.*
Me (Option 2): ‘I do supercomputer simulations of space plasmas.’
Someone: ‘Oh that sounds cool, tell me more!’

This work should be the ultimate answer to the question what I have been doing in the last four years. I would like to take this opportunity to apologise to everyone who has had to endure my impromptu lectures on plasma physics, space weather or high-performance computing, be it on the plane, at the climbing wall, in a bar or in your living room. I also wish to express to you my gratitude, as I learn something every time I am asked to explain my research.

I have been working since 2011 at the Finnish Meteorological Institute (FMI), initially as a summer student, then as a master’s student and since 2012 as a doctoral student. My first thanks go to Prof. Hannu Koskinen, who originally made sure I was accepted by the University of Helsinki, then hooked me with his plasma physics lectures and later transmitted my application to the Vlasiator project for a summer job. My thesis work was done under the supervision of the Vlasiator team leader Res. prof. Minna Palmroth and in close collaboration with Dr. Dimitry Pokhotelov, Dr. Sebastian von Alftan and Dr. Arto Sandroos, all four of whom formed my team of close-range advisers (*lähibiohjaajat*). Many thanks for your guidance when it was needed and your trust that I would do what was necessary when given some academic freedom.

My thanks also go to the extended Vlasiator team, including Dr. Urs Ganse, Dr. Liisa Juusola and Dr. Riku Järvinen at FMI, Dr. Heli Hietala at the University of California in Los Angeles, Prof. Rami Vainio at the University of Turku and Dr. Markus Battarbee at the University of Central Lancashire. Furthermore, I will keep fond memories of our office 2B21 even after the moves to come. Dr. Chandrasekhar Anekallu, Dr. Ilja Honkonen, Dr. Gábor Facskó, Dr. Dimitry Pokhotelov, Otto Hannuksela and, last but not least and for the longest time, Sanni Hoilijoki all had to put up over the years with my swearing at my screen in various languages and with my sprawling green plants. Thanks
for your patience and the good moments spent in this office or further afield.

Although the cover of this book bears my name alone, it is the sum of work done with numerous colleagues. Thanks to my co-authors mentioned above and Dr Andris Vaivads, Dr Olga Gutynska, Dr Lynn Wilson, Dr Brian Walsh, Dr David Sibeck and Prof. Steve Milán. I am indebted to the pre-examiners of this thesis, Prof. Mats André and Prof. James Drake, who despite the admittedly short notice took the time to read and comment thoroughly my text, which significantly improved as a consequence. I would like to thank in advance Dr Benoît Lavraud, who accepted to travel to Finland during the possibly worst season and be the opponent of this thesis. I am honoured that these three great scientists agreed to examine my work.

All of this in Finland would not have been possible without the support, confidence and love of my family and friends, even when we flew to an unknown country without confirmed accommodation and, as it turned out, without a confirmed study place either. Thanks to all of you on either side of the Baltic, and first and foremost to my wife Kathrin and our child, who still enjoys warm protection from the snow as of this writing, but will attend the public defence, continuing a tradition I started some 26 and a half years ago.

My work at FMI was funded by several projects of the Academy of Finland as well as the European Research Council starting and consolidator grants QuESpace and PRESTISSIMO awarded to Minna Palmroth. I also acknowledge financial support from the Magnus Ehrnrooth foundation, the Finnish national Doctoral Programme in Astronomy and Space Physics, the University of Helsinki Chancellor’s travel grant and the University of Helsinki Doctoral programme in Particle Physics and Universe Sciences. With their help I was able to attend several schools and conferences and to visit colleagues, which allowed me to present my work and get to know the community.

No animals were intentionally harmed during the writing of this thesis. However, it gave the opportunity to explore cafés and bars in central Helsinki. Some 18 pieces of cake, one piece of quiche, one crème brûlée, 2 ice creams and one bagel, as well as 8 cups of coffee, 3 iced teas, one hot chocolate, 4 beers, one Picon-beer and one cidre were consumed at Café DaJa, Café Panik, Kakkugalleria, Brooklyn Café, Il Birrificio, Apéro, Löyly and Kaffeehaus Röntgen.

Git statistics show that in over 130 commits, the \LaTeX{} source grew by over 2400 lines and shrunk by almost 300 lines. Dear reader, you have been saved over 10% of material. I hope you won’t regret opening this book or file, it has entertaining stuff for the layperson and the expert alike. I wish you an enjoyable read!

Yann Pfau-Kempf
Helsinki, November 2016
Publications

This thesis consists of an introduction and four papers. They have not been included in prior theses and are published in international peer-reviewed journals. The topic and the contribution of the author of this thesis are summarised here for each paper.

Paper I. Wave dispersion in the hybrid-Vlasov model: Verification of Vlasiator


The dispersion of low-$\beta$ plasma wave modes is simulated with Vlasiator and compared to the solution computed by the Waves in Homogeneous, Anisotropic Magnetized Plasmas (WHAMP) code [Rönnmark, 1982, WHAMP, 2016].

The author developed the additional parts of the Vlasiator code needed to perform the simulations of this paper and the analysis tools to process the simulation data as well as the WHAMP output. The author wrote the paper under the guidance of the co-authors.

Paper II. Vlasiator: First global hybrid-Vlasov simulations of Earth’s foreshock and magnetosheath


This technical paper presents the algorithms employed in Vlasiator and gives numerical and physical results of local and global simulations.

The author ran the simulation and performed the analysis presented in Section 4 of Paper II. The author took part in preparing, performing and analysing the simulation presented in Section 5 of the paper. The author wrote Section 4 and contributed to the writing of the rest of the paper as well.
Paper III. Ion distributions in the Earth’s foreshock: Hybrid-Vlasov simulation and THEMIS observations


This paper compares the ion velocity distributions obtained in the terrestrial foreshock of global magnetospheric simulations to the classic types of velocity distributions measured by the THEMIS spacecraft. The mechanism leading to the loss of gyrotropy of the ions interacting with the foreshock waves and the limitations of the hybrid-Vlasov approach are also discussed.

The author performed the study with the help of the co-authors, took part in the preparation and performance of the simulation presented and performed its analysis with the help of the co-authors. The author wrote all parts of the paper directly pertaining to modelling.

Paper IV. Evidence for transient, local ion foreshocks caused by dayside magnetopause reconnection


A scenario is introduced by which dayside reconnection generates magnetosheath and subsequently bow shock perturbations, which in turn lead to the generation of local foreshock-like ion beams reflected off the bow shock and travelling with the bow shock perturbation. Ground-based magnetometer and ionospheric radar as well as spacecraft observations support the scenario.

The author took part in the preparation and performance of the simulation presented and performed its analysis. The author performed the spacecraft data analysis with help of the co-authors and wrote the paper, with the exception of the sections on ground-based magnetometer and ionospheric radar measurements.
Variables, symbols and acronyms

This thesis uses the International System of units (SI units). The subscripts e and i stand for electrons and ions respectively, where applicable.

Variables

\[ \mathbf{r} = (x, y, z) \quad \text{Position vector and components/coordinates} \]
\[ \mathbf{v} = (v_x, v_y, v_z) \quad \text{Velocity vector and components/coordinates} \]
\[ \mathbf{a} = (a_x, a_y, a_z) \quad \text{Acceleration vector and components/coordinates} \]
\[ t, \omega \quad \text{Time, angular frequency} \]
\[ \mathbf{k}, \mathbf{k}^\prime \quad \text{Wave vector, wave number (= 2\pi times inverse wavelength)} \]
\[ \mathbf{B}, \mathbf{B} \quad \text{Magnetic field and its magnitude} \]
\[ \mathbf{E}, \mathbf{E}, \phi \quad \text{Electric field and its magnitude, electric potential} \]
\[ \mathbf{V}, V \quad \text{Bulk velocity of plasma and its magnitude (speed)} \]
\[ \mathbf{j}, j \quad \text{Current density and its magnitude} \]
\[ \mathbf{P}, P \quad \text{Pressure tensor, scalar pressure} \]
\[ \mathbf{T}, T \quad \text{Temperature tensor, scalar temperature} \]
\[ n, \rho_m, \rho_q \quad \text{Number density, mass density, charge density} \]
\[ m, m_i \quad \text{Mass, proton mass (≈ 1.673 \cdot 10^{-27} \text{ kg})} \]
\[ q_s \quad \text{Electric charge of species s} \]

Constants and symbols

\[ i \quad \text{Imaginary unit} \]
\[ \frac{D}{Dx}, \frac{\partial}{\partial s} \quad \text{Total, partial derivative with respect to } s \]
\[ \nabla, \nabla_{(r)}, \nabla_{(v)}, \nabla_{(r,v)} \quad \text{Del operator, del operator along the } \mathbf{r}, \mathbf{v}, (\mathbf{r}, \mathbf{v})-\text{coordinates} \]
\[ \Delta l \quad \text{Small increment of variable } l \]
VARIABLES, SYMBOLS AND ACRONYMS

\( q \)  
Elementary charge \((\approx 1.602 \cdot 10^{-19} \text{C})\)

\( c \)  
Speed of electromagnetic waves in vacuum \((= 299,792,458 \text{ m s}^{-1})\)

\( \mu_0 \)  
Permeability of vacuum \((= 4\pi \cdot 10^{-7} \text{T m A}^{-1} \approx 1.257 \cdot 10^{-6} \text{A}^{-2} \text{kg m s}^{-2})\)

\( \varepsilon_0 \)  
Permittivity of vacuum \((= (\mu_0 c^2)^{-1} \approx 8.854 \cdot 10^{-12} \text{A}^2 \text{kg}^{-1} \text{m}^{-3} \text{s}^4)\)

\( k_B \)  
Boltzmann constant \((\approx 1.381 \cdot 10^{-23} \text{m}^2 \text{kg s}^{-2} \text{K}^{-1})\)

\( \oplus, R_{\oplus} \)  
Earth, Earth radius \((\approx 6400 \text{ km})\)

\( \mathcal{O} \)  
Link/reference to a source code file

\( \mathcal{M} \)  
Music suggestion for the chapter

\( \mathcal{Y} \)  
Drink suggestion for the chapter\(^1\)

\( \mathcal{F} \)  
Food suggestion for the chapter

**Acronyms**

**CFL condition**  
Stability condition on discretised numerical algorithms which forbids propagation faster than one cell length per time step. Named after Courant, Friedrichs, and Lewy [1928].

**1D, 2D, ...**  
One, two, ... -dimensional or one, two, ... dimensions

**IMF**  
Interplanetary magnetic field

**MHD**  
Magnetohydrodynamic(s)

**PIC**  
Particle-in-cell

---

\(^{1}\)The appropriate intake of water (or coffee) all along is assumed.
Chapter 1

Introduction

1 part of pastis, 4–5 parts of blood orange juice, 1/2 part of lemon juice, ice cubes

Good French saucisson, like duck with green pepper

Some good apéritifs

Kind of Blue

Miles Davis

Before setting off to explore the arcane world of hybrid-Vlasov plasma simulations of the interaction of the solar wind with the Earth’s magnetosphere, some gear is required. This introductory chapter provides the traveller with a simple but sturdy overview of plasma, the Earth’s magnetosphere, parallel computing and the importance of modelling in space physics. Such equipment is necessary for the journey.

The variety of interesting and relevant plasma phenomena warrants their introduction in a separate Chapter 2. Chapter 3 then fleshes out the broad physical approaches used to simulate space plasmas before Chapter 4 presents the hybrid-Vlasov model Vlasiator in more detail. The author’s major code contributions to Vlasiator are the topic of Chapter 5 while his main scientific contributions are summarised in Chapter 6. Conclusions and a vision for the future in Chapter 7 complete the first part of the thesis, which is followed by the reprinted Paper I, Paper II, Paper III and Paper IV.

1.1 What is plasma?

This question must be answered right away, since the present thesis is all about plasma. Remarkably, the word plasma is the same in virtually all languages, unlike a number of other common physical terms, hence no luck in finding comparative help there. In Ancient Greek, πλάσμα means something formed
or moulded, such as a clay figure [Liddell and Scott, 1940], which led nineteenth century botanists and anatomists to call the shapes they saw through their microscopes plasma membranes or plasma cells. Later plasma became the liquid bathing cells, from the inside like cytoplasm or the outside like blood plasma.

Around 1927, one of the forefathers of plasma physics, Irving Langmuir, looked with his team for a catchy phrase to designate electric discharges in vacuum vessels they were investigating, as they felt the topic was to become important. A part of the discharge ‘acted as a sort of sub-stratum carrying particles of special kinds, like high-velocity electrons [...] and ions of gas impurities. This remind[ed] him of the way blood plasma carries around red and white corpuscles and germs’ [Mott-Smith, 1971]. Indeed the term plasma caught on and is still used to designate a state of matter where electrons and ions are separated from each other and have quite some more freedom to move about than in any other classical state of matter, solid, liquid or gas.

Plasmas are found in many contexts, from fluorescent lamps to plasma sterilisers, from supernova remnants to stellar interiors. Therefore it is hard to pin plasma down even with ranges of parameters such as temperature or composition. At the most general level, plasma is usually characterised by three features:

- Plasma consists of free electric charges interacting via electric and magnetic forces;
- Plasma exhibits collective behaviour arising from these interactions;
- Plasma is quasi-neutral: it does not carry a global net electric charge.

This is enough for theoreticians to set out and write down equations describing an arbitrary plasma. But in order to apply any such equation to solve a specific problem, one needs to delve into particulars of a given plasma and give up a certain amount of generality in exchange for some degree of tractability.

Textbooks covering basics of plasma physics include those by Bellan [2006], Bittencourt [2004], Boyd and Sanderson [2003] and Koskinen [2011]. Chapter 2 is dedicated to the presentation of the plasma phenomena studied by the author in Papers I to IV.

1.2 The solar wind, the Earth’s magnetosphere and why they matter

Exactly like all other stars, the one closest to us, the Sun, is a massive ball of hot plasma of about $2 \cdot 10^{30}$ kg mass [United States Naval Observatory, 2016]. The fine balance of forces in the vicinity of the solar surface is such that a tenuous stream of plasma continuously escapes from the Sun’s gravitational pull and
expands into interplanetary space at typically supersonic speeds [e.g. Chapter 1, Koskinen, 2011]. The mechanisms heating and accelerating this plasma in the solar corona, to form what is called the solar wind, are not fully understood and the subject of active research [e.g. Marsch, 2006].

What is better known, thanks to the larger number of available observations, are the structure and characteristics of the solar wind near Earth. As was predicted in the pioneering work by Alfvén, Biermann and Parker in the 1950s, the solar wind is not solely the result of the expansion of the solar corona and its acceleration to speeds of the order of 1000 km/s. A critical aspect is that the solar wind plasma is threaded by and carries away a fraction of the magnetic flux emerging from the solar interior to the Sun’s surface. While the accelerated plasma is expanding radially outwards, the magnetic field is still rooted in the rotating Sun, giving rise to the Archimedean spiral structure of the interplanetary magnetic field (IMF) named the Parker spiral to honour one of its main theoreticians [Biermann, 1951, 1957, Alfvén, 1957, Parker, 1958]. The overall structure of the IMF is not a perfect spiral though, owing to the solar activity. Faster streams of solar wind catch up with slower solar wind, or the odd coronal mass ejection resulting from eruptions near the Sun’s surface hurls a cloud of plasma and a magnetic structure out into interplanetary space. The solar wind has been studied quite extensively in the last decades with a fleet of spacecraft like the Ulysses probe which observed the Sun and the solar wind for the first time out of the ecliptic plane [Wenzel et al., 1992], the SOlar and Heliospheric Observatory (SOHO) continuously picturing the Sun at various frequencies for over 20 years [Domingo et al., 1995] or the Advanced Composition Explorer (ACE) which has served as the Earth’s primary upstream solar wind monitor at the solar-terrestrial L1 Lagrange point for almost 20 years [Stone et al., 1998].

What happens when the fast solar wind hits celestial bodies lurking in the darkness? There is no simple answer to this question, as many factors such as the presence or absence of a magnetic field, an atmosphere or conductivity of the body affect the result. The following only covers the solar wind interaction of magnetised planets with an atmosphere like Earth. The texts by Luhmann [1995b] and Walker and Russell [1995] are a good starting point for the reader interested in a deeper introduction to the interaction of unmagnetised and magnetised bodies with the solar wind.

Planets with a strong internal magnetic field offer a large interaction cross-section to the solar wind and the IMF, while on the inside their field dominates the plasma behaviour within the volume of their magnetosphere. Even though the tenuous plasma making up a magnetosphere and its magnetic field are difficult to image, a few hundred years of ground-based measurements, six decades of spacecraft data and a wealth of model simulations have given a good idea of the general features of at least the terrestrial magnetosphere. The work of
which this thesis is a part is also a contribution to building a more comprehensive and detailed understanding of the Earth’s magnetosphere. The following paragraph aims at giving a generic representation of a magnetosphere despite being written with the terrestrial example and Figure 1.1 in mind.

The magnetic field of a magnetised planet forms an obstacle to the incoming supersonic electrons and ions as they feel the Lorentz force. The raging solar wind flow has to bow to the laws of physics and form a shock to be diverted around the planet\(^1\). In analogy with more down-to-Earth fluid dynamics, this shock is called the bow shock. It forms the interface with the upstream, pristine solar wind and has a roughly paraboloid shape. The bow shock produces a steep increase in plasma density, temperature and magnetic field intensity, as well as a sharp drop in speed, diverting the flow down the flanks of the obstacle. The surface along which the plasma flows, and which it can only cross in exceptional circumstances, is called the magnetopause. The region of shocked solar wind plasma and draping IMF between the bow shock and the magnetopause is the magnetosheath. Owing to the shape of the magnetic field of the planet, which is very close to a dipole field, the magnetopause presents two funnel-shaped cusps towards the polar regions. The magnetopause is the limit of the actual magnetosphere defined as the volume where the planetary magnetic field controls plasma behaviour, although, much as in the present thesis, ‘magnetosphere’ is often (mis)used to describe the whole structure including the shock and sheath. On the nightside, the magnetic field is stretched into the long magnetotail, a roughly cylindrical extension of the magnetosphere consisting of two lobes of very low density plasma, each connected to the polar region of one hemisphere of the dipole. In the inner magnetosphere, where the magnetic field is closest to dipolar, plasma is trapped by the magnetic field configuration. Oppositely-drifting positive ions and negative electrons give rise to the equatorial ring current around the planet. Deep in the magnetosphere, high-energy electrons can be trapped for periods significantly longer than their travel time around the planet. These regions of intense radiation are called the radiation belts of the planet. To complete the sketch, the upper reaches of the atmosphere of the planet – if it has one, obviously – are partially ionised most notably by the solar ultraviolet radiation, forming a conducting ionosphere. There is much to be said about interesting geophysics happening in the ionosphere and lower down, too, but for the sake of not leading the reader – and before them, the author – too far astray, this is not covered here. The texts by Luhmann [1995a] and Carlson and Egeland [1995] give a taste of ionospheric physics.

One structure deserves a special mention in the context of the present work. Although by definition shocks preclude the transport of information upstream by the means of waves, plasma physics offers a quirk which allows

\(^1\)See Section 2.3 for more details on plasma shocks.

\(^2\)See Section 2.4 for more details on magnetic reconnection.
Figure 1.1 – Overview of the terrestrial magnetosphere in the polar plane, as obtained with the hybrid-Vlasov model Vlasiator. Background colour: plasma density. Grey lines: magnetic field lines. Arrows: plasma bulk velocity, length scaled proportionally. The constant solar wind and a southward IMF with an angle of 45° flow in from the right side (black uniform region). The plasma is compressed and the flow diverted by the bow shock, beyond which the plasma flows in the magnetosheath (red and yellow region) over the cusps of the geomagnetic field. The magnetic field is close to dipolar in the inner region and stretched to form the magnetotail on the night-side. The magnetotail is characterised by a low plasma density (green and purple region). Magnetic reconnection² in the tail plasma sheet strongly accelerates plasma, magnetic islands are formed in the reconnection exhaust and expelled from the simulation box (structure in the left-most part of the magnetotail). Ions reflected by the bow shock generate instabilities causing waves to grow and form the foreshock upstream in the region where the IMF is almost parallel to the shock normal direction (black and blue perturbed region south of the bow shock.).
particles to travel back from a shock. A charged particle encountering a shock begins to gyrate with a shorter radius than before because of the stronger downstream magnetic field. Provided the particle has enough energy after half a gyration when its velocity vector is pointing upstream again, and if the angle $\theta_{Bn}$ between the shock normal direction and the magnetic field is small enough so that the particle does not immediately gyrate back into the bow shock, that very particle can escape into the solar wind. This new particle population upstream of the bow shock is forming the aptly-named foreshock, and by virtue of a variety of plasma interactions, the foreshock is pervaded by a zoo of waves and turbulence. More on that follows in Chapters 2 and 6.

In addition to being a fascinating topic of fundamental science – what is the near space environment like and what processes are at play there? – the study of the interaction of the solar wind with planets and other bodies is a key building block of space weather research. Space weather describes the conditions in space which affect the Earth’s environment and have consequences on man-made systems as well as on life in orbit and on the ground [United States National Research Council, 1997]. This covers the problems for life and satellite electronics posed by the radiation environment in space, or the beautiful polar lights enchanting Arctic and Antarctic nights. But solar eruption-induced ionospheric perturbations can also disturb or prevent the propagation of radio waves critical to satellite positioning systems and to communication for industry or rescue operations in remote regions. Perturbations to the geomagnetic field can lead to induced currents in the ground and in long artificial conductors such as power networks and pipelines, causing in the worst cases large-scale blackouts. In between, humans travelling at high altitudes and latitudes for example on an intercontinental flight across the polar regions during episodes of higher particle precipitation into the atmosphere, are subject to higher doses of radiation.

The effects of space weather are complex and multi-faceted, and in most instances they are observed as they occur or predicted only at very short notice. A thorough understanding of the physics at work at all levels from the Sun to the ground via the solar wind and its interaction with the magnetosphere is the condicio sine qua non by which scientists and stakeholders might one day be able to forecast space weather reliably. Such a forecast would enable to make life- and asset-saving decisions in due time and reduce the degree to which our modern society is vulnerable to potential strong space weather events.

### 1.3 Parallel computing

Many scientific problems and the equations used to solve them are too large and complex to be solved with a finite number of pens, a realistic amount of paper and a manageable number of workers in a time of the order of magnitude of
the work contract duration of a doctoral student. Unfortunately, the legendary
infinite open-space office full of graduate students with pocket calculators has
not been accepted for funding due to budget cuts and ethical reasons, among
others. And in any case, the students have better things to do than perform
the job of a processor. Not only that, but a growing subset of these large
scientific problems consists of tasks that are too heavy for a single computer
to tackle, be it because the memory space accessible to the processor is too
limited or because the result would come in too late to be useful. Luckily
however, funding agencies have not been as reluctant to build analogues of
that office by putting together increasing numbers of processors into massive
supercomputers. This section introduces key techniques used to set multiple
processors to work on the same task.

The easiest case is when the job to be parallelised consists of a number of
independent smaller tasks, such as the chunk-wise processing of a large dataset.
Individual tasks do not need to know what the others are doing and only report
their result at the end. The programmer has nothing more to do than to decide
how to divide the dataset among processors and how each task reports its result.
This category are the embarrassingly parallel problems.

The next technique is called shared-memory programming. Modern pro-
cessors often consist of several computing cores accessing the same physical
memory. This allows a program to be split into several threads each running
on a different core and handling its own chunk of the task. The program-
ner’s job here is more involved as they have to decide among others when
the program should create or merge threads, and to make sure that threads do
not overlap and try to access the same memory location at the same time, a
situation called race condition. The best-known tool for shared-memory par-
allelisation of programs written in C/C++ or Fortran is the OpenMP library,
that provides compiler instructions defining the code sections to be threaded
and the scope or access rules of variables, for example [Dagum and Menon,

If the problem at hand is too large to even fit on a multi-core processor
and shared-memory parallelisation is thus not appropriate any more, the task
needs to be split across multiple processors each with its own memory. This
is distributed-memory computing: each processor has its own local memory
and executes its own instance of the program, called a process. Domain decom-
position is a classic example of distributed-memory parallelisation, by which
each process is assigned a part of the computational domain to work on. If a
process needs to get information that it does not have locally, such as a value in
a neighbouring part assigned to another process, the programmer needs to im-
plement the communication between processes themselves. A very commonly
used tool is the Message Passing Interface (mpi), a C/C++ and Fortran library
defining functions for example sending and receiving messages between indi-
vidual processes or performing collective operations like variable reductions
across or communication among all processes [MPI Forum, 2004].

Vector processing is a technique which dominated supercomputing from the 1970s to the 1990s but fell out of fashion as conventional microprocessors became less expensive, before gaining momentum again in recent years. The idea behind vectorisation is to use a single processor to perform operations on multiple numbers at the same time. A processor’s arithmetic logic unit can only process one pair of numbers at a time, like summing \( A \) and \( B \), but the memory line of the unit is usually larger than the number representation format used by the program. By appropriately organising the memory layout and computations, the logic unit can therefore process several pairs of numbers simultaneously in one clock cycle, such as four pairs of 64-bit numbers with a logic unit using 256-bit vectors. A variety of tools are available to vectorise programs, which depend heavily on the processor targeted. Common ones for Intel’s x86 processors are the MMX, SSE, AVX and FMA sets of instructions [Intel, 2016]. The programmer’s main task is to arrange data in memory in such a way that the compiler can easily replace single-number operations by equivalent vectorised operations.

In general, none of the programming techniques described here is the only good solution. Some problems can be solved equally well at small scales with shared- or distributed-memory programs. However, modern supercomputers typically consist of tens or even hundreds of thousands of processors each comprising up to several tens of cores [Strohmaier et al., 2016], making a combination of shared- and distributed-memory techniques as well as vectorisation unavoidable in high-performance computing applications, in order to leverage as much of the theoretically available power as possible.

1.4 The role of simulation in space physics

The two historic pillars of science have been theory and experiment, fostering the progress of knowledge through their mutual influence. Experiments and measurements produce new information, which theories try to interpret and explain. As an example, based on the repeated observations of strong geomagnetic perturbations during intense displays of aurorae borealis, early investigators such as Størmer and Birkeland formulated hypotheses trying to explain experimental datasets [Stern, 1989, and references therein]. They especially suggested that electrons could be trapped in the terrestrial magnetic field in near-Earth space.

Space physics can be defined as the study of space as far out as has been reached by spacecraft taking in-situ measurements. This is not a very precise definition and the boundaries to adjacent fields like planetary sciences when studying the environments of other planets, geophysics when studying our Earth’s magnetosphere, or solar physics when considering that the Sun’s surface and interior cannot be reached in-situ, to name but a few, are by no means
strict and hermetic. Using this definition nevertheless, and with the notable exception of geomagnetic and auroral observations, which had been ongoing for more than a century already, space physics took off as an active field of research with the advent of the space age in 1957–1958 as direct measurements in space became feasible. Some of the earliest observations were that space is not empty but filled with radiation. The Soviet Sputnik 2 – better known for carrying the dog Laika into space but also equipped to measure radiation – and the American Explorer 1 were the first to measure the van Allen radiation belts, or to do any measurements in space for that matter [van Allen et al., 1958, Chernov and Yakovlev, 1958]. These experiments confirmed that the radiation belts indeed consist of trapped electrons, a picture further bolstered when atomic bombs were detonated in orbit, populating new regions with energetic electrons in the process [Christofilos, 1959].

Space physicists are, in some respect, better off than other colleagues because they do have local measurements available, something astrophysicists, cosmologists or deep-Earth geophysicists will never enjoy. Yet almost sixty years after Sputnik 1 – aeons in terms of marketing and product life cycles – it is still no small feat to send probes into orbit or beyond, and their price tag puts a strong limit on the number of spacecraft any state, agency or institute can afford to build and launch. One thing space physicists are helpless with, on the other hand, is the reproducibility of their experiment: in space chances are slim that the same conditions are encountered the next time the spacecraft flies past (if ever), in order to try to nudge the detectors or tweak the detection ranges to get a better measurement. One option here is to fly several spacecraft in a constellation to get simultaneous and spatially distinct samples of the same events. This has been very popular in space physics, with successful examples including the Cluster four-spacecraft mission providing plasma and magnetic field measurements in the Earth’s magnetosphere [Escoubet et al., 2001], the Time History of Events and Macroscale Interactions during Substorms mission [THEMIS, Angelopoulos, 2008] with initially five spacecraft in Earth orbit and later two of them forming the Acceleration, Reconnection, Turbulence, and Electrodynamics of the Moon’s Interaction with the Sun constellation around the Moon [ARTEMIS, Angelopoulos, 2011] or the four Magnetospheric Multi-Scale (MMS) observatories reaching for the first time down to electron kinetic scales to study energy transfer in the Earth’s magnetosphere [Burch et al., 2016a]. Each of these constellations brought about fabulous leaps in physical understanding, especially when the key point very often is telling apart spatial and temporal structures, but inherently the measurements are still taken in no more than a handful of moving points in the vast emptiness of space. In other words, the big problem for space physicists is to get the big picture when all that is available are time series of point-wise measurements.

The obvious next step in trying to provide a context to sparse spacecraft measurements is to seek the support of mathematical models, which usually
form the basis of physical theories. As the following Chapter 3 shows though, the equation sets needed to describe somewhat accurately something as complex as, say, the interaction of electrons and protons flying at supersonic speeds with a large, static magnetic dipole, are such that solving them requires the help of computers. As in most areas of the physical sciences, numerical modelling has become the third pillar of research in space physics. Numerical models, which include simulations, are essential both to theory and experiment. By choosing the physical ingredients mixed into a model, the relative importance of various processes can be assessed, improving the theoretical insight. And by allowing to look at the whole system evolving in time, numerical simulations are invaluable in helping to grasp the global dynamics in relation to spacecraft data. Numerical modelling has also become instrumental in the development of instruments – pun intended – for new spacecraft missions. Simulation data can for example be processed according to planned instrument parameters in order to define a priori resolutions, sensitivities and accuracies sufficient to confirm observationally phenomena only simulated hitherto. Whence it can be said that the traditional paradigm of scientific research relying on the interplay of theory and experiment has to be extended to include numerical modelling, as all three are critically supporting and complementing each other to facilitate the building of knowledge.

Having acquired the basic tools if they were not yet at hand, the reader can now set out for the travel. The work underpinning and constituting the present thesis lies at the crossroads of all topics brushed in this introduction, namely the realms of plasma physics, especially applied to study the Earth’s magnetosphere, and the lands of supercomputers, of which the power can be harnessed for the good of mankind and the progress of science.
Chapter 2

Plasma phenomena

Good wine like a dry white alsace or a

tannic red bordeaux

Collective effects arising from complex

interactions:

Die Kunst der Fuge

Johann Sebastian Bach

Haitian Fight Song; or All the Things You C#

Charles Mingus

Carmen Fantastique

Edson Cordeiro and the Klazz Brothers

What is meant by collective behaviour arising from plasma interactions? Plasma physics combines the joys of electrodynamics, statistical physics and fluid dynamics to conjure up a fascinating world of possibilities. It takes textbook upon

lecture series to even cover the basics, wherefore this Chapter 2 only describes

those aspects important within the scope of this work including Papers I to IV

and abstains from complete mathematical derivations. The author found the

books by Koskinen [2011, Chapters 1 to 11] and Bittencourt [2004] particularly useful when studying basic plasma physics, but plenty of further valuable

literature exists.

2.1 Characteristic length and time scales

The electric potential of a single charged particle in vacuum is given by the

Coulomb potential,

$$\phi = \frac{q}{4\pi \varepsilon_0 r^2}.$$  

(2.1)
where $r$ is the radial distance from the particle\(^1\). As explained in Chapter 1, plasma consists of free electric charges but it is usually quasi-neutral, which means that charges tend to rearrange in order to cancel large-scale electric fields. As a consequence, if a test particle is introduced in a plasma, charges are displaced in order to retain quasi-neutrality. Assuming that the particles of species $s$ are in equilibrium, their density obeys

\[
n_s = n_{0s} \exp\left(-\frac{q_s \phi}{k_B T_s}\right)
\]  

(2.2)

and it can be shown that the potential $\phi$ is modified by the presence of the plasma to become

\[
\tilde{\phi} = \phi \exp(-r/\lambda_D)
\]  

(2.3)

where

\[
\lambda_D = \sqrt{\frac{\epsilon \phi k_B}{\sum_s n_{0s} q_s^2 / T_s}}
\]  

(2.4)

is the decay length scale of the potential, known as the Debye length. The plasma effectively shields the test particle so that its potential is exponentially weakened with distance. This whole consideration only holds if there are enough particles within a Debye length of any other particle, and if the system size $L$ is much larger than the Debye length:

\[
n_{0}^{-3} \ll \lambda_D \ll L.
\]  

(2.5)

This condition provides a criterion to determine whether a system can be considered an actual plasma, although the strength of the ‘much smaller than’-signs is left to the appreciation of the beholder.

Pondering further the fate of a quasi-neutral electron–ion plasma perturbed by a test charge, the most fundamental plasma oscillation, known as the Langmuir oscillation, can be understood intuitively. A local perturbation introduced in the electron density causes a restoring Coulomb force exerted by the ions, which do not react as fast with their large inertia. Therefore the electrons oscillate around their equilibrium position, at the so-called electron plasma frequency

\[
\omega_{pe} = \sqrt{\frac{n_e q^2}{m_e \epsilon_0}}.
\]  

(2.6)

Another fundamental frequency belonging to any species $s$ in a magnetised plasma is the frequency at which the particles gyrate around the magnetic field owing to the Lorentz force acting perpendicularly to $\mathbf{B}$. This frequency is

\(^1\)All variables, constants and symbols are defined in the section starting on p. xv.
variously termed the gyrofrequency, Larmor frequency or cyclotron frequency
\[ \omega_{cs} = \frac{q_s B}{m_s}. \]  \hfill (2.7)

Both these frequencies are expressed in rad/s. The inverse of the gyrofrequency is the gyropoint period or Larmor period \( 2\pi / \omega_{cs} \).

These frequencies are also related to two length scales useful in characterising the spatial scales at which phenomena occur. The radius of the gyromotion of a particle travelling at speed \( v_\perp \) perpendicular to \( B \) – the gyroradius or Larmor radius – is given by
\[ r_{Ls} = \frac{v_\perp}{\omega_{cs}} = \frac{m_s v_\perp}{q_s B}, \]  \hfill (2.8)

while the inertial length of a particle in a plasma can be defined as the effective gyroradius of that particle flying at the Alfvén speed (in the case of electrons, this is the electron Alfvén speed; see the next section for the definition of the Alfvén speed):
\[ d_s = \frac{c}{\omega_{pe}}. \]  \hfill (2.9)

The Larmor radius and inertial length are collectively designated as the kinetic scales, above which mostly fluid-like phenomena occur but which need to be resolved if effects related to the individual particle motion – kinetic effects – are to be taken into account.

### 2.2 Waves and instabilities

A great deal of what makes plasma interesting, and more complex than fluid dynamics or electrodynamics alone, is encapsulated in the intertwined concepts of plasma waves and instabilities. When solving plasma equations, one can often write solutions in terms of plane wave modes such as
\[ \mathbf{B} = B_0 e^{i(k \cdot r - \omega t)}, \]  \hfill (2.10)

which describes the magnetic field of a mode with a wave vector \( \mathbf{k} \) and a frequency \( \omega \). To the enjoyment of the theoretical plasma physicists, a variety of cases arise where the frequency is not purely real and can be written as \( \omega = \omega_r + i\gamma \), so that when rewriting Equation (2.10) the purpose gets clear:
\[ \mathbf{B} = B_0 e^{\gamma t} e^{i(k \cdot r - \omega_r t)}. \]  \hfill (2.11)

If \( |\gamma| \ll \omega_r \), what remains is essentially a plane wave \( \mathbf{B} \approx B_0 e^{i(k \cdot r - \omega_r t)} \). If \( \gamma < 0 \) the real exponential term damps the wave. On the other hand, if \( \gamma > 0 \), the amplitude grows exponentially: the wave becomes an instability. The study of
plasma waves and instabilities theoretically, observationally and in simulations forms a large part of plasma physics, as they are a key mechanism by which energy is transported or transferred from particles to fields and vice-versa.

The characteristic frequencies introduced in Section 2.1 are crucial when studying wave propagation in plasma. A wave with a frequency close to $\omega_{ps}$ or $\omega_{cs}$ can resonate with the plasma oscillation or the gyromotion of the particles of species $s$. In that process, energy from the wave is transferred to the resonating particles and as a result the wave is damped.

Plasma waves can be classified in several ways depending on their direction of propagation with respect to the ambient magnetic field, their polarisation, whether they are compressive or not or whether they are longitudinal or transverse perturbations. Classic large-scale and low-frequency wave modes, that is at wavelengths longer than the plasma species’ gyroradius and inertial length and at frequencies lower than the species’ gyrofrequency and plasma frequency, are

- the ion sound wave, a compressive perturbation of the plasma density similar to the sound wave from fluid/gas dynamics, propagating with a speed $v_s = \sqrt{k_B T_c/m_i}$;
- the Alfvén wave, a non-compressive transverse perturbation of the magnetic field propagating in directions close to parallel to the magnetic field direction with a speed $v_A = B/\sqrt{\mu_0 \rho_m}$;
- the magnetosonic wave, a compressive perturbation of the plasma density and the magnetic field strength propagating close to perpendicularly to the magnetic field with a speed $v_{ms} = \sqrt{v_s^2 + v_A^2}$.

When studying plasma waves analytically, it is common to describe them with the dispersion relation expressing $\omega$ as a function of $k$. The phase velocity of the wave is then defined as

$$v_p = \frac{\omega}{k},$$

which gives the velocity of the wave fronts. $\hat{k} = k/k$ is the unit vector pointing in the direction of the wave $k$, thus perpendicularly to wave fronts. The group velocity, which gives the velocity of wave groups or packets and the direction in which the wave transports energy, is defined as

$$v_g = \frac{\partial \omega}{\partial k}.$$  

In the low-frequency regime at hand, the sound, Alfvén and magnetosonic waves introduced above are not dispersive, meaning that their dispersion re-
2.2. WAVES AND INSTABILITIES

The relation is approximately linear:

\[ v := v_g = v_p = \frac{\omega}{k} \Leftrightarrow \omega = kv. \]  

(2.14)

The more complex behaviour of these waves at higher frequencies is the subject of Paper I and is addressed in more detail in Section 6.1.

As a stroll into any conference or a glimpse through the papers published by any journal relating to plasma physics can convince the reader, charged particles in magnetic fields are a treacherous and vicious lot not waiting for anyone or anything in particular to get unstable and wreak havoc in the Sun’s chromosphere, the Earth’s magnetosphere or one’s pet fusion reactor. Understanding and predicting or controlling plasma instabilities is the purpose of many endeavours. Plasma instabilities can be traditionally sorted into two broad categories, namely macroinstabilities and microinstabilities. The former occur on large scales with respect to the relevant kinetic scales while the latter relate to details of particle motion and behaviour at small scales. This separation is somewhat artificial nevertheless, as kinetic effects can cause large-scale instabilities as well. A few of the general instability types are reviewed in the following paragraphs.

Some classic macroinstabilities or fluid instabilities have counterparts in fluid dynamics. For example, the Kelvin-Helmholtz instability causes the interface between two plasma domains which have a strong relative shear velocity to ripple and form large vortices eventually mixing plasma from both domains. It has been observed to happen very often at the Earth’s magnetopause where the magnetosheath plasma picks up speed and flows along the flanks [e.g. Kavosi and Raeder, 2015].

Other macroinstabilities are intimately tied to magnetised plasmas. In the case of the mirror instability, a previously homogeneous volume of magnetic field grows an alternating pattern of high and low magnetic field intensity. The plasma tends to be trapped in the pockets of lower magnetic field intensity and a characteristic succession of anticorrelated peaks and troughs of magnetic field and plasma density forms. This instability is well-known to occur in the sub-solar region of the terrestrial magnetosheath [e.g. Souček et al., 2008, Hoilijoki et al., 2016]. A sibling of this instability is the firehose instability. Both these instabilities are the result of the plasma temperature being different along the magnetic field direction and across.

There is a large variety of microinstabilities involving particular conditions for only ions or electrons or combinations thereof. What is always behind a microinstability is the presence of some form of anisotropy in the plasma, be it that one species does not flow in the same direction as another, that a beam propagates separately from the bulk population, or that plasma is hotter along the magnetic field direction than across, to name but a few examples. The
Bunemann instability for instance arises when the electrons flow with a different bulk velocity than the ions, while the bump-on-tail and many other beam instabilities arise when a portion of the particle population flows separately from another. Beam instabilities are ubiquitous in solar and space physics, an example being type II radio emissions near the Sun [Ganse et al., 2012].

Overall, instabilities can only occur if there is a source of energy for them. A bulk velocity shear can trigger the Kelvin-Helmholtz instability, whereas beam instabilities draw their energy from the relative velocity of the various populations. Identifying the energy source helps to pin down the instability and predict its evolution. The energy has to come from somewhere: if the source of the instability is drained, the instability does not grow any further. This can take place in several ways: the actual source of kinetic energy or anisotropy can have stopped, or the system can be reconfigured so extensively by the grown instability that there is no possibility for it to happen any more. For example, the Kelvin-Helmholtz instability ceases once both domains have completely mixed.

2.3 Shocks

As soon as characteristic wave speeds have been defined, it is allowed to ask what happens if an object travels through a medium faster than a given wave\(^2\). And the answer is literally shocking.

Analytically, a shock forms as the solution of a transport or propagation equation when the slope becomes infinitely steep at some point, which means that the previously continuous and differentiable solution becomes discontinuous at that point and partial solutions have to be sought on either side of the discontinuity.

Phenomenologically, information about a potential downstream obstacle cannot reach the upstream medium by means of a given wave, if the medium is flowing faster than the characteristic speed of that wave. As a common example, an observer listening ahead of a supersonic aircraft cannot hear it: they cannot obtain information about the incoming aircraft by means of sound waves. Reversing the frame of reference – a mandatory skill for anyone who wishes to study shocks, the air flowing supersonically cannot know about the aircraft sitting still further down the path. In order to divert the flow around the obstacle, a shock forms, which is nothing else than a steepened wave front standing against the flow. It builds an interface between the upstream medium, flowing unwittingly at high speed, and the downstream medium, which is slowed down below the wave speed and diverted around the obstacle in a layer called the sheath. It has to be noted that an obstacle is not mandatory for a shock

\(^2\)Care is required in the case of electromagnetic waves in vacuum, but this definitely exceeds the scope of the present work.
2.3. SHOCKS

to form. A fast flow catching up with a slower one forms a shock front if the speed difference exceeds the characteristic speed. Energy conservation requires that the kinetic energy be transformed, which usually means, in classical fluids or plasma, that the medium is compressed and heated up when crossing the shock.

The more general evolution of all plasma parameters across a shock can be derived from mass, momentum, energy and magnetic flux conservation. The results are called the Rankine-Hugoniot jump conditions, which relate in a simple and closed form the upstream and downstream density, pressure, velocity and magnetic field components. The shock conditions are fully determined by giving either the upstream or downstream conditions and solving the Rankine-Hugoniot conditions. The interested reader is directed to Section 11.2 of the book by Koskinen [2011] or for example to the works by Burgess [1995] and Burgess and Scholer [2015] for a more complete treatment of this topic, only the evolution of the magnetic field is to be discussed here a little more.

Because of the conservation of magnetic flux and due to the slowing down of the plasma flow at a supermagneto sonic shock, the magnetic field has to be compressed too by the shock. However, assuming that the shock surface is planar and the upstream magnetic field uniform, it follows that the downstream magnetic field is also uniform. Since $\mathbf{B}$ is solenoidal (Equation 3.3), the immediate consequence is that the component of the magnetic field normal to the shock surface is constant across the shock, in other words, the compression of the magnetic field only affects the components of $\mathbf{B}$ tangential to the shock.

Key figures to describe a shock are the Mach numbers $M$, defined as the ratio of the inflow or upstream speed $V$ to the characteristic wave speeds $c_w$:

$$
M = \frac{V}{c_w},
$$

meaning that a shock has several Mach numbers like the sonic, Alfvénic and magneto sonic Mach numbers related to the eponymous wave modes presented earlier. Another important set of shock parameters are the angles between the inflow velocity $\mathbf{V}$, the upstream magnetic field $\mathbf{B}$ and the unit vector $\mathbf{n}$ orthogonal to the shock. The most relevant one in the present thesis is the angle between the upstream magnetic field and the shock normal

$$
\theta_{Bn} = \cos^{-1} \left( \frac{\mathbf{B} \cdot \mathbf{n}}{B} \right).
$$

Indeed, as already pointed out in Section 1.2 and as illustrated by Figure 2.1, the value of $\theta_{Bn}$ conditions the fate of many a charged particle colliding with a shock. A particle hitting a shock gyrates more tightly around the compressed magnetic field direction in the shock layer than it did upstream and after half a gyration it can be turned back to the upstream direction. At a
Figure 2.1 – Behaviour of a charged particle encountering a supermagnetosonic shock. The shock causes a compression of the magnetic field, thus the gyroradius of the particle is shorter downstream of the shock.

(a) In the (quasi)perpendicular case ($\theta_{Bn} \approx 90^\circ$), all particles cross the shock and escape downstream.

(b) In the quasiparallel case ($\theta_{Bn} \approx 0^\circ$), some particles bend back and escape upstream, forming the foreshock population.

quasiperpendicular shock, where $\theta_{Bn} \approx 90^\circ$, the next half gyration returns the particle towards the shock and it is transported through the shock layer and further downstream since $\mathbf{B}$ is essentially parallel to the shock plane (upper part in Figure 1.1 and Figure 2.1a). Conversely, at a quasiparallel shock where $\theta_{Bn} \approx 0^\circ$, an inflowing particle with suitably high energy can be reflected at the shock and escape back along the magnetic field direction upstream of the shock, forming a counter-streaming plasma population in the foreshock region (lower part of the upstream region in Figure 1.1 and Figure 2.1b). This new population is a source of free energy favouring the growth of waves, thus the perturbation of the subsequently incoming solar wind and foreshock particles, and eventually the perturbation of the bow shock itself once the foreshock perturbations advected by the solar wind reach and cross the bow shock. Chapter 6 addresses foreshock phenomena in more detail as foreshocks are the central topic of Paper III and Paper IV. Finally, let it be mentioned that good introductory texts on shocks in magnetised plasmas are Chapter 11 in the book by Koskinen [2011] and the works by Burgess [1995] and Burgess and Scholer [2015].

2.4 Magnetic reconnection

Within the formalism of magnetohydrodynamics (MHD), a set of equations describing plasma as a magnetised fluid (see Section 3.3), it can be demonstrated that a given small volume of plasma always remains on the same magnetic field
2.4. MAGNETIC RECONNECTION

line, while the topology of the magnetic field is preserved. This is often referred to as plasma being frozen-in to the magnetic field. In this framework, the flow of plasma and magnetic field are closely-knit as the plasma drags the fields along and stays topologically connected to the same field line. At large scales, MHD has been shown to be a handy tool to describe plasmas such as the solar wind or some astrophysical contexts (see Section 3.3 and references there). Qualitatively, the frozen-in theorem is also very convenient for plasma physicists to apprehend and get a feeling for the possible plasma behaviour in a certain configuration. It is nevertheless important to note that the whole edifice of MHD relies on neglecting all kinetic effects at small spatial and short time scales.

A typical configuration where MHD hits its limits forms when two domains with antiparallel magnetic field flow towards each other. Owing to the frozen-in principle, nothing more should happen but a plasma and magnetic field pile-up, unless some flow channel for the plasma regions around each other is found. In reality, the behaviour of plasma is greatly modified, in particular because the magnetic field goes through zero between both domains. In a volume of the approximate size of the relevant kinetic scales, called the diffusion region, charged particles are demagnetised and follow ballistic trajectories, only maybe accelerated by electric fields, until they exit the region and feel some magnetic field again. As they have a larger Larmor radius and inertial length, ions are demagnetised at larger scales than electrons; the electron diffusion region is embedded in the ion diffusion region (see Figure 2.2).

In addition to particles being demagnetised, a topological reconfiguration of \( \mathbf{B} \) takes place in the diffusion region, so that field lines initially only threading either side of the interface end up connected with a field line from the other side. This happens in more or less complex geometries, the simplest being the X-line topology represented in Figure 2.2. In that essentially 2D case, the whole process of magnetic reconnection is organised around a single magnetic null, the X-point at the centre of the diffusion region. This process is energetically favourable: the initial \( \mathbf{B} \) configuration is more stressed than after reconnection. The excess magnetic energy is transferred to the plasma, which is accelerated within and near the diffusion region through the relaxation of the magnetic field lines and expelled at or near the inflow Alfvén velocity. As can be expected, reconnection seldom occurs in a purely 2D configuration and its topology can become complex. The picture of reconnection drawn in this section corresponds to antiparallel reconnection. Magnetic reconnection can also occur in cases where only a component of both magnetic domains is antiparallel, a case known as component reconnection. The reviews by Zweibel and Yamada [2009], Yamada et al. [2010] and Pontin [2011] as well as textbooks such as the classics by Biskamp [2000] and Priest and Forbes [2007] are a good start for the reader willing to wander into the amazing maze of magnetic reconnection.
Figure 2.2 – 2d sketch of magnetic reconnection. The black lines are tangent to the magnetic field. Plasma flows in from the upper and lower edges towards the X-line. Ions decouple from the magnetic field within the ion diffusion region (IDR) while electrons are demagnetised in the electron diffusion region (EDR – not to scale). Plasma is accelerated in the process and escapes towards the exhaust regions.

By allowing the topological reorganisation of antiparallel magnetic domains, releasing energy along the way and accelerating the inflowing plasma, magnetic reconnection plays a key role in many dynamic systems such as solar eruptions and, unsurprisingly, in the terrestrial magnetosphere. Under southward IMF, reconnection takes place at the magnetopause in the subsolar region. This dayside reconnection increases the total magnetic flux in the magnetosphere and eventually triggers reconnection in the current sheet separating both lobes of the magnetotail. This interplay of dayside and nightside reconnection drives the Dungey cycle, which completes by field lines and plasma travelling towards the dayside again through the inner magnetosphere [Dungey, 1961]. Magnetic reconnection is the focus of current magnetospheric research, especially since the launch of MMS [e. g. Burch et al., 2016b, Cassak, 2016]. Dayside reconnection is also a key ingredient in the phenomenon reported for the first time in Section 6.4 and in Paper IV.
Chapter 3

Space plasma simulation

Plasma is a quasi-neutral bunch of electrically charged particles interacting collectively, but how can that be modelled mathematically? In principle, plasma is only about the equation of motion of point-like charged particles in electromagnetic fields. In practice, there are too many particles and the relevant time and spatial scale ranges to cover are too wide. This is why a number of approaches exist, varying in degree of physical accuracy and inversely with the computational weight. There is no free lunch: a lightweight method has less physical detail while a physically completely accurate model cannot be computed in a reasonable time.

The task is, in many ways, similar to sports reporting. In theory, one could model the *Tour de France* down to the physiological parameters of each cyclist, the pedal phase and tire pressure of the bicycles, the micrometeorological conditions and the exact 3D profile of the road including surface roughness and cracks. But no one can afford that.

The emphasis here lies on the physical ideas underpinning common plasma models used in space physics, not on their discretisation for efficient computing. The reader is referred to further literature describing numerical plasma simulation methods [e.g. Matsumoto and Sato, 1985, Hockney and Eastwood, 1988, Büchner et al., 2003] and to Chapter 4 of this thesis detailing the algorithms of the hybrid-Vlasov model Vlasiator.
3.1 Particles in cells, a rough picture

The equation of motion of a particle subjected to the Lorentz force is given by Newton’s second law of motion

\[ \mathbf{a} = \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) . \]  

(3.1)

It can easily be integrated once the initial position and velocity of the particle are known. The electric and magnetic fields can be computed with the help of Maxwell’s equations

\[ \nabla \cdot \mathbf{E} = \frac{\rho_{\text{q}}}{\varepsilon_0} , \]  

(3.2)

\[ \nabla \cdot \mathbf{B} = 0 , \]  

(3.3)

\[ \nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t} , \]  

(3.4)

\[ \text{and } \nabla \times \mathbf{B} = \mu_0 \left( \mathbf{j} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) , \]  

(3.5)

since with the position and velocity of all the particles, the charge density \( \rho_{\text{q}} \) in Gauss’s law (3.2) and the current density \( \mathbf{j} \) in Ampère’s law (3.5) are defined in every position as well. Thus for set initial conditions, one can compute the fields and the motion of each particle in these ambient fields, and then iterate through time.

The basic plan is fairly simple, but unfortunately it gets out of hand very rapidly with a growing number of particles in the system. Even with a small magnetosphere of \( 20^2 \times 100 R_E^3 \) and a low ion and electron density of \( 1 \text{ m}^{-3} \), one ends up with an impossible \( 2 \cdot 10^{25} \) particles to handle. However, closer inspection of Equation (3.1) indicates that particles with the same charge-to-mass ratio obey the same equation, which suggests one way of reducing the number of particles. Suitably similar particles could be grouped into larger macroparticles with an equivalent \( q/m \). If the macroparticles are judiciously chosen to sample well the total population, the amount of computations would be automatically reduced without affecting the modelled physics.

This forms the idea behind particle-in-cell (PIC) methods: replace the actual particles with much fewer representative macroparticles and solve their equation of motion in electric and magnetic fields. From that, \( \mathbf{E} \) and \( \mathbf{B} \) are computed, using the local charge and current densities. This is usually not done for each macroparticle position separately but the effect of particles is accumulated and averaged within, and the field values are computed for discrete cells of a grid to simplify the process – hence the -in-cell part of the name.

Here the astute plasma modeller is confronted with one of the major hurdles of PIC algorithms. For the purpose of computing the fields, a macroparticle has to be weighted according to the proper number of ions or electrons it represents, which is simple enough to do. Additionally though it may be useful
3.1. PARTICLES IN CELLS, A ROUGH PICTURE

to assign each macroparticle some spatial shape instead of a point-like location, so that its influence is not restricted only to the closest grid cell and does not jump discretely from one cell to the next while propagating through the grid. This somewhat extended macro-particle can then be thought of as sampling a small volume instead of a Dirac $\delta$-like point in the $(r, v)$ phase space. The right choice of such a shape function allows to simplify the expressions involved when interpolating the fields to the grid and this is one of the keys to developing a successful PIC model [e. g. Hockney and Eastwood, 1988, Markidis et al., 2010, Kilian et al., 2013].

Another hurdle is the inclusion of collisions. Coulomb collisions are elastic binary collisions at long range due to the electric interaction. The divergence of the Coloumb force at short range is avoided since the effects of macroparticles is averaged over the grid cells to compute the fields. Furthermore in many cases plasmas in space physics are so tenuous that head-on particle collisions are virtually nonexistent: neglecting collisions means simulating a collisionless plasma, which is a perfectly valid assumption in many contexts. In regimes where collisions are important, such as in the Earth’s ionosphere, their direct treatment is possible but resource-consuming, so that other methods such as Monte-Carlo approaches have been developed.

Given the relative conceptual and mathematical simplicity of the PIC approach, in combination with the possibility of almost arbitrarily scaling the computational load of a task by cranking up or down the allowed number of macroparticles, PIC simulations have been popular and successful in plasma research for decades, be it in global magnetospheric research [e. g. Savoini et al., 2013, Karimabadi et al., 2014], local studies [e. g. Daughton et al., 2011] or more exotic topics like supernova remnants [e. g. Ohira et al., 2009] or pulsar magnetospheres [e. g. Philippov et al., 2015].

While it is the major strength of PIC models to use macroparticles, it is also their largest drawback. The reduced number of particles introduces statistical noise affecting the sampling quality of the modelled particle populations at the microphysical level. On macroscopic scales, the variables computed from the particle populations (density, velocity, pressure etc.) are affected by some level of statistical noise too. With clever weighting and interpolation techniques the impact of the noise can be diminished, but ultimately the only possibility to improve the sampling is to increase the number of modelled macroparticles and with that obviously the computational cost. The latter scales linearly with the number of particles, but the random noise level scales inversely with the square root of that number, so that the improvement does not come cheap. Eventually the question boils down to how few macroparticles the modeller – or the referee! – is comfortable with whilst still trusting the physical results.

Is the Tour de France enthusiast satisfied with a teamwise coverage of the day’s stage, or would they like a more detailed report, at a higher cost?
3.2 Particle distributions in higher dimensions

Explicitly pushing around all particles in self-consistent electromagnetic fields is computationally too expensive and PIC methods can suffer from statistical noise. Instead of following particles, one could try to study the behaviour of a function \( f \) giving the density of particles in space. Particles are uniquely identified by giving their position \( \mathbf{r} = (x, y, z) \) and velocity \( \mathbf{v} = (v_x, v_y, v_z) \) as a function of time \( t \), which means that \( f \) has to be defined in a 6D phase space \( (\mathbf{r}, \mathbf{v}) \).

In other words, this distribution function describes how many particles are going in which direction (v-coordinates) at any point (r-coordinates). Chemical, nuclear or high-energy processes whereby particles would split, bind, fuse, or otherwise appear or vanish do not fall into consideration – \( f (\mathbf{r}, \mathbf{v}, t) \) is strictly conserved in time:

\[
\frac{Df (\mathbf{r}, \mathbf{v}, t)}{Dt} = 0. \quad (3.6)
\]

Were it not, there would have to be source or sink terms on the right-hand side. Writing out the total time derivative yields

\[
\frac{\partial f}{\partial t} + \frac{\partial (\mathbf{r}, \mathbf{v})}{\partial t} \cdot \nabla_{(\mathbf{r}, \mathbf{v})} f = 0
\]

\[
\Leftrightarrow \frac{\partial f}{\partial t} + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_{(\mathbf{r})} f + \frac{\partial \mathbf{v}}{\partial t} \cdot \nabla_{(\mathbf{v})} f = 0
\]

\[
\Leftrightarrow \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \mathbf{a} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \quad (3.7)
\]

which is nothing more than the advection equation of a conserved density \( f \), albeit in 6D instead of the 3D many are more acquainted with. The acceleration due to the Lorentz force has not changed since Equation (3.1), so that the evolution equation for the distribution function \( f_s (\mathbf{r}, \mathbf{v}, t) \) of a particle species \( s \) can be given as

\[
\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \frac{\partial f_s}{\partial \mathbf{r}} + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = 0. \quad (3.8)
\]

It has to be noted that \( \mathbf{v} \) is not the single-particle velocity as in Equation (3.1) any more but the velocity space coordinate in our 6D phase space. Equation (3.8) is the Vlasov equation in its full splendour.

*Caveat lector*, a significant amount of physical generality has now been surreptitiously traded in for a great deal of mathematical simplification: this equation takes into account the Lorentz force but completely neglects collisions. The Vlasov model has no difficulty with that though and happily provides a description of collisionless plasmas. And since there are no particles to speak of in this formalism, it is no easy thing to include the treatment of collisions.
3.2. PARTICLE DISTRIBUTIONS IN HIGHER DIMENSIONS

Wrapping the variations of \( f_s \) due to collisions into a term on the right-hand side of the Vlasov equation gives

\[
\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \frac{\partial f_s}{\partial \mathbf{r}} + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = \left( \frac{\partial f}{\partial t} \right)_C,
\]

which is essentially equivalent to the well-known Boltzmann equation from statistical physics, applied here for plasma. Without going into details, the collisional term is usually of such a mathematical form that its computation is rather expensive. Such term has not yet been included in large-scale Vlasov-based space plasma models. Considering the physical systems targeted in the future, their relevance will increase and it is likely that efforts to include computationally more efficient collisional terms in Vlasov-based space plasma simulations will be made. Such effort has already been undertaken in the field of gyrokinetic simulations [see Section 3.4 below and e.g. Abel et al., 2008, Barnes et al., 2009]. Although collisions are unimportant in the studies pertaining to this thesis, it is not so in adjacent fields. Be it in stellar plasmas, ionospheres or in fusion reactors: many a plasma context requires to include collisions for a sensible treatment.

A further term in the equations is being neglected within the scope of this work: gravity. It would be included alongside the Lorentz force in the acceleration term, both in \textsc{pic} and in Vlasov methods. Its effects are not foreign to space plasma physics: the structure of stars, the expansion of the solar wind or the dynamics of the Earth’s ionosphere for example are all affected or even dominated by gravitation. However in the present work its influence can safely be disregarded.

As in \textsc{pic} methods, the electromagnetic fields are needed to propagate the solution in time, and as in \textsc{pic} methods too, the charge and current densities are computed in each simulation cell and fed into Maxwell’s equations (Equations 3.2–3.5). Since the particle distribution function \( f_s (\mathbf{r}, \mathbf{v}, t) \) expresses the number density of particles of species \( s \) at a time \( t \) in six dimensions, that is in a volume \( d^3r d^3v = dx dy dz d\nu_x d\nu_y d\nu_z \) of the \((\mathbf{r}, \mathbf{v})\)-space, the total number of particles \( N_s \) is given by

\[
N_s = \int_{(\mathbf{r})} \int_{(\mathbf{v})} f_s (\mathbf{r}, \mathbf{v}, t) d^3r d^3v,
\]

which defines the normalisation of \( f_s \). Following that principle, the (spatial) number density \( n_s \) is easily obtained as

\[
n_s (\mathbf{r}, t) = \int_{(\mathbf{v})} f_s (\mathbf{r}, \mathbf{v}, t) d^3v.
\]

This is the zeroth-order velocity moment of \( f_s \). The first-order moment

\[
\mathbf{\Gamma}_s (\mathbf{r}, t) = \int_{(\mathbf{v})} \mathbf{v} f_s (\mathbf{r}, \mathbf{v}, t) d^3v
\]
is the flux vector of particles at each point $\mathbf{r}$, which yields the bulk velocity $\mathbf{V}_s$ when normalised by the particle density:

$$\mathbf{V}_s(\mathbf{r}, t) = \frac{\Gamma_s(\mathbf{r}, t)}{n_s(\mathbf{r}, t)} = \frac{\int_{(v)} v f_s(\mathbf{r}, \mathbf{v}, t) d^3\mathbf{v}}{\int_{(v)} f_s(\mathbf{r}, \mathbf{v}, t) d^3\mathbf{v}}.$$  \hspace{1cm} (3.13)

The current density is then simply $\mathbf{j}_s = q_s \Gamma_s = n_s \mathbf{V}_s$. One step further lies the pressure tensor $\Psi_s$, the second-order velocity moment of $f_s$ (note the tensor product inside the integral):

$$\Psi_s(\mathbf{r}, t) = m_s \int_{(v)} (\mathbf{v} - \mathbf{V}_s) (\mathbf{v} - \mathbf{V}_s) f_s(\mathbf{r}, \mathbf{v}, t) d^3\mathbf{v},$$  \hspace{1cm} (π)

which can be used to define the temperature tensor $\Xi_s$ with $\Xi_s = n_s k_B \Psi_s$. The latter two become the scalar pressure $P_s$ and temperature $T_s$ in the case of an $f_s(\mathbf{r}, \mathbf{v}, t)$ spherically symmetric in $\mathbf{v}$.

Indeed, the Vlasov equation is great to simulate collisionless magnetised plasmas and a lot of exciting plasma physics such as wave propagation and instabilities can be readily modelled with it [e. g. Eliasson and Shukla, 2004, Umeda et al., 2012]. One of the major advantages when modelling $f_s$ instead of using randomly-distributed macroparticles in PIC methods is the absence of statistical sampling noise. But the astute plasma modeller probably started scratching her/his head somewhere around Equation (3.7) and wondering whether this is not going to consume a tremendous amount of computation and a humongous space in memory and on disk. Here is namely the crux of Vlasov methods: modelling a large system with good resolution in all six space and velocity dimensions means going for a huge number of sampling points. That’s why propagating $f_s$ turns out to be at the same time a major disadvantage with respect to PIC simulations: whereas the latter can live with a relatively low number of macroparticles per grid cell – there may be noise but the physics is essentially correct –, a decent sampling of $f_s$ in 3D velocity space requires a minimum resolution which cannot be too coarse or the moments obtained are grossly incorrect. Historically, Vlasov-based simulations have thus been used mainly to model reduced systems in 1D or 2D in space and velocity. To simulate larger systems with the Vlasov equation, top-tier supercomputing resources were and are still needed [e. g. PRACE, 2015].

Back on the Tour de France, knowing the density and velocity distribution of cyclists for each square metre of road throughout the day is certainly comprehensive, but maybe too much information?

### 3.3 Magnetohydrodynamics

The previous Section 3.2 introduces the concept of the particle distribution function and its moments, but the resulting equations such as Boltzmann’s
equation (3.9) are still on the computationally heavy side. Yet the plasma theoretician is by no means at the end of her/his wits, more simplifications can be undertaken to proceed.

One very popular set of equations are the magnetohydrodynamic (MHD) equations. They can be built as an extension to the Navier-Stokes equations including electric and magnetic fields, or derived from Boltzmann’s equation. The steps taken in the latter case are the following:

1. Consider the plasma as a single fluid by summing the equations and moments over the various particle species;

2. Take the zeroth, first, second, ... velocity moments of Boltzmann’s equation and combine with Maxwell’s equations (Equations 3.2–3.5) to obtain a set of equations describing the dynamics of the plasma moments (density, velocity, pressure, ...) and the electric and magnetic fields;

3. Choose an equation of state providing an appropriate closure of the system, in the case of MHD usually \( P = P_0 (n/n_0)^\gamma \) where \( \gamma \) is the polytropic index.

This procedure is not very involved and has been treated in detail before [e. g. Kempf, 2012], hence it is not written out in this thesis. The key assumptions and simplifications underlying the three steps above are to be emphasised nevertheless:

1’. In a single fluid, no effects related to multiple species or multiple populations are included;

2’. Only moments are included in the equations, thus neglecting all kinetic effects;

3’. The polytropic equation of state entails the assumption that a scalar, isotropic pressure suffices to describe the plasma, instead of a full tensor. Closing the equation system at this level excludes higher moments such as the third-order heat flux tensor.

It is also assumed along the way that the electron mass can be neglected with respect to the ion masses \( (m_i \gg m_e) \). The resulting system of equations takes up a few more lines when writing it out, but it is computationally much lighter
CHAPTER 3. SPACE PLASMA SIMULATION

to solve than the above-mentioned PIC and Vlasov/Boltzmann methods:

\[ \frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{V}) = 0 \]  
\[ \rho_m \left( \frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla \right) \mathbf{V} - \mathbf{j} \times \mathbf{B} + \nabla P = 0 \]  
\[ P = P_0 \left( \frac{n}{n_0} \right)^\gamma \]  
\[ \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \]  
\[ \nabla \times \mathbf{B} = \mu_0 \mathbf{j}. \]

This system thus comprises a continuity equation (3.15), a momentum equation (3.16), the equation of state (3.17) and the Maxwell–Faraday and Maxwell–Ampère equations (3.18 and 3.19). In the Maxwell–Ampère equation, it has to be noted that the displacement current is neglected, thus excluding light waves from the model (compare with Equation 3.5). The last equation closing the system is Ohm’s law, which provides a relation between the electric and magnetic fields. It is derived by manipulating the momentum transport equation. In its generalised form, it takes the shape

\[ \mathbf{E} + \mathbf{V} \times \mathbf{B} = \frac{j}{\sigma} + \frac{1}{q n_e} \mathbf{j} \times \mathbf{B} + \frac{1}{q n_e} \nabla \cdot \Psi_e + \frac{m_e}{q^2 n_e} \frac{\partial \mathbf{j}}{\partial t} \]  

where \( \sigma \) is the plasma’s electrical conductivity. With the assumption \( m_i \gg m_e \), one can readily drop the last term. Dimensional analysis shows that the second and third terms on the right-hand side are only important at spatial scales close to the ion kinetic scales, from which follows that they can be neglected in a fluid approach. Hence, the usual form of Ohm’s law in MHD is

\[ \mathbf{E} + \mathbf{V} \times \mathbf{B} = \frac{j}{\sigma} \]  

and in the case of very conductive plasma, it is simplified even further to

\[ \mathbf{E} = -\mathbf{V} \times \mathbf{B}. \]

What is left to work with in this description? The plasma has been reduced from a multi-species system of interacting electrons and ions with arbitrary velocities to a single fluid with a scalar pressure and no kinetic effects whatsoever. At best, MHD should only be fit to describe plasmas of homogeneous composition, in thermal equilibrium and exclusively at large time and spatial scales. Despite these theoretical limitations, MHD has found a wide use especially in space physics. Comparisons to observations have proven that although in principle MHD should not even properly apply, its results are reliable in many cases
and it can be applied to a lot of systems. Examples include simulations of the Sun and its cyclical variability [e. g. Käpylä et al., 2012], simulations of the solar wind and coronal heating [e. g. Evans et al., 2012], models of the solar wind propagation such as the Wang-Sheeley-Arge-Enlil model used for real-time solar wind forecasting [Parsons et al., 2011], global magnetospheric models such as the Grand Unified Magnetosphere-Ionosphere Coupling Simulation [GUMICS-4, Jahnunen et al., 2012], the Lyon-Fedder-Mobarry model [LFM, Lyon et al., 2004], the Block-Adaptive-Tree-Solarwind-Roe-Upwind-Scheme [BATS-R-US, Tóth et al., 2012] or the Open Geospace General Circulation Model [openGGCM, Raeder et al., 1998], as well as astrophysics with the Athena code [Stone et al., 2008] and even cosmology [e. g. Dolag and Stasyszyn, 2009]. Of course this list is by far not exhaustive.

Describing the Tour de France by just telling the route and the timing of the whole peloton is indeed a bit rough of a description, but isn’t that what most are interested in?

### 3.4 Hybrids and other combinations

As has become apparent in the first sections of this chapter, the choice of a model to solve a given plasma physical problem is often a matter of striking the right balance between the computational and the physical complexity. The more physics a model describes, the heavier it is to deploy. Conversely, a simple model might not include enough features to give a satisfying solution. The variety of mathematical approaches leading to equations describing plasma offers an endless playground to seek different models. A few of the more famous trails venturing into this vast landscape are indicated in this section.

Step 3 on the road to MHD in Section 3.3 requires to choose a closure to the system of equations in the form of an equation of state for the pressure. Indeed, taking increasing moments of Boltzmann’s equation produces a chain of equations expressing the $n^{th}$ moment of $f$ as a function of the moment of order $n + 1$. This chain can be truncated at any order with adequate assumptions. In addition to this, the summation over all species undertaken in step 1 can be skipped in order to take moments over electrons and ions separately and then couple the resulting equations by means of Maxwell’s equation. Models derived with this method are called multi-fluid as each species is described as a separate fluid. Common models are the five- and ten-moment two-fluid methods, where protons and electrons as well as their moments $(n, V, P)$ and $(n, V, \Psi)$ respectively are retained. This amounts to 5 respectively 10 moments ($\Psi$ is a symmetric $3 \times 3$ tensor) being taken into account for each species. More such two-fluid multi-moment methods exist, as the review of 5-, 8-, 10-, 13- and 20-moment methods by Schunk [1977], for example, shows. Being computationally lighter than PIC or Vlasov/Boltzmann methods, they are a seducing option including more kinetic physics than MHD [e. g. Wang et al., 2015].
As pointed out by Schunk [1977], the multi-fluid methods require that \( f \) be of a form which can be easily integrated; in practice it is assumed to be a Maxwellian. This restricts the validity of the models when situations arise where \( f \) is anisotropic, as is often the case in magnetised plasmas, where \( B \) and the Lorentz force naturally define a preferential direction. Going from MHD towards more complexity, a set of single-fluid equations with both parallel and perpendicular pressure \( P_\parallel, P_\perp \) has been first derived by Chew, Goldberger, and Low [1956] and is commonly known as the CGL equations. Going from Vlasov’s equation towards less complexity, another approach is to assume that \( f \) is symmetric around the magnetic field direction – gyrotropic – so that the azimuthal degree of freedom around \( B \) can be integrated over. This results in a different class of models, gyrokinetics, particularly popular in fusion plasma research [e.g. Dannert and Jenko, 2005, Angioni et al., 2009].

Many a plasma modeller’s headache originates in the disparity of scales between species. The ratios of the characteristic lengths and frequencies are equal to the ratio of the masses or its square root:

\[
\frac{\omega_{ce}}{\omega_{ci}} = \frac{\omega_{pe}}{\omega_{pi}} = \frac{r_{Li}}{r_{Le}} = \frac{d_{i}^{2}}{d_{e}^{2}} = \frac{m_{i}}{m_{e}}
\]

which in the best case with protons is approximately equal to 1836 (or its square root, defined to be 42 for mnemotechnic reasons). Thus during the time a proton leisurely surveys its large Larmor circle, a puny little electron flying at the same speed spins 1836 times around its 1836 times shorter path. Resolving electron and ion kinetics in a larger system encompassing fluid scales involves a scale separation in time stepping and spatial resolution which is not easily achievable, even with supercomputers.

Especially with this handicap in mind, it is worth questioning whether any of the requirements can be relaxed. The easiest is to relinquish the global nature of the simulation and not aim for hundreds of ion gyroperiods and inertial lengths – a restriction often not compatible with the nature of the research considered. The alternative is to ponder the need for detailed electron and ion physics simultaneously. The class of hybrid models emerges from this reasoning, describing ions kinetically while relying on a simpler fluid model for the electrons. Undoubtedly, the inherent loss of kinetic electron physics has to be justified, but the unmistakable bonus of such a hybrid is that the minimal resolution requirements correspond to the ion and no longer to the electron scales.

The most widely-used hybrid models combine PIC ions with MHD equations to close the system. They have been successfully applied to model magnetospheres in global contexts including ion kinetics [e.g. Lin and Wang, 2005, Omidi et al., 2005, Blanco-Cano et al., 2006, Karimabadi et al., 2014]. A completely equivalent scheme is obtained by replacing the PIC ions with a velocity

\[
\frac{\omega_{ce}}{\omega_{ci}} = \frac{\omega_{pe}}{\omega_{pi}} = \frac{r_{Li}}{r_{Le}} = \frac{d_{i}^{2}}{d_{e}^{2}} = \frac{m_{i}}{m_{e}}
\]
distribution. Due to the computational challenge it represents to simulate large systems with the hybrid-Vlasov method, it is not yet as common as hybrid-

A number of efforts are undertaken though and give excellent results [e. g. Valentini et al., 2007, Perrone et al., 2014, Pokhotelov et al., 2013]. Given that the rest of this thesis presents in depth the hybrid-Vlasov model Vlasiator, this very brief introduction of hybrid models shall stop here.

Concluding this chapter on plasma modelling, it should be mentioned that numerous other stratagems can be devised and combined to achieve specific goals.

To alleviate the pains of large mass ratios, PIC methods can be run with mass ratios lower than 1836 to reduce the gap between electron and ion scales. In the works cited above, the ratio is 50 [Karimabadi et al., 2014] respectively 84 [Savoini et al., 2013]. In this way, more electron physics is retained than with fluid electrons in a hybrid model.

In large, heterogeneous systems such as the magnetosphere or the propagation of eruptions from the Sun’s surface to the Earth, it is conceivable that the same level of physical detail is not necessary throughout the simulation. Consequently, a lot of work has been invested in attempts to couple fluid and kinetic models, providing a large-scale context with MHD or multi-fluids and detailed kinetics in selected regions of interest [e. g. Lapenta et al., 2013, Rieke et al., 2015, Tóth et al., 2016].

The present chapter attempts to give an overview of physical approaches that can be taken individually or combined to simulate plasmas. Each has its advantages and inconveniences and the choice is largely problem-specific. The second aspect, which would take an equally detailed chapter to treat if not a full textbook, is the discretisation of the algorithms to ensure a swift computation of the solution. This comprises both fundamentally different mathematical strategies to solve differential equations and more or less flexible alternatives regarding the efficient implementation and execution on (super)computers. It is not covered at all in this work, however the next Chapter 4 is focussing on the physical and technical options taken to develop the hybrid-Vlasov model Vlasiator and achieve its set goal of exceeding the limits of MHD in the simulation of the Earth’s magnetosphere.
Chapter 4

Vlasiator

 мер Take Five
Dave Brubeck

 мер Long Island Iced Tea – tequila, vodka, rum, triple sec, gin and cola
A true 60s drink

The year is AD 2016. Global magnetospheric physics is entirely occupied by MHD and PIC. All? Not quite! For a research group of indomitable Finns still holds out against the world and claims that hybrid-Vlasov global magnetospheric simulations are possible [Palmroth et al., 2013, Pokhotelov et al., 2013, Sandroos et al., 2015, Palmroth et al., 2015, Hoilijoki et al., 2016, 2017, Paper II, Paper III and Paper IV].

This chapter describes Vlasiator as it stands in 2016. The main features of the algorithm are presented without however going into implementation details. The source code’s git repository can be browsed, forked and cloned at https://github.com/fmihpc/vlasiator, current information on the code, data, science and team can be found at http://vlasiator.fmi.fi. References to files in the Vlasiator repository such as the main file vlasiator.cpp are given relative to the base directory of the source in the text and linked to https://github.com in the pdf version of the thesis.

The next Chapter 5 covers the author’s major code contributions pertaining especially to the field solver and the boundary conditions.

4.1 The hybrid-Vlasov model

Vlasiator treats H$^+$ ions (protons) kinetically through Vlasov’s equation (3.8):

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial r} + \frac{q}{m_i} (E_v + v \times B) \cdot \frac{\partial f}{\partial v} = 0. \tag{4.1}$$
CHAPTER 4. VLASIATOR

For simplicity from here on $f$ denotes the ion distribution function unless mentioned otherwise. The electric field $E_v$ is computed in a fashion similar to MHD using the following generalised Ohm’s law:

$$E_v = -V_i \times B + \frac{1}{qn_i} j \times B$$

$$\Leftrightarrow E_v = -V_i \times B + \frac{1}{\mu_0 q n_i} (\nabla \times B) \times B,$$  \hspace{1cm} (4.2)

where the second step is derived from Ampère’s law (3.5) again under the assumption that the variation of $E$ in time is slow, yielding $j = \nabla \times B/\mu_0$.

The time evolution of the magnetic field is computed using Faraday’s law (3.4), which can be rewritten to yield the increment of the magnetic field $\Delta B$ as a function of the increment in time $\Delta t$ and the electric field $E_v$:

$$\Delta B = -\Delta t \cdot \nabla \times E_v.$$  \hspace{1cm} (4.3)

The electric field computed with the generalised Ohm’s law (4.2) can be used in this equation, in which case $E_f = E_v$. Another option is to use a yet more simplified Ohm’s law

$$E_f = -V_i \times B,$$  \hspace{1cm} (4.4)

where the Hall term $j \times B/qn_i$ has been neglected. When using the latter electric field (4.4) to compute the magnetic field increment (4.3), the resulting scheme is a Hall-less hybrid algorithm. The Hall term cannot be neglected in $E_v$ (4.2) used in Vlasov’s equation (4.1) though, otherwise the ion momentum equation is violated and no force is exerted on the ions [Karimabadi et al., 2004]. The importance of the Hall term in plasma wave dispersion is the subject of Paper I and is discussed in Section 6.1 of this thesis. Paper II and Paper III however show that on global scales, a Hall-less $E_f$ is sufficient to capture the relevant physical phenomena, at least as long as ion kinetic scales are not properly resolved.

This set of equations makes Vlasiator a hybrid-Vlasov model, which can be Hall-less if $E_f$ from the simplified Ohm’s law (4.4) is used to propagate $B$ in time. The whole algorithm can be summarised as follows. The superscripts $n$ and $n + 1$ are used to designate the value of variables at time $t_0$ and after one time step at time $t_0 + \Delta t$, respectively. Assuming that the initial values $f^n$ and $B^n$ are known, which means that the moments of $f^n$ such as $V^n_i$ are also known, $E^n_v$ and $E^n_f$ are computed with Ohm’s law (4.2) and (4.4) in the Hall-less case:

$$E^n_v = -V^n_i \times B^n + \frac{1}{\mu_0 q n^n_i} (\nabla \times B^n) \times B^n,$$  \hspace{1cm} (4.5)

$$E^n_f = -V^n_i \times B^n.$$  \hspace{1cm} (4.6)
4.2. DISCRETISING A 6D SPACE

Vlasov’s equation (4.1) can be rewritten in terms of increments as well, yielding:

\[ f^{n+1} = f^n + \Delta f^n \]

\[ = f^n - \Delta t \left( \mathbf{v} \cdot \frac{\partial f^n}{\partial \mathbf{r}} + \frac{q}{m_i} (E^n_{\mathbf{r}} + \mathbf{v} \times B^n) \cdot \frac{\partial f^n}{\partial \mathbf{v}} \right). \quad (4.7) \]

Finally, \( B^n \) is advanced using Equation (4.3)

\[ B^{n+1} = B^n + \Delta B^n \]

\[ = B^n - \Delta t \cdot \nabla \times E^n. \quad (4.8) \]

With equations (4.5) to (4.8) the state at \( t_0 + \Delta t \) is known and the same steps can be used to progress from \( f^{n+1} \) and \( B^{n+1} \) to \( f^{n+2} \) and \( B^{n+2} \), and so forth.

4.2 Discretising a six-dimensional space

The position and velocity spaces are discretised separately in Vlasiator. \( \mathbf{E}, \mathbf{B}, n_i, \mathbf{V}_i \) and \( \psi_i \) as well as a few housekeeping variables are stored for each spatial cell on a uniform Cartesian grid. This grid is 1D, 2D or 3D depending on the simulated case. Alongside these variables, the 3D velocity distribution function at every spatial cell position \( \mathbf{r} \) is discretised on a uniform Cartesian grid in velocity space and stored in that spatial cell.

The storage and handling of this data in the cells is performed using the Distributed Cartesian Cell-Refinable Grid (DCCRG) library [Honkonen et al., 2013, DCCRG, 2016]. No grid refinement is used. DCCRG decomposes the position space into domains distributed among the MPI processes when distributed-memory parallelisation is activated. All detail of inter-process communication and bookkeeping of ghost cells, which hold values on one process corresponding to a neighbouring cell solved by another process, is delegated to DCCRG. Load balancing, ensuring that even in an inhomogeneous simulation each process handles approximately the same amount of data, is delegated further by DCCRG to the Zoltan library [Devine et al., 2002, Boman et al., 2016].

The phase-space cells store the volume average \( \bar{f} \) of the distribution function over their 6D volume:

\[ \bar{f} = \frac{1}{\Delta^3 r \Delta^3 v} \int_{\text{cell}} \mathbf{d}^3 r \mathbf{d}^3 v f(\mathbf{r}, \mathbf{v}, t), \quad (4.9) \]

where \( \Delta^3 r \Delta^3 v = \Delta x \Delta y \Delta z \Delta v_x \Delta v_y \Delta v_z \) denotes the phase-space integration volume, and \( \Delta x, \Delta y, \) etc. are the cell dimensions. For simplicity, \( f \) stands for the volume average of the distribution function over a 6D cell from here on.

In the early version of Vlasiator used in Paper I, Paper III and other works [Palmroth et al., 2013, Pokhotelov et al., 2013], and as explained in detail in
Paper II, $f$ is propagated forward in time by calculating fluxes at every cell face in all dimensions. The nature of Vlasov’s equation gives a simple form to the spatial $F_r = (F_x, F_y, F_z)$ and velocity $F_v = (F_{vx}, F_{vy}, F_{vz})$ fluxes:

\[
F_r = vf, \quad (4.10)
\]

\[
F_v = \frac{q}{m_i} \left( \mathbf{v} - \mathbf{V}_i + \frac{1}{\mu_0 q n} \nabla \times \mathbf{B} \right) \times \mathbf{B} f, \quad (4.11)
\]

as can be readily seen from Equation (4.7). Because $f$ is a volume average, it is propagated using a volume-averaged $\mathbf{B}$ computed with a divergence-free reconstruction method [Balsara, 2009, details on the field solving and reconstruction algorithms are given in Section 4.3]. The propagation of $f$ is obtained by incrementing it by the fluxes in and out of the cell in all directions:

\[
f^{n+1} = f^n - \sum_{j \in \{x,y,z, v_x, v_y, v_z\}} \frac{\Delta t}{\Delta j} \left[ F^e_j (j + \Delta j) - F^e_j (j) \right]. \quad (4.12)
\]

The fluxes are computed by means of a 3D wave propagation algorithm solving the Riemann problem of a propagating medium with a discontinuity at the cell interface [LeVeque, 1997, Langseth and LeVeque, 2000]. By construction, the scheme (4.12) is conserving mass except at the simulation domain boundaries.

Despite being a robust and reliable algorithm, this method is limited by a Courant-Friedrichs-Lewy (CFL) condition making it numerically unstable if $f$ is propagating further than into the next cell in one time step [Courant et al., 1928]. This restricts the time step length $\Delta t$:

\[
\Delta t < \min \left( \frac{\Delta x}{v_x}, \frac{\Delta y}{v_y}, \frac{\Delta z}{v_z}, \frac{\Delta v_x}{a_x}, \frac{\Delta v_y}{a_y}, \frac{\Delta v_z}{a_z} \right). \quad (4.13)
\]

This type of restriction is common to many algorithms modelling the propagation of a field in a fixed Eulerian grid.

To circumvent the time step limitation posed by the algorithm’s CFL condition, a semi-Lagrangian algorithm was implemented for the propagation of $f$. This is the method used in Paper IV, other works [Palmroth et al., 2015, Hoilijoki et al., 2016, 2017] and more recent production runs.

Lagrangian (and semi-Lagrangian) methods rely on the property that the solution $g(\mathbf{r}, t)$ to the advection equation

\[
\frac{\partial g}{\partial t} + \mathbf{V} \cdot \nabla g = 0, \quad (4.14)
\]

where the velocity field $\mathbf{V}$ is known, is of the form

\[
g(\mathbf{r}, t) = g(\mathbf{r} - (t - t_0) \mathbf{V}, t_0). \quad (4.15)
\]
4.2. DISCRETISING A 6D SPACE

Put in words, $g$ at time $t$ and position $r$ is the same as $g$ at the previous time $t_0$ at the location which flowed with the velocity $V$ to reach $r$ at time $t$, namely $r - (t - t_0) V$. This solution is unconditionally stable, provided that $V$ is known everywhere at all times. When using this approach to solve an advection equation numerically, often $V$ is not known \textit{a priori} as it is being computed as well. The method can however be applied stepwise by computing $V^n$ from $g^n$ and using it to compute

$$g^{n+1}(r) = g^n(r - \Delta t V^n).$$

Using this formal solution to follow the propagation of a small volume of the fluid $g$ results in a Lagrangian method. It can become impractical because of the inherent deformation of the Lagrangian volumes with the flow, and the possible accumulation of such Lagrangian volumes in some regions of the simulation leaving other regions badly covered. Semi-Lagrangian approaches use the stepwise solution (4.16) in combination with a fixed Eulerian grid instead, to retain a meaningful coverage of the simulation domain. The decisive aspect immediately stands out: the starting point of the flow $r - \Delta t V^n$ needed to obtain the value $g^{n+1}(r)$ is unlikely to be located exactly at a Eulerian grid point where the values $g^n$ are known. Semi-Lagrangian algorithms rely on interpolating $g^n$ from the fixed Eulerian grid points to the location $r - \Delta t V^n$ to obtain the values $g^{n+1}$. The choice of the interpolation technique is aimed at preserving the monotonicity of solutions and conserving $g$. Even though the semi-Lagrangian method is numerically stable for any $\Delta t$, the convergence towards a sensible solution is of course paramount and the length of the time step is guided by the fastest relevant dynamics in the problem considered.

One of the beautiful aspects of Vlasov’s equation is that the acceleration of $f$ in velocity space (4.11) is nothing more than a rotation around $B$, the gyromotion, in the frame where $-V_i + \nabla \times B / (\mu_0 q n) = 0$. Hence the semi-Lagrangian transformation to accelerate $f$ is constructed from a translation, a rotation around $B$ and a translation back to the original frame, while the spatial translation is by definition just that. Essentially, the algorithm amounts to mapping $f$ from the original grid to another rotated and/or translated grid.

Vlasiator uses a 3D conservative semi-Lagrangian scheme developed originally for weather prediction models [Slice3D, Zerroukat and Allen, 2012] and adapted to the separate spatial translation and velocity space acceleration of $f$, similarly to the split of the flux propagation method in Equations (4.10) and (4.11). Slice3D is based on the use of volume averages of the variable to propagate. The scheme is dimensionally split: the interpolation and mapping to the transformed grid is performed one dimension at a time, which corresponds a translation and a shear transformation of the grid in each dimension. The interpolation scheme uses piecewise quartic polynomials [Piecewise Quartic Method (pQM), White and Adcroft, 2008] for the acceleration and piecewise quadratic polynomials [Piecewise Parabolic Method (pPM), Colella and Woodward, 1984] for the translation part of the propagation respectively. A
Piecewise Linear Method (PLM) is also implemented in Vlasiator and any combination of PQM, PPM and PLM is allowed for acceleration and translation. The combination of PQM for acceleration and PPM for translation is found to be the best trade-off between the quality of the interpolation results and the computational weight of the interpolation steps in Vlasiator. The implementation of the SLICE3D algorithm is available as a standalone library [lib-slice3d, 2016].

The geometric simplicity of the transformations allows to use the external library Eigen to perform the geometrical transformations [Eigen, 2016]. In the absence of a formal CFL condition, the configurable parameter for the Vlasov solver acceleration is the maximum rotation angle around \( B \), following physical considerations. The SLICE3D algorithm does have limitations, for example a rotation of exactly 90° is not possible, but they are beyond the range of sensible physical limits. For the translation part of the Vlasov solver, a CFL limit is set to prevent the propagation of \( f \) further than into the neighbouring cells. This is done so that the communication of ghost cell information only needs to be performed over a fixed neighbourhood of one cell depth.

### 4.3 The field solver

The field solver propagating \( E \) and \( B \) uses the upwind constrained transport method proposed by Londrillo and Del Zanna [2004]. Its foremost quality is its conserving the divergence of \( B \) to numerical accuracy [as shown in the author’s previous work, Kempf, 2012], therefore conserving the solenoidality of \( B \) provided the initial and boundary conditions do not introduce divergence. This property relieves from devising a divergence-cleaning method removing spurious numerical monopoles appearing in algorithms which do not ensure \( \nabla \cdot B = 0 \) down to numerical accuracy.

On the spatial grid, the electric field components are stored as values averaged along the cell edge pointing to the corresponding direction while the magnetic field components are stored as values averaged over the cell face orthogonal to the corresponding direction, as illustrated in Figure 4.1. The use of such a staggered grid makes the formulation of a discretisation scheme of the field propagation equations (4.5) and (4.8) particularly intuitive. The advantage is clear when reformulating Equation (4.8) in integral form:

\[
\Delta \vec{B} = -\Delta t \nabla \times \vec{E} \Leftrightarrow \int_{\text{face}} \Delta \vec{B} \cdot \vec{dS} = -\Delta t \oint_{C} \vec{E} \cdot \vec{ds},
\]

where \( \vec{dS} \) is the unit vector pointing orthogonally to the cell face and \( \vec{ds} \) is the unit vector pointing along the directed path following the edges around the face right-handedly. Dropping the integration symbols as the face and edge averaging on the staggered grid is implicitly understood, the increment in \( B_z \)
for the cell located at coordinate indices \((i, j, k)\) can be written out as

\[
\Delta B_z \, dx \, dy = -\Delta t \left( (E_{x(i,j,k)} - E_{x(i,j+1,k)}) \, dx + (E_{y(i+1,j,k)} - E_{y(i,j,k)}) \, dy \right),
\]

where \(dx\) (respectively \(dy, dz\)) is the cell length in \(x\) (respectively \(y, z\)) direction and similar expressions hold for \(B_x\) and \(B_y\). In this form, \(\Delta \mathbf{B}\) is analytically the curl of \(\mathbf{E}\) and thus \(\nabla \cdot \Delta \mathbf{B} = 0\) by construction. The computation of \(\mathbf{E}\) using Equation (4.5) is correct if edge-averaged values for \(\mathbf{V}_i\) and \(\mathbf{B}\) are used. In Vlasiator, \(\mathbf{V}_i\) and \(\mathbf{B}\) are interpolated to and averaged along the edges using the divergence-free reconstruction method proposed by Balsara [2009]. As mentioned in Section 4.2, the same reconstruction scheme is used to compute the volume-averaged \(\mathbf{B}\) needed to propagate \(f\) with Vlasov’s equation (4.7).

The field solver is subject to a CFL condition limiting its stability if the time steps are too long. In principle, no plasma wave is allowed to travel faster than one cell width per time step, and this is even reduced to half a cell width in the field solver used [Londrillo and Del Zanna, 2004, Paper II]. The relevant wave modes are added to the flow speed to determine the maximally allowed time step. The magnetosonic speed has to be taken into account, as well as the whistler wave speed, a mode only present when the Hall term is present in Ohm’s law [e.g., Paper I]. This mode is dispersive and has a higher group and phase speed with increasing \(k\). The whistler speed on a discrete grid is

\[
\nu_w = \frac{2\pi v_A^2 m_i}{dx qB},
\]

meaning that the limiting time step length, which depends on \(dx/\nu_w \sim dx^2\), decreases quadratically with the resolution of the simulation grid in regions where the whistler speed dominates.
4.4 From six to seven dimensions: time stepping

Having described how the various solver parts work individually, it is time to explain how they are working together. Section 4.2 shows that the spatial translation $S_T$ and the acceleration in velocity space $S_A$ of $f$ are separated. Vlasiator uses the method of Strang [1968] to split the propagation so that $f^n$ can be understood as the result of $n$ applications of a time-stepping operator:

$$f^n = \left[ S_T \left( \frac{1}{2} \Delta t \right) S_A (\Delta t) S_T \left( \frac{1}{2} \Delta t \right) \right]^n f^0. \quad (4.20)$$

Consecutive applications of $S_T \left( \frac{1}{2} \Delta t \right)$ can be combined into a single propagation of length $\Delta t$, so that the acceleration and translation operations are shifted with respect to each other by half a time step, a method called the leapfrog algorithm. It is initialised by half a time step of translation

$$f_r \left( t_0 + \frac{\Delta t}{2} \right) = S_T \left( \frac{\Delta t}{2} \right) f$$

and all further steps are advancing with a complete $\Delta t$:

$$f_v \left( t + \Delta t \right) = S_A (\Delta t) f_r \left( t + \frac{1}{2} \Delta t \right) \quad (4.22)$$
$$f_r \left( t + \frac{3}{2} \Delta t \right) = S_T (\Delta t) f_v \left( t + \Delta t \right) \quad (4.23)$$

where the subscripts $r$ and $v$ indicate the state of $f$ after a translation and an acceleration step, respectively. This Strang splitting scheme is used with both the wave-propagation-based and the semi-Lagrangian Vlasov solvers.

The field solver propagates $E$ and $B$ between the translation and acceleration step. The simplest option is a first-order Euler step using $n^n$, $V^n$, and $E^n$ to compute $B^{n+1}$ and then $E^{n+1}$. A more accurate algorithm using a second-order Runge-Kutta method also exists. As this was implemented by the author of the present thesis, it is the subject of its own detailed Section 5.1.

Dynamic plasma simulations require the time step length to be adapted during the simulation. The range of allowed CFL numbers – the ratio between the distance travelled by the fastest wave mode and the cell length – is configured by the user where applicable. For the semi-Lagrangian Vlasov solver, a maximum rotation angle of the velocity distribution is set and a maximum CFL number of 1 is enforced to make sure that $f$ only flows into the nearest neighbours, facilitating the handling of ghost cell communication. A check is run at each time step during the simulation to ensure the limits are not exceeded. If the time step would be going over (or below) the admissible range, $f$ is translated half an old time step backwards and half a new time step forward again, after which the simulation is run with the new time step.
4.5 Making kinetic global magnetospheric simulations possible

Some of the tools and methods which allow Vlasiator to model the terrestrial magnetosphere in a global context using the hybrid-Vlasov approach deserve a special mention in this section.

It is important to track the efficiency of any high-performance computing application, both at the global level and at more fine-grained levels. An offspring of the Vlasiator development is the Parallel Hierarchical Profiler phiprof library [phiprof, 2016]. It uses a simple interface based on start and stop functions to time the execution of the enclosed code region. It supports MPI parallelisation and OpenMP threading, an arbitrary number of nested levels and reports the results in ASCII format. Good profiling is a cornerstone as detailed knowledge of the performance of individual code sections allows targeted optimisations, which are the most beneficial.

On the physical side, the three major solver parts – field solver, Vlasov translation and Vlasov acceleration – have very different computational characteristics. The field solver only handles E and B with the knowledge of \( n \), \( V_i \) and \( \Psi_i \), so 13 scalars per spatial cell and their derivatives, which involve only the directly neighbouring cells. The Vlasov solver has to work with the 6D distribution function, but whereas translation by essence involves flows from and into the spatially neighbouring cells, acceleration is local to the spatial cell. This means for example that the field solver and the Vlasov acceleration involve updating of the neighbouring cell information to ensure all processes are up to date while acceleration can proceed without such an update.

Since the field solver is relatively lightweight, acceleration involves less communications, and to allow the use of longer translation time steps, the acceleration and field solvers can be subcycled. In order not to break the CFL condition imposed by the whistler wave speed, the field solver can execute a number of first-order time steps. Similarly, in order not to accelerate the velocity distribution by an exaggerated amount in one step, the acceleration can be done in several subcycles. Again, this is not imposed by a CFL limitation but by the largest angle the person in charge of the simulation feels comfortable rotating \( f \) around \( B \) in velocity space without compromising the physical results. The main concerns are the coupling to the other solver parts when \( f \) rotates by large amounts in one step and the introduction of larger interpolation artefacts with increasing rotation angles. The maximum angle is set to 44° in the large magnetospheric simulations presented in this work, split in at most 2 to 4 subcycles occurring exclusively in the region close to the inner boundary where the dipole field is strongest.

In global magnetospheric simulations, the need also arises to treat spatial derivatives of the magnetic field with care to ensure a good quality of interpo-
lations and reconstructions even in regions with strong gradients, such as in the vicinity of the dipole’s origin. Vlasiator splits the magnetic field into two components, a constant background field and the perturbed field component added to the background. The background field, its derivatives and its volume-averaged values are computed analytically upon initialisation and stored for later use, while the field solver only computes the perturbed part of the magnetic field which is then added to the background field. The background field is required to be curl-free so that computations pertaining to the Hall term do not include the background field. This method reduces significantly the inaccuracy of the field interpolations near the dipole’s origin in magnetospheric simulations, thus resulting in better stability of the model.

The single biggest gain in computational efficiency in the development of Vlasiator is the application of the concept of sparse velocity space. It stems from the observation that a simple velocity distribution function, such as a Maxwellian with a thermal speed of the order of 100 km/s drifting at 500–1000 km/s, occupies a much smaller volume than the whole velocity space possibly reached within a magnetospheric simulation, which can easily be ±2000 km/s in every direction. By setting an adequate phase space density threshold $f_{\text{min}}$, it is possible to disregard a sizeable fraction of velocity space (in practice upwards of 90% of it) and neither propagate nor store $f$ where $f < f_{\text{min}}$. The overhead incurred by controlling the propagation of $f$ and (de)allocating memory for the phase space regions becoming (in)significant, which is of the order of 10% of the computation time, is largely compensated by the reduction in memory usage and computations. Vlasiator is the first Vlasov model using such an approach, to the knowledge of the team. To stretch one last time the Tour de France metaphor, it pays off not to report on the empty sections of the road.

In practice, velocity space cells are organised in memory in blocks of $4 \times 4 \times 4$ cells. If $f > f_{\text{min}}$ in any of the cells of a block $\Lambda$, the neighbouring blocks in all six dimensions are retained so that $f$ can propagate from $\Lambda$ in any direction. If a block becomes significant because its $f$ increases above $f_{\text{min}}$, all its neighbours are created if they were not yet allocated. Conversely, a block whose neighbours all lie below $f_{\text{min}}$ is deallocated. Although this mechanically breaks the conservation of $f$ inherent to the Vlasov equation and thus the conservation of mass, if $f_{\text{min}}$ is sufficiently low and the velocity space resolution not too coarse, the mass losses are negligible [see Paper II, for a quantification using the wave-propagating Vlasov solver]. Figure 4.2 illustrates the principle with a simple velocity distribution taken from a simulation. The parts of $f$ above $f_{\text{min}}$ are in full colour in the right panel, while the parts retained in neighbouring cells and blocks are greyed out. The block 4 for example is not stored as it is not neighbouring to any block with $f > f_{\text{min}}$.

In recent simulations posterior to the ones presented in Papers I to IV, adaptive schemes are introduced to optimise the threshold $f_{\text{min}}$. Instead of a
constant value throughout the simulation box and for the whole run, $f_{\text{min}}$ can be made to vary proportionally to the spatial cell’s density within a certain range. This allows to better target cold and low-density regions like the magnetotail without increasing excessively the load of hot and dense regions like the magnetosheath.

On the technical side, it comes as no surprise to the reader that all the parallelisation techniques presented in Section 1.3 are leveraged in Vlasiator. As explained in Section 4.2, the DCCRG library is used to perform domain decomposition. The up to 3D position space is decomposed into domains assigned to MPI processes communicating the necessary ghost cell information to each other. As this inter-process communication yields some overhead, classic staggered communication schemes are used, in which the simulation cells which have all the necessary neighbour information perform their computations while the process boundary cells are exchanging neighbour information.

The next level of performance when scaling to larger systems, especially on supercomputers where each node comprises tens of cores, is reached through hybrid parallelisation. OpenMP threading is used to process logically separate parts of the simulation domain, each thread handling typically one spatial cell or one velocity space block at a time, as this entails distinct memory accesses thus avoiding race conditions. The optimal ratio of threads per process is a sensitive parameter depending both on the architecture of the nodes and the communication network, and on the geometry of the simulation, as the proportion of spatial cells located at process boundaries varies in 1D, 2D or 3D.

Vectorisation is mostly applied to velocity space computations. The most intensive parts of velocity space computations are vectorised so that single
arithmetic operations are performed on eight numbers simultaneously. This is the fundamental reason for grouping the velocity space cells in blocks of $4 \times 4 \times 4$. Fog’s vector class library is used in Vlasiator, as it allows a relatively easy implementation of vectorised operations [Fog, 2016]. Substantial performance gains are observed when using vectorised operations with respect to the non-vectorised version.
Chapter 5

Major code contributions

🌱 • Ardbeg Supernova 2015
• Ardbeg Ten Years Old
🎶 L’Apprenti sorcier

Paul Dukas

No simulation can exist without substantial investments in sometimes long and often frustrating development and debugging tasks. Vlasiator is no exception. It is the growing and ripening fruit of many years of work by a dedicated team combining the skills and knowledge of the fields of space plasma physics and high-performance computing.

From the research funding perspective, initial and continuing model development is arguably one of the most expensive budget posts in modelling projects. From the same perspective, a project’s progress is often measured by the number of peer-reviewed publications resulting from it. Yet the extended description of a model and its implementation is often not deemed to be worth a scientific publication of its own.

The author of this thesis is also the author of significant parts of the Vlasiator source code, which justifies the emphasis put both on physical and computational aspects in the previous chapters. Some code developments are original work and critical to the successful modelling of the terrestrial magnetosphere at large scales. This chapter serves the dual purpose of acknowledging the behind-the-science technical achievements and providing a reference to future generations of Vlasiator users and developers.

5.1 Second-order field solver time stepping

Along with improvements in the Vlasov solver’s spatial and time accuracy, the field solver time stepping has been improved from a simple first-order accurate Euler method to a second-order Runge–Kutta method. The Runge–Kutta
algorithms are a large family of methods to solve differential equations using intermediate solution steps. As a reminder, the basic time progression of the field solver computes

\[ \mathbf{B}^{n+1} = \mathbf{B}^n - \Delta t \nabla \times \mathbf{E}^n, \]

\[ \mathbf{E}^{n+1} = -\mathbf{V}_{i}^{n+1} \times \mathbf{B}^{n+1} + \left( \nabla \times \mathbf{B}^{n+1} \right) \times \mathbf{B}^{n+1} / \left( \mu_0 q n_i^{n+1} \right). \]

Taking advantage of the fact that \( f \) is propagated forward using the leapfrog algorithm, two values of the moments can be interpolated from the moments computed after translation and acceleration respectively. \( n_i^n, \mathbf{V}_i^n, \mathbf{B}_i^n \) are computed before the translation and the acceleration step, while a set of moments \( n_{i}^{n+1/2}, \mathbf{V}_i^{n+1/2}, \mathbf{B}_i^{n+1/2} \) can be computed after the translation and before the acceleration [Eq. (20) and (21) in Paper II, Valentini et al., 2007]. With these intermediate moment values, one can extend the field solver time propagation to obtain a second-order accurate Runge-Kutta method:

\[ \mathbf{B}^{n+1/2} = \mathbf{B}^n - \frac{1}{2} \Delta t \nabla \times \mathbf{E}^n, \]

\[ \mathbf{E}^{n+1/2} = -\mathbf{V}_{i}^{n+1/2} \times \mathbf{B}^{n+1/2} + \left( \nabla \times \mathbf{B}^{n+1/2} \right) \times \mathbf{B}^{n+1/2} / \left( \mu_0 q n_i^{n+1/2} \right), \]

\[ \mathbf{B}^{n+1} = \mathbf{B}^n - \Delta t \nabla \times \mathbf{E}^{n+1/2}, \]

\[ \mathbf{E}^{n+1} = -\mathbf{V}_{i}^{n} \times \mathbf{B}^{n+1} + \left( \nabla \times \mathbf{B}^{n+1} \right) \times \mathbf{B}^{n+1} / \left( \mu_0 q n_i^{n} \right). \]

Then \( \mathbf{E}^{n+1} \) and \( \mathbf{B}^{n+1} \) are used for the next acceleration and translation steps.

Despite its improving the diffusive and stability properties of the field solver, the introduction of the Hall term in Ohm’s law brings changes which mean that the second-order accurate time stepping of the field solver is seldom if ever applied in magnetospheric simulations. The second-order time stepping is used in Paper I, Paper II and Paper III, but not any more in Paper IV or by Palmroth et al. [2015] and Hoilijoki et al. [2016, 2017] for the reason explained in the next section.

### 5.2 The Hall term in Ohm’s law

It is as major extension of the physics described to add the Hall term in Ohm’s law in a hybrid kinetic model as it is when going from ideal to Hall MHD.

The Hall term introduced in Vlasiator’s Ohm’s law and used in Paper I is of low spatial accuracy. It uses the spatial derivatives of the face-averaged \( \mathbf{B} \) components to compute \( \mathbf{j} = \nabla \times \mathbf{B} / \mu_0 \) and the volume-averaged \( n_i \) to compute the Hall term \( \mathbf{j} \times \mathbf{B} / (q n_i) \), which is directly added to the edge-averaged \( \mathbf{E} \) components.

The second-order accurate Hall term developed in 2013–2014 relies on the same divergence-free reconstruction method used elsewhere in the field solver,
for example to compute the volume-averaged \( B \) components for the Lorentz force in the Vlasov solver. This method was proposed by Balsara [2009]. Legendre polynomials serve as the function basis to interpolate \( B \) at any point within a spatial cell with the knowledge of the face-averaged components and their spatial derivatives. The first polynomials \( P(x) \) are given by

\[
P_0(x) = 1; \quad P_1(x) = x; \quad P_2(x) = x^2 - \frac{1}{12}; \quad P_3(x) = x^3 - \frac{3}{20}x.
\] (5.7)

Their scaling is adapted to a local coordinate system in which the cell spans \([-1/2; 1/2]\) in every dimension. Equations (7)–(9) by Balsara [2009] list expressions for \( B(x, y, z) \) anywhere within the cell up to fourth-order accuracy, in terms of the Legendre polynomials with coefficients built as linear combinations of the face-averaged components of \( B \) and their spatial derivatives to suitable order. The coefficients are listed in that paper’s Equations (10)–(14).

To compute \( j \times B \) for the Hall term, \( B \) is reconstructed to third-order accuracy using up to second derivatives, yielding a second-order accurate \( j = \nabla \times B/\mu_0 \). It is then crossed with a second-order accurate \( B \). The resulting expressions are integrated along the cell edges to obtain edge-averaged expressions for the components of \( j \times B \) which can be directly added to the edge-averaged \( E \) computed with the upwind constrained transport method from Londrillo and Del Zanna [2004]. The symmetry of the local coordinate system and of the Legendre polynomial basis functions results in significantly simplified linear combinations of the reconstruction coefficients, as odd terms are zeroed through the edge-averaging.

The initial analytic calculations are performed using the wxMaxima sheet \&doc/fieldsolver/Hall_term/HallTermComputations_dxdydz.wxm, as a fair amount of algebra is involved despite the symmetries. The resulting strings of coefficients can be converted to C source code with a bash script similar to \&doc/fieldsolver/Hall_term/codifyCoeffs.sh. That is then included in the source code of the second-order accurate Hall term in \&fieldsolver/1dz_hall.cpp.

Among the physics added to a hybrid or MHD model by including the Hall term in Ohm’s law, one notable phenomenon are the whistler waves. As explained in Section 4.3, this puts a constraint on the maximal allowed time step for the field solver to remain stable. Equation (4.19) shows that regions of low density or high magnetic field drive down the time step, as does spatial resolution.

One of the methods described in Section 4.5 to make global magnetospheric simulations affordable is to subcycle some solvers to allow for longer global time steps. If the field solver time step is significantly smaller than the Vlasov solver limits, and if the configuration allows it, the field solver switches from second-order time stepping (Section 5.1) to a sequence of first-order time steps using the basic Euler method.
As the Vlasov propagation is not applied, the moments of $f$ are not updated during the substeps. It is left to the appreciation of the user how sensitive and sensible the approximation is in the case at hand. It is straightforward to configure how many subcycles are allowed and the model automatically adapts the time step if the limits would be exceeded. In the case of global magnetospheric simulations lasting hundreds of seconds or tens of gyroposters, the choice is usually made to allow up to a few tens of subcycles until the Vlasov solver limit is reached. Thus the simulation time step is typically of the order of $20 - 50$ ms instead of $1$ ms or less. This does somewhat affect physics at short time scales as $f$ does not follow the fields as tightly, but it is deemed a reasonable measure to enable longer or better-resolved simulations within the same computational budget. In particular, since ions decouple completely from the whistler wave at high frequencies the impact is reduced. Additionally, the focus in the global simulations lies on the large scales. Also, typical file sizes only permit to save a snapshot of fields and moments with a frequency of one or two files per simulated second, which anyway does not allow to appreciate evolutions on the scale of milliseconds.

### 5.3 System boundary conditions

Once the solvers are numerically and physically stable and working as intended, the biggest can of worms affecting the life of model developers are the boundary conditions determining the behaviour of the simulation at its edges. The simplest is of course to use periodic conditions as in Paper I, but this is not feasible when modelling a magnetosphere in the solar wind. The term ‘system boundary conditions’ is used in Vlasiator to distinguish them from process boundaries which are of importance among others in staggered MPI communication schemes.

From the coding point of view, the system boundaries are classes inherited from the base SysBoundaryCondition class defined in "sysboundary/sysboundarycondition.h". These system boundary classes contain functions performing a variety of tasks. They define whether a spatial cell pertains to the boundary type, set the initial state of cells belonging to the boundary type and they set the boundary conditions for $E$, $B$, $f$ and their derivatives. The class mechanism easily allows extension by coding a new class with the relevant members or by inheriting from an existing class to specialise some tasks.

From the physical point of view, global magnetospheric simulations require at least the following basic system boundary types: a solar wind inflow boundary, outflow boundaries for the other outer faces of the simulation domain and some form of inner boundary to shield the origin of the dipole magnetic field.

Due to the properties of the solvers used in Vlasiator, and in particular the staggered grid method employed in the field solver, two layers of simulation
cells form a system boundary. The outer layer is by definition two cells away from the inner domain and does not contain velocity space blocks, it is sufficient to set the moments in these cells. The fields also need to be set ‘by hand’ in the outer cells. The inner layer of cells stores velocity space blocks, moments and fields. While \( f \), the moments and \( \mathbf{B} \) need to be handled explicitly by the system boundary class, \( \mathbf{E} \) is computed self-consistently by the field solver in these cells.

The easiest system boundary implemented in Vlasiator is the inflow boundary. It initialises the two layers of cells with a Maxwellian \( f \) and its moments and sets \( \mathbf{B} \), according to the parameters passed in an input ASCII file. Nothing is done subsequently to these cells so that they ensure a constant inflow boundary condition. A feature allows to reinitialise this boundary when continuing the simulation from a restart file, mimicking a dynamic change such as the inflow of a shock front parallel to the inflow boundary. The behaviour of the inflow boundary and some of the code base are taken over from the \textsc{gumics-4} global magnetospheric MHD model [Janhunen et al., 2012].

The increasing dipole magnetic field close to the origin limits drastically the time step length in typical magnetospheric simulations. Furthermore, when approaching the Earth, the ionosphere is the latest region along the way where the conditions deviate so much from a fully ionised, collisionless plasma that some other treatment is required. This is why an inner boundary is needed when performing global magnetospheric simulations.

The inner boundary in Vlasiator emulates a perfect conductor. The magnetic field from the inner cells neighbouring to each boundary cell is averaged, and only the perturbed \( \mathbf{B} \) components normal to the boundary’s surface are retained in the boundary cell. The electric field in the outer boundary cells, in this case the ones towards the origin of the dipole field, is identically set to zero. As for \( f \), the initial Maxwellian distribution is retained throughout the simulation. The geometry of the inner boundary can be chosen to be a square, a circle or a diamond (a circle respectively in the 1-norm, the Euclidean norm and the \( \infty \)-norm), though the regular circular boundary has been preferred for all large-scale runs.

The major reason presiding to the admittedly rudimentary current choice of an inner boundary in Vlasiator is precisely the simplicity of the scheme: Occam’s razor teaches that it is preferable to use such a simple solution as long as it works sufficiently well. The focus on the outer regions – magnetotail, magnetopause, magnetosheath, bow shock and foreshock – as well as the radius at which the boundary is usually set – \( 30,000 \text{ km} \approx 5 R_\oplus \) – so far warrants this choice. Implementing a more physical inner or ionospheric boundary, if not done very carefully and thoroughly, is likely to introduce worse artefacts and instabilities to the model than it would solve issues.
The other outer boundaries of the simulation box have somewhat hastily been named ‘outflow’ boundaries [Paper III] by the then younger, eager student who wrote the initial system boundary code and looks back today. They are however not technically ensuring outflow, they merely perform a copy [Paper II]. $f$ and the moments as well as $B$ are copied from the nearest inner cell in the simplest implementation. More recent developments to the code in \verb|osysboundary/outflow.cpp| aim at reducing the undesirable effects the copy condition can have, in particular self-replicating phenomena. They allow to choose whether an outer boundary performs a classic copy condition, a copy condition limiting the value of $f$ to avoid self-replication, or keeps static conditions.

It is not unique to Vlasiator that clean outflow boundaries are difficult to achieve. In *pic* algorithms particles can be drawn from a random distribution matching either the neighbouring cells or a prescribed one such as a Maxwellian, in order to compensate for the particles flowing out of the simulation domain. In Vlasiator, owing to the necessity to fully describe $f$, the solution is less obvious and a plain copy condition can be insufficient.

Another difficulty resides in the propagation of waves through the boundary. While ideally the boundary should be transparent to all wave modes and not reflect any energy back into the domain, let alone trigger wave modes propagating inwards, it is practically impossible to achieve this. Simulation techniques developed in some areas allow to match a single mode and modify its dispersion relation such that it is extinguished at the boundary, or to damp exponentially modes propagating in one precise direction [e.g. Berenger, 1994, Oskooi and Johnson, 2011]. The opposite approach, trying to cleanly transmit waves through the boundary, has also been successfully applied in some contexts [e.g. Liao et al., 1984, Wagner and Chew, 1995]. Unfortunately, none of these methods can be applied to a continuous spectrum of waves and directions.

The development of the parts of the model presented here, along with a flurry of smaller tasks and contributions to the development of analysis and visualisation tools, have formed a substantial fraction of the work done by the author of these lines, although it does not directly appear as such in scientific publications and presentations.
Chapter 6

Verification, validation and new scientific results

Champagne Cuvée spéciale
Daniel Étienne, Cumières

Windows of the World – Played by the
Yorkshire Building Society Band
Peter Graham

Beyond the pure technical testing [Kempf, 2012, among others], a sustained effort of verification and validation punctuated the youth of Vlasiator. Palmroth et al. [2013] provide a verification of an early form of the Vlasov solver in a test-Vlasov simulation where $f$ is propagated in fields obtained from a GUMICS-4 global MHD run. Another verification step comparing the propagation of plasma waves in Vlasiator to the results of the WHAMP solver is presented in Paper I (Section 6.1). The technical Paper II includes verification through the study of an ion beam instability (Section 6.2), and the validation against observed ion foreshock velocity distributions introduced by Pokhotelov et al. [2013] and in Paper II is developed in depth in Paper III (Section 6.3). This strong footing paves the way towards new scientific contributions in a variety of areas, such as the propagation of foreshock waves [Palmroth et al., 2015], the effects of magnetosheath waves on dayside reconnection [Hoilijoki et al., 2017] or the discovery of new foreshock phenomena (Paper IV and Section 6.4).

6.1 Wave dispersion

The behaviour of the solvers in some of the simpler propagation benchmarks such as advection or acceleration tests [Kempf, 2012] is essentially prescribed by the technical design of the algorithms. The propagation of plasma waves, on the other hand, is a phenomenon emerging from the physics described by
the equations and not directly modelled in Vlasiator. This makes the analysis of wave dispersion a powerful tool to evaluate the physical correctness of the model or detect otherwise subtle and inconspicuous bugs. It is usually done by studying 2D \((\omega, k)\) plots resulting from the discrete Fourier transformation of the \((t, x)\) dataset extracted from a 1D periodic simulation.

For anything but the simplest cases tractable by hand, the Waves in Homogeneous, Anisotropic Magnetized Plasmas (WHAMP) code [Rönmark, 1982, WHAMP, 2016] is a reference numerical tool in the space plasma physics community for problems related to wave propagation and instabilities. WHAMP solves the linearised plasma dispersion equation numerically and yields parameters such as the wave frequency and growth rate when provided the physical parameters of the plasma to consider and the wave vector coordinates to probe.

Paper I reports the comparison of plasma wave dispersion in Vlasiator with the WHAMP results. The simple test setup consists of 1D periodic simulations of weakly perturbed plasma, initialised with a Maxwellian velocity distribution and small perturbations to the density and bulk velocity. The angle of propagation of the waves with respect to \(\mathbf{B}\) is set by changing the angle of the background magnetic field with respect to the box orientation. The Fourier-transformed dataset is then directly comparable to the wave modes found with WHAMP. To take into account the lack of electron physics in Vlasiator, the electron temperature is set to a very small value in WHAMP, suppressing the ion-acoustic wave mode which requires \(T_e \gg T_i\).

As expected from theoretical derivations, the wave modes present in the Hall-less hybrid-Vlasov model correspond to the MHD modes throughout the range covered; the Alfvén wave in the quasiparallel direction and the magnetosonic wave in the quasiperpendicular direction.

With the Hall term included in Ohm’s law, the wave modes tend towards the MHD modes only in the low frequency and long wavelength limit. In the quasiparallel regime at high \(k\), the Alfvén wave splits into the left-hand mode resonating with the ion gyromotion at \(\omega = \omega_{ci}\), and the right-hand whistler mode. In the quasiperpendicular regime, the ion Bernstein modes appear in addition to the magnetosonic mode.

The excellent match of the Vlasiator dispersion plots with the WHAMP output is a significant verification result strengthening the confidence in the validity of the hybrid-Vlasov approach, at the same time as it emphasizes the need for sufficiently rich physics in Ohm’s law to obtain a model departing significantly from MHD.

### 6.2 Ion beam instabilities

The Mr. Hyde personality of plasma waves are the plasma instabilities: as explained in Section 2.2, waves and instabilities are two aspects of the same phenomenon. The correct description of plasma instabilities can be considered an
even more stringent verification test of simulation models than wave dispersion is. The foremost characteristic of an instability to get right is the growth rate, but the instability threshold and its saturation amplitude are also important for the physical correctness of the results.

Beam instabilities, triggered by the presence of at least two distinct particle populations in velocity space, are by definition excluded from single fluid models. They are however critical in many contexts, such as the interaction of the terrestrial magnetosphere and the solar wind in the quasiparallel bow shock and foreshock region. This is why the ion/ion right-hand resonant beam instability, triggered by the presence of a fast, cool beam – the beam speed is higher than the Alfvén speed, the core population thermal speed and the beam thermal speed – is selected in Paper II as an additional verification test. As the results in Section 4 of Paper II show, the dependency of the instability growth rate on the beam density is well reproduced by Vlasiator. Other aspects of the comparison to the analytical result are too dependent on characteristics of the simulation setup and the algorithm to be matched accurately. In particular, the 1D box of 10 $R_{\oplus}$ length constrains the wave number of the growing mode to one not necessarily matching the fastest-growing mode in an unbounded system.

An important feature of the Hall-less regime in Vlasiator emerges from the set of verification studies in Paper I, Paper II and by Pokhotelov et al. [2013]. Although the Hall-less wave dispersion simulations show that for weakly perturbed Maxwellian velocity distributions, only the MHD wave modes propagate, the Hall-less local and magnetospheric simulations show that ion kinetic effects such as ion beam instabilities are described by the model even without the Hall term in Ohm’s law. This is consistent with the fact that the scaling of the Hall term makes it anyway very small if ion kinetic scales are not spatially resolved.

Building on the verification results, new physical understanding is gained through the combined use of WHAMP and Vlasiator to study the evolution of ion beam generated waves and instabilities. A long-standing problem was posed by the observation that the foreshock ultra low frequency (ULF) waves propagate obliquely with respect to the IMF, despite their growth rate being maximised along the magnetic field direction in linear theory [e.g. Eastwood et al., 2005a]. Taking advantage of the global context offered by the simulation to map the spatial evolution of the ion beam speed and density across the foreshock, Palmroth et al. show that the refraction index of the foreshock ULF waves evolves in a way explaining the oblique propagation of the waves even in the case of a quasi-radial IMF orientation which precludes the $E \times B$ effects previously assumed to explain the observations [Palmroth et al., 2015].

### 6.3 Ion foreshock populations

Vlasiator matching the theoretical predictions of wave dispersion and beam instability growth is a milestone in verifying the code, but it is only at best
as good as the theory used in comparison. Ultimately, validation against experimental data is the only relevant yardstick to establish confidence in and credibility of Vlasiator simulation results.

The ion foreshock is a prime ion kinetic phenomenon absent from global MHD models and the abundance of spacecraft measuring particle velocity distribution functions makes the foreshock a good target for validation studies. Currently operating spacecraft missions visiting the Earth’s foreshock and observing ions include Geotail [Nishida, 1994], Cluster [Escoubet et al., 2001] and THEMIS/ARTEMIS [Angelopoulos, 2008, 2011]. They were preceded by a number of missions since the 1970s, so that the morphology of the foreshock region and of the particle distributions found there are well-known observationally [see e. g. the reviews by Fuselier, 1994, Eastwood et al., 2005b].

As Pokhotelov et al. [2013] and Paper II show, Vlasiator produces clean and well-resolved ion velocity distributions including the backstreaming foreshock populations. In Paper III, the types of ion velocity distributions found in the foreshock region of the simulation are matched with THEMIS observations of the same distribution types. Additionally, the simulated spatial evolution of the beam density and velocity in the foreshock reproduce well observational statistics. This, along with the related analysis of the ULF wave characteristics in Paper II and especially by Palmroth et al. [2015], are validation studies showing that Vlasiator results are in good correspondence with observations.

Being a global model, Vlasiator provides a spatial and temporal context which enables to better understand the point-wise spacecraft measurements and propose new interpretations. The simulation analysis in Paper III concludes that in Vlasiator simulations, the very gyrotropic field-aligned ion beams generate ULF waves via the ion/ion right-hand resonant instability, which in return interact with the beam population and cause it to gradually lose its gyrotropy further downstream from the foreshock edge. This confirms the temporal and spatial evolution from gyrotropic to gyrophase-bunched ions in the foreshock previously deduced from observations [Meziane et al., 2001].

### 6.4 Local foreshocks

The goal of the concerted development, verification and validation effort is naturally a simulation tool allowing to produce new scientific knowledge relevant to space physics and possibly space weather activities. In the four years preceding the writing of this thesis, Vlasiator has reached the maturity necessary to enable ground-breaking progress in a wealth of areas of plasma physics. The phenomenon presented in Paper IV and explained in this Section is a discovery made on the basis of Vlasiator global magnetospheric simulations and confirmed in part by ground- and space-based observational data. It combines in a unique way all the plasma physical phenomena introduced in Chapter 2 and in Paper I, Paper II and Paper III.
Figure 6.1 is a graphic representation of the scenario explained in the next paragraphs based on the simulation presented in Paper IV.

When the IMF has a southward component, this very component is antiparallel to the northward geomagnetic field in the subsolar region of the magnetopause. Magnetic reconnection occurs in such cases. In this process of magnetic field topological reconfiguration, magnetic energy is released in the form of plasma accelerated out of the reconnection region at sometimes supermagnetosonic speeds. At the subsolar magnetopause, the variability of reconnection in space and time causes the formation of magnetic structures such as flux tubes propagating poleward along the magnetopause. These structures were first observed by spacecraft and are called flux transfer events (FTEs). In 2D simulations, FTEs take the form of magnetic islands. Vlasiator 2D polar plane simulations of the terrestrial magnetosphere under southward IMF show that reconnection is active throughout and produces a flurry of magnetic islands accelerated away from reconnection X-points [Hoilijoki et al., 2017]. Magnetic islands are well-visible in Figure 6.1a and b.

Owing to their momentum and magnetic structure, the islands are preceded (followed) by bow (stern) magnetosonic wave fronts which propagate throughout the magnetosheath, as shown in Figure 6.1b and c. In the case of larger magnetic islands, the bow wave fronts steepen so much that they drive a foreshock-like ion beam ahead of them in the sheath. Once a wave front hits the bow shock, the pressure balance is displaced upstream in favour of the increased magnetosheath pressure and the shock surface bulges out with respect to its unperturbed position (Figure 6.1d).

This affects the shock geometry so that $\theta_{Bn}$ can become favourable to ion reflection off the bow shock in a very localised patch at the foot of the bow shock bulge. A field-aligned ion beam exactly similar to those observed at the foreshock upstream edge propagates into the solar wind. An example is in Figure 6.1d and e. The major difference with respect to the ‘regular’ foreshock is that this ion beam is very localised – as wide as the region where $\theta_{Bn}$ allows reflection at the bow shock – and travels along with the original wave front. Observationally, this means that a transient field-aligned ion beam is observed for a duration of about 1 min without any perturbation of the IMF or the solar wind. Due to the swift travel of the ion beam, instabilities do not have time to grow and a foreshock-like beam is observed without the ULF waves usually associated with foreshock observations.

In a global simulation, such a complex chain of processes is readily visible and can be analysed under all aspects, in an extended spatial and temporal context. Its comprehensive observational confirmation is a different story altogether. Ideally, spacecraft observing dayside reconnection and FTEs at the magnetopause, fast-mode wave fronts in the magnetosheath and transient ion beams upstream of the bow shock as well as monitoring the upstream solar wind are
Figure 6.1 – Illustration of the local foreshock scenario [Paper IV]. (a) Density plot: dayside reconnection generates magnetic islands. (b) Density plot: strong islands push bow (black dashed) and heck (white dash-dotted) wave fronts into the magnetosheath. (c) Profiles along the white arrow: the fronts are fast-mode waves. (d) Parallel ion temperature plot: the fronts push out the bow shock (white continuous line) and modify the shock normal direction (white continuous segment) with respect to the unperturbed bow shock (dashed). An ion beam is emitted (upstream structure). (e) Ion velocity distribution at the white cross: a typical field-aligned ion beam is observed upstream of the bow shock and ‘regular’ foreshock.
required. In spite of the extensive fleet of (multi-)spacecraft missions dedicated
to space physics, the coincidence of a useful spacecraft constellation during a
suitably stable period of southward IMF is a rare event, taking into account
that solar wind or IMF perturbations as well as the too close vicinity of the
bow shock or the foreshock from the beam-observing spacecraft have to be
excluded.

Paper IV presents the scenario identified in simulations as well as observa-
tions supporting it. On August 30, 2004 during a long interval of stable south-
ward IMF, Geotail observed transient ion beams in the solar wind well upstream
of the bow shock and the foreshock. No suitable spacecraft was observing the
magnetopause or the magnetosheath simultaneously, but ground-based magnet-
tometer and radar ionospheric observations give an indirect proof of dayside
reconnection and FTEs occurring during the event.

The scenario leading from steady southward IMF to the occurrence of tran-
sient ion beam signatures upstream of the bow shock and distinct from the
‘regular’ foreshock is new in several respects. It represents a new mechanism
generating a time-dependent behaviour of the bow shock in the absence of
driving upstream conditions. It is distinct in nature and consequences from the
likes of shock reformation in the quasiperpendicular region or the dynamics
of the quasiparallel bow shock region including the ‘regular’ foreshock. It fur-
thermore offers an interpretation of transient foreshock observations in cases
where no obvious IMF or solar wind perturbations could otherwise account for
the ion beam formation. Finally, it suggests the existence of steep fast mag-
netosonic wave fronts in the magnetosheath related to dayside reconnection, a
phenomenon hitherto not confirmed by in-situ observations and only seldom
reported in literature before [e. g. by Sibeck and Omidi, 2012, who report slow-
mode waves though].

The work presented in this chapter demonstrates that Vlasiator is now estab-
lished as a reliable and well-performing tool. Its unique and unprecedented
capability of modelling the Earth’s magnetosphere in a global context with
the kinetic hybrid-Vlasov model fosters advances in space physics and promises
breakthroughs in adjacent fields of plasma physics.
Chapter 7

Conclusions and outlook

Something sweet and homemade
Symphony No. 9 in E minor, ‘From the new world’, especially the fourth movement
Allegro con fuoco

Antonín Dvořák

The present thesis, comprising the pages preceding these lines and the four peer-reviewed journal articles included hereafter, condenses the essence of the work achieved by its author within the Vlasiator project. The new kid on the block Vlasiator is established as a landmark of kinetic plasma simulations and the results obtained by the team are pushing the frontiers of space physics.

The first years of using Vlasiator have widened the horizon of magnetospheric simulations, sometimes against the prevalent intuition. The equations can lead to believe that the hybrid-Vlasov system, especially in its Hall-less form, would not depart significantly from fluid models as long as kinetic scales are left unresolved. The results speak for themselves, showing kinetic phenomena even at coarser resolutions. It has also been repeatedly astonishing to see how supposedly local, kinetic, microphysical phenomena interact and combine to eventually affect global processes. Most of this may seem natural with the wisdom of hindsight, but it was definitely not generally envisioned beforehand.

Despite their covering most of the global system, magnetospheric Vlasiator simulations are still restricted to two spatial and three velocity dimensions due to the sheer size of the runs. A straightforward extension of the simulation box in the third dimension with the same resolution and using the current algorithms would right away exceed the limits of manageability even on top-tier supercomputers – hundreds of millions of core-hours and hundreds of terabytes of memory are a fantasy of the future. The development roadmap is therefore unequivocally aiming at enabling 6D kinetic magnetospheric simulations.

The eager scientist increases the resolution of the runs whenever better
performance is available. The weight of the run for a given system does not decrease in that way, however. The order of magnitude of improvements necessary to reach global 6D modelling at kinetic resolution will not be attained unless drastic mesh coarsening is applied to some regions of the simulation. Adaptive mesh refinement in velocity space combined with the existing sparsity technique should also be deployed in order to efficiently track the distribution. This joint strategy resulting in a truly 6D adaptive mesh is the only realistic path towards fully global hybrid-Vlasov magnetospheric simulations.

The implications are far-reaching for the code base. Neither the Vlasov solver nor the field solver support non-uniform meshes for now. Hence extending their functionality or replacing them partially is a sizeable investment. Alongside this fundamental development, a major step has to be taken to improve the physical quality of the magnetospheric model. The simple, perfectly conducting inner boundary is sufficient for 5D simulations but certainly too unrealistic for the future. The choice of an adequate ionospheric model will be guided by the experience gathered with global MHD and PIC models, but its coupling to the hybrid-Vlasov model is a challenge never faced before. These plans entail thorough testing, verification and validation throughout.

The development of Vlasiator is not solely focussing on the requirements of 6D magnetospheric simulations. It has gone hand-in-hand with technological advances from the onset.

Leveraging the full power of supercomputers is only possible by targeting today the cutting-edge technology of tomorrow. This is why the code will not simply be adapted to meet physical goals, but performance gains will be sought from new computing paradigms, too. Emerging possibilities include extreme threading on architectures offering hundreds of processing cores per node and the development of algorithms which can run simultaneously on classic cores and graphical processing units (GPUs) on the same nodes. The key is to keep in sync with the pulse of the computing industry.

Another technological driver in the search for improved space plasma models is the rich array of available ground-based and in-situ measurements. Only a few years ago, Vlasiator ion distributions sported a better resolution than spacecraft data. The ball was returned swiftly to the modelling camp at the latest following the launch of the Magnetospheric Multi-Scale constellation (MMS) and its high-cadence, high-resolution datasets in 2015 [Burch et al., 2016a]. Spacecraft technology is also driving pre-flight model development, when it comes to modelling future instruments and deciding on their optimal parameters.

Of course the purpose of Vlasiator is not to provide work to the cooling systems of high-end supercomputers, nor is it to produce ever better-resolved pretty pictures (although the former has happened and the latter are extremely important!). Scientific progress in space physics and plasma modelling is the
compass giving the direction.

Even though Vlasiator is rooted in magnetospheric science, its capabilities extend much wider. Hybrid-Vlasov modelling is shedding light onto a variety of other topics such as particle acceleration and turbulent energy dissipation in the solar wind. The overarching ambition, besides a better first-principle understanding of the physics at play, is to build a comprehensive picture of the solar-terrestrial interaction. This is a prerequisite to the reliable modelling and operational forecasting of the space weather chain.

Modelling also belongs to the toolkit of modern space technology research and development. What the high cost of in-orbit or even laboratory testing prohibits, can in part be achieved by simulations, especially in the proof-of-concept phase. In the case of Vlasiator, the simulation support offered to studies of novel space propulsion techniques, namely the electric sail and the Hall thruster, results in a fruitful and mutually beneficial collaboration.

Regardless of a certain bias in the present thesis, plasma physics is not the private turf of space scientists. The long-term dream of exploiting nuclear fusion as a safer and cleaner source of energy to quench mankind’s thirst for growth and expansion relies largely on – or is hampered by the lack of – the theoretical and technical mastery over hot and unstable plasmas. In that field as well, theory and experiment need the complement of modelling, in particular beyond the classic MHD and PIC frameworks. Developing and applying Vlasiator to solve scientific riddles in fusion plasma research opens up new perspectives which will undoubtedly multiply the return on investment.

The core business of Vlasiator remains nonetheless kinetic simulations of the Earth’s magnetosphere. The development of hybrid-Vlasov models to simulate space plasmas at large scales in the last decade has brought new insights in ion physics that are not accessible to particle-based methods. Yet the advent of satellite missions such as MMS and the planning of projects like the Turbulence Heating Observer [THOR, Vaivads et al., 2016] are already pushing the hybrid simulation community out of its comfort zone and down the turbulent cascade to smaller scales. Due to the electron and ion scale separation, the realistic, fully kinetic simulation of the terrestrial magnetosphere is impossible – something said of hybrid-Vlasov magnetospheric simulations not long ago still. Since impossible n'est pas français, the quest of innovative approaches to modelling the magnetosphere including electron physics in addition to kinetic ions is naturally the next journey to embark on. Should a hybrid model include multi-moment or even gyrokinetic electrons? Are subgrid-scale or heuristic modelling, let even the buzzword deep learning appear here, paving the road to success? Ideas abound and the territory is almost completely uncharted. If there be dragons, the time is ripe to coax them out and tame them.

---

1Impossible is not French.
Bibliography


R. M. Evans, M. Opher, R. Oran, B. van der Holst, I. V. Sokolov, R. Frazin, T. I. Gombosi, and A. Vásquez. Coronal heating by surface Alfvén wave


L. Wang, A. H. Hakim, A. Bhattacharjee, and K. Germaschewski. Comparison of multi-fluid moment models with particle-in-cell simulations of


Paper I

Wave dispersion in the hybrid-Vlasov model: Verification of Vlasiator

© 2013 AIP Publishing LLC.

Reprinted from
*Physics of Plasmas*, Volume 20, Number 112114, November 2013, doi:10.1063/1.4835315,
with the permission of AIP Publishing.
Wave dispersion in the hybrid-Vlasov model: Verification of Vlasiator

Yann Kempf,1,2,a) Dimitry Pokhotelov,1,2 Sebastian von Alfthan,1 Andris Vaivads,3
Minna Palmroth,1 and Hannu E. J. Koskinen1,2
1Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
2Department of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland
3Swedish Institute of Space Physics, Box 537, 751 21 Uppsala, Sweden

(Received 9 October 2013; accepted 13 November 2013; published online 26 November 2013)

Vlasiator is a new hybrid-Vlasov plasma simulation code aimed at simulating the entire magnetosphere of the Earth. The code treats ions (protons) kinetically through Vlasov’s equation in the six-dimensional phase space while electrons are a massless charge-neutralizing fluid [M. Palmroth et al., J. Atmos. Sol.-Terr. Phys. 99, 41 (2013); A. Sandroos et al., Parallel Comput. 39, 306 (2013)]. For first global simulations of the magnetosphere, it is critical to verify and validate the model by established methods. Here, as part of the verification of Vlasiator, we characterize the low-β plasma wave modes described by this model and compare with the solution computed by the Waves in Homogeneous, Anisotropic Multicomponent Plasmas (WHAMP) code [K. Römmark, Kiruna Geophysical Institute Reports No. 179, 1982], using dispersion curves and surfaces produced with both programs. The match between the two fundamentally different approaches is excellent in the low-frequency, long wavelength range which is of interest in global magnetospheric simulations. The left-hand and right-hand polarized wave modes as well as the Bernstein modes in the Vlasiator simulations agree well with the WHAMP solutions. Vlasiator allows a direct investigation of the importance of the Hall term by including it in or excluding it from Ohm’s law in simulations. This is illustrated showing examples of waves obtained using the ideal Ohm’s law and Ohm’s law including the Hall term. Our analysis emphasizes the role of the Hall term in Ohm’s law in obtaining wave modes departing from ideal magnetohydrodynamics in the hybrid-Vlasov model. © 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4835315]

I. INTRODUCTION

The exponential growth in available computing power has made hybrid and fully kinetic plasma simulations increasingly feasible for a variety of space plasma applications. Non-exhaustive examples include hybrid particle-in-cell (hybrid-PIC) simulations of planetary environments1,2 and magnetospheres,3–6 full-PIC studies of magnetic reconnection,7,8 and local hybrid-Vlasov simulations of wave-particle interactions in the solar wind,9 or full-Vlasov simulations of the Kelvin-Helmholtz instability.10 Vlasiator is a new self-consistent hybrid-Vlasov simulation code in which the ions (protons) are treated kinetically via Vlasov’s equation and electrons are a massless charge-neutralizing fluid. It is based on a robust finite volume method, which has been optimized for the modeling of the entire magnetosphere of the Earth (http://vlasiator.fmi.fi).11,12 To our knowledge, Vlasiator is the first hybrid-Vlasov code to allow simulations on this scale. An important advantage of the hybrid-Vlasov model with respect to (hybrid-)PIC methods is the absence of noise related to the low number of particles representing the distribution function, as hybrid-Vlasov algorithms propagate the full distribution function in the six-dimensional phase space using Vlasov’s equation. The uniform sampling in velocity space provides a description of the distribution function with a quality comparable to spacecraft measurements. However, the six-dimensional representation of the distribution function in the hybrid-Vlasov approach makes the memory and computing requirements high, even for modern massively parallel supercomputers.

The hybrid-Vlasov scheme is a relatively new approach to computational plasma physics at large scales because of the aforementioned need of computing resources. Hence it is critical to provide basic benchmarks to assess the possibilities and limits of the model in terms of its physical features. As part of the verification of Vlasiator, we perform local simulations to investigate the propagation of low-β plasma waves in the hybrid-Vlasov model by studying wave dispersion in a variety of cases. The dispersion of the left- and right-hand polarized modes as well as the ion-acoustic waves propagating parallel to the magnetic field has been used before to verify a hybrid-Vlasov simulation code.13 We extend this test to include more plasma wave modes in all propagation directions ranging from parallel to perpendicular to the magnetic field and present dispersion surface plots. The cases included here are in a range of parameters relevant to understand global magnetospheric simulations. Thus, this study helps bolstering the interpretation of current and future large-scale simulations of the magnetosphere of the Earth in 2+3 and 3+3 spatial and velocity dimensions. We compare the results obtained with Vlasiator to the solutions calculated using the Waves in Homogeneous, Anisotropic Multicomponent Plasmas code (WHAMP, https://github.com/irfu/whamp), which solves the linearized kinetic dispersion equation numerically.14,15 The simple and robust solution approach of WHAMP starting from the general dispersion equation and its strong

---

a)Electronic mail: yann.kempf@fmi.fi
establishment as a tool for the determination of wave dispersion in homogeneous plasmas make it an ideal tool to verify Vlasiator results.

Since the central interest of Vlasiator lies in simulating the entire magnetosphere of the Earth, we focus on wave modes at spatial and temporal ion scales. In ideal magnetohydrodynamic (MHD) theory, the only possible wave modes are the Alfven wave, which is a shear electromagnetic mode propagating at all angles except perpendicularly to the magnetic field, and the fast and slow magnetosonic modes, which are compressional electromagnetic modes propagating at an oblique angle with respect to the magnetic field. The dispersion equation of the shear Alfven wave is given by

$$\frac{\omega}{k} = v_\text{A}\cos \theta,$$

where $\omega$ is the angular frequency, $k$ the wave number, $v_\text{A} = B/\sqrt{\rho_0 m_i}$ ($B$ is the magnetic field intensity, $\rho_0$ is the permeability of vacuum, and $\rho_i$ is the mass density) is the Alfven speed and $\theta$ is the angle between the wave vector $k$ and the magnetic field vector $B$. The fast $(+)$ and slow $(-)$ magnetosonic wave dispersion equation is given by

$$\left(\frac{\omega}{k}\right)^2 = \frac{1}{2} \left( v_s^2 + v_h^2 \right) \pm \frac{1}{2} \sqrt{ \left( v_s^2 + v_h^2 \right)^2 - 4 v_h^2 v_s^2 \cos^2 \theta },$$

where $v_h = \sqrt{k_B T/m_i}$ ($k_B$ is the Boltzmann constant, $T$ is the temperature, and $m_i$ is the ion mass) is the sound speed. In a low-$\beta$ plasma, we have $v_\text{A} > v_h$. For parallel propagation ($\theta = 0$), the Alfven and fast magnetosonic modes coincide and the mode propagates at $v_\text{A}$, whereas the slow magnetosonic mode becomes the sound wave propagating at $v_h$. In the perpendicular case ($\theta = \pi/2$), the Alfven and slow mode cannot propagate and only the fast magnetosonic mode propagating at magnetosonic speed $v_\text{MS} = \sqrt{v_h^2 + v_s^2}$ subsists. These modes are dispersionless.

Multi-fluid and kinetic theories describe more plasma wave modes. The left- and right-hand polarized modes (L- and R-mode hereafter) propagate along or quasi-parallel to the magnetic field. They couple to the ion and electron gyro-motion, therefore they resonate at the ion and electron gyro-frequency, respectively. In the low-frequency, small wave number limit both modes converge towards the ideal MHD Alfven mode.

The last wave modes in the range of frequencies and wavelengths of interest in this paper are the ion Bernstein modes, which are electrostatic ion-cyclotron resonances (quasi-)perpendicular to the magnetic field. Their dispersion equation can be expressed in terms of infinite sums of Bessel functions in the hybrid-Vlasov model.19 A thorough review of plasma wave dispersion surfaces obtained with WHAMP at ion and electron scales (including the modes introduced in this section) is given by Andrê.18

This paper is organized as follows. In Sec. II, we present Vlasiator and WHAMP as well as the simulation setup used with Vlasiator. Then, we present our results in Sec. III. We show the importance of the Hall term in Ohm’s law, and we study dispersion curves for parallel, perpendicular and oblique propagation as well as dispersion surfaces. We discuss the results in Sec. IV and conclusions are drawn in Sec. V.

II. MODEL AND METHODS

A. The hybrid-Vlasov model in Vlasiator

The solver for Vlasov’s equation implemented in the hybrid-Vlasov model of Vlasiator is based on a three-dimensional wave propagation algorithm19,20 applied separately to translation in position space and acceleration in velocity space. It is self-consistently coupled to a field solver21 which uses divergence-free magnetic field reconstruction.22 We only introduce here the solver features relevant to this study.

Vlasiator solves Vlasov’s equation,

$$\frac{\partial}{\partial t} f(r,v,t) + v \cdot \nabla f(r,v,t) + a \cdot \nabla v f(r,v,t) = 0,$$

where $r$ and $v$ are the spatial and velocity coordinates, $t$ is time, $f(r,v,t)$ is the six-dimensional phase-space density of ions with mass $m$ and charge $q$, and acceleration $a$ is due to the Lorentz force

$$a = \frac{q}{m} (E + v \times B),$$

in which $E$ is the electric field and $B$ is the magnetic field.

In the hybrid-Vlasov model, Vlasov’s equation is coupled to Maxwell’s equations. The displacement current is neglected in the Ampère-Maxwell law and the equations take the form

$$\nabla \times E = -\frac{\partial}{\partial t} B,$$

$$\nabla \times B = \mu_0 j,$$

where $j$ is the total current density. Note that the equation $\nabla \cdot B = 0$ is respected by the field propagation algorithm of Vlasiator by construction, provided the initial conditions are divergence-free.21,22

Ohm’s law describes the relationship between the electric and the magnetic field. It is needed to close the hybrid-Vlasov system of equations, when updating the magnetic field using Faraday’s law (Eq. (5)). In the present study, Vlasiator is using the ideal Ohm’s law supplemented by a Hall term (rightmost term in Eq. (7)) with first-order spatial accuracy,

$$E = -V_i \times B + \frac{1}{\rho_i} j \times B. \quad (7)$$

The ion charge density $\rho_i$ and (ion) bulk velocity $V_i$ are obtained from velocity moments of the distribution function, $j$ is computed using Eq. (6). This effectively represents electrons as a massless, charge-neutralizing fluid. The rest of the solvers in Vlasiator retain second-order spatial accuracy throughout in smooth cases. In cases with strong spatial gradients—in position or velocity space alike—flux limiters
effectively reduce the spatial accuracy of the solvers in order to preserve the numerical stability of the scheme.

Note that although the Hall term in the computation of the electric field can be neglected using the ideal Ohm’s law, it has to be retained in the electric field input into the Lorentz force (Eq. (4)) in order to model bulk forces on the ions. This makes Vlasiator a proper Hall-less hybrid model when the Hall term in Ohm’s law is not used.\textsuperscript{23}

B. The WHAMP code

WHAMP is a code solving the linear analytic dispersion equation of waves in magnetized plasmas. It can include several populations with differing number density, mass, temperature, loss cone, anisotropy and drift parameters for anisotropic Maxwellian distributions.\textsuperscript{14,15,18} A general form of the plasma wave dispersion equation is $D(\omega, k) \cdot E(\omega, k) = 0$, where $D$ is the dispersion tensor and $E$ is the wave electric field. Solutions can be found by equating the determinant of the dispersion tensor to zero, $|D(\omega, k)| = 0$. WHAMP solves this in a linearized form using a Padé approximant to ensure a fast computation yet good approximation of the result.

The WHAMP interface is designed to take in the plasma parameters initially and then query the solution point by point in the $(k-\omega)$ space. Given an initial $(k_x, k_y)$ point WHAMP tries to find a wave mode close-by and returns the frequency, wave vector, and growth rate of the mode among others. A script querying WHAMP systematically is used to ease comparison with the full dispersion plots and surfaces obtained with Vlasiator following the methods presented in Sec. II C.

When comparing WHAMP and Vlasiator results, the electron temperature $T_e$ in WHAMP is set to a small value in order to mimic the absence of electron pressure gradient effects in our hybrid-Vlasov model. The electron temperature is ignored in Vlasiator, effectively suppressing in the model the ion-acoustic wave for which $T_e \gg T_i$ must hold.

C. Simulation setup and processing

The simulation setup of Vlasiator in the present study consists of a one-dimensional spatial domain along the $x$-axis, constraining the wave vector to be in that direction, with fully periodic boundary conditions. The angle between the magnetic field and the wave vector is defined by setting the magnetic field orientation with respect to the simulation box. The initial conditions are uniform up to small random perturbations in the number density and bulk velocity. The velocity distribution is isotropic and Maxwellian, therefore excluding waves growing from anisotropy-driven instabilities. The system relaxes and no forcing is applied during the simulation.

A two-dimensional space-time $(x-t)$ dataset is formed by saving the spatial profile of a bulk variable at every (constant) time step. The total run time is typically several ion gyroperiods. The data is first windowed along the time dimension using a Hamming window to reduce the noise induced by the abrupt start and stop of the time series. Then it is subjected to a discrete Fourier transformation to produce a $(k-\omega)$ dispersion plot.

Dispersion surfaces are computed by extracting the points above a threshold in each $(k-\omega)$ dataset to retrieve the dispersion branches for multiple angles. The extracted data is interpolated and re-sampled on a Cartesian coordinate grid to obtain smoother dispersion surfaces more readily comparable with the dispersion surfaces from WHAMP.

The simulation and plasma parameters of all Vlasiator simulations presented in this paper are given in Table I. WHAMP uses exactly the same parameters as input, except for the electron temperature as discussed above.

III. RESULTS

A. The Hall term in Ohm’s law and parallel propagation

In ideal MHD, Ohm’s law takes the form $\mathbf{E} = -\mathbf{V} \times \mathbf{B}$, where $\mathbf{V}$ is the bulk velocity and the Hall term $j \times \mathbf{B}/\rho_q$ has been omitted with respect to Eq. (7). The absence of spatial scales in this form of Ohm’s law in the hybrid-Vlasov model prevents the kinetic coupling of the ions to any wave mode. There are thus no resonances and the wave modes present are the non-dispersive ideal MHD wave modes. In the parallel case this means that the only wave mode one can observe in the dispersion plot is the ideal MHD Alfvén wave (Eq. (11)), as illustrated in Figure 1 on the left-hand side ($\theta = 0.001$). All other cases hereafter were obtained using the Hall term in Ohm’s law.

Using $\mathbf{j} = \nabla \times \mathbf{B}/\rho_q$ [Eq. (6)], the Hall term can be expressed as $(\nabla \times \mathbf{B}) \times \mathbf{B}/(\rho_q \mathbf{V})$. It becomes apparent in this form that through the current density, spatial derivatives of the magnetic field and thus spatial scales are introduced. One consequence is that wave dispersion can occur in the hybrid-Vlasov model and one departs from the ideal MHD description.

The dispersion plot for the same parameters but with the Hall term included is presented on the right-hand side of Figure 1. As expected, the L-mode has a resonance at the fundamental ion gyrofrequency. The R-mode on the other hand is not expected to resonate at $c_\omega \rho_q$ (beyond the upper edge of the figure) because of the lack of electron physics in our hybrid-Vlasov model. At low frequencies and low wave numbers, the two modes converge towards the ideal MHD dispersion.

<table>
<thead>
<tr>
<th>Simulation parameters</th>
<th>Dispersion surface/parallel/oblique</th>
<th>Perpendicular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation domain length</td>
<td>$2.5 \times 10^8$ m</td>
<td>$5.0 \times 10^8$ m</td>
</tr>
<tr>
<td>Number of spatial cells</td>
<td>10 000</td>
<td>10 000</td>
</tr>
<tr>
<td>Velocity space resolution</td>
<td>$4.0 \times 10^4$ m/s</td>
<td>$4.0 \times 10^4$ m/s</td>
</tr>
<tr>
<td>Time step</td>
<td>0.001 s</td>
<td>0.025 s</td>
</tr>
<tr>
<td>Plasma parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ion (proton) number density</td>
<td>$1.0 \times 10^6$ m$^{-3}$</td>
<td>$1.0 \times 10^6$ m$^{-3}$</td>
</tr>
<tr>
<td>Ion (proton) temperature</td>
<td>$1.0 \times 10^8$ K</td>
<td>$1.0 \times 10^8$ K</td>
</tr>
<tr>
<td>Magnetic field intensity</td>
<td>$5.0 \times 10^{-8}$ T</td>
<td>$1.0 \times 10^{-8}$ T</td>
</tr>
<tr>
<td>Plasma $\beta$</td>
<td>$1.4 \times 10^{-3}$</td>
<td>$3.5 \times 10^{-2}$</td>
</tr>
</tbody>
</table>
Alfvén wave. Without the Hall term in Ohm’s law, the simulation result matches the ideal MHD theory. With the Hall term, the simulation result matches the dispersion obtained by WHAMP. The reason for the slight discrepancy in the R-mode at higher frequencies is analyzed in Sec. III C.

B. The ion Bernstein modes in perpendicular propagation

In the perpendicular propagation case ($\theta = 1.57$) shown in Figure 2, the dispersion plot exhibits the ion Bernstein modes, which are relatively weak but correspond well to the Vlasiator solution. The magnetosonic mode is non-dispersive in Vlasiator whereas it bends towards the lower-hybrid plateau in WHAMP (top of the plot and beyond), the first few ion Bernstein modes are visible but relatively weak.

WHAMP solution. The non-propagating mode at $\omega_{ci}$ and its harmonic at $2\omega_{ci}$ are due to the initial random perturbations fluctuating at the ion gyrofrequency throughout the simulation domain at all spatial scales. The strongest mode in the plot is the magnetoacoustic mode, which is non-dispersive in Vlasiator. It shows dispersion in the WHAMP solution and bends towards the lower-hybrid plateau (top of the plot and beyond), a feature not reproduced by Vlasiator again due to lacking electron physics in the hybrid-Vlasov model.

C. Dispersion surfaces and oblique propagation

Following the tradition established by André to study wave dispersion using dispersion surfaces in the $(k_x, k_z, \omega)$ space, we produced the dispersion surfaces for Vlasiator as well. The comparison of the Vlasiator dispersion surfaces with the WHAMP solution is shown in Figure 3. The WHAMP solution is plotted regardless of the damping or the relative amplitude of the modes. This explains the smaller extent of the L-mode surface (lower surface) as well as the absence of the ion Bernstein modes in this regime in the Vlasiator results.

In order to convey a better impression of the relation between the Vlasiator and WHAMP solutions, the dispersion for oblique propagation at $\theta = 0.3$ is presented in Figure 4. A first observation is that Vlasiator and WHAMP consistently show a strong damping of the L-mode beyond $k \cdot r_L \approx 0.18$; the signal weakens in Vlasiator and WHAMP does not pick the mode any more. Another observation is that while at low frequencies and long wavelengths both solutions overlap completely, the R-mode does not match exactly at higher $\omega$, in a similar way as was observed for the parallel propagation case in Sec. III A. Increasing the spatial resolution of the Vlasiator simulation improves the situation, in that the Vlasiator R-mode is closer to the WHAMP solution for a longer range. This indicates that the spatial accuracy of the code, and especially the accuracy of the Hall term, which is currently of lower order than the Vlasov and field solvers, is responsible for this discrepancy.

IV. DISCUSSION

In this study, we show that the wave dispersion results from Vlasiator are closely matched by the WHAMP solution
in all propagation directions. The good correspondence between these two essentially different approaches to solving Vlasov’s equation, namely the linearized kinetic theory and the hybrid-Vlasov simulation, is an indicator of the good quality of the Vlasiator results. They clearly show that the Hall term in Ohm’s law is critical to go beyond ideal MHD in terms of wave modes described by the hybrid-Vlasov model. Introducing spatial scales through the derivatives of the magnetic field, the Hall term in Ohm’s law makes wave dispersion possible. In particular the L- and R-modes split and the L-mode, coupling to the ion gyromotion, resonates at the ion gyrofrequency. In addition to this, the ion Bernstein modes are a feature which arises purely from the kinetic description of magnetized plasma. Two major consequences and expected shortcomings of the limited electron physics of the hybrid-Vlasov model are that the R-mode should not resonate at any frequency, and that the magnetosonic mode does not couple to electrons to form the lower-hybrid plateau at the ion-electron lower-hybrid frequency \( \omega_{\text{LH}}^2 = \omega_{\text{pi}}^2 + \omega_{\text{ci}}^2 / (1 + \omega_{\text{pe}}^2 / \omega_{\text{ci}}^2) \) in quasi-perpendicular propagation.

This work is a major step in the verification of Vlasiator and it provides vital insight into the model’s wave modes. The aim of the development of Vlasiator is to provide the first self-consistent hybrid-Vlasov model able to simulate the entire magnetosphere of the Earth including ion-kinetic effects. Plasma wave modes are a key feature which should be described accurately by the model and properly understood in order to interpret global magnetospheric simulations. It is important to note that the splitting of wave modes introduced by the Hall term in Ohm’s law also occurs on temporal and spatial scales comparable to and longer than the ion scales. These long scales are of primary interest in global magnetospheric simulations using a hybrid code in which the emphasis is placed on ion kinetics. Therefore differences are expected to appear between Hall-less simulations of the magnetosphere and their counterparts using the Hall term in Ohm’s law, even if the ion gyroradii were not resolved. Adding the Hall term to Ohm’s law in Vlasiator is expected to make the described physics richer but suitable numerical accuracy and stability in global magnetospheric simulations will only be achieved with a Hall term of appropriate spatial accuracy. Realistic results have already been obtained nevertheless in successful Hall-less magnetospheric simulations, comparing favorably to observed phenomena.\(^{24}\)

V. CONCLUSIONS

By studying the dispersion of plasma waves in the hybrid-Vlasov model of the new Vlasiator code and comparing it to the linearized solution computed with WHAMP, we provide a key verification benchmark for Vlasiator. The match between the two fundamentally different approaches is excellent, even with a Hall term in Ohm’s law of first-order spatial accuracy. This study underlines the importance of using a sufficiently rich Ohm’s law in the hybrid-Vlasov model to obtain a kinetic description of waves departing from ideal MHD, even on temporal and spatial scales typically associated with ideal MHD. The future addition of
terms to Ohm’s law such as the electron pressure gradient term is expected to play a significant role as well in introducing richer electron physics, the lower-hybrid plateau for example. This will improve the quality of the model and the variety of physical phenomena included in it, even more so in the perspective of implementing adaptive mesh refinement in Vlasiator to resolve ion-kinetic scales in selected regions.

ACKNOWLEDGMENTS

We acknowledge that the results of this research have been achieved using the PRACE (Partnership for Advanced Computing in Europe) Research Infrastructure resource Hermit based in Germany at the High Performance Computing Center Stuttgart (HLRS).

The Quantifying Energy circulation in Space plasmas (QuESpace) project, in which Vlasiator is developed, has received funding from the European Research Council under the European Community’s seventh framework programme (FP-7/2007-2013/ERC) agreement No. 200141-QuESpace. The work of Y.K., D.P., S.A., and M.P. has been supported by the Academy of Finland.

Paper II

Vlasiator: First global hybrid-Vlasov simulations of Earth’s foreshock and magnetosheath

© 2014 Elsevier Ltd.

Vlasiator: First global hybrid-Vlasov simulations of Earth’s foreshock and magnetosheath

S. von Alfthan\textsuperscript{a,n}, D. Pokhotelov\textsuperscript{a,b}, Y. Kempf\textsuperscript{a,b}, S. Hoilijoki\textsuperscript{a,b}, I. Honkonen\textsuperscript{a,b}, A. Sandroos\textsuperscript{a}, M. Palmroth\textsuperscript{a}

\textsuperscript{a} Earth Observation Unit, Finnish Meteorological Institute, Helsinki, Finland
\textsuperscript{b} Department of Physics, Faculty of Science, University of Helsinki, Helsinki, Finland

\textbf{ABSTRACT}

We present results from a new hybrid-Vlasov simulation code, Vlasiator, designed for global magnetospheric simulations. Vlasiator represents ions by a six-dimensional distribution function propagated using a finite volume approach. The distribution functions are self-consistently coupled to electromagnetic fields with electrons modeled as a charge-neutralizing fluid. A novel sparse representation of the distribution function reduces the computational demands of the problem by up to two orders of magnitude. The capabilities of the code are demonstrated by reproducing characteristics of the ion/ion right-hand resonant beam instability, as well as key features of the collisionless bow shock and magnetosheath in front of the Earth’s magnetosphere in global five-dimensional (two in ordinary space, three in velocity space) simulations. We find that Vlasiator reproduces the ion velocity distribution functions with quality comparable to spacecraft observations.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Global modeling of the near-Earth space environment is important for several reasons. First, global models provide information on plasma processes in space and time, complementing in situ observations and giving them a context. Second, accurate modeling results of the near-Earth environment give feedback to instrument developers, so that improved instrumentation can be implemented for future missions. Third, simulations can also be used to monitor quantities that are difficult to measure using instrumentation; for example global energy flow between the solar wind and the magnetosphere (Palmroth et al., 2003, 2004). Finally, if computed in real-time, global simulations can be used to monitor space weather conditions, giving warnings for different stakeholders.

The above reasons for developing global simulations all indicate that the simulation needs to be numerically accurate and the model should describe the relevant physical phenomena. The most popular solution so far has been to implement the magnetohydrodynamic (MHD) equations in a global computer simulation. The level of numerical accuracy has been discussed mainly in terms of the order of the solution (see the review by Ridley et al., 2010), while the question remains as to whether the MHD model itself is rich enough to be used as a basic tool for interpreting the system. MHD has proven to give a good description on large scales and where plasma can be described with one temperature (e.g. Janhunen and Palmroth, 2001), while for example in the inner magnetosphere its power to explain physical phenomena rapidly decreases. This is also true for the ion foreshock which is the focus of the simulations presented in this paper. Beyond-MHD ion kinetics need to be modeled to describe collisionless interactions of shock-reflected ion populations with the solar wind (Scholer et al., 1993; Eastwood et al., 2005b).

Increasing computational resources enables new demanding approaches to be considered, either based on coupling dedicated models together (Tóth et al., 2012; Glozer et al., 2013), or by using a more appropriate set of equations to be solved throughout the system. The coupled approach may be quicker to implement if the dedicated models exist. On the other hand, the interfaces between the models may be problematic and in some cases the coupled approach does not yield a better description of the entire near-Earth space (e.g. Honkonen et al., 2013b). Therefore simulations based on an improved description of plasma throughout the system are a promising choice for developing new global models.

In the hybrid approach electrons are modeled as a fluid and ions are described by a kinetic model. Typically hybrid approaches are hybrid particle-in-cell (PIC) simulations, where ions are...
modeled with macroparticles for which plasma kinetic equations are solved (Winske et al., 2003). Such efforts have been dedicated to other solar system bodies (Brecht and Ferrante, 1991; Kallio and Janhunen, 2002) as well as to the Earth’s magnetosphere in two spatial dimensions (Lin, 2003; Omidi et al., 2005; Blanco-Canó et al., 2006; Omelchenko and Karimabadi, 2012) and occasionally in three spatial dimensions (Lin and Wang, 2005). This approach may yield solutions with an undesirable level of numerical noise, especially when looking at ion velocity distribution functions. Noise can be reduced by increasing the number of ions launched in the simulation, but the computational demands increase rapidly as the level of noise only goes down with the inverse of the square root of the number of particles.

Another choice is not to model the ions as particles, but to use Vlasov’s equation and model directly the evolution of the six-dimensional ion distribution function (three in ordinary space and three in velocity space) yielding a hybrid-Vlasov approach. Hybrid-Vlasov simulations require massive amounts of memory and computations to propagate the ion distribution function. Due to the computational complexity, the global hybrid-Vlasov approach has not received a lot of attention in the past, while local simulations have been implemented (Valentini et al., 2007, 2010; Eliasson and Shukla, 2007). On the other hand, when successful on a global scale, such a simulation will give an improved description for the ion distribution function and remove the numerical noise present in the previous hybrid-PIC approaches.

This paper presents the technical description and magnetospheric simulation results of Vlasiator (http://vlasiator.fmi.fi/), a self-consistent hybrid-Vlasov simulation code. While Vlasiator is inherently six-dimensional, here we report simulations in a five-dimensional global setup, comparable to Omidi et al. (2005), reporting kinetic properties of the collisionless bow shock in the ecliptic plane in two dimensions. This paper follows Palmroth et al. (2013), presenting first test-particle simulation results in the global scale on a coarse six-dimensional grid. An experimental version of Vlasiator that runs on graphics processing units has also been presented earlier (Sandroos et al., 2013). A verification study of the code in simulations of plasma wave dispersion was presented by Kempl et al. (2013) and a short presentation of first magnetospheric simulation results is given in Pokhotelov et al. (2013).

The paper is organized as follows: first, we describe the numerical scheme for the self-consistent hybrid-Vlasov approach in Section 2. In Section 3, we describe the sparse representation of the distribution function which enables global simulations on contemporary supercomputers and discuss the numerical diffusion. A local simulation of the ion/ion right-hand resonant beam instability is presented and compared to analytic results in Section 4. We then present the global magnetospheric case in Section 5, where we simulate the Earth’s magnetosheath and ion foreshock for three different interplanetary magnetic field (IMF) conditions in two ordinary space dimensions in the ecliptic plane, and three dimensions in velocity space. Our results present the first uniformly discretized ion velocity distribution functions with quality comparable to spacecraft measurements within the Earth’s ion foreshock (e.g. Kis et al., 2007) and in the magnetosheath (e.g. Souček and Escoubet, 2011). We will show that the ion foreshock includes the kinetic features present in Omidi et al. (2005), while mirror mode structures in the magnetosheath display characteristics present in Southwood and Kivelson (1993), Souček et al. (2008), promising fruitful avenues for further global simulations.

2. Numerical scheme

The fundamental description of charged particle motion in an electromagnetic field is given by Vlasov’s equation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{a} \cdot \nabla f = 0,$$

(1)

where \( \mathbf{r} \) and \( \mathbf{v} \) are the spatial and velocity coordinates, \( f(\mathbf{r}, \mathbf{v}, t) \) is the six-dimensional phase-space density of a particle species with mass \( m \) and charge \( q \), and acceleration \( \mathbf{a} \) is given by the Lorentz force

$$\mathbf{a} = \frac{q}{m} (\mathbf{E}_{\text{ion}} + \mathbf{v} \times \mathbf{B}),$$

(2)

where \( \mathbf{E}_{\text{ion}} \) is the electric field affecting ions and \( \mathbf{B} \) is the magnetic field. Vlasov’s equation is a six-dimensional advection equation. An inherent assumption in Vlasov’s equation is that the system is collisionless, which is a good approximation in tenuous space plasmas.

The bulk parameters of the plasma, such as the ion charge density \( \rho_i \) and current density \( \mathbf{j} \), are obtained as velocity moments of the ion velocity distribution function

$$\rho_i = q \int d^3v f(\mathbf{r}, \mathbf{v}, t),$$

(3)

$$\mathbf{j}_i = q \int d^3v \mathbf{v} f(\mathbf{r}, \mathbf{v}, t).$$

(4)

These also give the bulk velocity of ions,

$$\mathbf{V}_i = \mathbf{j}_i / \rho_i.$$  

(5)

Vlasiator is a self-consistent hybrid-Vlasov code implementing the Hall-less hybrid model introduced by Karimabadi et al. (2004), with ions described by a distribution function propagated according to Vlasov’s equation. The magnetic field is updated using Faraday’s law:

$$\nabla \times \mathbf{E}_{\text{OM}} = - \frac{\partial}{\partial t} \mathbf{B},$$

(6)

and the system is closed by an ideal Ohm’s law giving the electric field:

$$\mathbf{E}_{\text{OM}} = - \mathbf{V}_i \times \mathbf{B},$$

(7)

where \( \mathbf{V}_i \) is the ion bulk velocity given by Eq. (5). In the Lorentz force entering Vlasov’s equation (1) the electric field is given by

$$\mathbf{E}_{\text{ion}} = - \mathbf{V}_i \times \mathbf{B} + \frac{1}{\rho_i} \frac{\partial}{\partial t} \mathbf{j},$$

(8)

where the second term on the right-hand side is the Hall term. The Hall term needs to be included, otherwise the momentum equation is violated and no bulk force is exerted on the ions. The total current density \( \mathbf{j} \) is obtained from Ampère–Maxwell’s law where the displacement current has been neglected:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}.$$  

(9)

2.1. Ion propagation

We propagate the distribution function with a finite volume method (FVM). In FVM schemes the simulation domain is covered with a mesh consisting of a finite number of cells. In Vlasiator we have split the full six-dimensional mesh into ordinary and velocity space. In ordinary space we use a three-dimensional Cartesian mesh. We use the term spatial cell for cells in the ordinary space
mesh. Each spatial cell contains the field variables \((B, E)\). They are stored on a staggered grid such that the face average of the magnetic field is stored on the faces and the edge average of the electric field is stored on the edges (Londrillo and del Zanna, 2004). Each spatial cell also contains a three-dimensional Cartesian velocity mesh where each velocity cell contains the volume average \(\bar{f}\) of the distribution function over the ordinary space volume of the spatial cell, and the velocity space volume of the space mesh cell:
\[
\bar{f} = \frac{1}{V} \int \delta r \delta v \bar{f}(r, v, t),
\]
where \(\delta r = \Delta x \Delta y \Delta z\) and \(\delta v = \Delta v_x \Delta v_y \Delta v_z\) denote the phase-space integration volumes, and \(\Delta x, \Delta y, \Delta v_x, \Delta v_y, \Delta v_z\) are the sizes of the cell in each coordinate direction. Hereafter and in Eq. (10) we are omitting the coordinates \(r, v, t\) in \(\bar{f}\) and \(\bar{f}\) whenever the meaning is clear. The volume average \(\bar{f}\) is propagated forward in time by calculating fluxes at every cell face in each of the six dimensions. In the case of Vlasov’s equation the spatial \((F_x, F_y, F_z)\) and velocity \((F_{vx}, F_{vy}, F_{vz})\) fluxes take on particularly simple forms,
\[
F_x = v_x f, \quad F_y = \frac{q}{m} \left( v_y - \frac{v \cdot B}{\rho_s} \right) \times B, \quad F_z = v_z f,
\]
where we have substituted \(\rho_s \) and \(j\) by Eqs. (8) and (9) in the Lorentz force. \(B\) is a volume average calculated from face averages using divergence-free reconstruction polynomials given in Balsara et al. (2009). The propagation of \(\bar{f}\) is given by
\[
\bar{f}(t + \Delta t) = \frac{\Delta t}{\Delta x} \left[ F_x(x + \Delta x) - F_x(x) \right] - \frac{\Delta t}{\Delta y} \left[ F_y(y + \Delta y) - F_y(y) \right] - \cdots - \frac{\Delta t}{\Delta v_z} \left[ F_{vz}(v_z + \Delta v_z) - F_{vz}(v_z) \right].
\]
By construction the FVM scheme in Eq. (13) guarantees the conservation of mass except at the boundaries of the simulation domain.

We further split Eq. (13) into the spatial translation \(S_x\) and acceleration \(S_a\) operators. \(S_x\) propagates the distribution function in time in ordinary space using the fluxes in Eq. (11), and \(S_a\) propagates the distribution function forward in time in velocity space using the fluxes defined in Eq. (12). Both operators propagate the distribution function with a three-dimensional second-order accurate wave-propagation method (Leveque, 1997; Langseth and Leveque, 2000). The wave-propagation method is based on solving one-dimensional Riemann problems at the cell faces and applying flux limiters to suppress oscillations. Waves emanating from the Riemann problems are further split by solving Riemann problems in directions transverse to the wave propagation, and end up modifying fluxes on faces in transverse directions. Each cell face can contribute to the flux in up to 18 adjacent cells. We only add non-zero contributions to adjacent fluxes by identifying the affected cells based on the acceleration or velocity direction.

We use Strang (1968) splitting to propagate the six-dimensional distribution with second-order accuracy in time
\[
\bar{f}(t + N\Delta t) = \left[ S_f \left( \frac{\Delta t}{2} \right) S_x(\Delta t) S_a \left( \frac{\Delta t}{2} \right) \right]^N \bar{f}(0).
\]
We use the leap-frog algorithm for the propagation so that two translation half-steps are combined into one full timestep. To start the simulation chain, we first translate the distribution function forward by half a timestep,
\[
\bar{f}'(t + \frac{1}{2}\Delta t) = S_f \left( \frac{\Delta t}{2} \right) \bar{f}(t).
\]
After this, subsequent applications of the acceleration and translation operators propagate the distribution function forward by one timestep \(\Delta t\) from \(\bar{f}'(t + \frac{1}{2}\Delta t)\) to \(\bar{f}'(t + \Delta t)\),
\[
\bar{f}'(t + \Delta t) = S_a(\Delta t) \bar{f}'(t + \frac{\Delta t}{2}).
\]
\[
\bar{f}'(t + \Delta t) = S_x(\Delta t) \bar{f}'(t + \frac{\Delta t}{2}).
\]

The algorithm is stable as long as the timestep fulfills the Courant-Friedrichs-Lewy (CFL) condition, so that the Courant number is at most 1. The Courant number is for each dimension given by
\[
C = \frac{u \Delta t}{\Delta x},
\]
where \(u\) corresponds to the velocity; \(v_x, v_y, v_z\) in ordinary space and \(a_x, a_y, a_z\) in velocity space. As corresponds to the size of the simulation cell; \(\Delta x, \Delta y, \Delta z\) in ordinary space and \(\Delta v_x, \Delta v_y, \Delta v_z\) in velocity space. The timestep is dynamically adjusted so that it fulfills the CFL condition and is not shorter than necessary. The maximum velocity and the acceleration in all spatial cells are tracked and the maximum Courant number in the whole system is computed. If it falls outside of a user-specified range of Courant numbers (here 0.7–0.8) the timestep is changed so that the maximum Courant number falls into the middle of the range. We also compute the maximum timestep allowed by the field solver and set the timestep according to that solver if it is lower. To change the timestep we propagate the distribution function using a half-timestep propagation to a non-mixed state, change the timestep to its new value, and return to the leap-frog integration using another half-step. The timestep is only modified frequently so the overhead is small. This dynamic timestep also greatly helps with performance as we use optimal timesteps throughout the simulation.

In simulations with strong local fields leading to high acceleration, e.g., Earth’s dipole magnetic field, the timestep is limited by the acceleration \(S_a\) operator. To enable longer (global) timesteps we split the propagation in velocity space into shorter substeps, where the length of each substep is set according to the CFL condition so that the Courant number of the substep is in the middle of the allowed range. The magnetic field is constant during the substepping, but the effective electric field does change as we recompute the bulk velocity \(\bar{V}\) in Eq. (12) after each substep. By substepping the acceleration operator we can keep the global timestep reasonable (of the order of 50 ms in global simulations), which minimizes the total number of timesteps in a simulation. The length of the substep is computed on a cell-by-cell basis, thus substepping only happens close to Earth in magnetospheric simulations. When we substep, the propagation of the distribution function is not second-order accurate in time. As we only do it close to Earth, the effect should be negligible in regions that we investigate in this work. Also, as each timestep introduces additional numerical diffusion (see Section 3.2) the effective diffusion is lower due to the smaller number of timesteps.

2.2. Field propagation

In Vlasov the propagation of fields and ions is self-consistent. The fields couple to the Vlasov solver through the magnetic field which is used in the Lorentz force for the acceleration operator, and the Vlasov solver on the other hand needs the values of \(\rho_s\) and
which are computed from the distribution function. The magnetic field is propagated using the algorithm by 
and is described in the previous section. All reconstructions between volume-, face- and edge-averaged values follow Balsara et al. (2009). The face-averaged values of \( \mathbf{B} \) are propagated forward in time using the integral form of Faraday’s law (Eq. (6))

\[
\frac{\partial}{\partial t} \int \mathbf{B} \, d\mathbf{A} = - \oint_{\partial} \mathbf{E} \cdot d\mathbf{S},
\]

where the integral on the left-hand side is taken over a cell face and the line integral is evaluated along the contour of that face. Once \( \mathbf{E} \) has been computed based on Ohm’s law (Eq. (7)), it is easy to propagate \( \mathbf{B} \) using a discretized version of Eq. (19). When computing each component of \( \mathbf{E} \) on an edge, the solver computes the candidate values on the four neighboring cells of the edge. In the magnetosonic case, when the plasma velocity exceeds the speed of the fast magnetosonic wave mode, the upwinded speed, upwinded value from one of the cells is used. In the magnetosonic case the value is computed as a weighted average of the electric field on the four cells and a diffusive flux is added to stabilize the scheme.

For the time integration we use a second order Runge-Kutta method. To propagate the field from \( t \) to \( t + \Delta t \) we need to propagate \( \mathbf{B} \) and \( \mathbf{V} \) values at both \( t \) and \( t + \Delta t/2 \). With the leapfrog algorithm we do not have a real value for the distribution function at these times; computing it would be expensive. Following Valentini et al. (2007) we compute first-order accurate interpolations for the distribution function and use these to compute the required values:

\[
\tilde{f}(t) = \frac{1}{2} f(t) + \frac{1}{2} f(t + \Delta t) \]

(20)

\[
\tilde{f}(t + \frac{\Delta t}{2}) = \frac{1}{2} f(t) + \frac{1}{2} f(t + \Delta t).
\]

(21)

The field solver also contributes to the dynamic computation of the timestep. With our Ohm’s law the fastest characteristic speed is the speed of the fast magnetosonic wave mode and we use that speed to compute the maximum timestep allowed by the field solver. For the field solver we use Courant numbers between 0.4 and 0.5, as higher values cause numerical instability of the scheme (Londrillo and Zanna, 2004). In Vlasiator the magnetic field has been split into a perturbed field updated during the simulations, and a static background field. The electric field is computed based on the total magnetic field and all changes to the magnetic field are only added to the perturbed part of the magnetic field. The background field must be curl-free and thus the Hall term can be computed based on the perturbed part only. This avoids numerical integration errors arising from strong background field gradients. In magnetospheric simulations the background field consists of the Earth’s dipole, as well as a constant IMF in all cells.

3. Numerical implementation

Vlasiator is a hybrid-Vlasov code that is targeting global magnetospheric simulations. Simulating global events using a static discretization requires on the order of 1000 cells in each ordinary space dimension, and on the order of 100 cells in each velocity space dimension, giving in total \( 10^{12} \) phase-space cells for the full six-dimensional distribution function. As a first step we have targeted five-dimensional simulations (two spatial coordinates, three velocity coordinates) comprising on the order of \( 10^{12} \) phase-space cells. To enable these simulations we have developed a sparse representation of the distribution function that reduces the number of cells by up to two orders of magnitude (see Section 3.1). Additionally Vlasiator has to be able to efficiently utilize large supercomputers and clusters. To achieve good scalability we have implemented a two-level parallelization scheme. On the first level we use the MPI-based DCCRG grid library (Honkonen et al., 2013a) to implement the parallel three-dimensional ordinary space grid. The second parallelization level is done by threading the computation of each MPI process using OpenMP. In the field solver we have threaded computation by threading all major loops over spatial cells. When propagating the ion distribution function we have threaded the computation over the velocity space grid. This approach gives us good scaling to tens of thousands of cores. The large global simulations presented here have been performed on 16,384 cores on Hermit (Cray XE6), based in Germany at the High Performance Computing Center Stuttgart (HLRS).

3.1. Sparse velocity grid

The main computational load in a hybrid-Vlasov simulation is due to the six-dimensional distribution function. A typical ion population is localized in velocity space, for example a Maxwellian distribution, and a large portion of velocity space is effectively empty. A key technique for enabling large-scale global simulations in Vlasiator is the sparse velocity grid where the empty cells do not exist. As we neither propagate nor store the effectively empty velocity cells we save memory, I/O and computational resources. We have divided the velocity grid into velocity blocks comprising \( 4 \times 4 \times 4 \) velocity cells. The sparse representation is done at this level: either a block exists with all of its 64 velocity cells, or it does not exist at all. In the sparse representation we define that a block has content if any of its 64 velocity cells has a density above a user-specified threshold value. A velocity block exists if it has content, or if it is a neighbor to a block with content in any of the six dimensions. Neighbors are here defined in the sense of the minimum neighborhood needed by the solvers. In velocity space we include all 26 nearest neighbors, while in ordinary space we include up to second-nearest neighbors in the 6 face-normal directions and nearest neighbors in all other directions (Sandros et al., 2013). We can optionally also add the second-nearest neighbors in velocity space, but typically this is not done. The neighbors are included as the distribution function in contact blocks needs neighboring velocity cells to which it can flow, otherwise the extent of the distribution function could not change.

Analytically Vlasov’s equation conserves mass, as does our propagation algorithm in the limit of an infinite velocity space. A sparse velocity grid does not preserve all moments of the distribution function perfectly, but the impact is minimal with low threshold values. The losses come both from fluxes that flow out through the outer faces of the grid, as well as losses when a block is removed if it does not fulfill any existence criteria. Of these losses, the loss through the outer faces in velocity space dominates.

To quantify the losses we take the dispersion test presented in Kempf et al. (2013) and simulate it over five ion gyroradii with different threshold values. Additionally we also test the effect of changing the neighborhood size for block creation in velocity space. This decreases the losses as the outer boundary faces are farther away from the significant parts of the distribution function. In Fig. 1 the relative losses are plotted as a function of the threshold relative to the maximum velocity cell distribution function value for the initial state. In the range of relative thresholds between \( 10^{-4} \) and \( 10^{-5} \) the losses are of the order of 0.001%
of the total number density; larger thresholds show much larger losses while further reduction of the threshold value has very minor impact. Given the computational gain of using the sparse velocity space description, the sacrifice of mass conservation at such a low level is deemed beneficial. By adding an extra layer of blocks, we can further decrease losses by one order of magnitude. Fig. 1 shows the absolute value of the losses; in the usable range for the thresholds the losses can also be negative which implies that matter is created. The reason for this behavior lies in the shape of the distribution function. The use of flux limiters to ensure stability even when high gradients are present causes the distribution function to form a region of negative density on its outer edge, and the negative losses that occur for small threshold values are due to the boundary being inside this region.

In Fig. 2 the fill ratio, meaning the fraction of blocks which exist out of the maximal velocity space, is shown. Using a relative threshold lower than $10^{-5}$, or including second-nearest neighbors, only moderately decreases losses, while the number of blocks and thus the computational load increases significantly. In the simulations presented here we have used a threshold close to $10^{-5}$ and only included nearest neighbors as that gives a good balance between accuracy and performance.

The amount of saved resources depends on the actual limits of the velocity space, but for our simulations the code is at least an order of magnitude faster and the reduced memory load also enables us to run on less cores than otherwise. In the global magnetospheric simulation presented in Section 5 the fill ratio is on average 1.0%, reducing memory and computational load by up to 100 times. Typically up to 10% of simulation time is spent in computations related to the sparse representation, which is a small cost compared to the performance benefit. The code is also more robust as we can use a very large velocity space and there is no risk of the distribution function reaching the boundary.

### 3.2. Diffusion

Numerical diffusion in velocity space artificially enhances the thermal pressure of plasma during gyromotion. The numerical diffusion can be reduced by increasing the resolution in velocity space (e.g. Umeda et al., 2010). Various flux/slope limiters can be used to mitigate the numerical diffusion of hybrid-Vlasov schemes (Sandroos et al., 2013). Here the monotonized central (MC) limiter (van Leer, 1977) has been used since it does not disturb the shape of the original Maxwellian distribution. More aggressive limiters such as superbee or Sweby (Roe, 1986; Sweby, 1984) can reduce the diffusion more efficiently even for low velocity resolution but their applicability for kinetic plasma simulations is questionable as they tend to distort the shape by flattening the top and steepening the edges of the distribution function.

In Fig. 3 we have plotted the artificial increase in temperature for a Maxwellian ion distribution function with a bulk velocity $V_i = 500 \text{ km/s}$, number density $n_i = 10^6 \text{ m}^{-3}$ and temperature $T_i = 10^5 \text{ K}$. This distribution is propagated in a magnetic field of strength 5 nT with a timestep of 0.05 s. These parameters are chosen to match the values we use for the solar wind in one of the global simulations in Section 5. The system is homogeneous in ordinary space and the diffusion stems purely from gyromotion in velocity space. For our solver which is second order accurate in space, the numerical heating is reduced by a factor of almost 4 by doubling the velocity resolution. Due to limited computational resources the choice of velocity resolution is a compromise between feasibility of the simulations and the level of artificial heating. In this work the choice was 20 km/s for the global simulations. At this resolution the temperature of the solar wind impinging on the different parts of the bow shock does not differ substantially. In the 45° IMF simulation (see Section 5), the plasma hitting the nose of the shock has been propagated 5500 timesteps and has a temperature of $4.8 \times 10^5 \text{ K}$, while plasma hitting the outermost flanks has been propagated 14,000 steps and has a temperature of $8.4 \times 10^5 \text{ K}$. The temperature is significantly higher...
than at the solar wind inflow boundary (100,000 K), but the relative increase across the foreshock region is deemed acceptable.

3.3. Boundary conditions

Both in ordinary and velocity space the simulation has a maximum extent, and the outer faces of the simulation box have a boundary condition. In velocity space the boundary condition is an outflow condition, and if the distribution function hits the limit of the velocity space the simulation loses matter. To avoid this, the maximum velocity limits are set to large values based on a conservative estimation of the largest possible velocities for the particular simulation case.

In ordinary space Vlasiator supports periodic boundaries used in the dispersion test in Section 3.1 and in the z-direction in the global simulation. It also supports the three distinct boundaries found in a magnetospheric simulation: (1) sunward boundary of the simulation from which the solar wind flows into the simulation, (2) inner boundary of the simulation which surrounds the Earth at a user-defined distance from the Earth's center and (3) outflow boundary conditions.

At the solar wind boundary a static Maxwellian ion velocity distribution is set in all cells, and that distribution function is unchanged throughout the simulation. The temperature and velocity of the solar wind are set through the width and location of the Maxwellian distribution. Vlasiator does not have divergence removal, therefore we should not introduce magnetic field divergence at the solar wind boundary. To set an IMF we simply add a constant magnetic field to the background field in all cells including the boundary cells. The dipole has a much stronger value close to the Earth, thus the additional contribution of the IMF is small there. At the inner boundary the distribution function is set to a stationary state and the perturbed magnetic field is set to zero. This inner boundary is far from being realistic with respect to the actual inner magnetosphere and can be viewed as a first implementation that enables studies of phenomena, such as the ion foreshock, which are located far from the inner boundary. At the outflow boundary each cell copies from its nearest normal neighbor cell the distribution function, and its magnetic field value. This allows plasma to flow out of the system.

4. Ion/ion right-hand resonant beam instability

Under favorable IMF conditions the Earth’s bow shock reflects incoming particles which will stream back along the magnetic field, forming the ion foreshock (e.g. Eastwood et al., 2005b). This can lead to ion/ion beam instabilities which grow and generate waves in the foreshock region. We investigate the ion/ion right-hand resonant beam instability in the regime relevant to the global magnetospheric simulations presented in Section 5 and in Pokhotelov et al. (2013).

When the thermal speeds $v_{th}$ are small compared to the beam speed $v_b$, we are in the cool beam regime. When additionally the Alfvén speed $v_A$ is small $v_A \ll v_b$, the maximum growth rate $\gamma_m$ of the right-hand resonant ion/ion beam instability can be given as

$$\frac{\gamma_m}{\omega_G} \approx \left( \frac{n_e}{2n_b} \right)^{1/3},$$

where $\omega_G$ is the ion cyclotron frequency, $n_b$ is the beam number density, $n_e$ is the core number density and $n_z = n_e + n_b$ is the electron or total ion number density (Gary, 1978, 1991; Gary et al., 1984). According to Gary (1978) this approximation is matching numerical results best in the range $10^{-3} < n_b/n_e < 10^{-1}$. In addition the results in Gary et al. (1984) show that the wave number and the growth rate of the fastest-growing mode also depend on the beam speed relative to the Alfvén speed.

The simulation used to compare with this analytic result is one-dimensional with fully periodic boundary conditions. The simulation box has a length of 10 Earth radii ($R_E$), 70 spatial cells and 20 km/s velocity space resolution, to match the parameters of the simulations presented in Section 5. The simulation box is aligned with the uniform magnetic field $B_z$ of 5 nT. The orthogonal magnetic field component $B_y$ only has small random perturbations (<1% amplitude). The ion distribution consists of a core Maxwellian with $n_e = 1.0 \times 10^6$ m$^{-3}$ number density, $1.0 \times 10^9$ K temperature and 500 km/s bulk velocity along the magnetic field as well as a beam with $1.0 \times 10^8$ K temperature and 500 km/s bulk velocity in the opposite direction. The beam density is varied between $5.0 \times 10^5$ m$^{-3}$ and $2.0 \times 10^4$ m$^{-3}$. We have $v_{th} = 41$ km/s, $v_k = 109$ km/s and $v_b = 1000$ km/s, thus we are in the regime of Eq. (22) ($b_k < v_b \approx v_b$).

In the simulation, the fastest-growing mode has the wave number $m = 5$, that is a wavelength of 2 $R_E$ and a period close to 30 s (or a frequency of $\omega_m \approx 0.44 m_b$). In Fig. 4 we show the analytic result from Eq. (22) for the growth rate of the ion/ion right-hand resonant beam instability together with the Vlasiator simulation results. It is obvious that the growth rates are not as high as the analytic maximum rate. This arises for one part from the fact that the wave length is constrained to be an integer fraction of the box size. Additionally, Eq. (22) is derived under the assumption that $\gamma_m \approx \omega_m$, which is not a very accurate approximation in the present case. However the dependency of the growth rate on the beam-to-total number density ratio is reproduced quite accurately. The comparison with the top right and bottom left panels in Fig. 1 of Gary et al. (1984), which are the two closest to the case simulated here, shows that their growth rates in numerical computations based on linear Vlasov theory are lower than the analytic result too and close to our results.

5. Global magnetospheric simulations

The global magnetospheric simulation is set up to cover the region of near-Earth space from the inner boundary (defined at 6 $R_E$ radial distance) to the solar wind thus covering the dayside part of the Earth’s magnetosphere, the magnetosheath, the bow shock and the foreshock region. The Geocentric Solar Ecliptic (GSE) coordinate system is used with its x-axis pointing from the Earth towards the Sun, its y-axis in the ecliptic plane pointing towards dusk and its z-axis perpendicular to the ecliptic plane. Table 1 specifies the parameters for the three different simulations

![Fig. 4. Growth rate of the ion/ion right-hand resonant beam instability. The analytic solution corresponds to Equation (22), the simulation results correspond to the parameters presented in Section 4.](image-url)
presented in this section. The simulation box is shifted in each simulation to capture the ion foreshock region. The spatial resolution of the simulations is $0.13 \, R_E$, while the velocity space has a resolution of 20 km/s, and its range in all three velocity dimensions is from $-2000 \, \text{km/s}$ to $2000 \, \text{km/s}$. The spatial resolution was limited by the computational demands of the approach, and it is coarser than the ion skin depth and the ion gyroradius. The Earth’s magnetic dipole is directed along the $z$-axis, and has a strength of $8.0 \times 10^{22} \, \text{Am}^2$. The simulations are set up with constant typical solar wind conditions given in Table 1. In the three simulations the IMF is pointing at $0^\circ$, $30^\circ$, and $45^\circ$ angle with respect to the $x$-axis. To initialize the simulation, the IMF components are set over the entire simulation box (see Section 3.3).

Fig. 5 presents an overview of the $45^\circ$ IMF orientation simulation, representing a typical Parker spiral condition. The bow shock has two distinct regions: the quasi-perpendicular shock region in the dusk sector and the quasi-parallel shock region in the dawn sector. The ion foreshock boundary is clearly seen as a line separating the regions of quasi-parallel and quasi-perpendicular

<table>
<thead>
<tr>
<th>IMF Angle</th>
<th>Solar wind $B_n$ (nT)</th>
<th>$V$ (km/s)</th>
<th>$n_i$ (cm$^{-3}$)</th>
<th>$T_i$ (K)</th>
<th>Simulation box</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0^\circ$</td>
<td>5.0</td>
<td>700</td>
<td>2</td>
<td></td>
<td>$-10 \text{ to } 70$ \ $-40 \text{ to } 40$</td>
</tr>
<tr>
<td>$30^\circ$</td>
<td>4.0</td>
<td>700</td>
<td>2</td>
<td></td>
<td>$-10 \text{ to } 70$ \ $-60 \text{ to } 40$</td>
</tr>
<tr>
<td>$45^\circ$</td>
<td>5.0</td>
<td>500</td>
<td>1</td>
<td></td>
<td>$-20 \text{ to } 40$ \ $-67 \text{ to } 52$</td>
</tr>
</tbody>
</table>
bow shock. While the quasi-perpendicular bow shock is unable to reflect particles into the solar wind, the quasi-parallel bow shock is characterized by reflected populations of ions accelerated by the shock and streaming back against the incoming solar wind. In the simulation, this region is characterized by ion distributions known as intermediate type distributions (Fuselier, 1995) with the incoming solar wind (Maxwellian distribution) combined with a cap-shaped population of backstreaming ions. This type of ion distributions is typically associated with quasi-monochromatic foreshock waves routinely observed in the ion foreshock known as 30 s waves (Le and Russell, 1994) believed to be generated by the backstreaming shock-energized ion population via ion/ion resonance interaction (Gary, 1991). Multi-spacecraft analysis using Cluster satellites shows these waves to be compressional magnetosonic modes with periods close to 30 s and wavelengths of ~1 RE propagating both parallel as well as obliquely with respect to the background magnetic field with angles up to 30° (Russell, 1988; Eastwood et al., 2004, 2005a).

Fig. 5 shows an example of an intermediate type ion distribution (point D), and a ring distribution (point B) at the foreshock boundary. The backstreaming ion populations appear to be more diffused deeper in the foreshock near the bowshock (point C), which is generally consistent with observational statistics (Fuselier, 1995). Associated wave structures are visible throughout the dawn (quasi-parallel) region of the foreshock. The magnetosheath in this simulation is formed behind the bow shock and consists of highly anisotropic shock-energized plasma. A typical ion distribution from the magnetosheath (point A) is presented in Fig. 5. It appears as a gyrotropic bi-Maxwellian with substantially higher temperature in the direction transverse to the background magnetic field. Such ion distributions are known to be subject to the mirror mode instability as well as to the ion-cyclotron instability (Southwood and Kivelson, 1993). Mirror mode structures with spatial scales of few ion inertial lengths are routinely observed across the Earth’s magnetosheath (Souček et al., 2008) and characterized by an anti-correlation between plasma density and total magnetic field. Magnetosheath structures seen in Fig. 5 demonstrate strong anti-correlation between total magnetic field and density and can be interpreted as large scale mirror modes (Pokhotelov et al., 2013). Smaller scale ion-cyclotron Alfvénic waves also routinely observed in the magnetosheath do not appear in this simulation, likely due to insufficient spatial resolution.

Fig. 6 shows an overview of the two other global simulations presented in this paper, the top row showing the radial and the bottom row the 30° case. The left panels give the number density, while the right panels show the magnetic field color-coded for the simulation in question. The white lines refer to Fig. 7, while the

---

Fig. 6. The ion foreshock simulated with Vlasiator in two spatial and three velocity dimensions for 0° IMF orientation (top row) and 30° IMF orientation (bottom row). Left: number density (cm⁻³); right: magnetic field magnitude (nT). The white dots refer to the position for which Fig. 8 is computed.
dots refer to Fig. 8. For the radial case, the foreshock forms within a large area surrounding the bow shock, extending from the subsolar point to the flanks. The wave structure within the foreshock appears similarly compressional as in Fig. 5, and also exhibits a structure along the direction perpendicular with respect of the IMF. There are both parallel and oblique waves within the foreshock as also evident in observations (Russell, 1988; Eastwood et al., 2004, 2005a). In the 30° case, on the other hand, the foreshock forms in the dusk part of the flank, as expected from the IMF direction and the waves are again compressional.

To investigate the wave properties in more detail, in Fig. 7 we present the number density and magnetic field along a cut parallel to the magnetic field within the foreshock in each of three simulations. Fig. 7a, b refers to the radial case, 7c, d to the 30° case, and 7e, f to the 45° case, respectively. Fig. 7 indicates that in each simulation, the waves appear to be quasi-monochromatic compressional oscillations. In the radial case, the wave magnitude is about 10% of the background value for both the density and the magnetic field, consistent with theoretical predictions for the fast magnetosonic perturbations in the MHD limit (Le and Russell, 1994). The wavelength is about 1–2 RE. For the 30° case (Fig. 7c and d), the wave magnitude in magnetic field is again about 10% while the density perturbation is somewhat larger, about 15–20%, reflecting possibly the small changes in the input solar wind conditions (Table 1) compared to the radial case. The 30° case exhibits also a larger wavelength, above 3 RE, than compared to the radial case, and the number of waves along the cut through the parallel direction is therefore smaller. The 45° case shows smaller and larger amplitude waves depending on the location within the foreshock, and the amplitudes are around 10%. The wavelength is about 1–2 RE. Interestingly, the wave pattern close to the bow shock is different in nature compared to the compressional waves...
appearing around 15 Re distance from the bow shock surface. In all three cases, the wave patterns steepen towards the bow shock surface.

To investigate the wave period, we present in Fig. 8 one-dimensional data taken from a fixed position in space, plotted against time of the simulation. Fig. 8a and b is for the radial case, while Fig. 8c, d and e, f is for 30° and 45° cases, respectively, and again number density and magnetic field are presented. Fig. 8 corroborates the compressional nature of the wave field, while interesting deviations from this are also present (e.g., in the 30° case around 980 s, where the density and the magnetic field are uncorrelated). The periods of the waves range from around 20 s in the radial case to about 30 s in the 30° case and slightly above 30 s in the 45° case.

6. Conclusions and summary

In this paper we present a self-consistent hybrid-Vlasov simulation code, Vlasiator, that has been designed for global magnetospheric simulations. It uses a second-order accurate finite volume method to solve Vlasov’s equation and supports full six-dimensional cases (three spatial and three velocity coordinates). The method uses the Strang splitting to separate propagation in ordinary and velocity spaces, and is based on a three-dimensional wave-propagation algorithm. The field solver is a second-order accurate upwind constrained transport method, which is divergence-conserving by construction.

Vlasiator is parallelized with MPI and OpenMP to scale to high-end supercomputers in order to run the simulations on length scales comparable to the Earth’s magnetosphere. An important
algorithmic improvement has been done to increase performance:  
the distribution function is described with a sparse grid in velocity  
space. Additionally the velocity space propagation is substepped,  
as the strong magnetic fields close to Earth are limiting the  
timesteps to very small values. These techniques improve the  
performance of the simulation by two to three orders of  
magnitude.

Overall, numerical modeling of Vlasov’s equation with finite  
volume methods looks very promising. The results presented here  
as well as in Kempf et al. (2013) show that Vlasiator is robust, and can be  
used for global five-dimensional simulations (two in ordinary  
space, three in velocity space). The results suggest that Vlasiator  
is able to reproduce the key features of solar wind–magnetosphere  
interactions such as the collisionless bow shock, the ion foreshock  
and the magnetosheath. Characteristics of the backstreaming ion  
populations and associated electromagnetic waves are in agreement  
with the properties of ion velocity distribution functions and  
compressional magnetosonic waves propagating both obliquely  
and parallel with respect of the background magnetic field  
typically observed in the Earth’s ion foreshock region (Russell,  
1988; Eastwood et al., 2004, 2005a). Our results show that the  
foreshock waves have amplitudes around 10–20% with respect to  
the background value, depending on the simulation in question.  
Observational results also indicate that the foreshock waves can  
appear with a variable amount of compression (see e.g. a review  
Russell, 1988). The wavelengths vary from 1 Rs to above 3 Rs while  
the periods vary from 20 s to above 30 s, indicating a good match  
to observations (Eastwood et al., 2005a). Earlier hybrid simulations  
of the ion foreshock dynamics (Omidi et al., 2005; Blanco-Canó  
et al., 2006) reproduced similar quasi-monochromatic wave  
structures across the ion foreshock that, however, had mainly non-  
compressive Alfvénic nature, in contrast to the results shown here.  
Other types of wave structures known to exist in the ion foreshock,  
such as steepen nonlinear shear waves (Le and Russell, 1994)  
believed to be associated with gyrating ion distributions  
(Fuselier, 1995; Meziane et al., 2001) have been simulated earlier  
by Omidi et al. (2005) and Blanco-Canó et al. (2006) but do not  
appear in our simulations, possibly due to the simulation box  
not being large enough for the high-amplitude nonlinear structures to  
evolve before the foreshock structures get adveced by the solar  
wind flow into the bow shock. The absence of shocklets can be also  
attributed to the simplified ideal Ohm's law used here (e.g.,  
absence of the Hall term and electron pressure gradient term). In  
the magnetosheath region the simulations presented here as well  
as in Pokhotelov et al. (2013) reproduce highly anisotropic bi-  
Maxwellian distributions and associated mirror mode structures.  
Such anisotropic distributions and large scale mirror mode structures  
are routinely observed deep in the magnetosheath (Soucek and  
Escoubet, 2011; Soucek et al., 2008). Other instabilities that are  
responsible for the generation of smaller scale ion cyclotron  
waves observed in the magnetosheath are not resolved in the  
current simulations due to insufficient spatial resolution.

The hybrid-Vlasov model is similar to a hybrid-PIC model, and  
our global simulations demonstrate qualitative agreement with  
global simulations of the collisionless bow shock with hybrid-PIC  
codes. The most striking difference with respect to hybrid-PIC  
simulations is that hybrid-Vlasov velocity distribution functions  
appear as uniformly discretized functions similar to those seen in  
experimental data in contrast to spiky distributions derived from  
hybrid-PIC simulations. The more tenuous parts of phase space are  
well resolved in this model. This also means that quantities  
derived from the distribution function, such as density, also exhibit  
no visible noise in real space. The drawback of the approach is that  
it is computationally more demanding when run at comparable  
spatial resolution. In a particle simulation one typically has only a  
few hundreds of particles per spatial cell, while our simulations  
have on average more than $10^3$ velocity cells to properly resolve  
the distribution function. Compared to fluid approaches (MHD)  
the approach is computationallly orders of magnitude more  
demanding, but on the other hand the Vlasiator can simulate kinetic  
physics that cannot be reproduced with fluid models.

The current simulations represent the first use of a hybrid-  
Vlasov scheme for global magnetospheric simulations and thus no  
attempt was made to reach the ion gyroscales or the ion inertial  
scales in ordinary space. This was mostly due to the computational  
demands of the version of the code used for this work. These  
simulations required a few million processor core hours. Such  
approaches would also require implementation of a generalized  
Ohm’s law with Hall and electron pressure gradient terms of  
appropriate accuracy to properly resolve the physics at these  
scales. However we show that a number of well-known features of  
the collisionless bow shock and the ion foreshock can be  
simulated using the ideal Ohm’s law and without resolving all  
the kinetic scales. The velocity space resolution (20 km/s) was  
selected to limit numerical diffusion, and to properly resolve  
details in the magnetosheath distributions and the backstreaming  
populations (Fig. 5). The resolution is insufficient for properly  
resolving the initial solarwind, and numerical diffusion heats the  
solar wind up to a temperature of $4.8 \times 10^7 \text{K}$ at the nose of the  
bowshock. With a higher initial temperature the solar wind would  
also have been properly resolved throughout the system.  

The aim for Vlasiator is to perform similar simulations at  
resolutions comparable to ion inertial length scales, with improved  
physical description of the system. The physical description  
will be improved by adding further terms to Ohm’s law as well as  
by adding a time-varying solar wind to enable event simulations.  
Ion inertial length scales are on the order of 200 km and this  
implies a four- to five-fold increase in ordinary space resolution  
comparing to the simulations presented here. There are multiple  
avenues for further increasing the performance of Vlasiator  
to enable these simulations, and ion inertial scale simulations appear  
feasible. Some options include (1) semi-Lagrangian solver for  
propagating the distribution function (Zerroukat and Allen, 2012)  
allowing longer time-steps and thus removing the need to substep  
velocity space propagation, combined with higher order recon- 
structions (White and Adcroft, 2008) to reduce the diffusivity of  
the approach, (2) adaptive mesh refinement in all six dimensions  
(Arslanbekov et al., 2013) which could enable fully six-dimen- 
sional simulations with kinetic scales in selected regions of space.  
(3) sub-cycling the propagation in regions of space close to Earth  
and finally (4) continue to scale the code to even larger super-  
computers. It can be noted that the fastest available machines at  
the moment have a peak performance more than 200 times larger  
than the resources used here, and the performance of these  
machines continues to follow an exponential growth path.  

The performance and the physical results prove that Vlasiator is  
a viable platform for global magnetospheric physics, and even at  
this early stage it promises key benefits in the form of noise-free  
velocity distribution functions where even the tenuous regions are  
well-defined.
Acknowledgments

The Quantifying Energy circulation in Space plasmas (QuESpace) project has received funding from the European Research Council under the European Community’s seventh framework programme (FP-7/2007-2013/ERC) agreement no. 200141-QuESpace. The work is supported by the Academy of Finland under project nos. 138599 and 267144.

We acknowledge that the results of this research have been achieved using the PRACE Tier-0 Research Infrastructure resource Hermit based in Germany at the High Performance Computing Center Stuttgart (HLRS) and the PRACE Tier-1 Research Infrastructure resource Abel owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR) and operated by the Department for Research Computing at USIT, the University of Oslo IT department.

References


Paper III

Ion distributions in the Earth’s foreshock: Hybrid-Vlasov simulation and THEMIS observations

© 2015 American Geophysical Union.

Ion distributions in the Earth’s foreshock: Hybrid-Vlasov simulation and THEMIS observations

Yann Kempf1,2, Dimitry Pokhotelov2,3, Olga Gutynska4, Lynn B. Wilson III4, Brian M. Walsh4, Sebastian von Alfthan1, Otto Hannuksela1,2, David G. Sibeck4, and Minna Palmroth1

1Earth Observation Unit, Finnish Meteorological Institute, Helsinki, Finland, 2Department of Physics, Helsinki University, Helsinki, Finland, 3Mullard Space Science Laboratory, University College London, Dorking, UK, 4NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, 5Space Sciences Laboratory, University of California, Berkeley, California, USA

Abstract We present the ion distribution functions in the ion foreshock upstream of the terrestrial bow shock obtained with Vlasiator, a new hybrid-Vlasov simulation geared toward large-scale simulations of the Earth’s magnetosphere (http://vlasiator.fmi.fi). They are compared with the distribution functions measured by the multispacecraft Time History of Events and Macroscale Interactions during Substorms (THEMIS) mission. The known types of ion distributions in the foreshock are well reproduced by the hybrid-Vlasov model. We show that Vlasiator reproduces the decrease of the backstreaming beam speed with increasing distance from the foreshock edge, as well as the beam speed increase and density decrease with increasing radial distance from the bow shock, which have been reported before and are visible in the THEMIS data presented here. We also discuss the process by which wave-particle interactions cause intermediate foreshock distributions to lose their gyrotropy. This paper demonstrates the strength of the hybrid-Vlasov approach which lies in producing uniformly sampled ion distribution functions with good resolution in velocity space, at every spatial grid point of the simulation and at any instant. The limitations of the hybrid-Vlasov approach are also discussed.

1. Introduction

The supermagnetosonic solar wind flow impinging the Earth’s magnetic field creates the bow shock. At this boundary, the plasma is compressed and heated while slowing down to submagnetosonic flow speeds. In fluid theory no information can travel upstream of a shock, but kinetic processes can cause particles to be reflected back off a collisionless shock and propagate upstream along the magnetic field lines. The upstream region magnetically connected to the bow shock where reflected particle populations can interact with the solar wind population is called the foreshock [e.g., Eastwood et al., 2005a].

The foreshock can be separated into two regions conditioned by the ability of reflected particles to escape upstream from the bow shock or not. In the region where the angle $\theta_{\text{in}}$ between the interplanetary magnetic field (IMF) and the bow shock normal direction is too high, the reflected particles do not have enough energy to propagate upstream along the IMF lines and they are immediately advected back into the shock by the solar wind. In this region the foreshock is restricted to the collisionless bow shock foot [e.g., Paschmann et al., 1982]. When $\theta_{\text{in}}$ is low enough, reflected ions with sufficient energy parallel to the IMF can propagate upstream into the solar wind and generate waves and instabilities [e.g., Bavassano-Cattaneo et al., 1983; Sanderson et al., 1983; Themens et al., 1985; Fuselier et al., 1987; Wilson et al., 2009, 2012]. In all but radial IMF conditions the solar wind flow has a component normal to the magnetic field line, which causes the edge of the foreshock to be deflected downstream toward the shock. This is much more pronounced for ions as they have a lower parallel velocity than electrons, and thus, the electron foreshock edge is almost along the IMF whereas the ion foreshock edge is bent toward the shock.

The instabilities triggered by the backstreaming ions have a finite growth time; therefore, there is a region at the foreshock edge where backstreaming ions are present, but the instabilities have not yet grown to generate significant ultralow frequency (ULF) waves. The region where ULF waves are present is called the ULF foreshock, and it is delimited in the upstream direction by the ULF foreshock boundary. A variety of ULF wave modes can develop in the foreshock, including 1 Hz waves, 3 s waves, compressive sinusoidal waves with periods of about 30 s, and steepened shocklets. See, for example, the review given by Le and Russell [1994].
Early observations led to a categorization of backstreaming ions into reflected and diffuse populations [Gosling et al., 1978; Bonifazi et al., 1980], but quickly a third category, intermediate populations, was identified [Bonifazi and Moreno, 1981]. The reflected ions, more commonly called field-aligned beams (FABS) in recent literature, exhibit a narrow pitch angle distribution collimated along the magnetic field direction and streaming back with respect to the solar wind. Diffuse ion distributions are characterized by a wide pitch angle distribution and a broad energy distribution. Intermediate distributions are field aligned similarly to the FABS, but they have a kidney-bean or crescent shape which is concave towards the solar wind core direction, in a cut plane including the mean backstreaming velocity and magnetic field vectors. Further types of ion distributions have also been observed. Gyrating ions have a narrow gyrotropic ring-shaped distribution centered on the magnetic field direction, and gyrophase bunched distributions have a nonzero mean velocity component orthogonal to the field direction. For a review of ion distribution types in the foreshock, see, for example, Fuselier [1994].

Simulations of collisionless shocks have been performed for a long time to help interpret the observations and to test theories. The dimensionality and complexity of collisionless shock simulations have increased in pace with the available computing power. Early simulations were one dimensional in space and used hybrid particle-in-cell (hybrid-PIC) algorithms [e.g., Leroy et al., 1982]. They soon moved to full particle simulations [e.g., Lembège and Dawson, 1987] and increased in dimensionality to 2-D hybrid- and full-PIC simulations [e.g., Lembège and Savoini, 1992; Savoini et al., 2010, 2013], gaining in complexity and detail of the physical description at large scales [e.g., Lin, 2003; Omidi et al., 2005; Blanco-Cano et al., 2006a] as well as for local kinetic processes [e.g., Ofman and Gedalin, 2013a, 2013b] and even reaching three spatial dimensions for hybrid-PIC simulations of the magnetosphere [e.g., Lin and Wang, 2005]. PIC-based methods suffer from the statistical noise inherent to the random sampling of the particle distribution injected in the simulation and the low number of particles that can feasibly be simulated. This statistical noise is practically unimportant when computing global quantities or moments of the particle velocity distributions, but it becomes more prominent when considering the velocity distributions themselves. Apart from spatial or temporal averaging, one way to alleviate this is to use algorithms based on Vlasov’s equation, which are, however, significantly more computationally intensive. As a consequence, hybrid-Vlasov models have been used for local foreshock simulations in one spatial dimension [e.g., Eliasson and Shukla, 2007], but they have not been considered feasible in this context in higher dimensions so far because of the lack of achievable resolution.

In this paper we present the ion distribution functions in the foreshock which have been simulated using the hybrid-Vlasov model Vlasiator [Palmroth et al., 2013; Pokhotelov et al., 2013; Kempf et al., 2013; von Alfthan et al., 2014] (see also http://vlasiator.fmi.fi). We use the simulation first presented by Pokhotelov et al. [2013], where the model was used to simulate the Earth’s magnetosheath, bow shock, and foreshock in two spatial and three velocity dimensions in the equatorial plane with a realistic strength of the magnetic dipole field. The hybrid-Vlasov approach ensures a uniform sampling of the ion distribution function in all spatial and velocity dimensions, as the full three-dimensional ion velocity distribution function is propagated in each real space cell. This relieves from the statistical noise due to the low number of superparticles that can be injected into hybrid- or full-PIC simulations, however, at a high computational cost. The unprecedented description allows to draw a complete snapshot picture of the ion distribution types encountered in the whole foreshock. These results are compared to the ion distributions measured by the Time History of Events and Macroscale Interactions during Substorms (THEMIS) multispacecraft mission [Angelopoulos, 2008], showing a good match of the simulation and the observed data, thus validating the novel Vlasiator code. This simulation also supports the view drawn from observations by Meziane et al. [2001] that foreshock ULF waves are the consequence of ion beam instabilities generated by backstreaming ion populations reflected by the shock and that the gyrophase bunched ion distributions seen in regions with strong ULF waves are the result of beam disruption by the waves.

The article is structured as follows: we introduce the simulation and the measurements used in section 2, then present the modeled gyrotropic and gyrophase bunched ion distributions in sections 3 and 4 before comparing them to spacecraft data in section 5. A discussion of the results is given in section 6, including a discussion of the limitations of the hybrid-Vlasov approach with respect to hybrid-PIC modeling, and conclusions are drawn in section 7.

KEMPF ET AL. FORESHOCK ION DISTRIBUTIONS
2. Methods
2.1. Global Magnetospheric Hybrid-Vlasov Simulation

The hybrid-Vlasov simulation code Vlasiator [von Alfhans et al., 2014] uses a finite volume approach to propagate the ion distribution function following Vlasov’s equation

\[
\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{a} \cdot \nabla f(r, \mathbf{v}, t) = 0,
\]

where \( r \) and \( \mathbf{v} \) are the spatial and velocity coordinates, \( f(r, \mathbf{v}, t) \) is the six-dimensional phase space density of a particle species with mass \( m \) and charge \( q \), and \( \mathbf{a} \) is the Lorentz force per unit mass

\[
\mathbf{a} = \frac{q}{m}(\mathbf{E} + \mathbf{v} \times \mathbf{B}).
\]

where \( \mathbf{E} \) is the electric field and \( \mathbf{B} \) is the magnetic field. Vlasov’s equation is coupled to Maxwell’s equations in which the displacement current in Ampère-Maxwell’s law (5) is neglected:

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},
\]

\[
\mathbf{v} \cdot \mathbf{B} = 0, \quad \text{and}
\]

\[
\nabla \times \mathbf{B} = \mu_0 \mathbf{j},
\]

where \( \mu_0 \) is the permeability of free space and \( \mathbf{j} \) the total current density. The system is closed with Ohm’s law, which takes the ideal magnetohydrodynamic (MHD) form

\[
\mathbf{E} = -\mathbf{V} \times \mathbf{B},
\]

where the ion velocity \( \mathbf{V} \), is obtained from moments of the ion distribution. This closure neglects electron inertia. Note also that the electric field in equation (6) for the field propagation does not include the Hall term \( \mathbf{j} \times \mathbf{B} / \rho_\text{H} \) (\( \rho_\text{H} \): charge density). This imposes a restriction on the possible wave modes in the model, as, for example, the parallel-propagating low-frequency wave modes do not split into the electron and ion whistler modes [Kempf et al., 2013].

The simulation used in this work covers the region of near-Earth space from the inner boundary (at 6 Earth radii \( R_\text{E} \) radial distance from the origin) to the solar wind, thus covering the dayside and some of the nightside magnetosphere, the magnetosheath, the bow shock, and the foreshock region. The Geocentric Solar Ecliptic (GSE) coordinate system is used \( (X_{GSE}, Y_{GSE}, Z_{GSE}) \), where \( X_{GSE} \) is the six-dimensional phase space density of the ion velocity \( \mathbf{V} \), obtained from moments of the ion distribution. This closure neglects electron inertia. Note also that the electric field in equation (6) for the field propagation does not include the Hall term \( \mathbf{j} \times \mathbf{B} / \rho_\text{H} \) (\( \rho_\text{H} \): charge density). This imposes a restriction on the possible wave modes in the model, as, for example, the parallel-propagating low-frequency wave modes do not split into the electron and ion whistler modes [Kempf et al., 2013].

The simulation is run with constant Maxwellian upstream solar wind conditions \( (V_{SW} = 500 \text{ km/s}, n_{SW} = 10^6 \text{ m}^{-3}, T_e = 10^5 \text{ K}) \) and an IMF of 5 nT magnitude pointing at an angle of 45° with respect to the \( X_{GSE} \) axis \( (\text{IMF} B_x = -B_y = 3.5 \text{ nT}) \), which corresponds to an outward Parker spiral condition. The simulation is initialized with a Maxwellian background population of density \( n_0 = 10^5 \text{ m}^{-3} \) and temperature \( T_e \) flowing at \( V_{SW} \), and the Earth’s dipole field as well as the IMF in the whole simulation volume. The solar wind flows in with the density \( n_{SW} \) from the upstream boundary \( (+X_{GSE}) \) while the initial \( n_0 \) density background flushes through the simulation box and the other three outer boundaries.
Figure 1. Color code: ion number density (m$^{-3}$) in the foreshock and magnetosheath regions of the simulation. Black line: foreshock edge, upstream of which the ion distribution consists only of the solar wind Maxwellian distribution. White plus signs and letters: location of the distribution function types presented in sections 3 and 4 and the corresponding Figures 2 to 10. Coordinates: $X_{\text{GSE}} - Y_{\text{GSE}}$ (horizontal-vertical) plane, scaled in Earth radii ($R_E$).

are outflow boundaries. The inner boundary has perfect conductor conditions for the fields and maintains a static Maxwellian ion distribution. We present in this study the overview and classification of ion distributions obtained in the foreshock after a simulation time of 1000 s ($\approx 76$ solar wind ion gyroperiods). This simulation took a few million core hours to compute with an average of 24,000 cores on Hermit at the High Performance Computing Center Stuttgart (HLRS), Stuttgart, Germany.

Figure 1 shows the foreshock region of the simulation. The color code gives the ion number density. The black line delimits the ion foreshock; upstream of it the ion distribution consists only of the solar wind Maxwellian distribution. In the region of the bow shock at values of $Y_{\text{GSE}}$ larger than $10 R_E$, $\theta_{\text{BS}}$ is so high that no ions can be reflected far off the bow shock. The region where $Y_{\text{GSE}} < 10 R_E$ is pervaded with backstreaming ions reflected by the bow shock, causing ion beam instabilities which generate compressive ULF waves, as showcased by the color-coded number density. These waves have periods of around 30 s and wavelengths of some $2 R_E$ (not shown). The foreshock waves are advected by the solar wind into the bow shock, making the shock surface much less smooth and steady than in the quasi-perpendicular region. The region with ULF waves ends at the ULF foreshock boundary in the upstream direction, where then begins a region of the foreshock where backstreaming ions flow but no waves have been generated yet due to the finite growth rate of the instabilities. That region ends at the ion foreshock edge marked by the black line. Downstream of the bow shock, marked by a steep increase in density, the magnetosheath is well developed in the simulation, but it is not covered in this work. The white crosses and letters refer to the locations of the ion distributions presented in sections 3 and 4.

2.2. THEMIS Measurements and Methods

We compare the simulation results with observations from the triaxial fluxgate magnetometer (FGM) [Auster et al., 2008] and electrostatic analyzer (ESA) [McFadden et al., 2008] on the THEMIS spacecraft. The magnetometer provides vector measurements of the magnetic field with 3 s time resolution. ESA provides three-dimensional ion distribution functions in the 5 eV to 25 keV range as rapidly as once every 3 s. We fit bi-Maxwellians to the various observed FABs to estimate their densities, perpendicular/parallel temperatures, and bulk velocities. This may not capture the full complexity of the ion distributions, but it has been found sufficient to compare with the simulated ion distributions presented in this work.

We use three methods to identify intervals when the THEMIS spacecraft were located within the foreshock. First, we calculate whether or not the observed interplanetary magnetic field orientation would connect an observer at the spacecraft location to the bow shock, the location of which is determined from the model
Figure 2. Simulated field-aligned beam (FAB) distribution, location A in Figure 1. (left) Three-dimensional isosurface plot and (right) two-dimensional projected isocontour plot. Axes scaled in 1000 km/s, Z axis parallel to the local magnetic field orientation.

of Jetáb et al. [2005] for the prevailing solar wind conditions. Second, we interpret intense high-frequency magnetic field and density perturbations as evidence for fast mode waves in the ion foreshock [Paschmann et al., 1979; Hoppe and Russell, 1983]. Finally, we use observations of suprathermal diffuse and backstreaming ions to identify the ion foreshock.

3. Gyrotropic Ion Distributions in the Simulation

In this and the next section we present the ion distribution function types seen in the hybrid-Vlasov simulation and illustrated in Figures 2 to 10. In all of these figures, on the left the phase space density isosurface at a value of $1.0 \cdot 10^{-15}$ m$^{-6}$ s$^{-3}$ is plotted in a coordinate system where the magnetic field is parallel to the vertical/Z axis of the velocity space. On the right, the distribution has been projected onto a Cartesian plane containing the parallel/Z direction and isocontour lines ranging from $1.0 \cdot 10^{-15}$ to $1.0 \cdot 10^{-9}$ m$^{-6}$ s$^{-3}$ are shown. The locations from which the example distributions were taken are shown in Figure 1 (labels A to I). We first focus here on the various types of near-gyrotropic ion distributions reproduced in the global magnetospheric simulation.

3.1. FAB/Ring Beam Distributions

Field-aligned beam distributions are characterized by a pitch angle distribution well collimated along the magnetic field direction. In the simulation they are present in a narrow region between the ULF foreshock boundary and the ion foreshock edge, where the instability due to the backstreaming population has not yet grown to generate ULF waves disrupting the gyrotropic distributions. An example of such a distribution function from location A in Figure 1 is shown in Figure 2.

At the upstream edge of this region, that is at the very outer edge of the ion foreshock, the distributions lack zero pitch angle ions and have a ring structure centered on the magnetic field direction, as shown in Figure 3 with the distribution from location B. Given their location at the edge of the ion foreshock, these ring...
beams originate in the quasi-perpendicular region of the bow shock. As discussed by Möbius et al. [2001] based on observations with the Cluster spacecraft, strong ring beam distributions of specularly reflected ions are observed in the quasi-perpendicular region. While these ions usually propagate downstream, a small fraction can have sufficient energy to escape upstream and form the ring beams shown here.

3.2. Diffuse Distributions

Diffuse distributions can be recognized by their pitch angle distribution and their temperature. The pitch angles cover more than $90^\circ$, and sometimes the distribution even surrounds the solar wind core completely. In addition to this, diffuse distributions are significantly wider and the energy falls off slower than for the core distribution. This type of distribution exists in the simulation in regions deep in the foreshock (i.e., far from the foreshock edge) within a few $R_E$ of the bow shock. Figure 4 shows a diffuse distribution from location C.

3.3. Intermediate Distributions

Between the two extremes of FAB and diffuse distributions one can identify a broad category of intermediate ion distributions. Although the transition is smooth, one can separate them from the distribution types presented above. On the one hand, intermediate distributions are distinct from FABs in that they have a curvature concave toward the solar wind core distribution. These distributions have been called crescent- or kidney bean-shaped [e.g., Thomsen et al., 1985; Kis et al., 2007] based on planar cuts or projections used in spacecraft and simulation data representations, but we prefer the term cap-shaped distributions, which obviously matches the three-dimensional distributions we show here (example in Figure 5, location D in Figure 1). On the other hand, intermediate distributions have only backstreaming ions (pitch angle $< 90^\circ$) unlike the diffuse distributions and their energy distribution is narrower, criteria which allow to separate both types.

Figure 4. Simulated diffuse distribution, location C in Figure 1. (left) Three-dimensional isosurface plot and (right) two-dimensional projected isocontour plot. Axes scaled in 1000 km/s, $Z$ axis parallel to the local magnetic field orientation.

Figure 5. Simulated intermediate distribution, location D in Figure 1. (left) Three-dimensional isosurface plot and (right) two-dimensional projected isocontour plot. Axes scaled in 1000 km/s, $Z$ axis parallel to the local magnetic field orientation.
4. Gyrophase Bunched Ion Distributions in the Simulation

All the distribution types presented so far have one common characteristic: they are very close to gyrotropic. Whenever this rotational symmetry with respect to the magnetic field direction is broken, one can term the distribution gyrophase bunched. We illustrate here some types of such distributions seen in the simulation.

4.1. Partial Ring Beam Distributions

At the upstream edge of the ring beam region is a narrow band with increasingly partial ring beam distributions until no beam remains in the solar wind. This is a purely geometrical effect due to the gyroradius of the ions in the ring beam [e.g., Schwartz et al., 2000]. The transition layer at the edge of the ring beam region is one gyrodiameter wide, and it can only be reached by ions in a fraction of the full $2\pi$ gyrophase range.

The width of the transition in the simulation is $\sim 3000$ km (not shown), consistent with the gyrating speed of the ring beams of $\sim 750$ km/s (see Figure 3) yielding a gyroradius of $\sim 1500$ km. The same geometrical effect is the cause of the simultaneous observation of FAB and gyrophase bunched distributions at the transition between a region of gyrophase bunched distributions and FABs [Meziane et al., 2004, especially Figure 4 therein].

This is a fundamental difference with the disrupted distributions presented and discussed in sections 4.3 and 6.3, respectively, for which the loss of gyrotropy is due to ULF wave interaction. Figure 6 taken from location E in Figure 1 shows a partial ring beam distribution.

4.2. Multiple Cap Distributions

In rare cases one can find multiple cap distributions in some location (Figure 7, location F). Either each cap can have the same symmetry axis or their rotational symmetry axes have different orientations. Such distributions appear in the central and downstream parts of the foreshock, far from the foreshock edge, but they are not restricted to a given region.
4.3. Disrupted Cap Distributions
The interaction of the backstreaming ion population with the ULF waves in the foreshock leads to a loss of gyrotrropy of the intermediate or cap distributions. Such disrupted cap distributions are seen in the simulation only in regions where ULF waves are present. The disruptions have a wide range, from the cap being off-centered with respect to the magnetic field direction to the formation of large gaps in gyrophase. Figures 8 and 9 (locations G and H) show two such stages of disruption. A particularly striking type of distributions seen in the simulation we termed spiral or corkscrew distributions owing to their characteristic spiraling shape around the magnetic field direction (Figure 10, location I). A discussion of the disruption of gyrotrropic distributions by the ULF waves is presented in section 6.3.

5. THEMIS Observations of Foreshock Ion Distributions
This section presents examples of ion distributions observed with THEMIS within the foreshock.

5.1. Distribution Functions Under Similar Upstream Conditions
To facilitate comparisons with the simulation results, we selected temporally separated events in which the observed solar wind parameters resembled those used in the simulations. In particular, we selected distributions observed when the IMF had a Parker spiral orientation, i.e., lay near the ecliptic plane and made an angle of \( \sim 45^\circ \) with the \( X\text{GSE} \) axis. The projections of THEMIS locations onto the \( X\text{GSE}-Y\text{GSE} \) plane are shown by colored symbols in Figure 11. Solid lines define estimated ion foreshock edge locations for each distribution; corresponding IMF orientations are plotted in dashed lines.

Table 1 lists the solar wind parameters for five THEMIS B and C events satisfying this condition. Figure 12 shows the corresponding ion distributions. Figure 12 (top panels) show 2-D contours of constant phase space density, while Figure 12 (bottom panels) show cuts through the distribution functions along velocities parallel and perpendicular to the ambient magnetic field direction.
Figure 10. Simulated spiral distribution, location I in Figure 1. (left) Three-dimensional isosurface plot and (right) two-dimensional projected isocontour plot. Axes scaled in 1000 km/s, Z axis parallel to the local magnetic field orientation.

The high-density core of the distribution function situated near the origin represents the incident solar wind population. Particle populations at higher velocities represent various categories of reflected ion populations. The FABs presented in Figures 12a–12d are well separated from the solar wind populations. Each of these FABs exhibits a clear temperature anisotropy. The distribution seen in Figure 12e differs from the others in that it was taken near the terminator deep within the ion foreshock.

A comparison of the predicted and observed ion velocity distributions shows many similar features. The distributions presented in Figure 12 can be directly compared to the simulated FAB, ring beam, and diffuse distributions (Figures 2 to 4). The distributions 12b and 12d are intermediate, while the narrower 12a and 12c are FABs. The distribution 12e is more diffuse. We show the beam parallel drift velocity normalized to the solar wind speed versus the distance of the spacecraft from the ion foreshock edge in Figure 13.

5.2. Case Study: Radial Evolution of Distribution Functions

We also present a case study of an event on 10 July 2007 when all THEMIS spacecraft were in the ion foreshock and moving earthward. The projections of the spacecraft locations into the $X_{GSE}$-$Y_{GSE}$ plane are shown together with an estimated ion foreshock edge and modeled bow shock locations in Figure 14. The IMF direction is shown by dotted lines. Distribution functions taken in those locations from simultaneous measurements are presented in Figure 15.

By fitting bi-Maxwellians to the observed FABs, we estimate their parameters. Since all five spacecraft are at nearly the same distance from the ion foreshock edge, we present the dependence of beam density and
Table 1. Parameters of the THEMIS Events Selected to Correspond to the Simulated Solar Wind Conditions

| S/c | Date and Time   | IMF Cone | Angle | B\(_{\text{XYZ}}\) | \(|B|\) | \(V_X\) | \(N\) | Distance to Foreshock Edge (R\(_E\)) |
|-----|----------------|----------|-------|-----------------|-------|--------|------|-------------------------------------|
| 1   | 6 September 2008 0148 | THB | 43°  | [4.9; –3.5; 2.8] | 7     | –483   | 2    | 11.9                                |
| 2   | 6 September 2008 0200 | THC | 46°  | [4; –4.1; –0.04] | 6     | –454   | 3    | 9                                   |
| 3   | 6 September 2008 0201 | THB | 45°  | [4.1; –4.1; 0.4] | 6     | –472   | 3    | 11.9                                |
| 4   | 8 September 2008 1429 | THC | 41°  | [2.9; –2.5; –0.4] | 4     | –512   | 2    | 7.5                                 |
| 5   | 31 October 2008 1903 | THC | 47°  | [1.9; –1.3; 1.5] | 3     | –544   | 2    | 19.5                                |

\( THB, \) THEMIS B; \( THC, \) THEMIS C.

Figure 12. (a–e) THEMIS ion velocity distributions for the locations in Figure 11. (top) The contour plots show contours of constant phase space density (red is high). Colored arrows show the Sun direction (magenta) and solar wind flow (black). \( B_0 \) is the unit vector in the direction of the magnetic field. (bottom) Cuts of distribution functions along coordinate axes with one-count level (green).
parallel drift velocity at each spacecraft normalized to the solar wind density and speed, respectively, at the radial distance of the spacecraft in Figure 16.

We observe a decrease of the beam density associated with an increase in the beam parallel drift velocity. Analysis of the distributions in Figure 15 reveals that this increase in the beam velocity is caused by the loss of slower ions at larger distances from the bow shock, leading to a relatively greater portion of fast ions and thus to an increase in the mean drift velocity.

6. Discussion

6.1. Beam Speed Versus Foreshock Edge Distance

As pointed out by Pokhotelov et al. [2013] about this simulation, the foreshock ion distributions show the property that the backstreaming velocity relative to the incoming solar wind velocity is highest at the upstream foreshock edge and decreases gradually the deeper (i.e., farther away from the foreshock edge) one goes into the foreshock, at the same time as the distributions seen change from ring beams/FABs to more diffuse distributions. Bonifazi and Moreno [1981] reported average ratios of beam to solar wind speed of 2.0, 1.75, and 1.18 for FABs, intermediate, and diffuse distributions, respectively, and more recently Meziane et al. [2004], Eastwood et al. [2005b], and Kis et al. [2007] reported similar values from Cluster observations.
Figure 15. (a–e) THEMIS ion velocity distributions for the locations of Figure 14. Same format as in Figure 12.

In Figure 17b we show the ratio of the backstreaming population’s bulk speed to the solar wind speed in the simulation’s foreshock region, along with the foreshock edge’s and the presented distributions’ locations. The solar wind core population is considered to be all velocity space cells within a sphere encompassing the population in velocity space, while the beam population consists of all velocity space cells outside of that sphere.
Figure 16. Radial dependence of beam parameters for the distributions of Figure 15: (a) beam density and (b) beam parallel drift velocity normalized to solar wind speed. THEMIS A (THA): black square; THB: red triangle; THC: green asterisk; THD: light blue plus sign; THE: blue diamond.

The bulk speeds are then calculated as moments taken over each population separately. The solar wind and beam speeds are taken in the simulation frame, which can be considered similar to a spacecraft frame as a spacecraft velocity would be much smaller than the solar wind and beam speeds seen in the foreshock. In Figure 13 we show the same speed ratio as a function of the distance to the foreshock edge obtained from the THEMIS distributions at the locations of Figure 11.

It is obvious that indeed the fastest beam speeds occur close to the foreshock edge and that deeper in the foreshock the backstreaming ions are slower. This is related to the reflection efficiency of the bow shock, which depends on $\theta_{\text{sh}}$: at the foreshock edge $\theta_{\text{sh}}$ is large while still enabling reflection, resulting in high reflected beam speeds whereas deeper in the foreshock $\theta_{\text{sh}}$ and the reflected beam speed both decrease. The relative beam speed as a function of the foreshock edge distance obtained in the simulation corresponds well with the observational results reported earlier and those presented here.

6.2. Beam Density and Speed Versus Radial Distance

Another point of comparison we present here is the dependence of the ion beam density and speed on the radial distance from the bow shock. The observational case study in section 5.2 shows that the beam density relative to the solar wind core population density is decreasing by a factor of $\sim 10$ while the distance from the bow shock increases by a few $R_E$. This is consistent with previous observations. For example, Bonifazi et al. [1983] reported in a statistical study that the density of the diffuse ion population decreases by a factor of $\sim 3$.
within $3 R_e$, and more recently Kis et al. [2004] reported energy-dependent e-folding distances of the diffuse ion density ranging from 0.5 $R_e$ at 11 keV to $\sim 3 R_e$ at $\sim 30$ keV, which is comparable to the scales observed in Figure 16a. In Figure 17a we show the simulated ratio of the beam density to the core density. The beam and core populations have been separated in the same way as explained above for the speed ratio. The simulated beam densities decrease similarly by a factor of $\sim 10$ within a few $R_e$ from the bow shock while they do not decrease as strongly any more further out radially in the foreshock.

The observed beam speed as a function of radial distance is shown in Figure 16b. A similar increase of beam speed with radial distance is visible in the simulation (Figure 17b) in the region of FAB distributions close to the foreshock edge (within $\sim 10 R_e$), but it disappears deeper in the foreshock. Both the beam density decrease and beam velocity increase with increasing distance from the source have been observed in THEMIS data and reported by Wilson et al. [2013] for FABs upstream of short long-amplitude magnetic structures (SLAMS) in the foreshock. Although these FABs are from a different source, the processes leading to their formation are very similar.

Our results demonstrate the good correspondence between the simulation and observations of the dependence of the backstreaming beam density and to a lesser extent of the beam speed on the radial distance from the bow shock.

6.3. Evolution of Backstreaming Populations in the Foreshock

The current work presents the types of ion distribution functions in the terrestrial foreshock as simulated with the hybrid-Vlasov code Vlasiator and compares them to spacecraft observations. Although we do not aim to explain the spatial and temporal evolution of the distributions nor to identify all formation mechanisms, some leads for further studies can be identified.

Gurgiolo et al. [1983] proposed a mechanism by which a part of the incoming solar wind distribution is specularly reflected. It is gyrating and thus nongyrotropic initially, it propagates and develops into a spiral (or helical) and multipeaked structure in phase space [Gurgiolo et al., 1983, Figures 2 and 3]. The similarity with spiral distributions seen in our magnetospheric simulation led us to investigate whether this mechanism is a possible source. We tried to reproduce Gurgiolo et al.’s mechanism in one-dimensional local simulations, where a solar wind and a lower-density beam distribution (both isotropic and Maxwellian) counterstream. It was not possible to obtain distributions similar to that proposed model, neither with field-aligned nor with gyrating beams, even with perpendicular speeds of the same order as the core-to-beam drift speed.

We performed an additional local simulation to investigate the mechanisms leading from gyrotrropic intermediate distributions to nongyrotropic ones. A Maxwellian core distribution and a Maxwellian backstreaming beam with 1% of the core density, both with densities and temperatures similar to the ones obtained in the global simulation, are propagated in a one-dimensional box with periodic boundary conditions.

The run is initialized with only the main longitudinal magnetic field component. The wave-generating beam instability grows self-consistently. In order for the instability to grow in a shorter run time, random noise with a maximum amplitude of 0.5% of the longitudinal magnetic field strength is included in the transversal directions at the initial state. A lower level of seed noise does not affect the result, only resulting in a longer growth time of the instability to the same amplitude.

Figure 18 shows the result of this local simulation. Figure 18 (top left) shows the initial distribution function set throughout the simulation box. Figure 18 (top right) shows the distribution function, and Figure 18 (bottom) shows the transversal magnetic field components after $\sim 10$ gyroperiods.

The initially isotropic beam distribution interacts with the growing wave. It loses its gyrotropy and becomes spiral shaped. We also performed simulations with anisotropic bi-Maxwellian backstreaming beams mimicking foreshock cap distributions, but it was found that the results are robust and do not deviate significantly from the simpler case of an isotropic Maxwellian beam.

This local simulation leads to distributions similar to the gyrophase bunched ones seen in the ion foreshock of the global magnetospheric simulation. Additionally, the local instability simulations presented by von Alfthan et al. [2014] show that the source of the compressive ULF waves seen in the foreshock of the simulation is the ion/ion right-hand resonant beam instability. We therefore come to a conclusion similar to the one drawn by Meziane et al. [2001] from the observation of gyrating ion distributions and ULF waves in the terrestrial foreshock. The more likely scenario is
Figure 18. Local simulation of the loss of gyrotropy of a backstreaming beam through wave interaction. (top) Three-dimensional ion distribution function isosurface plots, axes scaled in 1000 km/s, Z axis parallel to the local magnetic field orientation. (left) Initial state and (right) state after \( \sim 10 \) gyroperiods. (bottom) Transversal magnetic field components after \( \sim 10 \) gyroperiods.

1. The ULF waves are generated by the backstreaming ring beam/FAB distributions via the ion/ion right-hand resonant beam instability with the solar wind core distribution;
2. The beam is disrupted by wave-particle interaction leading to the formation of increasingly strongly gyrophase bunched distributions.

It will be the subject of future work to identify and better characterize the mechanisms at play in the foreshock, using test or tracer particles within the hybrid-Vlasov simulation to track the origin of various distribution function features.

6.4. Performance of the Hybrid-Vlasov Model

It should be pointed out that the global magnetospheric simulation gives an idealized picture of the foreshock in several respects. The ideal MHD Ohm’s law closure of the hybrid-Vlasov model limits wave-particle interactions so that less wave modes are present. For example, the low-frequency parallel-propagating MHD Alfvén wave mode is not split into the electron and ion whistler modes [Kempf et al., 2013]. It is, however, not a concern because the spatial resolution of the global magnetospheric simulation would not allow to resolve the whistler modes.

The limited resolution in space and in velocity, dictated by the computational size of the problem, also limits the phenomena that can be resolved. For instance, the full steepening of foreshock ULF waves into shocklets or short large-amplitude magnetic structures (SLAMS) at quasi-parallel shocks seen in PIC simulations [e.g., Scholer et al., 2003; Tsubouchi and Lembege, 2004] is not reproduced in our simulation. The foreshock waves do grow in amplitude toward the bow shock but not strongly enough to form shocklets or SLAMS. This is the result of a lack of spatial resolution to resolve steep profiles down to ion scales and of too strong numerical diffusion arising from the low resolution.

As explained in section 2, the distribution function is not saved below a set phase space density threshold. This threshold is, however, more than 6 orders of magnitude below the solar wind core population’s peak phase space density, an amplitude comparable to the range covered by the THEMIS observations presented (Figures 12 and 15 herein). As shown by von Alfthan et al. [2014], the resulting departure from exact mass conservation is kept at a reasonably low level. The present work shows that this does not impact adversely the features of the ion foreshock velocity distributions either.
In this run the smooth upstream conditions differ from the observed turbulent and dynamic solar wind conditions: all fluctuations seen in the foreshock in the simulation (in Figure 1 as well as in Pokhotelov et al. [2013]) originate in the ion/ion beam instabilities due to the ions that are reflected at the bow shock and stream back along the magnetic field. The absence of solar wind turbulence here facilitates the propagation of reflected ion beams with little perturbation until their disruption by the instability-generated ULF waves. Thus, this simulation does not properly model the inherent statistical noise and turbulence level of real solar wind plasma. In that sense, the velocity distribution functions are overly idealized in comparison to spacecraft data. The good correspondence of the simulation and observations, however, indicates that kinetic physics at large spatial and time scales is sufficient to interpret the features of the ion distributions and their evolution in the foreshock depicted in this work.

To the extent possible, comparisons with ion distributions in hybrid- or full-PIC simulations of bow shocks show agreement with the distribution types and locations in this work. For example, distribution b) in Figure 2 of Savoini et al. [2013] has a lower beam speed farther away from the foreshock edge than distribution a). They can, respectively, be compared to the intermediate and FAB distributions (Figures 5 and 2 herein). The latter even exhibits a ring structure similar to Figure 3 herein. The ion distributions in Figure 3 of Blanco-Cano et al. [2006a] and distribution A in Figure 2 of Blanco-Cano et al. [2006b] also compare well with the same FAB and intermediate distributions from Vlasiator. Distribution B in the latter could be compared to the disrupted distributions (Figures 8 and 9 herein), distributions C and D are more difficult to associate. In Figure 2 of Wang et al. [2009] the distributions are rather perturbed, so that they can be associated with the strongly disrupted distribution (Figure 9 herein). Distribution B also resembles a diffuse distribution (Figure 4 herein) since it shows particles streaming ahead of the core population. Although the resolution is not fully sufficient, distribution A shows features reminiscent of the spiral or multiple cap distributions (Figures 7 and 10 herein).

The novelty of the hybrid-Vlasov approach lies in the high-resolution ion velocity distributions obtained at every spatial cell and at any instant in a snapshot fashion without the statistical noise inherent to hybrid- or full-PIC simulations. This allows detailed studies of processes affecting the ion distributions as there is no need for time integrations or integration volumes larger than one grid cell to obtain detailed ion distributions. A downside of this method is that the origin of a particular part of a distribution function cannot be traced directly unless analyzing carefully distributions at all places and times, which is impractical because of limited disk storage capacities. The development plans for Vlasiator include using tracer particles in future simulations to circumvent this inconvenience.

The major drawback of the hybrid-Vlasov model in comparison to hybrid-PIC is the computational size. Recent large-scale simulations such as those presented by Karimabadi et al. [2014] are run with a few hundred particles per spatial cell (200 in that study), typically on a few hundred cores. On the other hand, the velocity space distribution in each spatial cell in the magnetospheric simulation in this work consisted of an average of approximately 200,000 velocity space cells and it used up to 24,000 cores. With these significant differences in mind, it is clear that hybrid-PIC and hybrid-Vlasov simulations have respective strengths and weaknesses, making them complementary tools for plasma physics research.

7. Conclusions

By demonstrating that the code reproduces well the types of ion distribution functions observed in the terrestrial ion foreshock, this work gives an observational validation of Vlasiator. Field-aligned beams and ring beams are seen at the upstream edge of the foreshock, while deeper in the foreshock (i.e., farther away from the foreshock edge), intermediate distributions dominate, which get eventually disrupted by the ULF waves present to become gyrophase bunched. We find that our simulation supports the scenario proposed by Meziane et al. [2001], in which shock-reflected field-aligned beams generate ULF waves by a resonant ion/ion beam instability, which in turn disrupt the beam or intermediate distributions and cause them to become gyrophase bunched. Diffuse distributions are present close to the bow shock in the deep foreshock regions.

The decrease of the backstreaming beam speed with increasing distance to the foreshock edge, which is related to the reflection efficiency of the bow shock decreasing with decreasing $\theta_{bac}$, has been reported observationally before [Bonifazi and Moreno, 1981; Meziane et al., 2004; Eastwood et al., 2005b]. It is well reproduced by Vlasiator and readily seen in the spacecraft data presented here. The decrease of the beam density and
increase of the beam speed with increasing radial distance from the bow shock also correspond well between the simulation and THEMIS observations.

The computational cost of hybrid-Vlasov simulations implies significant limitations, in terms of dimensionality as well as the spatial and velocity resolutions. By nature of the model, the history of the distribution functions cannot be followed as easily as with particles either. Nevertheless, the simulated ion distributions shown in this work are in good agreement with those produced by other codes and especially PIC algorithms [e.g., Blanco-Cano et al., 2006a; Savoini et al., 2013], which provides a verification of Vlasiator. Despite the large computing requirements, it is clear that a major strength of the hybrid-Vlasov approach lies in the outstanding quality of the ion distribution functions obtained.

Acknowledgments

Y.K. acknowledges fruitful discussions with H.E.J. Koskinen and B. Lembège. We acknowledge that the results of this research have been achieved using the PRACE (Partnership for Advanced Computing in Europe) Tier-1 Research Infrastructure resource Hermes based in Germany at the High Performance Computing Center Stuttgart (HLRS) and the PRACE Tier-1 Research Infrastructure resource Abel owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR) and operated by the Department for Research Computing at USIT, the University of Oslo IT department. The Quantifying Energy circulation in Space plasmas (QuESpace) project, in which Vlasiator was initially developed, has received funding from the European Research Council under the European Community’s Seventh Framework Programme (FP-7/2007-2013/ERC). The work of Y.K., D.P., S.A., O.H., and M.P. was supported by Science and Technology of Finland. The work of D.P. has been received funding from the European Research Council under the European Community’s Seventh Framework Programme (FP-7/2007-2013/ERC) and especially PIC algorithms [e.g., Blanco-Cano et al., 2006a; Savoini et al., 2013], which provides a verification of Vlasiator. Despite the large computing requirements, it is clear that a major strength of the hybrid-Vlasov approach lies in the outstanding quality of the ion distribution functions obtained.

References


Michael Balikhin thanks the reviewers for their assistance in evaluating this paper.
Paper IV

Evidence for transient, local ion foreshocks caused by dayside magnetopause reconnection

© 2016 Authors.

Evidence for transient, local ion foreshocks caused by dayside magnetopause reconnection

Yann Pfau-Kempf1,2, Heli Hietala3, Steve E. Milan4, Liisa Juusola1, Sanni Hoilijoki1,2, Urs Ganse5,1, Sebastian von Alfthan6, and Minna Palmroth1

1Earth Observation, Finnish Meteorological Institute, Helsinki, Finland
2Department of Physics, University of Helsinki, Helsinki, Finland
3Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, USA
4Department of Physics and Astronomy, University of Leicester, Leicester, UK
5Department of Physics and Astronomy, University of Turku, Turku, Finland
6CSC, IT Center for Science, Espoo, Finland

Correspondence to: Yann Pfau-Kempf (yann.kempf@helsinki.fi)

Received: 20 May 2016 – Revised: 31 August 2016 – Accepted: 17 October 2016 – Published: 4 November 2016

Abstract. We present a scenario resulting in time-dependent behaviour of the bow shock and transient, local ion reflection under unchanging solar wind conditions. Dayside magnetopause reconnection produces flux transfer events driving fast-mode wave fronts in the magnetosheath. These fronts push out the bow shock surface due to their increased downstream pressure. The resulting bow shock deformations lead to a configuration favourable to localized ion reflection and thus the formation of transient, travelling foreshock-like field-aligned ion beams. This is identified in two-dimensional global magnetospheric hybrid-Vlasov simulations of the Earth’s magnetosphere performed using the Vlasiator model (http://vlasiator.fmi.fi). We also present observational data showing the occurrence of dayside reconnection and flux transfer events at the same time as Geotail observations of transient foreshock-like field-aligned ion beams. The spacecraft is located well upstream of the foreshock edge and the bow shock, during a steady southward interplanetary electric field and in the absence of any solar wind or interplanetary magnetic field perturbations. This indicates the formation of such localized ion foreshocks.

Keywords. Interplanetary physics (planetary bow shocks) – magnetospheric physics (magnetosheath; solar wind–magnetosphere interactions)

1 Introduction

The super-Alfvénic solar wind impinging upon the geomagnetic field is slowed down and diverted around the Earth by the bow shock which forms upstream of our planet. Most of the plasma is abruptly compressed and heated by the shock while being transported downstream into the magnetosheath. There, it flows along the magnetopause surface, which delimits the magnetosphere, that is, the magnetic cavity in which the Earth is situated.

Fluid theories such as ideal magnetohydrodynamics imply that no wave or matter can travel upstream from a shock. However, it is well-known observationally and explained by kinetic plasma theory that, given a high enough Mach number and a small enough angle $\theta_{B-n}$ between the upstream magnetic field ($B$) and the shock normal direction ($n$) ($\theta_{B-n} \lesssim 40–60^\circ$, e.g. Greenstadt et al., 1980; Schwartz et al., 1983), a fraction of the incoming ions is reflected by the shock surface and streams back along the magnetic field direction. The region where such a backstreaming ion population exists is called the ion foreshock. It is the stage for a variety of plasma beam instabilities generating waves and has been studied observationally as well as in simulations for several decades (e.g. Bavassano-Cattaneo et al., 1983; Sanderson et al., 1983; Thomsen et al., 1983; Fuselier et al., 1987; Le and Russell, 1992; Eastwood et al., 2005b; Burgess et al., 2012; Wilson et al., 2013). The value of $\theta_{B-n}$ allowing the reflection of particles is dependent on several factors,
among which most notably is the assumed reflection mechanism, as the results of Greenstadt et al. (1980), for example, show.

The bow shock under steady solar wind conditions is generally assumed to be a simple surface such as a paraboloid with a dawn–dusk asymmetry primarily due to the Earth’s orbital motion. This follows from fluid dynamical considerations, numerical simulations as well as statistical studies of spacecraft observations. The inherently local nature of spacecraft measurements compared to the spatial scale of the bow shock, even in the case of spacecraft constellations such as Cluster (Escoubet et al., 1997), Time History of Events and Macroscale Interactions during Substorms (THEMIS; Angelopoulos, 2008) or the Magnetospheric Multi-Scale (MMS; Burch et al., 2016a) missions, does not allow us to determine the global shape of the bow shock surface at a given instant in time. Statistical studies are the method of choice (e.g. Merka et al., 2005; Meziane et al., 2014). Thus it is also usually assumed that under steady conditions, the ion foreshock is located in the solar wind volume magnetically connected to the bow shock surface region where \( \theta_{B-n} \lesssim 50^\circ \). Depending on the orientation of the interplanetary magnetic field (IMF), this can be one or two compact regions in space.

Another tenet of ideal magnetohydrodynamics is the conservation of magnetic field line topology, which is a good assumption on large scales or in collisional plasma but breaks down on smaller scales when strong magnetic field gradients are present. Through the kinetic process of magnetic reconnection, antiparallel magnetic field lines rearrange their topology while strongly accelerating inflowing plasma out of the reconnection region (see, e.g., reviews by Zweibel and Yamada, 2009, and Treumann and Baumjohann, 2013, and references therein). A prime example of magnetic reconnection in near-Earth space occurs at the magnetopause in the subsolar region, when inflowing southward IMF reconnects with the northward-oriented geomagnetic field lines (e.g. Phan et al., 2000; Paschmann, 2008; Dunlop et al., 2011). This phenomenon drives global magnetospheric dynamics as first proposed by Dungey (1961), and therefore it is key to space weather studies (e.g. Cassak, 2016; Burch et al., 2016b).

The topological reconfiguration of magnetic field lines at the magnetopause can lead to the formation of well-delimited structures called flux transfer events (FTEs). The classic picture of an FTE is that of a magnetic flux tube connected both to the magnetosheath and the magnetosphere, but its topology can be more complex. FTEs were first observed by Russell and Elphic (1978, 1979) and Haerendel et al. (1978) (who termed the process magnetic flux erosion) and subsequently much studied in space and ground observations as well as simulations (e.g. Kawano and Russell, 1997; Wild et al., 2001, 2003; McWilliams et al., 2004; Fear et al., 2009; Eastwood et al., 2016; Hasegawa et al., 2016; Milan et al., 2016). FTEs travel downstream along the magnetopause with the magnetosheath plasma and are recognized by their signature in magnetic field measurements, such as the bipolar deflection of the component normal to their axis in the case of a flux rope or magnetic island (e.g. Omidi and Sibeck, 2007; Dorelli and Bhattacharjee, 2009; Sibeck and Omidi, 2012; Eastwood et al., 2012). Their signatures also include poleward moving auroral forms (PMAsFs) and their equivalent in radar observations named poleward moving radar auroral forms (PMRAFs), which result from poleward ionospheric flows. Oscillations can also be observed by ground magnetometers (e.g. Øieroset et al., 1996; Milan et al., 2000; Pang et al., 2009). Although their role is crucial in the solar wind–magnetosphere interaction, allowing plasma exchange along the reconnected magnetic field lines, FTEs have not so far been thought to be the direct cause of significant upstream effects.

In this work, we propose a scenario by which dayside magnetopause reconnection generates FTEs, which in turn cause steepening fast magnetosonic bow and stern waves to propagate throughout the magnetosheath. The increased pressure behind the steepened wave fronts causes the bow shock to bulge outward in an area travelling along the bow shock surface. The geometry at the edge of such a bulge can lead \( \theta_{B-n} \) to become smaller than \( \sim 50^\circ \) in a well-delimited region, detached from the “regular” foreshock, upstream of which a beam of reflected ions generates a local, transient and travelling foreshock. This chain of processes has first been identified in a two-dimensional polar-plane hybrid-Vlasov simulation of steady southward IMF interacting with an Earth-like dipolar magnetic field. The simulation was performed using the Vlasiator model (http://vlasiator.fmi.fi). We also present observational data supporting the scenario. Geotail spacecraft observations show the existence of short foreshock-like ion beams during steady southward IMF, in a region well-detached from both the bow shock and the regular ion foreshock region and in the absence of any IMF fluctuations, thus excluding a momentary transit of the spacecraft through the regular foreshock due to a change in IMF orientation. Simultaneously, the signatures of FTEs moving poleward are found in ground magnetometer and SuperDARN (Super Dual Auroral Radar Network) radar data.

Section 2 describes the simulation and the observational methods. The scenario of magnetopause–bow-shock–foreshock interaction is detailed in Sect. 3, while the ground and spacecraft observations are presented in Sect. 4. The results are then discussed in Sect. 5 before the conclusions are given in Sect. 6.

2 Methods

2.1 Hybrid-Vlasov simulation

The hybrid-Vlasov model Vlasiator has been developed with the aim of producing global magnetospheric simulations
of the Earth’s magnetosphere including kinetic physics beyond magnetohydrodynamics while avoiding the limitations due to the statistical sampling inherent to particle-in-cell approaches (von Alfthan et al., 2014). Vlasiator has been used to study amongst other things the magnetosheath and the foreshock in equatorial plane simulations of the terrestrial magnetosphere (Pokhotelov et al., 2013; Kempf et al., 2015; Palmroth et al., 2015; Hoilijoki et al., 2016). It solves Vlasov’s equation to propagate the ion (proton) velocity distribution function in up to three spatial and three velocity dimensions. The equation system is closed via Ampère’s and Faraday’s laws as well as a generalized Ohm’s law including the Hall term (see von Alfthan et al., 2014, and Palmroth et al., 2015, for more details).

The simulation used in this study is two-dimensional in the polar \( x-z \) plane and three-dimensional in velocity space. It covers both the dayside and the nightside magnetosphere. The spatial coordinates are similar to the Geocentric Solar Magnetospheric (GSM) coordinate system with the \( x \) axis pointing from the Earth towards the Sun and the \( z \) axis orthogonal to the \( x \) axis and parallel to the geomagnetic dipole field axis (no dipole tilt). We use a two-dimensional line dipole centred at the origin and scaled to match the geomagnetic dipole strength in the same way as is done by Daldorff et al. (2014). The steady solar wind has a proton density of \( 1 \times 10^6 \text{ m}^{-3} \), an inflow temperature of 0.5 MK and a velocity of \(-750 \text{ km s}^{-1}\) purely along the \( x \) axis. The constant and purely southward IMF has an intensity of 5 nT. The upstream boundary maintains a constant field and a Maxwellian velocity distribution; the three other outer boundaries have copy conditions ensuring proper outflow. The inner boundary, which is set at a distance of 30 000 km (\( \sim 5 \text{ Earth radii, } R_E \)) around the origin, enforces a static Maxwellian velocity distribution and perfect conductor field boundary conditions. The out-of-plane direction is treated periodically. The boundaries are located at 47 \( R_E \) from the origin in each direction. Since this study concentrates on dayside phenomena, the nightside is not shown in this work. The physical resolution is 300 km or 0.047 \( R_E \) or 1.3 solar wind ion inertial lengths and the velocity space extends from \(-4000 \text{ to } +4000 \text{ km s}^{-1}\) in all three dimensions with a resolution of 30 km s\(^{-1}\) or 0.33 solar wind ion thermal speeds. The phase space density sparsity threshold is \( 10^{-15} \text{ m}^{-6} \text{s}^{-3} \) (see von Alfthan et al., 2014, and Kempf et al., 2015, for details on the sparse phase space strategy used in Vlasiator).

The simulation has been run for over 1850 s or 140 solar wind proton gyroperiods, and it reaches a steady state on the dayside after less than 900 s or 70 gyroperiods. The bow shock and the magnetopause form as expected and there is a foreshock at high latitudes both in the Northern and Southern hemispheres. The magnetosheath is pervaded by anisotropy-driven waves, most notably mirror-mode waves as has been demonstrated by Hoilijoki et al. (2016). Figure 1 shows an overview of the simulation setup after 1150 s of simulated time. The magnetopause–magnetosheath–bow-shock structure is clear, and a large magnetic island can be recognized due to its high density at the position \((+6, -7) \text{ } R_E\); another one is in the southern cusp region and a series of smaller magnetic islands is visible at the dayside magnetopause northward of the equator.

![Figure 1. Colour code: plasma number density (protons m\(^{-3}\)) after 1150 s of simulation time. Contour lines: magnetic field lines. A large magnetic island is prominent at \((+6, -7) \text{ } R_E\); another one is in the southern cusp region and a series of smaller magnetic islands is visible at the dayside magnetopause northward of the equator.](image)

2.2 Spacecraft and ground measurements

We first use solar wind densities, velocities and the IMF one-minute averaged data from NASA/GSFC’s OMNI data
set accessed through CDAweb (Coordinated Data Analysis Web) to identify suitable intervals of stable solar wind and steady southward IMF conditions.

We also use in situ spacecraft measurements from Geotail in this study. The ion velocity distribution measurements are taken from the Low Energy Particle instrument (LEP; Mukai et al., 1994). During the event presented in Sect. 4.3, LEP was in the EA (energy-per-charge analyser) mode, which is not well-suited to properly measure the cold core solar wind ion population but does not impact the quality of suprathermal ion measurements. Editor-B data are available for that event, meaning that only two-dimensional velocity distributions are available. These projected distributions are produced using data from all three-dimensional channels (LEP instrument team, personal communication, 26 August 2016). They are provided in Geocentric Solar Ecliptic (GSE) coordinates ($x_{GSE}$-axis pointing from the Earth towards the Sun, $y_{GSE}$-axis in the ecliptic plane pointing towards dusk and $z_{GSE}$-axis perpendicular to the ecliptic plane). Magnetic field measurements are from the Magnetic Field Measurement fluxgate magnetometers (MFG; Kokubun et al., 1994).

The following measurements from the Wind spacecraft are used: solar wind data from its Solar Wind Experiment (SWE; Ogilvie et al., 1995), IMF data from its Magnetic Fields Investigation (MFI; Lepping et al., 1995), moments from its 3-D Plasma and Energetic Particle Analyzer (3-DP; Lin et al., 1995) as well as densities retrieved from the electron plasma frequency measured by the radio and plasma wave instrument ( WAVES; Bougeret et al., 1995). The following datasets from the Advanced Composition Explorer (ACE) spacecraft are used: IMF measurements from the Magnetic Fields Experiment (MAG; Smith et al., 1998) and ion moments from the Solar Wind Electron Proton Alpha Monitor (SWEPAM; McComas et al., 1998).

Ground-based ionospheric backscatter data from SuperDARN (Greenwald et al., 1995) as well as ground magnetometer data from the International Monitor for Auroral Geomagnetic Effects (IMAGE, http://space.fmi.fi/image, Tanskane, 2009) are used. Additionally, we use the electrojet activity auroral electrojet (AE) indices provided by the University of Kyoto through the World Data Center for Geomagnetism (Davis and Sugiura, 1966).

3 Magnetopause–bow-shock–foreshock interaction scenario

The scenario proposed in this work has been identified in the simulation presented in Sect. 2.1. We describe the scenario here in Sect. 3 in a narrative fashion and present the corresponding observations in Sect. 4. Limitations are discussed in Sect. 5.

Figure 2. (a) Close-up view of the large magnetic island from Fig. 1 travelling tailward along the magnetopause. Colour code: plasma number density, protons m$^{-3}$. Contour lines: magnetic field lines. Arrows: rotated coordinate system ($N, Y, T$) with $N$ normal to the magnetopause and $T$ parallel to it. (b) Magnetic field evolution at a virtual spacecraft located at the white cross in panel (a), in simulation (GSM) and (c) rotated coordinates. The grey vertical bar indicates the time of panel (a) and Fig. 1. The characteristic bipolar signature of the passing magnetic island is obvious in the $B_Y$ component.

3.1 Magnetopause reconnection

Under steady southward IMF, magnetic reconnection occurs typically along a line at the equator on the magnetopause (e.g. Trattner et al., 2007; Dunlop et al., 2011; Hoilijoki et al., 2014). In the present simulation, the position of the X-line is not stable in time and multiple reconnection sites can coexist at any given time on the magnetopause. Reconnected field lines form magnetic islands in the exhaust regions of reconnection sites, which grow and travel downstream (poleward) along the magnetopause. This continuously ongoing process is prominent in the animation provided in the Supplement. The magnetic islands can be seen as the two-dimensional equivalents of FTLEs, that is, cuts through an out-of-plane flux rope. A more detailed analysis of the propagation of the magnetic islands and the location and intensity of magnetic reconnection is the subject of a separate study.

Figure 2 shows such a magnetic island and time series of the magnetic field components seen at a virtual spacecraft over which the magnetic island flows. The magnetic
Field components are shown both in the simulation coordinates and in a coordinate system \((N, Y, T)\) rotated by 150° in the plane of the simulation so that \(N\) points in the direction normal to the magnetopause and \(T\) points along the magnetopause. The strong bipolar fluctuation in the \(B_N\) component is characteristic of the passage of a magnetic island.

### 3.2 Magnetosheath waves and bow shock perturbations

Figure 3a shows a magnetic island in the southern cusp region. The increased dynamic pressure of the magnetic islands with respect to the surrounding magnetosheath plasma drives bow waves ahead of the islands. These fast magnetosonic waves propagate throughout the magnetosheath and steepen to almost form fast forward shocks. In some cases, strong magnetic islands can also be followed by a fast reverse wave front, but these stern waves are less steep than the bow wave fronts. Both the bow and stern fast mode waves are visible in Fig. 3a. The Profiles of plasma density, velocity and temperature perpendicular to the magnetic field as well as the magnetic field intensity show clearly the steep correlated increase corresponding to the fast forward wave front in Fig. 4a–d. The forward wave fronts are steep enough to reflect ions much in the way a shock can accelerate ions to generate upstream foreshock populations. Figure 3b and c shows the two- and three-dimensional velocity distribution function isocontours at the location of the white cross in Fig. 3a. The accelerated ions are clearly visible ahead of the core population in the \(-V_x\) and \(-V_z\) direction. The structure is also readily visible in the profile of the temperature parallel to the field.

![Figure 3.](image)

**Figure 3.** (a) Example of bow (black dashed) and stern (white dash–dotted) fast wave fronts driven by a magnetic island (density peak at \((2, -9) R_E\); colour code: plasma number density; protons m\(^{-3}\); simulation time 1340 s). The bow wave accelerates particles ahead of it, as can be seen in the (b) two-dimensional projected isocontour and (c) three-dimensional isocontour plots of the ion velocity distribution (phase space density in s\(^3\) m\(^{-6}\); 3-D isocontour at \(1 \times 10^{-15}\) s\(^3\) m\(^{-6}\)) taken at the location of the white cross. The core population with very low drift velocity (blue, pink and grey isocontours, centre and top right part of the 3-D isocontour) is preceded by an accelerated population in the \(-V_x\) and \(-V_z\) direction. The white arrow shows the location of the profiles shown in Fig. 4.

![Figure 4.](image)

**Figure 4.** Cut across the bow wave front along the white arrow in Fig. 3 showing (a) the plasma density and (b) velocity, (c) the temperature parallel and perpendicular to the magnetic field, and (d) the magnetic field intensity. The correlated jump in all these parameters at the abscissa \(2 R_E\) characterizes the steep fast forward wave front. Note the fast wave signature in the perpendicular temperature while the parallel temperature is much more sensitive to the particle beam accelerated ahead of the wave front.
3.3 Local foreshocks

Figure 5a shows the detailed view of a small region of the bow shock south of the subsolar point. When the angle between the shock normal and the upstream magnetic field $\theta_{B-n}$ is lower than $\sim 50^\circ$, incoming particles with sufficient energy can be accelerated back upstream and form a foreshock. Consequently, when the bow shock bulge caused by the fast-wave-mediated pressure increase is pronounced enough, the region of the bulge with $\theta_{B-n} \leq 50^\circ$ is the source of a separate ion beam propagating upstream along the IMF direction. This is of course only the case when the bulge has not yet travelled into the region where the mean $\theta_{B-n}$ is allowing ion reflection anyway. The spatial extent in the direction parallel to the bow shock surface is limited to the corresponding patch of favourable $\theta_{B-n}$, and this localized foreshock travels along with the driving magnetosheath wave front until it merges with the regular foreshock further downstream. Due to its being a travelling and transient ion beam, the expected beam instabilities do not have time to grow to form ultra low-frequency (ULF) waves as in the regular foreshock (see, e.g., von Alfthan et al., 2014; Palmroth et al., 2015). This means that the typical ULF wave signature characteristic of the regular foreshock is absent from this structure.

Figure 5a shows an example of a localized foreshock driven by a magnetosheath wave. The colour code in the figure shows the temperature parallel to the magnetic field, which is sensitive to the presence of an ion beam. The white contour curve is set at a plasma density of $2 \times 10^6 \text{ m}^{-3}$ (twice the solar wind density), which highlights the position and shape of the bow shock. The dashed curve would indicate the approximate location of the bow shock were it not for the increased pressure due to a fast wave front in the magnetosheath. The continuous and dashed segments indicate the normal direction for each of these curves, $\theta_{B-n} = 41^\circ$ and $54^\circ$; thus the angle between both is $13^\circ$. (b) Two-dimensional projected isocontour and (c) three-dimensional isocontour plots of the ion velocity distribution (phase space density in $\text{s}^3 \text{ m}^{-6}$; 3-D isocontour at $1 \times 10^{-15} \text{s}^3 \text{ m}^{-6}$) at the location of the white cross. The field-aligned beam is prominent and directly comparable to Figs. 2 and 6 in the work by Kempf et al. (2015).
after the passage of a fast forward wave front. The continuous and dashed segments indicate the local direction normal to the respective curve. Their $\theta_{B-n}$ is respectively 41° and 54°. The ion beam is generated by solar wind ions reflected at the foot of the bow shock bulge where $\theta_{B-n}$ is favourable. Beyond $-3 R_E$ ($\pm$ GSM), the regular foreshock is also visible upstream of the bow shock as a region of increased parallel temperature. The fact that the density and parallel temperature increases due to the bow shock do not coincide, illustrates that the shock primarily heats plasma in the perpendicular direction. Isotropization of the velocity distribution happens further downstream in the magnetosheath.

The animation provided in the Supplement to this work shows that such local foreshocks occur both north and south of the ecliptic whenever favourable $\theta_{B-n}$ conditions are met at the foot of a bow shock perturbation.

Figure 5b–c shows the ion velocity distribution function corresponding to a field-aligned beam population, which is similar to the distribution expected at the edge of the regular foreshock upstream of the ULF foreshock boundary (Kempf et al., 2015). The density of the beam is of the order of 1% of the inflowing density as in the regular foreshock (not shown). Figure 6 shows an energy–time spectrogram for the tailward and southward sectors of the velocity distribution. The tailward part contains the cold solar wind core population, while the southward sector, in the direction of the field-aligned beam, indicates the typical signature expected when a local foreshock crosses an upstream spacecraft. The timescale of a local foreshock crossing is on the order of 1 min in the simulation (almost 70 s in Fig. 6), but this value can vary depending on the geometry of the event. Other factors affecting the observational signatures are discussed in Sect. 5.

4 Spacecraft and ground observations

In Sect. 3 we present a scenario based on a global hybrid-Vlasov simulation, in which dayside reconnection eventually leads to the formation of local, transient foreshock-like structures upstream of the terrestrial bow shock and outside of the region where the angle between the shock normal and the IMF ($\theta_{B-n}$) would normally be favourable for ion reflection. In this section, we present observational data from an event on 30 August 2004 which supports the interpretation of the simulation. While Geotail observed transient field-aligned ion beams in the solar wind upstream of the bow shock and the foreshock between 08:09 and 08:24 UT, ground-based SuperDARN radar data and IMAGE magnetometer data indicate that dayside reconnection was active and producing FTEs.

4.1 Upstream pristine solar wind conditions

The OMNI data set (from ACE) containing the upstream magnetic field, ion velocity, ion density and ion temperature on 30 August 2004 is plotted in Fig. 7. It shows that the IMF turned south at about 05:00 UT and $B_z$ remained strongly negative around $-10 \times 10^3$ nT for most of the day until about 22:00 UT. The velocity and temperature of the solar wind remained stable around 480–490 km s$^{-1}$ and $10^4$ K respectively between 08:00 and 19:00 UT, while $B_z$ slowly decreased from $-8$ to $-11$ nT and the density slowly increased from about 5 to $10^3$ cm$^{-3}$.

4.2 Ground observations

The strong southward $B_z$ component of the IMF is the cause of strong magnetic reconnection at the dayside magnetopause, which in turn is known to produce numerous FTEs (e.g. Kawano and Russell, 1997). Global activity indices clearly indicate ongoing magnetic reconnection during the event. The prolonged period of southward IMF triggered a geomagnetic storm and the increased levels of magnetic reconnection both on the dayside and the nightside are reflected in the AE indices, which started picking up between 05:00 and 06:00 UT and reached levels above AE = 500 nT after 08:00 UT.

Evidence for continuous FTE activity during the period 06:00 to 10:00 UT is observed by two SuperDARN radars in the Southern Hemisphere, presented in Fig. 8. The Kerguelen and Syowa East radars were observing backscatter from the pre-noon and noon region during this period. Figure 8a–c show the Kerguelen line-of-sight velocity data at 07:08, 07:32 and 08:22 UT. Three regions of backscatter are labelled A to C. In the polar cap (A), 1 km s$^{-1}$ flows away from the radar (antisunwards) are observed, 700 m s$^{-1}$ flows away from the radar (polewards) are seen entering the polar cap near noon (B), and 700 m s$^{-1}$ flows towards the radar (sunwards) are seen in the return flow region (C).
Earth rotates the look direction of the radar changes, but these backscatter regions continue to be observed. Range–time plots show that these backscatter regions are quasi-periodically pulsed with periods near 10–15 min, the expected signature of pulsed reconnection (e.g., Provan et al., 1998; Milan et al., 1999a, 2000; Wild et al., 2001). This is seen as poleward-moving enhancements in the backscatter power in the polar cap flows of region A in Fig. 8f. It is also observed as pulses of backscatter and flow in the return flow region, as first discussed by Milan et al. (1999b), that is, pulsed flows observed moving towards the Kerguelen radar (i.e., sunwards) in the return flow backscatter region C (Fig. 8e) and pulsing moving away from the Syowa East radar (also sunwards), in backscatter collocated with region C (Fig. 8i).

IMAGE magnetometers also observed signatures that could be interpreted as FTE activity. Figure 9a shows the ionospheric equivalent current density at 110 km altitude in the Northern Hemisphere at 08:15 UT. The equivalent current density was derived from 10 s IMAGE magnetometer data using spherical elementary current systems (SECS; Amm, 1997; Amm and Viljanen, 1999). Before applying the SECS method, a baseline was subtracted from the variometer data following van de Kamp (2013). The $J_{eq}$ data are presented as a function of Altitude Adjusted Corrected Geomagnetic Coordinates (AACGM; Shepherd, 2014) latitude and longitude, which at the given UT correspond to 09:05–11:09 magnetic local time (MLT). The plot has been rotated such that local noon is at the top. The plot shows eastward and equatorward equivalent current density vectors in the poleward part of the IMAGE field of view. If gradients of the ionospheric conductances are vanishingly small or aligned with the electric field in a large enough area, the equivalent current equals the Hall current, which flows antiparallel to the ionospheric $E \times B$ drift. According to Weygand et al. (2012), this is often a good approximation. Thus, the equatorward equivalent current in Fig. 9a may indicate poleward plasma flow entering the polar cap.

Figure 9b shows a $|J_{eq}|$ keogram, that is, latitude profiles of $|J_{eq}|$ along 105° longitude presented as a function of time.
between 06:00 and 10:00 UT. The vertical lines indicate the interval 08:09–08:24 UT during which Geotail observed the ion beam signature. The occurrence of the $|J_{eq}|$ intensifications observed by IMAGE between 70 and 75° latitude before about 09:00 UT roughly agrees in time with the Southern Hemisphere FTE signatures observed by SuperDARN. One of the intensifications occurred during the interval when Geotail observed the ion beam signature.

4.3 Geotail observations

On 30 August 2004, Geotail was located on the dayside of the Earth and upstream of the bow shock in the solar wind. Between about 08:00 and 08:30 UT, Geotail was
located at $(17.6, 23.0, -9.2) \, R_E$ in GSM coordinates and $(17.6, 24.5, 3.6) \, R_E$ in GSE coordinates. The location of Geotail with respect to a model bow shock and ion foreshock edge is shown in Fig. 10 (details of the bow shock and ion foreshock models are given in Sect. 5.1). Geotail/MGF observed stable IMF with $B_{GSM} = (7.5, -9) \, nT$, as shown in Fig. 11c. No perturbations of the magnetic field are seen which would modify its orientation, thus altering the magnetic connection to the bow shock and the location of the foreshock. Small-amplitude regular fluctuations are visible throughout, which coincide with Geotail’s nominal spin rate of 20 rpm. They are therefore likely a residual from the data calibration process.

The energy–time spectrograms from the Geotail/LEP instrument for the ions flying in the tailward and duskward directions are shown in Fig. 11a and b. Only two-dimensional projected ion velocity distributions in the ecliptic plane are available from LEP for this event, hence the choice of sectors. The tailward sector is dominated by the steady cold and dense solar wind core population just below 1 keV. Between about 08:09 and 08:24 UT (time delimited by the black dotted lines in Fig. 11) the duskward sector exhibits several occurrences of an ion population at energies of a few kilo-electronvolt reminiscent of the signature of foreshock field-aligned beam ions. The presence of a beam in this sector is consistent with the significant $B_y$ component of the IMF.

In contrast to regular foreshock field–aligned beams, this beam is transient and appears several times after 08:10 UT for one to a few minutes without correlation with changes in the magnetic field direction, as can be seen by comparing panels a–c in Fig. 11. Panel d shows the velocity and number density of the ions from the OMNI data set with a 1 min time resolution. The velocity is stable and varies only by about 1 %, while the density fluctuates between 4.5 and 5.5 cm$^{-3}$ but without correlating with the transient beam event. We choose the OMNI density and velocity data because the densities reported by Geotail/LEP do not seem to be consistent. This is likely due to the fact that LEP is in EA mode and not in SW (solar wind analyser) mode, which would have ensured a better measurement of the solar wind core population. To ensure that the choice of the OMNI data is sensible, we compare shifted ACE and Wind magnetic field mea-
Figure 11. Geotail/LEP energy–time spectrogram on 30 August 2004 between 07:48 and 08:47 UT, for the (a) tailward and (b) duskward sectors (±45° from the +x and +y direction respectively), extracted from the two-dimensional reduced velocity distribution in the GSE ecliptic plane. The tailward sector shows the cold and dense solar wind core population just below 1 keV. The duskward sector shows the signature of a transient beam whose density peaks several times between 08:09 and 08:24 UT (time delimited by the black dotted lines). (c) Geotail/MGF magnetic field in GSM coordinates. The magnetic field components and thus its orientation are stable during the event. The short-period oscillations coincide with Geotail’s nominal spin rate of 20 rpm. (d) OMNI plasma number density and velocity. The grey continuous bars indicate the time at which the velocity distribution shown in Fig. 12 is measured.

measurements to the Geotail/MGF data to check that the OMNI propagation algorithm is successful. We then check that the densities and velocities from ACE and Wind are similar to each other and to the OMNI dataset. Since this is the case and the OMNI values are similarly stable to the ACE (which was used to produce the OMNI dataset) and Wind data at a higher time resolution, we conclude that the OMNI dataset is reliable and sufficient here. LEP being in EA mode instead of SW affects the quality of the measurement for the core population but not for the beam, which is of prime interest here.

Figure 12 shows the projected two-dimensional velocity distribution in the $V_x$–$V_y$ GSE plane measured by the Geotail/LEP between 08:16:10 and 08:16:22 UT (time delimited by the grey continuous bars in Fig. 11). The solar wind core population is prominent in the lower left at $V_x \sim -400$ km s$^{-1}$, while a typical field-aligned beam flows back upstream along the magnetic field with positive $V_x$ and $V_y$ components. The black arrow points in the average direction of the magnetic field in the x–y GSE plane during the time interval, and the grey dots indicate all measurements taken at an 8 Hz cadence by the MGF instrument during the same time. Their close grouping once more indicates the stability of the magnetic field direction.

It is worth noting that while exactly similar magnetic field and solar wind conditions prevail in the 10 min preceding the event, no such field-aligned beam is seen before 08:10 UT. Additional ion beams are visible between 08:24 and 08:40 UT, but in their case the influence of magnetic field perturbations observed simultaneously cannot be conclusively ruled out.

5 Discussion

In Sect. 4 we present Geotail observations of transient field-aligned ion beams upstream of the Earth’s bow shock, while ground-based SuperDARN radar data and IMAGE magnetometer data show that pulsed dayside reconnection producing FTEs was occurring at the same time. This matches the observational signatures expected from the scenario drawn in Sect. 3 based on a global magnetospheric simulation. In this Sect. 5 we first investigate the position of Geotail with respect to the regular foreshock, and we then discuss the more general factors which might affect the interpretation of the simulation and the measurements.

5.1 Position of Geotail relative to the regular foreshock

It is important to ascertain that Geotail is not too close to the bow shock or to the foreshock. Indeed if it were in the vicinity of either, it could observe for example shock foot ion populations or the edge of field-aligned beam populations...
Figure 12. Coloured dots: Two-dimensional reduced ion velocity distribution in GSE coordinates measured by Geotail/LEP between 08:16:10 and 08:16:22 UT. The cold solar wind core flows with $V_x \sim -400 \text{ km s}^{-1}$, and the hotter field-aligned beam propagates in the opposite direction. Black arrow: averaged magnetic field direction in the $x-y$ GSE plane measured by Geotail/MGF during the interval. Grey dots: all individual Geotail/MGF measurements taken at 8 Hz cadence. The stability of the orientation of the magnetic field is remarkable.

from the foreshock, which would look very similar to the signature expected from a local foreshock.

We use the bow shock model from Merka et al. (2005) with the OMNI solar wind parameters and the Geotail/MGF observed magnetic field as inputs to determine the expected bow shock shape and position. As a first approach we determine the position of the expected ion foreshock edge. We trace the observed magnetic field to find the locus on the bow shock surface where $\theta_{B-n} = 60^\circ$, and then we trace the trajectories of ions reflected from there with twice the solar wind inflow speed in the solar wind rest frame. This is typical of field-aligned beams in the foreshock (e.g. Greenstadt et al., 1980; Eastwood et al., 2005a; Kis et al., 2007).

From this modelling we determine that Geotail is 3–5 $R_E$ away from the foreshock edge and 9–10 $R_E$ clear of the bow shock during the event between 08:09 and 08:24 UT, whence we could conclude that the spacecraft is well beyond reach of bow shock foot or foreshock edge ions. The result of this analysis is what is presented in Fig. 10.

However, as can readily be estimated from the observed velocity distribution shown in Fig. 12, the beam is significantly faster in the solar wind rest frame than twice the solar wind inflow velocity in the spacecraft frame. To get a better estimate of whether Geotail is within reach of the regular foreshock, we consider the trajectory of foreshock ions assuming adiabatic reflection at the bow shock (Schwartz et al., 1983; Liu et al., 2016). Adiabatic reflection and not specular reflection is assumed because it yields higher beam speeds and would thus be more likely to reach the spacecraft. Incoming ions at velocity $V$ impinge on the bow shock, which has a local normal vector $n$, and are reflected with a velocity

$$V' = -V + 2V_{HT},$$

where

$$V_{HT} = \frac{n \times (V \times B)}{n \cdot B}$$

is the de Hoffmann–Teller velocity of the bow shock (De Hoffmann and Teller, 1950) and $B$ is the IMF.

The validity of the assumption that ions are reflected adiabatically can be checked against the simulation. In Fig. 5a we have $\theta_{B-n} = 41^\circ$ at the foot of the local foreshock, the solar wind velocity is $-750 \text{ km s}^{-1}$ purely along the $x$ axis, and the IMF is constant and purely southward at 5 nT. With those parameters, Eqs. (1) and (2) yield $V_{sim} = (-750, 0, 1725.6) \text{ km s}^{-1}$. This does indeed correspond to the beam velocity as shown by the projected velocity distribution function in Fig. 5b, confirming the assumed adiabatic reflection process.

To check whether Geotail observed adiabatically reflected ions from the regular foreshock, we trace the observed beam from Geotail back to the bow shock along the IMF direction and invert Eq. (1) to reconstruct the solar wind velocity vector $V'$ that would yield the measured beam velocity with the assumed model bow shock. For the observed beam (Fig. 12), the resulting solar wind velocity vector would have the components $V' = (-557, -602, 317) \text{ km s}^{-1}$ in GSE coordinates, which is obviously not in accordance with the observed solar wind (Figs. 7b and 11d). Having thus ascertained that in these solar wind and IMF conditions the observed beam cannot have been reflected adiabatically from the modelled bow shock, we perturb the model bow shock until the adiabatically reflected ion trajectory matches the observed beam. When $n$ is rotated towards $B$ by an angle of $15^\circ$, the adiabatically reflected ion beam does indeed hit Geotail. This value is remarkably close to the angle of $13^\circ$ between the unperturbed and perturbed bow shock normals at the foot of the local foreshock in the global simulation (Fig. 5a). It has to be noted though that the beam velocities obtained with this approach do not agree well with the observed ones (reconstructed velocity $688 \text{ km s}^{-1}$, observed velocity $885 \text{ km s}^{-1}$ in the spacecraft frame), which means that the reflection and acceleration process and their geometry are probably more complex than the simple adiabatic reflection we assume here.

In summary this analysis demonstrates that Geotail is out of reach for adiabatically reflected field-aligned beam ions originating from the unperturbed foreshock under the prevailing solar wind and IMF conditions. By introducing an ad hoc local perturbation of the bow shock normal of $15^\circ$, we recover a beam direction consistent with the Geotail observations, which is similar to the bow shock perturbation seen in the global simulation.
5.2 Estimating the propagation direction of the perturbation

Assuming that three-dimensional velocity distribution measurements were available with a high cadence, it would be possible to estimate the direction and speed of propagation of the field-aligned ion beam. Indeed at the edge of an ion beam, non-gyrotropic partial ring or beam distributions are observed in a region of one gyrodiameter width (e.g. Schwartz et al., 2000; Kempf et al., 2015). Timing the transition from a partial to a full gyrotropic beam and back on each side of the beam passage across the spacecraft yields an approximate beam propagation speed since the gyrodiameter of the ions is known. Furthermore, the gyrophase of the ions at the very edge of the beam indicates on which side of the spacecraft the beam is located, so that the incoming and outgoing edges yield an estimate of the beam edge orientation. However the lack of knowledge of the shape of the beam complicates the matter to some extent. For the event presented in Sect. 4, such estimates are not feasible with the Geotail data available and the comparison of the event timings in the absence of a more detailed knowledge of the shape of the bow shock perturbation and beam is of no use.

5.3 Simulation model limitations

The main limitation of the simulation presented is its two-dimensionality. Due to this configuration, all inflowing magnetic flux is forced to reconnect at the magnetopause and cannot flow past without reconnecting, unlike in three dimensions. This forces magnetic reconnection to be strong and occur all the time at the magnetopause. Further, this likely means that the magnetic islands carry more momentum and thus drive stronger bow and stern waves into the magnetosheath than they would in three dimensions.

Additionally, the steady solar wind conditions in the simulation preclude any upstream turbulence, yielding a smooth bow shock and no more downstream turbulence than the anisotropy-driven wave activity in the magnetosheath. Therefore the magnetic-island-driven fast waves can propagate relatively unhindered in the magnetosheath and the localized field-aligned beam is also very prominent in the solar wind. More realistic turbulent conditions would certainly yield less conspicuous signatures.

Nevertheless, none of these limitations mean that the fully three-dimensional and turbulent case could not exhibit transient local foreshocks, they might only be more difficult to detect and distinguish from other sources of bow shock and foreshock perturbations or ion beams.

5.4 Observational limitations

The long chain of phenomena from the magnetopause through the magnetosheath and bow shock to the foreshock, constituting the scenario presented in this work, makes it daunting to observe the whole cascade of a single event in space and time. This would require the fortuitous availability of adequate measurement data firstly at the magnetopause to identify FTEs, secondly in the magnetosheath to single out steepened fast wave fronts, and thirdly upstream in a narrow region close to but definitely more than one ion gyroradius away form both the bow shock foot and the regular foreshock edge, all of this during a stable southward IMF stretch and in the absence of any magnetic field fluctuations which could either drive an ion beam or produce a regular foreshock crossing at the upstream spacecraft instead. No suitable spacecraft were located in the magnetosheath or at the magnetopause during the Geotail event presented above so that a direct observation of fast-mode magnetosheath wave fronts is not possible in this case.

As shown in Sect. 4, transient foreshock-like ion beams upstream but well-separated from both the bow shock and the foreshock are observed. At the same time, ground-based measurements confirm that dayside reconnection was occurring and producing FTEs propagating towards the poles. Without adequate magnetosheath observations, it is not possible to claim with certainty that the complete scenario mapped in Sect. 3 holds. Yet the observations are consistent with the first and the last part of the story, namely that while dayside reconnection is active and pulsed, a localized change in the bow shock shape causes localized ion reflection and the formation of a transient, local foreshock. It cannot be excluded that sources other than FTE-driven fast waves exist, but it is unlikely there would be distinctive features helping to tell them apart purely based on the observation of the ion beam without other measurements, from within the magnetosheath for example. Such putative sources could include unpredicted magnetosheath waves interacting with the bow shock or very localized solar wind transients not observed by the upstream spacecraft. Finally, we note that the presented scenario offers an alternative mechanism to explain transient foreshock encounters that may have been interpreted previously as foreshock skimmings due to a change in the magnetic connection of the spacecraft to the bow shock.

6 Conclusions

Global hybrid-Vlasov simulations of the terrestrial magnetosphere in the polar plane under constant southward IMF show that the two-dimensional equivalents of flux transfer events, formed by dayside magnetopause reconnection, drive steepened bow and stern fast-mode waves in the magnetosheath. The increased pressure behind the wave fronts causes the bow shock to bulge out, inducing favourable ion reflection conditions which can result in the generation of local, transient foreshock-like field-aligned ion beams upstream of the bow shock. The two-dimensionality of the simulation exacerbates the phenomena, but the scenario would be entirely similar in three dimensions.
Ground-based and spacecraft observations support this scenario. During an extended period of stable southward IMF, we observe ionospheric signatures of dayside reconnection and flux transfer events in SuperDARN radar data and IMAGE magnetometer data. Simultaneously, using Geotail magnetic field and ion velocity distribution measurements we observe the expected signature of an ion beam detached both from the bow shock and the regular foreshock and not linked to any upstream magnetic field fluctuation. Further observations especially in the magnetosheath are needed though to confirm that indeed fast-mode waves lead to bow shock deformations generating localized, transient field-aligned ion beams.

7 Data availability


Acknowledgements. The simulation was run on the Sisu supercomputer at the CSC – IT Center for Science, Espoo, Finland. We thank T. Mukai at ISAS, JAXA in Japan for providing Geotail/LEP data; S. Kokubun at STELAB, Nagoya University, Japan for providing Geotail/MGF data; A. Szabo at NASA/GSFC for providing Wind/MFI data; K. Ogilvie at NASA/GSFC for providing Wind/SWE data; R. Lin and S. Bale at UC Berkeley for providing Wind/3DP data; M. L. Kaiser at GSFC for providing Wind/WAVES data; N. Ness at Bartol Research Institute for providing ACE/MAG data; D. J. McComas at SWRI for providing ACE/SWEPAM data; and J. H. King and N. Papatashvili at AdnetSystems and GSFC for providing OMNI data, all through CDAnet. We also thank T. Nagai and Y. Saito for providing Geotail/MGF and Geotail/LEP data through DARTS at ISAS, JAXA in Japan. We thank the University of Kyoto for providing the AE electrojet indices through the World Data Center for Geomagnetism.

We acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Japan, South Africa, United Kingdom and United States of America. We thank the institutes who maintain the IMAGE magnetometer array.

Yann Pfau-Kempf, Sanni Hoiijoki, Sebastian von Alfthan and Minna Palmroth acknowledge financial support from the Academy of Finland under the project 267144/Vlasov. Part of this study was done by Yann Pfau-Kempf, Liisa Juusola, Urs Ganse, Sanni Hoiijoki and Minna Palmroth under the ERC CoG-682068-PRESTISSIMO project, a Consolidator grant to Minna Palmroth from the European Research Council. The work of Heli Hietala is funded by NASA contract NASS-02099. Steve E. Milan was supported by the Science and Technology Facilities Council (STFC), UK, grant no. ST/N000749/1. Urs Ganse acknowledges funding from the German Research Foundation Grant GA1968/1 and the Academy of Finland project 267186.

The topical editor, C. Owen, thanks two anonymous referees for help in evaluating this paper.

The Supplement related to this article is available online at doi:10.5194/angeo-34-943-2016-supplement.

Author contributions. Sebastian von Alfthan and Yann Pfau-Kempf designed and ran the simulation presented in this work. Heli Hietala helped in the analysis of the Geotail observations. Steve E. Milan provided the analysis and figure of the SuperDARN observations. Liisa Juusola provided the analysis and figure of the IMAGE observations. Yann Pfau-Kempf, Heli Hietala, Sanni Hoiijoki, Urs Ganse and Sebastian von Alfthan developed analysis tools used in this study. Minna Palmroth is the Principal Investigator of the Vlasiator team at the Finnish Meteorological Institute. Yann Pfau-Kempf led the analysis and prepared the manuscript and figures with the help of all co-authors.

References


