Numerically investigating the Hall current effect on magnetic reconnection

Stian Aannerud

Thesis submitted for the degree of
Master of Science in Astronomy

Institute of Theoretical Astrophysics
University of Oslo

May 2023

Abstract

Magnetic reconnection is a large part in driving solar outbursts, but how to explain their sudden bursts after periods of relative calm is still not fully known. One possible explanation is the often overlooked Hall effect, in which it becomes possible to explain reconnection through a build up of energy through separated currents. More focus has come on the Hall effect in recent years, as it seems it could be a powerful driver in this area.

We build upon the pre-existing solar MHD model Bifrost, and add in the Hall term in order to investigate this. Through a modification in Ohm's law we allow for current separation at inertial scales, and run tests for the effectiveness when compared with normal ideal MHD models. The goal of this thesis is to measure the reconnected flux from the Hall MHD model and see if it has a significant impact on reconnection rate.

Through following a tutorial laid out by Joseph D. Huba in his 2003 paper [3], we perform both a whistler wave validation test, and a reconnection setup. We found that the Hall effect indeed has a large impact on the amount of reconnected flux, which falls in line with previous investigations.

Acknowledgments

I would like to thank all who supported me through this process. Professor and advisor Boris Vilhelm Gudiksen for excellent guidance and advice, and great patience throughout. Thanks to my close friends and colleagues Marius B. Børvind and Dennis Fremstad for motivation and comradery throughout my studies. Also thanks to the University of Oslo, especially the Institute of Theoretical Astrophysics and the group Rosseland Centre for Solar Physics for taking me in and granting me the opportunity to work with them.

Contents

List of Figures

Chapter 1

Introduction

Shining bright in the sky, our sun is the largest and most astonishing part of our solar system. A giant molten ball of plasma constantly emitting light and energy in huge quantities, it is an immense force and brilliantly alluring. In addition to the constant stream of radiation sent out from its surface, it occasionally has even stronger, violent outbursts which send flowing supercharged streams of particles into space. These solar flares are abnormally strong, even for the sun, and a big topic of interest. However their mechanism is not yet fully understood. The surface of the sun is a big tangle of magnetic forces, constantly shifting and twisting. Sometimes two strong opposing fields start pushing against each other, vying for space, until they finally snap. The fields rapidly change shape, like rubber bands being pulled too far, and this pushes the surrounding plasma with tremendous force causing solar flares. Though it is the part where these fields lie against each other that is not very clear. As they push, and push against the other, what finally causes them to snap? They might keep pushing for a very long time, seemingly in a stable and unchanging position, but then suddenly letting loose all at once. It is theorised that there must be some sort of build-up of energy within these areas which we simply cannot see and some catastrophic event that sets the flare off.

This is where the Hall effect comes into play. It is a simple, yet oft overlooked, part of Ohm's law. It describes how a current of charged particles can generate an electric field. Usually this part is disregarded in solar physics, as we assume the plasma which makes up the sun to be very, very hot and the particle it consists of colliding often. That would mean that the particles are tightly bound and not separable. This leaves no room for any potential electric fields. However, it is possible that within the area of strong opposing magnetic fields, this might actually be possible. Due to the wildly different masses of electrons and ions, they are pushed around at different speeds, and just in such an area they might be pushed in just the right way to separate out. This would then produce an electric field. Energy would build up steadily, which after a while would be strong enough to cause the opposing solar magnetic fields to snap and reconnect.

Therefore I have in this thesis chosen to investigate whether this could be possible, by

implementing this Hall effect into the strong solar simulation code Bifrost, and studying its impact on reconnection rate. With access to modern high-tech computing power and a simulation built for speed, I have a great opportunity to investigate this problem here at the Institute of Theoretical Astrophysics at the University of Oslo. The Rosseland Centre for Solar Physics (RoCS) group at the institute has all the tools required to succeed. I will mainly be following the work by Joseph D. Huba [3], which goes through all the required steps to go from a normal magnetohydrodynamic solver like Bifrost, to a fully fledged Hall MHD setup. Configurations for validating our implementation through detecting unique Hall wave modes and a reconnection setup are both included and will be used as a guide.

We will start by covering background information required to understand these themes in the theory section. Here we start by covering the very basics through single particle motion 2.1. We do this because the underlying motion of electrons and protons explain the origin of more complex phenomena which are useful in our later descriptions. This naturally leads into the section on magnetohydrodynamics 2.2, where we introduce the large scale description of astrophysical plasma. As the name suggests it is a combination between hydrodynamics, fluids, and electromagnetic forces. It is the foundation to which we add our modification, the Hall effect 3.1. Having moved into the method section of the thesis, we cover where the Hall effect comes from and how it is applied. Despite the simplistic description, it gives rise to a couple more complex phenomena which we need to keep track of, its special whistler and Hall drift wave modes. Then we go into a all-you-need-to-know section on the Bifrost code which we use as a foundation in our simulation section 3.2, as well as how we implement our new addition numerically. In order to check that we have successfully done as we should, we describe how we test and compare our results against analytical solutions in the validation section 3.3. Here we look for the aforementioned whistler waves, and validate that they behave as we expect them to. If they do, then we know our Hall MHD model is correct. Having cemented that our code works as it should, we go through the real deal in the experiment and discussion section 4. We describe and perform a setup to simulate magnetic reconnection, as well as a method of measuring how much of an impact the Hall effect has. We summarize our findings in the conclusion 5, and add any possible improvement we could have done in the future work section 6.

Chapter 2

Theory

2.1 Single particle motion

A charged particle moving through electric and magnetic fields, $\mathbf{E}$, $\mathbf{B}$, get affected by the Lorentz force, $\mathbf{F}$, which can be linked to the particle acceleration using Newton's second law of motion

$$\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{2.1}$$

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{q}{m}(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{2.2}$$

where q, m, $\mathbf{u}$ is the charge, mass, and velocity of the particle respectively. Note how we can split the force in two separate parts: $q\mathbf{E}$ and $q\mathbf{u} \times \mathbf{B}$, meaning the contributions from the electric and magnetic fields are fully independent of each other. While the electric field provides a continious acceleration of the particle, the magnetic component of the force is relying on the particle having acquired a velocity already – whether this is from the electric field or another source is of no consequence.

2.1.1 Particle orbits

From the Lorentz force (2.1) we can see that magnetic forces always act perpendicular to the direction of movement. This causes a charged particle moving through the magnetic field to turn continuously in an orbital motion in the plane normal to the magnetic field. The magnetic Lorentz force only acts in this plane, meaning all movement parallel to the field does not affect the magnitude or direction of the magnetic Lorentz force. Therefore it is helpful to separate the particle velocity $\mathbf{u}$ into two components: Field parallel velocity $\mathbf{u}_\parallel$ such that $\mathbf{u}_\parallel \times \mathbf{B} = 0$ and field orthogonal velocity $\mathbf{u}_\perp$ which gives $\mathbf{u}_\perp \times \mathbf{B} = \mathbf{u} \times \mathbf{B}$. The magnetic orbit will have a radius r, known as the Larmor radius or gyroradius, and a gyrofrequency ω_c defined as: [7, p. 81-82]

$$r = \frac{u_\perp}{\omega_c} \qquad (2.3) \qquad\qquad \omega_c = \frac{qB}{m} \qquad (2.4)$$

As we can see the gyrofrequency and radius are effectively determined by the ratio between the magnetic field strength and particle mass. Comparing electrons and protons/positive ions for the same field setup, we can see that the heavier ions will have larger and slower orbits compared to the smaller and faster electron orbits due to their mass difference $m_p >> m_e$.

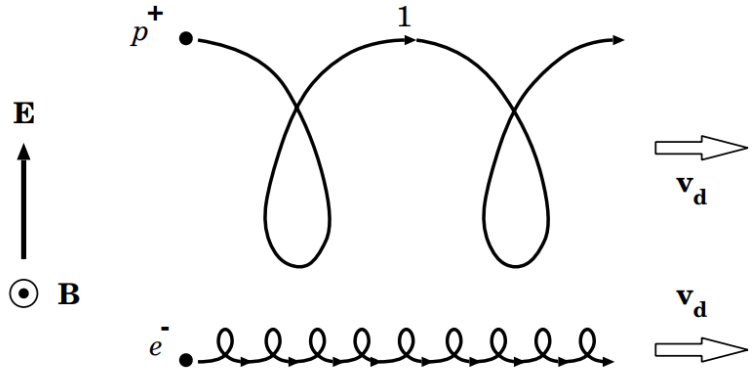

Figure 2.1: Illustration of the cycloid motion of a charged particle in a $\mathbf{E} \perp \mathbf{B}$ field configuration. From Somov [7].

In the presence of another external field accelerating the particle this orbit starts to drift, causing a cycloid motion as illustrated in Figure 2.1 for an electric field. The acceleration of the particle does not cause the speed of the particle to increase over time, as the contributions are cancelled when the particle reaches the opposite side of the cyclotron orbit and the forces are reversed. It does however cause the overall movement to *drift* with a velocity $\mathbf{v}_d$, defined as: [7, p. 84]

$$\mathbf{v}_d = \frac{1}{q}\frac{\mathbf{F} \times \mathbf{B}}{B^2} \qquad (2.5)$$

where the general force $\mathbf{F}$ is replaced with whichever external field affecting the particle. For an electric field this is the electric Lorentz force component $q\mathbf{E}$, and in a gravitational field it becomes $m\mathbf{g}$. The drift is normal to both the accelerating force and the magnetic field, and changes direction based on the particle charge in the case of an electric field. Note that this is only for the case where the external field is orthogonal to the magnetic field, as we established earlier all motion $\mathbf{u} \parallel \mathbf{B}$ is independent of the gyromotion in the plane perpendicular to $\mathbf{B}$.

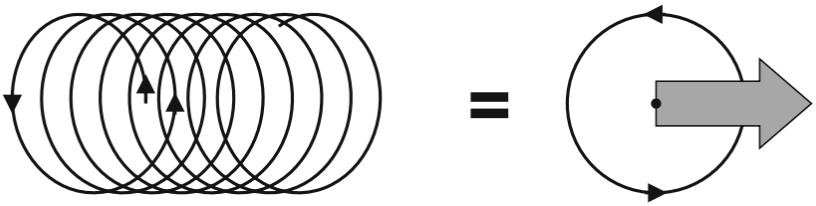

Figure 2.2: Illustration of the guiding centre approximation, from Piel [6].

In some cases the gyration of the charged particles can be ignored and an approximate description of the centre of the gyration motion be treated. The guiding centre approximation takes this approach, which requires that relevant scales are much larger than the gyroradius. That leaves only the circular gyromotion and a drift velocity equivalent to shifting point of reference to the centre of the gyromotion. An illustration of this shifted reference can be seen in Figure 2.2.

2.1.2 Plasma parameters

A collection of charged particles can be defined as a plasma. The plasma is characterised by a certain parameters which describes the nature of the plasma. The plasma frequency ω_p is one of the main characteristics of a plasma. A plasma is generally assumed to be quasi-neutral but if one imagines that the positively charged ions are collected in one klocation and this produces an electric field, then they will be accelerated away from this positive electric potential. The motion of all the ions will create a negative potential, making them reverse direction, so an oscilation will be set up. This oscillation will have a frequency which is called the plasma frequency. It has a different magnitude depending on the mass of the oscillating particle (ions and electrons). Another important plasma parameter is the Debye length. The Debye length is the length over which an electric field is neutralised by the surrounding charged particles of opposite charge. These parameters and a number of other parameters derived from these are shown below.

Summary of basic plasma parameters, expressed in terms of the appropriate charges q, masses M, densities n and temperatures T

$$\omega_p = \sqrt{\frac{q^2 n}{\epsilon_0 M}} \qquad \lambda_D = \sqrt{\frac{\epsilon_0 \kappa T}{q^2 n}} \qquad \omega_p \lambda_D = \sqrt{\frac{\kappa T}{M}} = u_{\mathrm{The}}$$

$$N_p \equiv n\lambda_D^3$$

$$\ell_c \sim \lambda_D N_p \qquad \tau_c \sim N_p/\omega_p \qquad v_c \sim \omega_p/N_p$$

Approximate expression for the resistivity ξ (with conductivity σ) of a fully ionized plasma

$$\xi \equiv \frac{1}{\sigma} = \frac{1}{N_p \omega_{pe} \epsilon_0}$$

The present summary does not include the corrections due to the Spitzer logarithm, and also some numerical factors are omitted.

2.2 Magnetohydrodynamics

Magnetohydrodynamics, abbreviated as MHD, is an approach and set of equations used to describe electrically conductive fluids, often in the presence of an external magnetic field. One such fluid is the astrophysical plasma which makes up the sun, giving MHD a notable presence within the field of solar physics. Compared to other plasma descriptions it is a relatively simple yet efficient method of understanding complex phenomena. It's efficient nature makes it a great candidate for numerical simulations such as our own used in this thesis.

Several assumptions are made when applying the base MHD description to plasmas regarding its physical properties and use cases. The plasma is assumed to be quasi-neutral in that there is an even amount of positive and negative charges and that they on the relevant scales neutralise each other making it very difficult to create large scale electric fields. Because of the plasma parameters, MHD is generally applicable only on large scales compared to the Debye length (see section 2.1.2). As the second part of its name implies, MHD also assumes that the the plasma can be described as a fluid. Fluid motions are dominated by collisions so MHD is only applicable on length scales much longer than the mean free path. The frequent collisions also ensures the positive and negative charges are mixed and can be treated as a single fluid. In the ideal case infinite conductivity is also given, letting electrons flow without any resistance.

Variations on MHD exist however, where each of the assumptions mentioned are removed. In the case where all the normal assumptions are used it is referred to as Ideal

MHD. It is often used in theoretical investigations, but is too simplistic to be applied in many real scenarios. Without resistivity there is no way the fluid can move across the magnetic flux direction. This is known as the frozen-in flux theorem, and it is similar to how the guiding centre approximation only allows independent particle movement along field lines. Loosening the criteria of infinite conductivity leads to resistive MHD, where some resistivity $\eta \neq 0$ lets the fluid flow across the magnetic field. For cases with magnetic reconnection this is crucial. We know that the theoretical resistivity in a plasma as it exists on the Sun is extremely low, but also observe that at crucial times, the resistivity of the plasma in the solar atmosphere must increase dramatically to allow for solar flares to occur. Another variation is Hall MHD, where we allow for charge separation between ions and electrons. We go into more detail on this variation in section 3.1.

At its core, MHD combines Navier-Stokes fluid dynamics and Maxwell's equations for electromagnetism. The fluid dynamics ensure physical quantities such as mass and momentum are conserved, while others describe the electric and magnetic forces which both affect the fluid and are created by the fluid. They are all expressed as partial differential equations, which have to be solved either analytically or numerically to get a description of plasma behaviour.

From fluid dynamics we get the mass continuity equation and the Cauchy momentum equation respectively:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{u} \tag{2.6}$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} = -\nabla \cdot (\rho \mathbf{u}\mathbf{u} - \tau) - \nabla P + \mathbf{J} \times \mathbf{B} + \rho \mathbf{g} \tag{2.7}$$

$$\tag{2.8}$$

Here ρ is the mass density, $\mathbf{u}$ is the fluid velocity, τ is the Cauchy stress tensor, P is the thermal gas pressure, $\mathbf{J}$ is the current density, $\mathbf{B}$ is the magnetic flux density, and $\mathbf{g}$ is the gravitational acceleration. The momentum equation (2.7) can take on different forms depending on the context in which it is used. Here we have added the Lorenz force as part of the equation, and often situations may neglect gravity and the stress tensor.

Two of Maxwell's equations make up the core of the electromagnetic section of MHD, Ampère's law and Faraday's law, along with Ohm's law:

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B} \tag{2.9}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{2.10}$$

$$\mathbf{E} = \eta \mathbf{J} - \mathbf{u} \times \mathbf{B} \tag{2.11}$$

Where μ_0 the the vacuum permeability/magnetic constant and η is the electrical resistivity. Note how Ampère's law is lacking the term $\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$ which is normally present,

where ϵ_0 is the vacuum permittivity/electrical constant. The vacuum constants have a relation $\mu_0\epsilon_0 = 1/c^2$, which means that this term is only important when violent changes in the electric field happens. The many charge carriers is assumed to be able to keep electric fields relatively small, so the term is assumed to be negligible in comparison to the other terms which allows us to ignore it. Often the three equations are combined into the single induction equation:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \frac{\eta}{\mu_0}\nabla^2 \mathbf{B} \tag{2.12}$$

This is done by finding the curl of Ohm's law, inserting for $\mathbf{J}$ from Ampère's law, and inserting for $\nabla \times \mathbf{E}$ from Faraday's law. The two terms on the right hand side are the inductive term and the diffusive term respectively.

Finally, in addition to the previous equations we need a relation between temperature and pressure to the other quantities. This is referred to as an equation of state (EOS), and can be expressed as a gas equation or an energy equation. While this could be any function, a typical choice could be the ideal gas law:

$$PV = nk_BT \tag{2.13}$$

Where V is volume, k_B is the Boltzmann constant, and T is the plasma temperature. In the perfect gas approximation we could also express this as $P = \rho(\gamma - 1)e$, where γ is the heat capacity ratio/adiabatic index and e is the specific internal energy of the plasma.

Chapter 3

Method

3.1 Hall

The Hall term appears as part of the generalized Ohm's law, which is as follows: [3]

$$\frac{m_e}{ne^2}\frac{\partial \mathbf{J}}{\partial t} - \frac{1}{ne}\nabla P_e = \mathbf{E} + \frac{1}{c}\mathbf{u}\times\mathbf{B} - \frac{1}{nec}\mathbf{J}\times\mathbf{B} - \mathbf{J}/\sigma \tag{3.1}$$

$$\tag{3.2}$$

Where $n = n_i = n_e$ is the plasma number density, P_e is the electron pressure, and σ is the electrical conductivity. Assuming we have length scales larger than the electron inertial length $L >> c/\omega_{pe}$, $L >> r_e$, where r_e is the electron Larmor radius, and large conductivity $\sigma \to \infty$ we can neglect the electron inertia, pressure, and conductivity terms. This lets us simplify Ohm's law to the expression:

$$\mathbf{E} = -\frac{1}{c}\mathbf{u}\times\mathbf{B} + \frac{1}{nec}\mathbf{J}\times\mathbf{B} \tag{3.3}$$

Here we have both the classic Ohm's law terms, as well as a new $\mathbf{J}\times\mathbf{B}/nec$ component. This is the Hall term. For further simplification we also introduce the Hall velocity V_H. It simplifies the expression as follows:

$$\mathbf{V}_H = -\frac{\mathbf{J}}{ne} \tag{3.4}$$

$$\mathbf{E} = -\frac{1}{c}\mathbf{u}\times\mathbf{B} - \frac{1}{c}\mathbf{V}_H\times\mathbf{B} \tag{3.5}$$

The Hall velocity can be thought of as the difference in velocity between the electrons and ions. Due to their much heavier mass, we can view the ions as a relatively motionless background compared with the electrons. Normally the electrons are bound to the ions, but at scales between the electron and ion inertial length the electrons are

locally uncoupled. The Hall velocity is then the electron velocity when compared to this background, subtracting away any potential large scale drift.

In areas with awaiting magnetic reconnection, where strong opposing magnetic fields push close to each other, the uncoupled electron flow induces an additional magnetic field orthogonal to the background field. An illustration of this can be seen in Figure 3.1 where an influx of electrons into a reconnection area generate new magnetic fields. This effect has been shown previously to have a significant effect on the reconnection rate [3], and we describe our experiment for numerically simulating this in section 4.1.

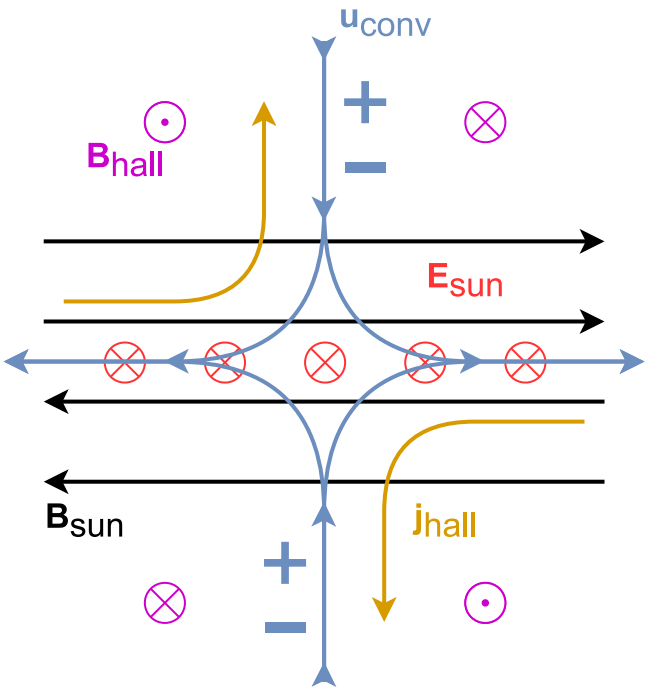

Figure 3.1: Illustration of relevant fields and velocities for the Hall effect in a reconnection site. The background magnetic field $\mathbf{B}_{\text{sun}}$ generates a background electric field $\mathbf{E}_{\text{sun}}$. Inflow of plasma $\mathbf{u}_{\text{conv}}$ leads to a current $\mathbf{j}_{\text{hall}}$. This current is controlled by the electrons due to the charge separation on inertial scales. Through the Hall effect a magnetic field $\mathbf{B}_{\text{hall}}$ is induced by the current.

Due to the added Hall effect, we get two new wave modes introduced. These waves are the whistler and Hall drift waves. They are much faster moving than all normal MHD waves, and will need extra consideration in numerical implementations to resolve properly. To do this we use their dispersion relations such that we can use their velocities to choose suitable time steps for numerical integration. We start by splitting the Hall contribution as part of the induction equation 2.12. Here we assume the limit where the Hall velocity dominates the normal background, such that $\mathbf{u} = \mathbf{u}_i + \mathbf{V}_H \approx \mathbf{V}_H$, where $\mathbf{u}_i$ is the background ion velocity. We neglect the $\eta\nabla^2\mathbf{B}$ term, due to our previous assumption that $\sigma \to \infty$, $\eta = 1/\sigma \to 0$:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V}_H \times \mathbf{B}) = -\nabla \times (\mathbf{J}/ne \times \mathbf{B}) \tag{3.6}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\frac{1}{ne}\nabla \times (\mathbf{J} \times \mathbf{B}) + \frac{1}{n^2 e}\nabla n \times (\mathbf{J} \times \mathbf{B}) \tag{3.7}$$

The dispersion relation for each wave mode is found by adding a perturbation in the magnetic field and solving for the wave frequency. We show the whistler wave dispersion solution method in section 3.3. The Hall drift waves are found similarly by perturbing the B_z component of the magnetic field, and we refer to Huba [3] for the full solution method. Each dispersion then is as follows:

$$\omega = \frac{k_z^2 B}{4\pi ne} = k_z V_A \left(\frac{k_z c}{\omega_{pi}}\right) \tag{3.8}$$

$$\omega = \frac{k_y B}{4\pi ne}\frac{1}{n}\frac{\partial n}{\partial x} = k_y V_A \left(\frac{c}{L_n \omega_{pi}}\right) \tag{3.9}$$

Where k is the wave number, V_A is the Alfvèn velocity, and $L_n = (\partial \ln n/\partial x)^-1$ is the density gradient scale length. Equations (3.24) and (3.9) are the whistler and drift dispersions respectively. Note the difference in dependence on the wave number k, where the whistler mode goes as the mode squared while the drift is linear. Details on how we implement these into our time step can be found in section 3.2.1.

3.2 Simulation

Our experiments are simulated using the Bifrost code [2], a highly modular 3D framework for solving the MHD equations as well as built-in full radiative transfer methods. Its core use area is in solar simulations, where user input initial conditions on a Cartesian grid of cells are advanced in time. Bifrost tracks four core quantities from which others are derived: The mass density ρ, momentum $\rho\mathbf{u}$, magnetic flux density $\mathbf{B}$, and internal energy per unit volume e. The partial differential equations describing these variables changes over time are implemented as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho\mathbf{u} \tag{3.10}$$

$$\frac{\partial \rho\mathbf{u}}{\partial t} = -\nabla \cdot (\rho\mathbf{u}\mathbf{u} - \tau) - \nabla P + \mathbf{J} \times \mathbf{B} + \rho\mathbf{g} \tag{3.11}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{3.12}$$

$$\frac{\partial e}{\partial t} = -\nabla \cdot e\mathbf{u} - P\nabla \cdot \mathbf{u} + Q \tag{3.13}$$

$$\mu\mathbf{J} = \nabla \times \mathbf{B} \tag{3.14}$$

$$\mathbf{E} = \eta\mathbf{J} - \mathbf{u} \times \mathbf{B}, \tag{3.15}$$

for velocity $\mathbf{u}$, stress tensor τ, gas pressure P, electric current density $\mathbf{J}$, gravitational acceleration $\mathbf{g}$, electric field $\mathbf{E}$, vacuum permeability/magnetic constant μ, and magnetic diffusivity η. The quantity Q added to the energy equation is a connecting term for other energy transfer systems, such as Bifrost's radiative transfer or thermal conduction. Note the different form of Faraday's law (3.12) and Ohm's law (3.15), where they are missing each a factor c from the usual Gaussian representation. Since the electrical field is only derived through the magnetic field, these factors are canceled for numerical ease. To fulfill the set of MHD equations some equation of state (EOS) is also required, such as the ideal gas law, and several possibilities are provided through the Bifrost code. We assume an ideal gas as the EOS for all experiments in this thesis.

To ensure stability and efficient time evolution, two explicit methods for timestepping are provided: The third-order Runge-Kutta scheme, and the third-order Hyman timestepping scheme [4]. For the Runge-Kutta scheme the timestep is split in several sub-steps which are then combined for an output third-order precision, whereas the Hyman timestepping uses a predictor-corrector scheme scaling to ensure Courant stability conditions, also reaching a third order precision. We use the Hyman method for all our tests here.

Calculations using spatial derivatives are optimized through using a staggered grid. Values for tracked quantities in the simulation are placed at explicit coordinates throughout the Cartesian grid, each point representing the cell which contains it. Naively one might place a cells values equally for all variables, but this can be computationally disadvantageous when involving derivatives. Bifrost by default uses a 6th order finite differential scheme, which causes the derivatives to shift in place by half a grid cell compared to the differentiated values. To avoid interpolating these values back in place where spatial derivatives are used in computations, we instead define our core quantities are initially shifted such that the derivatives fall into place where needed. This setup is called a *staggered grid*, and Bifrost's implementation is based on previous schemes by Nordlund & Galsgaard [5], though with opposite coordinate shifts. A representation of how variables are distributed by Bifrost in each cell can be seen in Figure 3.2.

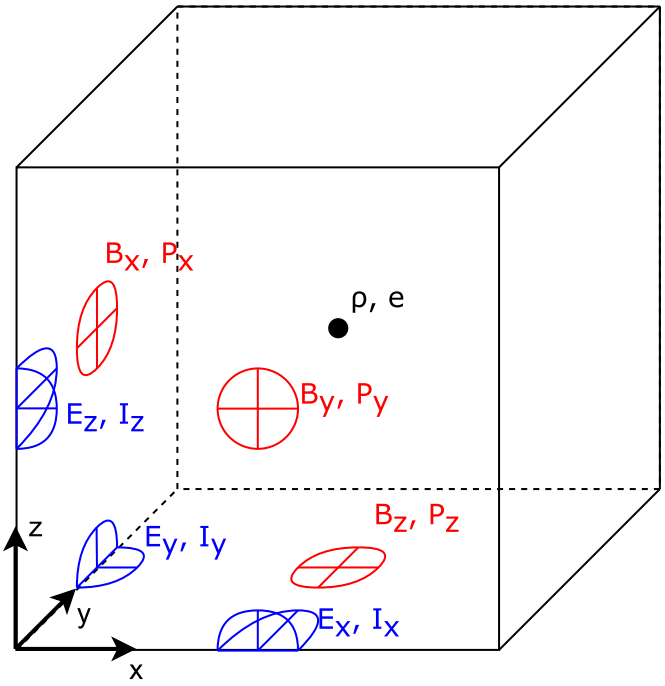

Figure 3.2: Schematic of how variables are staggered in comparison to each other in Bifrost. The density ρ and internal energy e are volume centered and marked as a black dot. The magnetic field $\mathbf{B}$ and momentum $\mathbf{P} = \rho\mathbf{u}$ components are face centered, where each component Bx, By, Bz have been moved half a step down along their respective axis x, y, z. This is different from the electric field $\mathbf{E}$ and current $\mathbf{I}$ which are edge centered, and have been moved half a step down along the other two axes (e.g. E_x has been moved half a step down y and z).

3.2.1 Hall implementation

In practice our Hall effect appears not as a part of Ohm's law, but rather an addition to Faraday's law (3.12). As the electric field is not a core quantity in the Bifrost simulation, rather a calculated quantity through the magnetic field, we need to connect the Hall term back to the magnetic field. This is done through Faraday's law as

$$\mathbf{E}_H = \frac{1}{ne}\mathbf{J} \times \mathbf{B} = -\mathbf{V}_H \times \mathbf{B} \tag{3.16}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} - \nabla \times \mathbf{E}_H, \tag{3.17}$$

where E_H is the electric field stemming from the Hall effect, and E consists of any other electrical fields. This is possible since the contributions from different electric fields are independent of each other, and the Hall effect simply appears as another addend. Note again the difference in a factor c from our previous Hall equations due to Bifrost's formulation. We also implement Courant criteria for the Whistler wave speed

V_{Hw} and Hall Drift wave speed V_{Hdw} [1] to ensure stability, as they are much faster than normal non-Hall speeds:

$$V_{\text{Hw}} = kV_A^2/\omega_i \tag{3.18}$$
$$V_{\text{Hdw}} = V_A^2/(L\omega_i) \tag{3.19}$$

Here k is the wave mode, V_A is the Alfvén speed, L is the characteristic length scale, and ω_i is the ion plasma frequency. As the we require stability for the fastest Whistler wave modes, we set $k = 2\pi/\delta x$ where δx is the spatial resolution of the numerical grid used. Whistler waves with higher wave numbers cannot be resolved, so are not considered for the timestep criteria either. The whistler waves are extremely fast. Compared to MHD wave speeds where the fastest are generally close to the Alfvén speed, whistler speeds are proportional to V_A^2. To satisfy the general Courant condition for these speeds would generally mean calculating several orders of magnitude more timesteps.

3.3 Validation

In order to validate our code and stress it as much as possible, we use a known analytical result for Hall MHD to compare with. Whistler waves are very far from standard MHD, so any weakness or flaw in the implentation of the Hall term should show up clearly in the propagation of the Whistler waves.

3.3.1 Whistler waves

We use a setup laid out by Huba [3] for detecting whistler waves. This is suitable for our purposes as the waves should come as a direct result of the extra Hall term, meaning we do not have to add any extra physics, and are easy to detect. The only requirement is that the simulation should be of length scales small enough and timescales short enough for the waves to be resolved.

First, we need to find an analytical expression for the wave dispersion to which we can compare our results, again following the method of Huba [3]. This is done by perturbing Faraday's law of induction, having inserted the Hall electric field. We assume a constant z-axis oriented background magnetic field $\mathbf{B} = B_0\hat{\mathbf{e}}_z$, and sinusoidal perturbations for x,y magnetic components $\delta B_x, \delta B_y \propto \exp(ik_z z - i\omega t)$:

$$\frac{\partial \delta \mathbf{B}}{\partial t} = -\frac{1}{ne}\nabla \times (\delta \mathbf{J} \times \mathbf{B}) \tag{3.20}$$
$$\omega \delta \mathbf{B} = \frac{1}{ne}(k_z B \delta J_x \hat{\mathbf{e}}_x + k_z B \delta J_y \hat{\mathbf{e}}_y) \tag{3.21}$$

Here (3.21) is the linearized version of the perturbed equation (3.20). We further expand this using Ampere's law, $\delta \mathbf{J} = c/(4\pi) \cdot \nabla \times \delta \mathbf{B}$, giving the connected equations:

$$\omega \delta B_x = -i \frac{c k_z^2 B}{4\pi n e} \delta B_y \tag{3.22}$$

$$\omega \delta B_y = i \frac{c k_z^2 B}{4\pi n e} \delta B_x \tag{3.23}$$

Which together are then solved to find the linear whistler wave dispersion:

$$\omega = \frac{k_z^2 B}{4\pi n e} = k_z V_A \left(\frac{k_z c}{\omega_{pi}} \right) \tag{3.24}$$

Note how the dispersion relation dependency on the ion plasma frequency again reiterates how the Hall effect is most relevant on small scales $L < c/\omega_{pi}$, where the phase velocity will exceed even the normal MHD Alfvén velocity.

For the numerical setup we assume a homogeneous density and a fixed spatial resolution. The magnetic perturbations in B_x and B_y are equal, but offset as sine and cosine waves respectively. Precise numbers for important quantities are listed as follows:

- Number density n, $n_0 = 10^{12}$ cm^{-3}.
- System size $L_z = 20$ cm, split into 120 points / 1D cells in z interval.
    * Size in x and y directions are irrelevant, for computational efficiency they are set to a single mesh point along each at $x = y = 0$.
- Plasma beta $\beta = 10^{-4}$.
- Background magnetic field $\mathbf{B} = B_0 \hat{\mathbf{e}}_z$, with strength $B_0 = 1000$ G.
- Perturbation size $\delta B = 10$ G, with x,y components

$$\delta B_x = \delta B \sin(2\pi m z / L_z) \tag{3.25}$$

$$\delta B_y = \delta B \cos(2\pi m z / L_z) \tag{3.26}$$

  for wave mode number m.

Finally the temperature is to be kept constant, though we are not able to fulfill this requirement as any resistive diffusion of the magnetic field is converted into heat in Bifrost. Since we hope that there will be very little diffusion in this test, we expect the effect to be noticeable but not severe as long as the wave is well resolved. The consequences of this are further discussed as part of our test results, though we should still expect to see numerical results which have correct relative sizes.

To compare our results with the analytical expression, we compute the frequency as function of wave mode for the numerical whistler waves. Tracking the perturbed variables B_x and B_y in time, we can perform a Fourier transform to extract the most prominent frequencies. The analytical dispersion relation (3.24) tells us which frequencies are to be expected for a given wave mode. Therefore we do multiple simulations for

different initial perturbation modes and see if the detected frequencies match this relation. We can clearly see that we should expect the frequency to go as the mode number squared, which holds up with the findings by Huba [3] as can be seen in Figure 3.3.

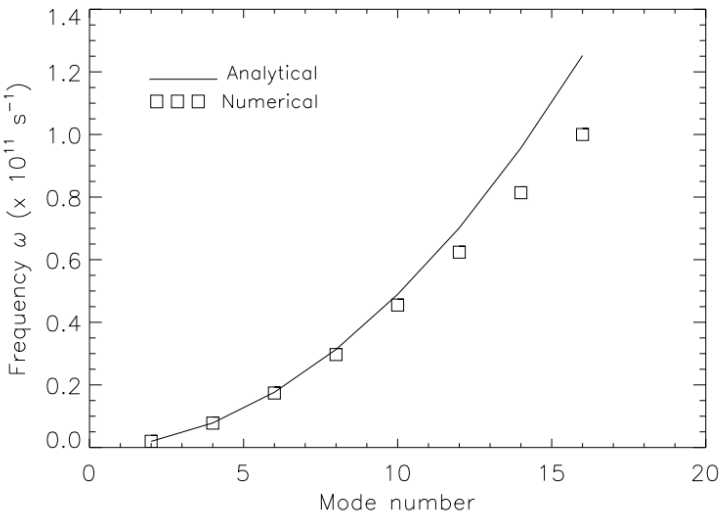

Figure 3.3: Comparison of analytical and simulated whistler wave frequency as function of mode number, as found by Huba [3].

After running a simulation with the setup of initial conditions described in section 3.3, we were able to detect oscillations in the magnetic field. We tracked the changes in magnetic field components B_x, B_y for all points in space over time, for all different mode number setups. An example of this can be seen in Figure 3.4, in which we have plotted the oscillations in B_y for some point in space as function of time. Note the slow increase in mean field strength with time. This is common for all points in space and for all wave modes and is likely due to an instability in the setup.

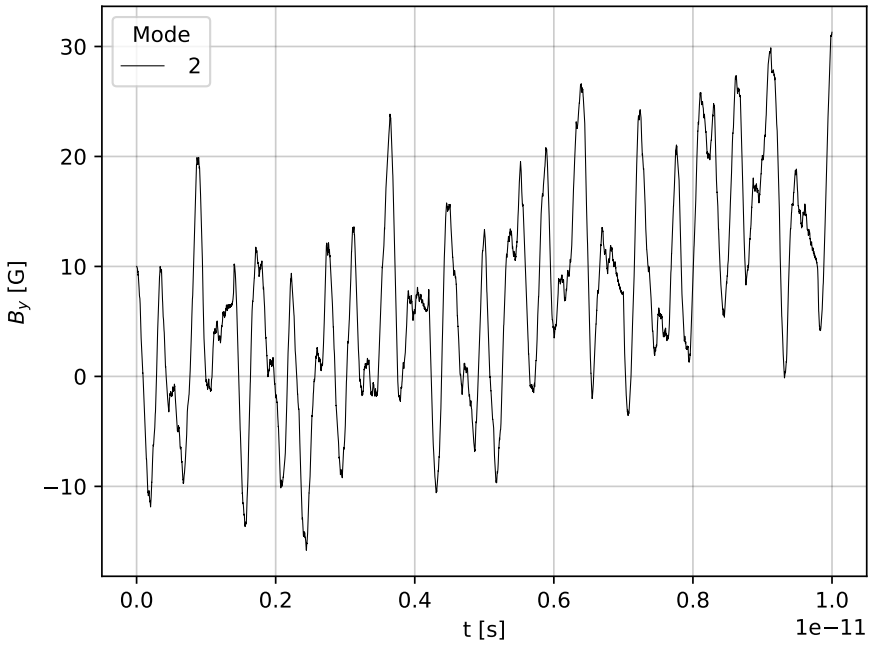

Figure 3.4: Plot of magnetic field component B_y for a point in space in the plasma, tracked over a time interval of 10^{-11} seconds. About 10^4 points in time are sampled, giving a resolution of 10^{-15} seconds between each point. A wave mode of 2 was chosen for this example as it has the clearest oscillation patterns over this interval.

The waves which we detected are a combination of various frequencies of oscillations, due to the many different overlapping MHD wave modes. In order to check if our whistler modes are part of these, we perform a discrete Fourier transform to break it down into frequency components. The result of this can be seen in Figure 3.5, where we have transformed the magnetic waves for a points in space for all mode numbers. Note the cutoff frequency at $f \approx 10^{12}$ where we only consider frequencies higher than this. This is because of the amount of lower frequency modes mixed in with the simulation. These are not resolved due to the short time interval sampled, where the discrete Fourier transform has a lower bound of $1/T = 10^{11}$ Hz. Likewise we have an upper bound of $N/T \approx 10^{15}$ Hz, as oscillations faster than this are not resolved.

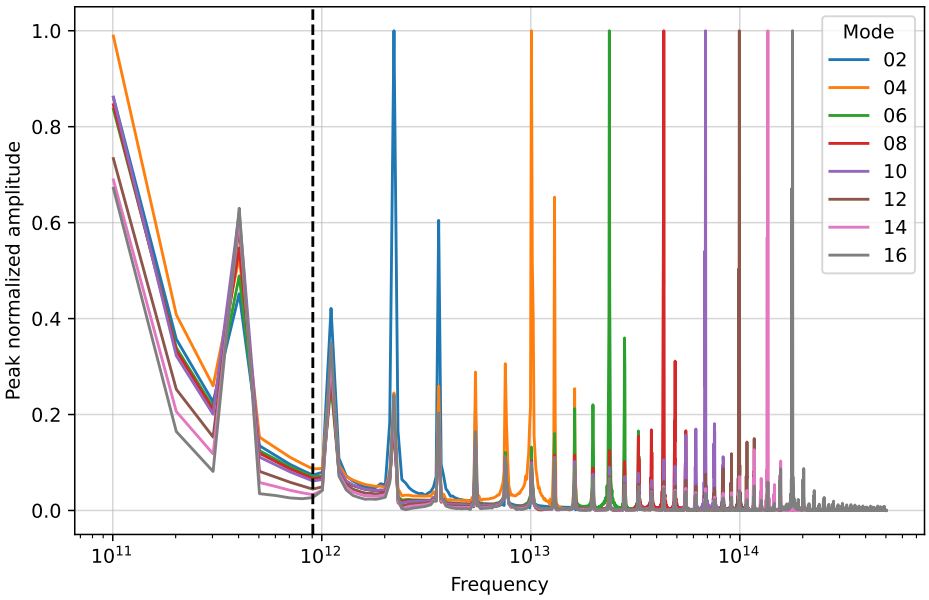

Figure 3.5: Plot of the Fourier transformed magnetic oscillations detected in the whistler wave experiment, for different wave modes. The vertical axis has been normalized by the maximum amplitude of each mode's frequency spectrum, in order to better compare the most prominent frequencies for each. Note the logarithmic frequency axis, as it spans five orders of magnitude. The vertical striped line at $f \approx 10^{12}$ Hz denotes our cutoff frequency, where we only consider frequencies larger than this.

These results correspond with the analytical expression, meaning we successfully managed to replicate the wave mode and frequency relation through our simulation. As can be seen in Figure 3.6 the detected whistler wave frequencies follow the same squared mode dependency as the analytical dispersion (3.24). Comparing our results to Huba's results shown in Figure 3.3, we can see that our frequencies are shifted up by a few orders of magnitude. However our analytical results back this up, meaning it is likely due to the deviations from the experiment description. The slight decrease in frequency as a function of mode number might be caused by our inability to keep the temperature constant. Our results still follow the expected quadratic growth, which seems to indicate they are still valid.

The analytical curve has had this extra constant multiplied to it, meaning Figure 3.6 does not completely represent our results. We have done this because we are confident this missing constant factor is not caused by a fault with our simulation, but rather the processing of the signal. We cannot accurately measure our error due to not having Huba's results in making Figure 3.3, but estimating the relative difference between the analytical and numerical results, our own curves differ by a factor of about 4π. This

is a very, very common factor in converting between SI and Gaussian/CGS systems of units, especially in an electromagnetic context. It is very likely that we have simply missed this in a step somewhere in our calculations, since all other parts of the results seem to be correct. Therefore we found it reasonable to reinstate this factor with the assumption that it should be there, though it is worth keeping this in mind when judging the accuracy of the simulation still since we have not been able to pinpoint the source.

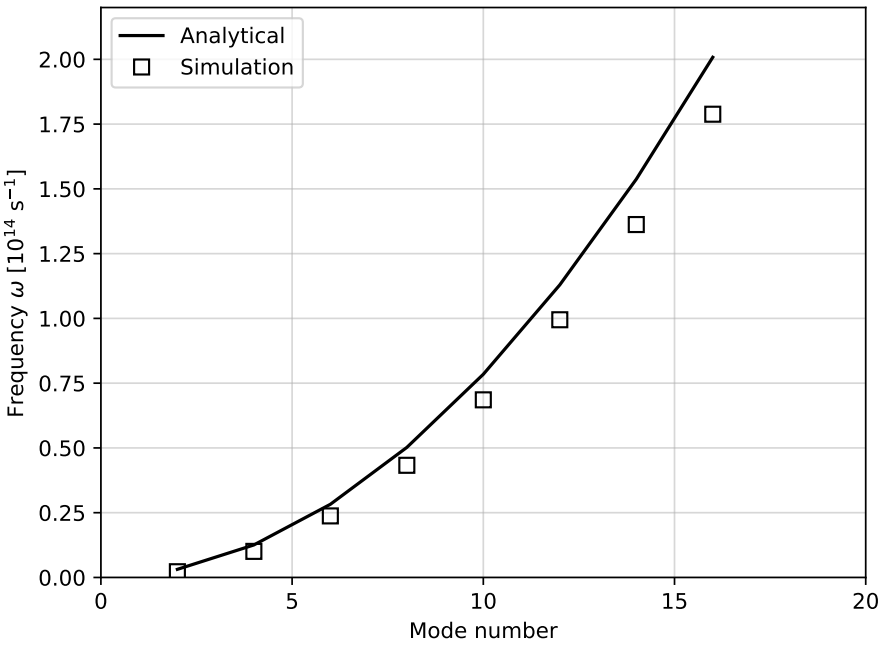

Figure 3.6: Plot comparing the analytical and simulated numerical results for whistler wave dispersion as function of wave mode. Numerical points are calculated for even mode numbers in range 2 to 16.

Despite some deviations, our results validate that our implementation of the Hall effect correctly results in whistler modes of correct ratio. Due to being able to replicate known analytical results, our Hall MHD simulation has reason to be trusted in other numerical experiments where we might not have an already known comparison. At least within reasonable use cases such as this thesis.

Chapter 4

Experiment and discussion

Implementing Hall-MHD into the solar atmosphere code Bifrost had the primary goal of investigating the Hall terms ability to speed up reconnection and possibly be the trigger for the fast energy releases observed in flares in the solar atmosphere. The solar atmosphere simulations are very complex and analysing such a simulation would be too time consuming. We instead wanted to investigate if a simple reconnection setup would actually show signs of a faster reconnection speed also when using Bifrost and the implementation used.

4.1 Magnetic Reconnection

To compare reconnection speeds we used a simple setup that would maek it possible to compare reconnection speeds both for Bifrost and another code but also between Bifrost with and without the Hall term implemented. We again use a setup provided by Huba [3], where we evolve a central magnetic island in time until we reach a reconnection state.

The experiment is situated in the 2D $x-y$ plane. As the Hall effect is only applicable on inertial scales, our unit of length is characterized by the speed of light divided by the plasma ion frequency c/ω_{pi}. The simulation area is a rectangle with side lengths $L_x = 25.6c/\omega_{pi}$ and $L_y = 12.3c/\omega_{pi}$ in x and y directions respectively. There are 100 grid points along x and 50 points along y. The z-axis is unimportant, especially since we have periodic boundary conditions for x and z. For y we have a zero-gradient boundary $\partial/\partial y = 0$.

The initial magnetic field is given by:

$$B_x(y) = B_0 \tanh(y/w), \text{ for } \omega = 0.5c/\omega_{pi} \tag{4.1}$$

Other magnetic components B_y, B_z are assumed to be zero. This magnetic field is then perturbed with a flux perturbation of the form:

$$\phi = -\phi_0 \cos(2\pi x/L_x)\cos(\pi y/L_y) \tag{4.2}$$

Where we set $\phi_0 = 0.1$. Huba does not specify the units of this flux perturbation however, and neither the magnetic field, so we will adjust this quantity later. It is also not specified how this field is applied as a perturbation. Therefore we have assumed it to be a vector potential where $\phi = \phi_z$. This means a perturbation to B_x and B_y would become:

$$\delta B_x = -\frac{\partial \phi}{\partial y} = -\phi_0 \frac{\pi}{L_y} \cos\left(\frac{2\pi x}{L_x}\right) \sin\left(\frac{\pi y}{L_y}\right) \tag{4.3}$$

$$\delta B_y = \frac{\partial \phi}{\partial x} = \phi_0 \frac{2\pi}{L_x} \sin\left(\frac{2\pi x}{L_x}\right) \cos\left(\frac{\pi y}{L_y}\right) \tag{4.4}$$

Where we then add these perturbation to the initial state of the system.

Other quantities are satisfied by balancing the magnetic and thermal pressure. We also have a characteristic time scale for the system through the Alfvèn time $\tau_A = (c/\omega_{pi})/V_{A0}$ where V_{A0} is the global Alfvèn velocity based on B_0. The inertial length c/ω_{pi} is defined by using the initial number density $n = n_0$, while the density itself has a lower floor of $n_b = n_0/5$. Finally the initial density n_0, temperature T, and magnetic field size B_0 is decided by balancing the pressures such that the magnetic beta equals one $\beta_0 = 8\pi n_0 T/B_0^2 = 1$.

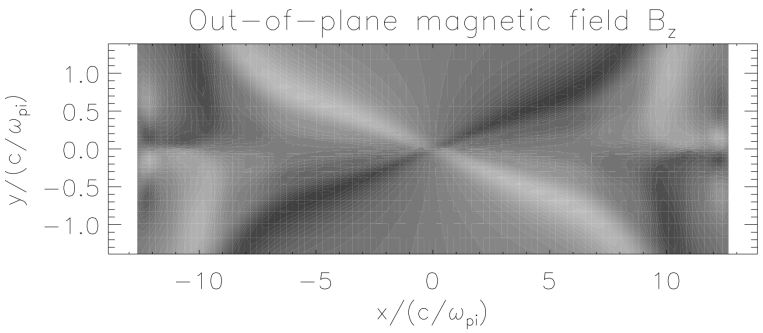

Figure 4.1: Magnetic field B_z component at time $t \approx 21\tau_A$, as shown by Huba [3]. Note the zoomed in view, where the y-axis only shows field in interval $\approx \pm 1.5$.

When comparing the expected results for Hall and non-Hall MHD in this experiment, the biggest difference the the magnetic field in the z direction. Without the Hall effect present there is no change in the magnetic field in this direction, and since we assume no initial field $B_{z0} = 0$ it would remain zero. Therefore this component is important to look at as it is a pure product of our added physics. In Figure 4.1 we have an image from Huba showing what the B_z magnetic component is expected to evolve into. This has happened after a time $t = 21\tau_A$, though we should not necessarily expect the exact same amount of time to pass due to our different simulations. We keep this as a goal for what our results should become.

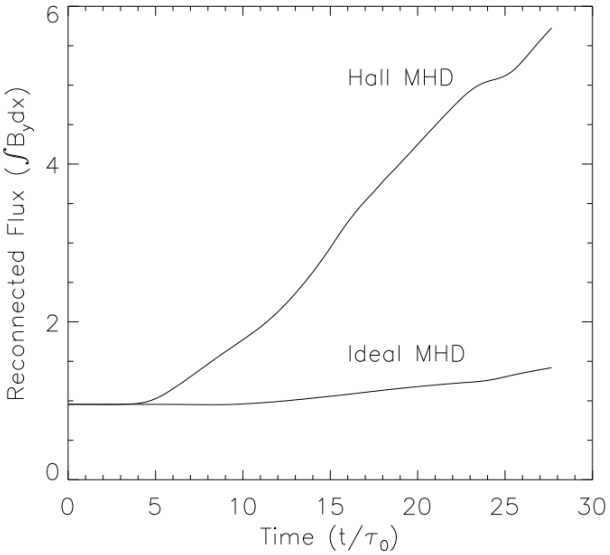

Figure 4.2: Comparison of reconnected flux as function of time between ideal and Hall MHD, as presented by Huba [3]. Note the near linear increase in flux for Hall MHD with time.

When choosing the density, temperature, and magnetic field strength, we based the quantities on solar corona-like quantities. A density of $n_0 = 3.06 \cdot 10^{10}\text{cm}^{-3}$ and field strength $B_0 = 10$ Gauss gives a temperature of $T = 10^6$ K, which is acceptable for the corona. Through testing we found the perturbation size of $\phi_0 = 0.1$ to be too small for any formation in reasonable time, so it is set instead to $\phi_0 = 1.0$. A plot of the initial magnetic field setup for B_x, B_y can be seen in Figure 4.3. We have no component in B_z for $t = 0$, and is therefore not included.

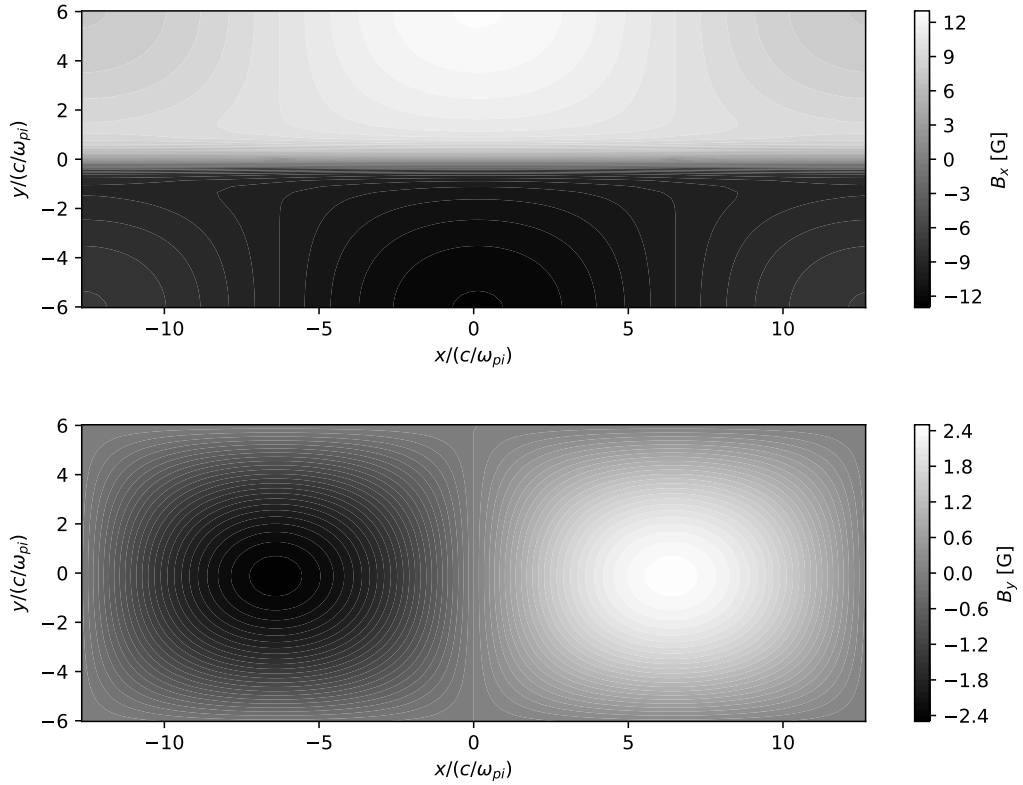

Figure 4.3: Initial state of the planar magnetic field. Top plot shows the B_x component, while the lower shows the B_y component. Note how B_x consists of both the background/initial field (4.1) as well as the perturbation (4.3), while the B_y field only stems from the perturbation (4.4).

In Figure 4.4 we show how the Hall magnetic field changes over time until reconnection. At time $t = \tau_A$ to $3\tau_A$ we get the initial induced field, which takes shape as four clear quadrants in an oblong shape. A neutral patch in the middle grows as the patches morph into sharper lines and connect to each other through the periodic boundary. Secondary fields appear inside the previously neutral patch as well as outside the surrounding lines for time $t = 4\tau_A$ to $6\tau_A$. Finally the central fields expand and reconnect at time $9\tau_A$.

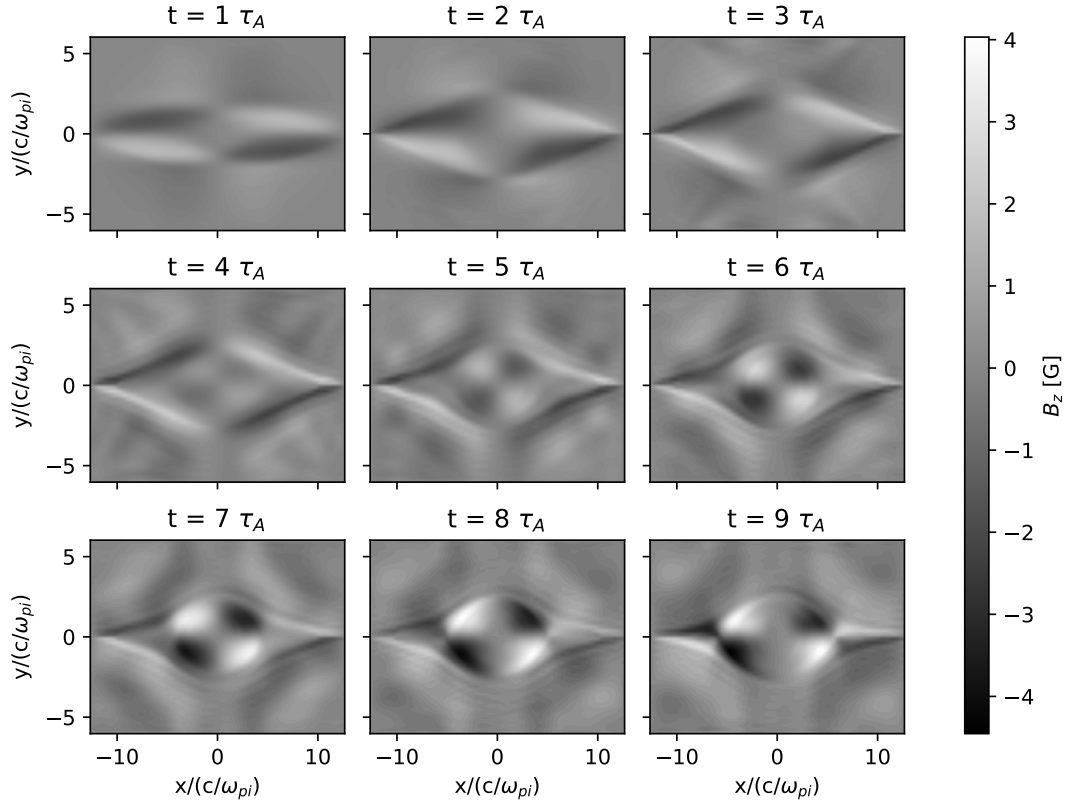

Figure 4.4: Plot showing the evolution of the Hall magnetic field as function of time for $t = 1\tau_A$ to $9\tau_A$. No plot is shown for $t = 0$, as there is no initial B_z field.

Looking closer at the final state, we can see resemblances to the results presented by Huba. We show a close-up view of the plot for $t = 9\tau_A$ in Figure 4.5, where we have cropped the y-axis to an interval $y \in [-1.5, 1.5]$. This is a similar state to the field from 4.1. While the central region does not connect as clearly in our plot, the left and right sections have the same field structure in terms of polarity, where we can see a checkerboard-like pattern. Note how our field connected much faster when compared with Huba. Without further testing we do not know for sure why this is, but a probable cause is our different choices for the initial conditions. As no values for the density, temperature, and magnetic field strength were provided, our choice might have impacted the reconnection rate. Especially the perturbation size ϕ_0 would likely affect the rate of propagation of the fields.

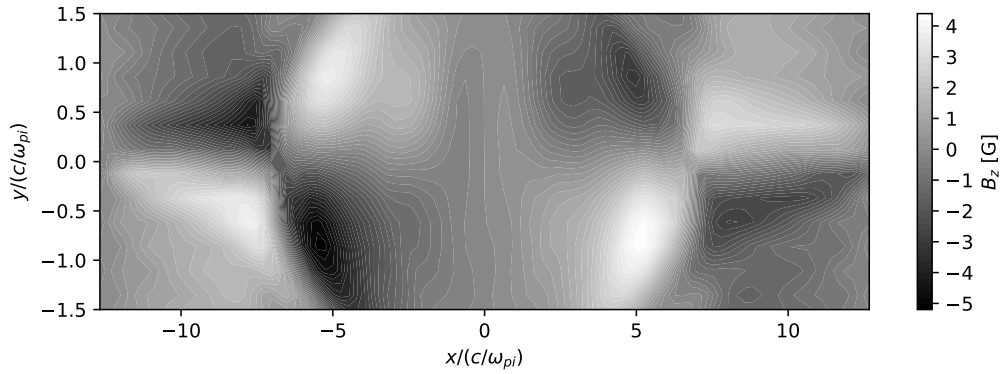

Figure 4.5: Close-up view of the B_z field at time $t = 9\tau_A$.

The initial magnetic field is almost purely in the x-direction except for the central pertubation. If reconnection takes place, the x-directed magnetic flux is converted to y-directed flux when passing through the central current sheet. In this case the x-boundary is periodic, which means that at most we will have a magnetic island forming in the current sheet. The y-directed flux is created by the creation of this magnetic island, so a measure of the reconnection rate is the amount of created magnetic flux in the y-direction per unit time. Huba shows this amount by integrating B_y and plotting it as a function of time in 4.2. Running a similar experiment should provide us with a similar result.

$$\Phi = \int B_y \mathrm{d}x \tag{4.5}$$

Where we do a line integral of B_y in x, getting the total change in B_y along this line. In Figure 4.2 Huba shows how the reconnected flux becomes much higher for Hall MHD compared with ideal MHD in their experiment. We also calculate this property in the same way.

Having calculated the reconnected flux for both with and without the Hall term, we clearly show that it has a major impact in this area. This is in line with what we saw presented by Huba. Comparing our plots in Figure 4.6 and Figure 4.2 we get a similar linear increase in the Hall flux, albeit more noisy, and a very similar comparative size between Hall and non-Hall flux at $t = 20\tau_A$ with each having roughly a $4:1$ ratio.

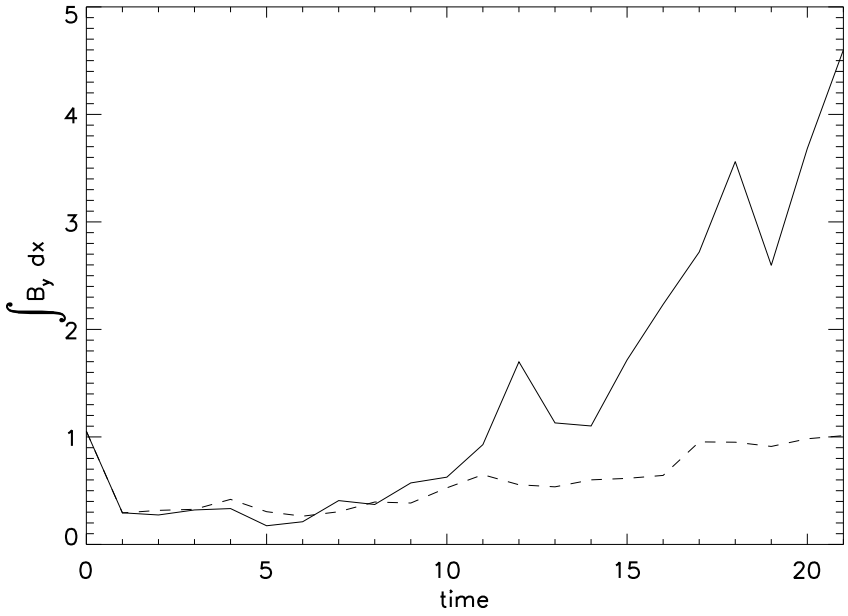

Figure 4.6: Plot of reconnected flux as a function of time for simulation including the Hall term (full) and excluding the Hall term (dashed). Time is in units of τ_A.

Chapter 5

Conclusion

We successfully managed to complete all our goals set out in this project. The implementation of the Hall effect into the Bifrost simulation was completed, taking into consideration all required steps of handling the fast whistler and hall drift modes. This was validated when we managed to detect whistler modes with the same relation as described by analytical results. Finally we performed a reconnection experiment and showed how the Hall effect led to a significant increase in reconnected flux.

Through this project I have learned a great amount of practical skills and information. I've learned about the Bifrost code, which gives large insight into high performance computation and how it might be applied, through using and modifying it myself. I've learned all about Hall MHD, both through its base MHD principles and equations, and also how small changes like this can lead to greatly different results. Through implementing and performing the whistler wave detection, I've learned to rigorously validate and test my results. And finally through following the tutorial set out by Huba [3] I've learned how to proceed with reproducing other people's work for comparative analysis.

Chapter 6

Future work

The implementation of Hall MHD into the Bifrost code makes it possible to test the effects of Hall MHD on top or possibly instead of the artificial numerical resistivity used in Bifrost. At present the artificial resistivity is based on energy conservation, minimal wave damping and code stabilisation. Most of the simulations that are performed with Bifrost code have a spatial resolution much larger than the inertial length of the ions, so in principle there should be no room for Hall MHD in the simulations typically performed with Bifrost. However, a number of researchers have mentioned Hall MHD as the answer to the catastrophic instability that leads to solar flares. The resistivity calculated from collision rates of particles cannot explain the fast and large energy release observed in solar flares, so some instability must happen inside the reconnection region. Hall MHD could be the answer. An instability should initiate on small scale and through an inverse cascade produce results on large scale. We could possible produce believable simulations without having the full range of scales all the way from the ion inertial scale to relevant solar atmosphere scales and in that way utilise the effects of Hall MHD in typical Bifrost simulations. A comprehensive investigation into the effects of Hall MHD with different resolutions in realistic atmosphere setups should be investigated to see how large an effect the Hall term actually has when the ion inertial range is not fully resolved, and what effect the tuning parameter v_H has. Such an investigation would be the obvious next step from a physical viewpoint.

The very small timestep required to resolve Whistler waves is a large concern. There are ways to save computation time while still resolving the small timescales required. To avoid using unnecessary extra computing power, it is possible to sub-cycle the Hall term. That is possible because there is such a large difference between the timestep required to resolve the Hall waves and the typical dynamical time of standard MHD. In reality that means that in a typical timestep required by the Hall waves, the rest of the terms in the PDEs are excedingly small. This allows us to take multiple steps in time for the Hall effect between each full update, giving a compromise between the efficiency of a larger time step and the detail of a smaller step.

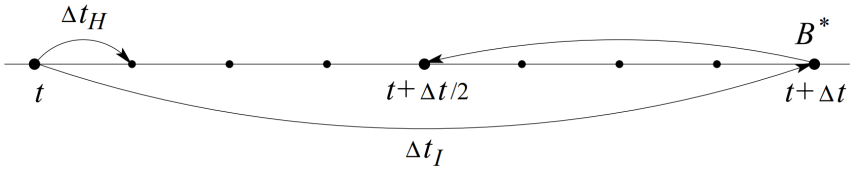

Figure 6.1: Schematic showing sub-cycling of the Hall term, from Huba [3].

While we could simply assume the slower variables to be constant as the Hall term is sub-cycled, we can gain some precision by using a more complex scheme. As with integration schemes there are many different varieties of sub-cycling schemes with differing orders of expected error, but a simple midpoint scheme detailed by Huba [3] would most likely be sufficient. The main idea is to calculate the non-Hall time step Δt and Hall time step Δt_H, then successively updating the magnetic field at Hall time until it reaches the normal step length at time $t + \Delta t$. Then this is averaged with the previous magnetic field from time t to get a midpoint value at time $t + \Delta t/2$. Finally a normal full step in time is taken using this averaged magnetic field from t to $t + \Delta t$. A schematic of this scheme can be seen in Figure 6.1.

Once these parameter studies of V_H and resolution scale and the increase in efficiency has been done, a full scale simulation should be undertaken. To compare the simulation with and without the Hall term employed could be interesting for several reasons. Both the total reconnected flux, the detailed and large scale topology of the field as well as the effect of the faster reconnection would be interesting to investigate.

Bibliography

[1] Thomas Gardiner. Algorithm and exploratory study of the Hall MHD Rayleigh-Taylor instability. Technical Report SAND2010-6747, 1008118, September 2010. URL: https://www.osti.gov/servlets/purl/1008118/, doi:10.2172/1008118.

[2] Gudiksen, B. V., Carlsson, M., Hansteen, V. H., Hayek, W., Leenaarts, J., and Martínez-Sykora, J. The stellar atmosphere simulation code bifrost - code description and validation. *A&A*, 531:A154, 2011. doi:10.1051/0004-6361/201116520.

[3] Joseph D. Huba. Hall Magnetohydrodynamics - A Tutorial. In Jörg Büchner, Manfred Scholer, and Christian T. Dum, editors, *Space Plasma Simulation*, volume 615, pages 166–192. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. Series Title: Lecture Notes in Physics. URL: http://link.springer.com/10.1007/3-540-36530-3_9, doi:10.1007/3-540-36530-3_9.

[4] J. M. Hyman. A method of lines approach to the numerical solution of conservation laws. In *Advances in Computer Methods for Partial Differential Equations - III*, pages 313–321, January 1979.

[5] Å Nordlund and Klaus Galsgaard. A 3D MHD code for Parallel Computers. *Technical manuscript*, pages 1–18, February 1995.

[6] A. Piel. *Plasma Physics: An Introduction to Laboratory, Space, and Fusion Plasmas*. Graduate Texts in Physics. Springer International Publishing, 2017. URL: https://books.google.no/books?id=9cA0DwAAQBAJ.

[7] B. V. Somov. *Plasma astrophysics*. Number v. 340 in Astrophysics and space science library. Springer, New York, NY, 2006. OCLC: ocm71782081.