

The nature of diffusive particle transport in turbulent magnetic fields

A Ivascenko
24790052

Thesis submitted for the degree *Philosophiae Doctor* in *Space Physics* at the Potchefstroom Campus of the North-West University

Promoter: Prof FA Spanier
Co-promoter: Prof M Bottcher

November 2016

Abstract

The transport of charged particles in turbulent plasma is a crucially important area of research in astrophysics, since it directly impacts our ability to interpret observations done with a major group of messenger particles. In order to understand these measurements of highly energetic charged particles – or cosmic rays – a comprehension of their interactions with the turbulent magnetic fields, which permeate the heliosphere, the interstellar and the intergalactic medium, is equally important as the understanding of their generation and acceleration at the sources, if not more so.

In this work, a numerical approach is taken to derive transport parameters for charged energetic particles in the heliosphere. A spectral incompressible MHD code is used to generate realistic turbulence in a self-consistent way. The properties of the turbulence are then probed by injecting test particles and analysing their propagation. New numerical analysis methods, that were developed to work especially well in strong turbulence scenarios, where classical methods and analytical solutions fail, are presented, together with their validation and transport parameter results obtained for various simulation setups. In most astrophysical scenarios the magnetic field fluctuations can be assumed to be much larger than the fluctuating electric fields $\delta\mathbf{B} \gg \delta\mathbf{E}$, consequently the predominant transport process is the change of the direction of the particle momentum relative to the magnetic field – or pitch angle μ – as opposed to the change of the absolute value of the momentum p , which is suppressed in comparison. Hence, the focus of the analysis is on the pitch angle diffusion coefficient $D_{\mu\mu}$. To demonstrate that the concept can also be applied to other quantities, results for the perpendicular spacial diffusion coefficient $D_{\perp}$ are derived and presented as well.

Additionally, an alternative method to generate turbulence in magnetised plasmas using Perlin gradient noise is described and its characteristics concerning particle transport are analysed and compared with the self-consistent MHD-turbulence in order to test its validity. Although the properties of the Perlin noise turbulence are not in complete agreement with MHD, the deviations can be neglected in specific cases (especially in strong turbulence) and are offset somewhat by the much lower computational effort.

Keywords

diffusion, heliosphere, magnetohydrodynamics (MHD), numerical simulation, particle transport, scattering, turbulence

Contents

Introduction	1
1 Phenomenology	3
1.1 The Sun	3
1.1.1 From core to surface	3
1.1.2 Solar atmosphere	4
1.1.3 Sunspots	5
1.1.4 Solar Wind	6
1.1.5 Interplanetary magnetic field	7
1.1.6 Heliospheric Turbulence	9
1.2 Solar energetic events	11
1.2.1 Solar Flares	11
1.2.2 Coronal mass ejections	13
1.2.3 Solar energetic particles	14
1.2.4 Particle acceleration in gradual SEP events	15
2 Theory	19
2.1 Plasma	19
2.1.1 Stochastic description of a plasma	20
2.1.2 Macroscopic variables	22
2.1.3 Macroscopic equations	24
2.1.4 Ideal MHD	25
2.1.5 Resistivity and incompressibility	27
2.1.6 Elsässer notation	29
2.2 Turbulence	33
2.2.1 Kolmogorov turbulence theory	35
2.2.2 MHD turbulence	37
2.3 Particle transport	45
2.3.1 Quasilinear theory	47
2.3.2 Kinetic wave-particle interaction	49
2.3.3 Wave generation with plasma instabilities	51
3 Methodology and Numerics	53
3.1 Plasma simulation	53
3.1.1 Pseudo-spectral description	53
3.1.2 Fourier transform implementation	55
3.1.3 Particle pusher	57

3.1.4	Code structure	58
3.2	Statistical analysis	64
3.2.1	Running diffusion coefficient	65
3.2.2	Kubo formalism	67
3.2.3	Diffusion equation approach	69
3.3	Perlin noise turbulence	73
4	Results	77
4.1	Method validation	77
4.1.1	Wave-particle interaction	77
4.1.2	Monte-Carlo validation	80
4.1.3	Running diffusion comparison	81
4.2	Pitch angle diffusion coefficients	88
4.2.1	Weak turbulence	88
4.2.2	Weak turbulence with higher resolution	91
4.2.3	Strong turbulence	93
4.2.4	Perlin noise results	97
4.3	Perpendicular diffusion coefficients	101
5	Summary	105
	Bibliography	107
	List of Publications	115
	Acknowledgements	117

Introduction

For many millennia man has looked up to the sky and wondered at what he saw. In recent history, the invention and perfection of the optical telescope showed us that the universe is much bigger and richer than we imagined, but it was only with the advent of multimessenger astronomy in the 20th century that we glimpsed the full extent of physical phenomena in the - suddenly much larger - universe, just waiting to be observed and understood. The start of that new era can arguably be dated with Victor Hess' discovery of cosmic rays in 1912. Hess' balloon experiments have shown an ionising radiation originating not from earth, but entering the atmosphere from above. In the 1930s first radio wave observations of an astronomical object were made by Karl Jansky. The 1940s laid the groundwork for X-ray astronomy, the 60s for gamma-ray astronomy as well as neutrino astronomy. Nowadays, countless detectors for all kinds of particles and all wavelength ranges of electromagnetic radiation are pointed outwards from the Earth, observing the multitude of phenomena in the universe.

This puts the astrophysicist in a rather unique position among his colleagues: instead of designing and building experiments himself he has to observe and interpret the countless configurations which the universe already produced on its own. This can be a mixed blessing since, on the one hand, we can study phenomena that are impossible to reproduce on Earth, but on the other hand, we can neither change the experiment nor our unique point of observation. So to be able to interpret the observations, beside the knowledge of the physical processes at the source, it is crucial to understand the transport processes taking place along the way between source and observer. Of the messenger particles mostly used in astronomy - photons, neutrinos, charged particles - the latter group is the one most prone to transport effects due to the electromagnetic interaction with the magnetic fields that permeate the entire universe. So, in order to understand phenomena, where highly energetic charged particles - or cosmic rays - are involved, a good theory of Galactic and heliospheric magnetic field configuration including turbulence is needed as well as a description of particle-field-interactions.

The importance of understanding particle propagation was demonstrated recently, when an anomalous excess in the leptonic cosmic ray spectrum was reported by the balloon-borne experiment ATIC in Chang et al (2008). The measurement was complemented by results from the PAMELA satellite claiming an unexpected rise in the positron spectrum above 10 GeV (Adriani et al, 2009, 2010) and later confirmed (at least qualitatively) in the lepton spectrum from the Fermi (Abdo et al, 2009; Ackermann et al, 2010) and AMS02 (Aguilar et al., 2014) satellites. This triggered a wave of explanations as to the source of those leptons, including nearby pulsars (Büsching et al., 2008; Grimani, 2009), supernova aftershocks, dark matter annihilation (Allahverdi et al., 2008) and decay among other, more exotic, sources. Unfortunately, most of the publi-

cations used rather simple transport models (if transport effects had been considered at all), so that, ultimately, the issue remains unresolved to this day, while some argue that due to propagation effects an exotic source of the excess leptons can not be distinguished from an astrophysical one.

Although this work focuses primarily on particle transport in the heliosphere, most of the principles and findings can be applied to interstellar and intergalactic charged particle propagation. The problem can be split up into two main issues: First, a comprehensive understanding of the magnetic field structure is needed. Besides the generation and properties of the large scale field in e.g. the heliosphere, this must also necessarily include a theory of turbulence in magnetised plasmas that is as complete as possible. Unfortunately, turbulence is one of the last unsolved problems of classical mechanics. Although there is a number of accepted theories, none of them describes all regimes of turbulence and the transitions between those regimes completely satisfactory. The second part of the problem is a description of particle interaction with magnetic field structures in plasma. This usually boils down to an analytic description of resonant wave-particle interactions in a linearised perturbative approach to plasma fluctuations or a statistical analysis of particle populations in numeric simulations. This work tries to combine both approaches to gain a more comprehensive picture of particle transport.

The structure of the work shall be briefly introduced at this point. The first chapter presents a short historic retrospection and the current state of the scientific knowledge concerning properties and phenomena of the Sun and the heliosphere. Instead of giving a broad and complete picture, we focus on details that are relevant to the characterisation and motivation of the specific numeric approach to the description of turbulent heliospheric plasma and particle transport therein taken in this work. Design decisions for the numerical magnetohydrodynamic (MHD) code GISMO, primarily used in this work, are motivated by observational and theoretical knowledge of the heliosphere. In the second chapter, the theoretical foundation for this work is presented. We discuss the basics of plasma theory and particle transport and present the current state of turbulence theories. The focus lies in the motivation and derivation of the Elsässer MHD formulation that is used in the spectral approach of GISMO. The particle transport part of the theory focuses on analytic cases that can be used to validate the numerical results. Chapter three has a twofold structure: first, the implementation details of the spectral incompressible MHD and particle code GISMO are presented and motivated. Secondly, the statistical analysis methods that were used in (and partially developed for) this work are discussed and analysed in detail. A short excursus at the end of the chapter describes an alternative turbulence simulation method. In chapter four, results obtained in this work are presented, prefaced by validation and test cases for the numerical approaches and analysis methods. The focus of the results section lies on particle transport parameters deemed to be the most important, namely the pitch angle diffusion coefficients $D_{\mu\mu}$ and the spacial perpendicular diffusion coefficient $D_{\perp}$. The final chapter summarises the findings, draws conclusions and points out areas of further research.

1 Phenomenology

This chapter provides a phenomenological description of the unique situation we find ourselves in at Earth's location in the solar system, including the delicate interplay of solar wind, turbulence, Earth's magnetic field and highly energetic particles from the Sun as well as from outside the heliosphere. After presenting a basic description of the heliosphere's properties, explaining the physical basics where necessary, it will outline the observational situation in the heliosphere that led to the modern level of understanding of its phenomena that arguably have a big influence on Earth and our daily lives.

In order to even mention all the phenomena that are observed on the Sun and in the heliosphere would go far beyond the scope of this work, so for the sake of brevity it will focus on topics relevant to particle transport in turbulent plasma, namely the generation and propagation of the solar wind, the nature of its turbulence, the origin and properties of high-energy particles encountered in the heliosphere and finally an overview of the space missions that helped deepen our knowledge of the solar system.

1.1 The Sun

Before coming to the ways the Sun is dominating the phenomena of the interplanetary medium, a brief overview of its structure and physical processes shall be given here.

1.1.1 From core to surface

The Sun is a fairly typical G-Type main sequence star. It has a mass of $m_{\odot} = 1.99 \cdot 10^{30}$ kg and a radius of $R_{\odot} = 6.96 \cdot 10^5$ km. At the core of the Sun, within about 25% of the solar radius from the centre, is where the proton-proton-fusion process takes place releasing about $3.846 \cdot 10^{26}$ W per second and sustaining a core temperature of about $1.5 \cdot 10^7$ K. The energy then propagates through the radiative zone, which extends from the core up to 70% of the solar radius, by radiative diffusion. Photons are constantly absorbed and re-emitted by hydrogen and helium ions and red-shifted to visible wavelengths in the process. The temperature drops by about two orders of magnitude over the extent of the radiative zone. Following next is the tachocline - a relatively thin transition layer between the uniformly rotating radiative zone and the differentially rotating convective zone. This leads to a very large shear stress tensor in the tachocline. It is hypothesized that the solar magnetic field is generated here by a dynamo effect. From $0.72 \cdot R_{\odot}$ up to the surface of the Sun the gas becomes cool and thin enough for convective heat transport to dominate over the radiative diffusion. This leads to extremely complex plasma motion with thermal cells heating up at the

tachocline, extending and travelling upward before cooling and dissolving at the surface, with the gas sinking back to the bottom of the convection zone, where the cycle starts anew. This leads to a highly granular structure at the surface of the convection zone with small and large granules representing Bénard cells close to the surface and reaching deep into the convection zone, respectively. Their temperature is about 100 K higher than that of the intergranular lanes. Above the convection zone the radiation reaches what is conventionally known as the surface of the Sun - the photosphere. Here practically all the energy transported from the Sun's core is absorbed and re-emitted as an almost perfect thermal black body spectrum with $T = 5777$ K. The temperature of the photosphere varies by a few thousand kelvin from top to bottom which leads to the phenomenon of “limb darkening” - the Sun disk appearing brighter in the centre than on the edge.

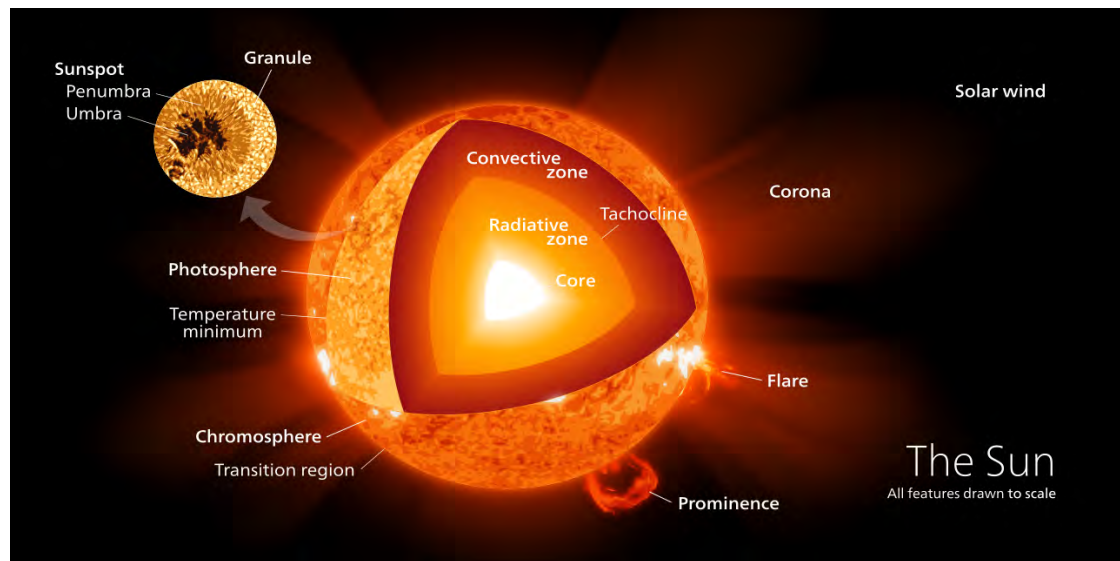


Fig. 1.1: A schematic drawing of the Sun's structure and the sites of selected solar phenomena. Source: Wikimedia Commons, Kelvinsong

1.1.2 Solar atmosphere

While the Sun's core provides the energy and lays the groundwork for the heliospheric phenomena, the really interesting phenomena happen in the atmosphere of the Sun. That the Sun has an atmosphere can be clearly seen during a total solar eclipse, with the chromosphere producing a colourful flash (which it derives its name from) at the beginning and end of the totality phase and the corona visible as a white halo during the period of totality. Early models assumed static, gravitationally stratified layers of chromosphere and corona. With time this picture changed to a much more dynamic and highly inhomogeneous mixing of the upper solar “layers” - the photosphere, chromosphere and

corona. Nevertheless, a distinction can be made by looking at the temperature in the different “layers”. Whereas the temperature in the photosphere falls towards the upper end (to a minimum of 4100 K), in the chromosphere it rises gradually up to $2 \cdot 10^5$ K. In the thin transition layer above the chromosphere there is another rapid rise in the temperature to 10^6 K, which corresponds to the lower end of the temperature range in the corona, the upper end being at about 10^7 K. This sharp increase in temperature is not completely understood yet, but is known to at least partially be caused by magnetic reconnection in the chromospheric and coronal plasma (Erdélyi and Ballai, 2007). Heating through Alfvén waves is another possible candidate (McIntosh et al., 2011).

1.1.3 Sunspots

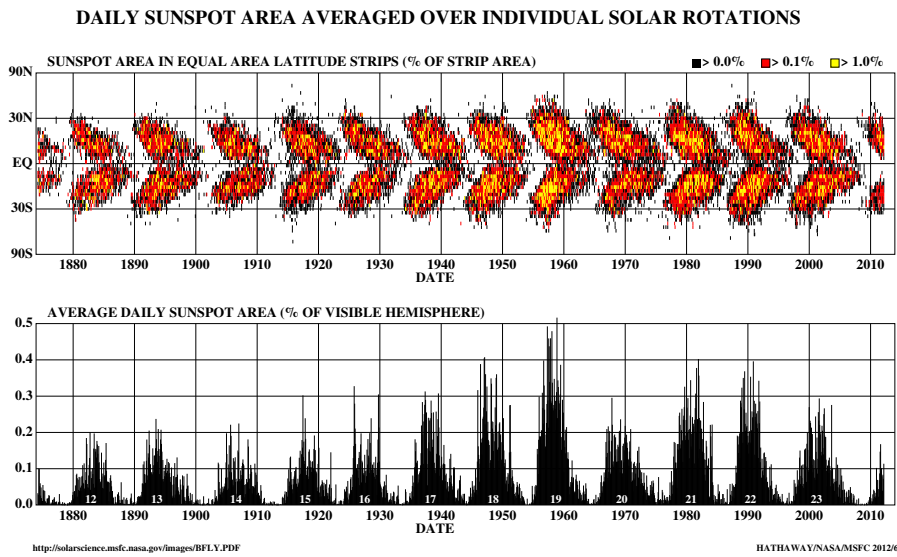


Fig. 1.2: Top: Butterfly diagram showing a sunspot pattern. Bottom: Fluctuation of the daily sunspot area over the 11-year Schwabe cycle. Source: NASA/ESA

First observations of the Sun as more than a static provider of sunlight can be dated back to the year 1610, when first telescopes were used by Thomas Herriot and Galileo Galilei to observe sunspots - transient darker areas on the Sun’s photosphere. More detailed observations of prominent sunspots and their periodic appearance led Johann Fabricius to the conclusion, that the Sun must be rotating. Samuel Schwabe later discovered the differential nature of the rotation and a periodicity in the generation of sunspots. At the beginning of the cycle sunspots appear at high solar latitudes ($\approx 30^\circ$). As the cycle progresses towards the maximum, more and more spots appear, their formation sites moving closer to the equator, resulting in the so-called “butterfly diagram” (Fig. 1.2). As the maximum is passed, sunspots begin to disappear until the minimum is reached and the 11-year cycle starts anew, with reversed magnetic polarity.

The origin of the sunspot cycle ultimately lies in the oscillation cycle of the Sun’s

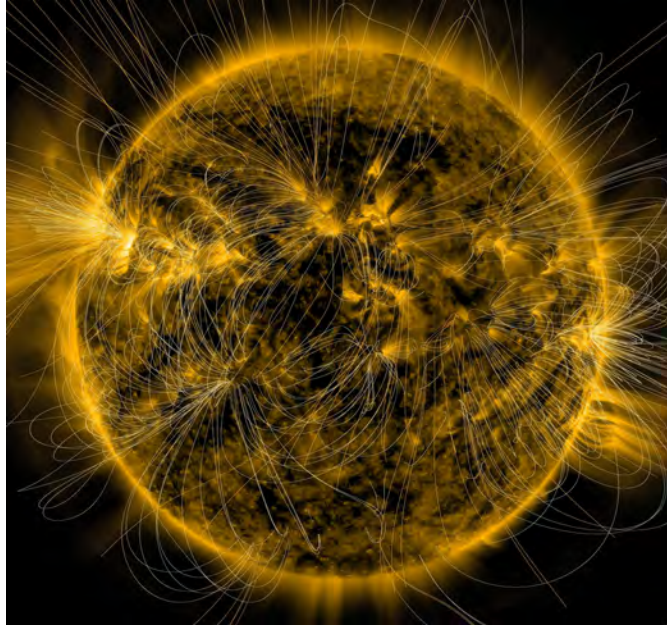


Fig. 1.3: A magnetic field line map of the Sun overlaid on a image in ultraviolet wavelengths. The map was created with the Potential Field Source Surface model. Source: NASA/SDO/AIA/LMSAL.

poloidal and toroidal magnetic field components. Figure 1.3 shows an impression of the solar magnetic field. At the maximum the poloidal dipolar field is at minimal strength, whereas the toroidal field generated through differential rotation in the tachocline is at its maximum. At this point plasma from the convection cells forces the toroidal field upwards through the photosphere and cools down, forming a pair of darker spots. This sometimes leads to the formation of a solar prominence - a thousand kilometres long loop-shaped plasma tube extending from the photosphere into the corona. In the declining phase of the solar cycle the energy shifts from the toroidal to the poloidal field and the sunspots start to disappear.

1.1.4 Solar Wind

A first hint of the wide-reaching influence of the solar dynamics came from observations of comets. Already in 1619 Kepler noticed that parts of comets' tails always pointed away from the Sun and had no curvature along the comet's trajectory, but had no explanation for the phenomenon at that time. From observations of solar flares and their connection to geomagnetic storms and auroral activity in the beginning of the 20th century it was suggested that there must be a constant quasi-neutral stream of charged particles flowing from the Sun. In the 1950s Biermann (1951, 1957) explained different deflections of the comet tail elements stating that the gas tail, composed mainly of CO^+ -Ions, always shows a strong uncurved deflection away from the Sun. He concluded this

to be consistent only with a continuous corpuscular outflow from the Sun unrelated to the prominent flares. Alfvén developed the idea further, stating that the flow must be magnetized plasma. The term “solar wind” was coined by Parker (1958), who noticed that the solar corona can’t be in a static equilibrium, but must either expand or collapse. His solution was the introduction of a supersonic solar wind, that was confirmed not one year later by in situ measurements made by the Soviet spacecraft Luna 3 and Venera 1.

Today we know from numerous measurements of satellite experiments like Wind, Ulysses, SOHO, SDO and Voyager (just to name a few), that there are in fact two main types - the slow and the fast solar wind - that differ in their characteristics, origin and variability.

The slow solar wind has an average velocity of 350 km/s, higher density, more variability and a more complex structure than the fast solar wind. Its composition is very similar to that of the solar corona. The exact origin and acceleration mechanism are still under investigation, but it appears to originate from smaller, less permanent structures in the “streamer belt” - mid latitudes around the solar equator - and seems to be correlated with the solar activity cycle. The main candidate for the source of the slow solar wind are helmet streamers: closed magnetic loops that connect regions of opposite polarity with electrons captured inside the field. They extend far into the corona, elongating to pointy tips with almost parallel magnetic field lines, and can be easily observed during a solar eclipse. Small blobs of plasma can be ejected from the flux tube at the tip, contributing to the slow solar wind flow.

The fast solar wind with an average velocity of 750 km/s and peak velocities above 10^5 km/s has a composition that closely matches that of the photosphere. It is thought to originate from coronal holes: areas of Sun’s corona with darker, colder, lower-density plasma. They form between strong magnetic field lines with the same polarity that exit from the Sun surface in near vicinity, for example when neighbouring sunspots have opposite magnetic polarity. This leads to open magnetic field lines along which particles can be accelerated and escape into space with relative ease. Observations in extreme UV/ soft X-ray show coronal holes forming predominantly at higher latitudes and near the poles during a solar minimum, which is also consistent with results of direct solar wind measurements by SWAN/SOHO (Bzowski et al., 2003).

1.1.5 Interplanetary magnetic field

The omnidirectionally streaming solar wind has a very important consequence for the interplanetary magnetic field. The solar wind plasma carries with it magnetic field lines radially outwards from the Sun (frozen-in flux). By the Sun’s intrinsic rotation the field is twisted into an Archimedean spiral, called the “Parker spiral” after Eugene Parker’s predictions of the solar wind and its associated phenomena in the 1950s (Parker, 1958). In the period of minimal solar activity the solar magnetic field is dominated by the poloidal dipole leading to two large polar holes with opposite polarities in the corona. The radial solar wind from these holes drags the frozen-in magnetic field outwards, such that near the equatorial plane a thin layer develops between the different magnetic polarities, where a current is induced according to Faraday’s law giving rise to the name

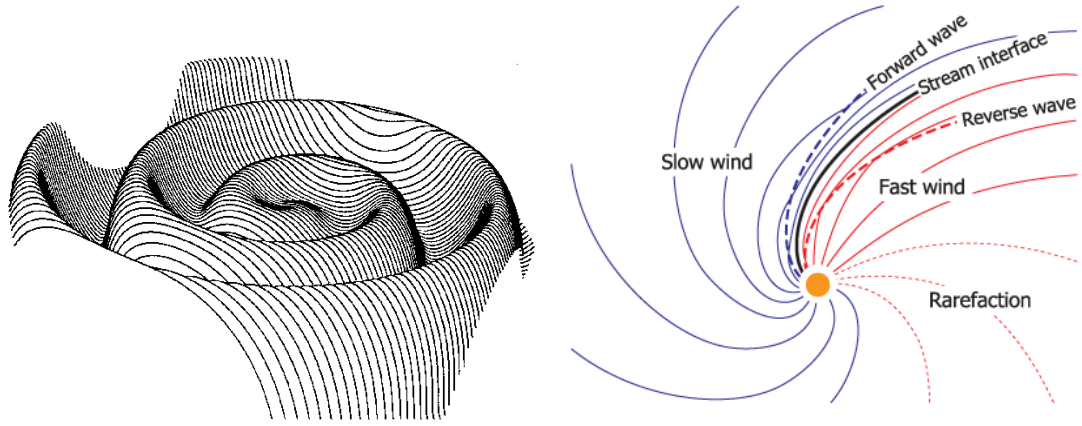


Fig. 1.4: Left: The “ballerina skirt” form of the current sheet layer separating magnetic fields of opposite polarity in the heliosphere (Jokipii and Thomas, 1981). Right: A sketch of the formation of a co-rotating interaction region. The fast solar wind (red) catches up to the slow solar wind (blue) and compresses the flow. A shock structure with a forward and reverse shock is formed.

“heliospheric current sheet” for this layer. The asymmetry of the coronal holes leads to the current sheet being not exactly in the equatorial plane, but either above or below it. As the Sun rotates, the sheet moves up and down, likening the whole structure to a ballerina’s skirt. The Earth can be above or below the skirt and depending on the direction, in which the magnetic field is pointing, is said to be in the “toward sector” or the “away sector”.

During a solar maximum the polar coronal holes become smaller, while small-scale opening and closing structures around the equator become more numerous. This makes the magnetic field structure much less regular and more variable. But even during the minimum it is far from regular.

As an example of complex interactions in the solar wind we take a look at so called co-rotating interaction regions or CIR. Since the slow solar wind component originates mainly from the equatorial regions and the fast component from the polar caps and since those regions rotate with different velocities due to the Sun’s differential rotation, the curvature of the fast wind spiral is smaller than that of the slow wind. Additionally the Parker spiral of the slow flow becomes wound more tightly further out from the Sun. This can lead to a situation where the fast solar wind “catches up” to the slow one and since both consist of ideal MHD plasma, they don’t easily mix. Instead the plasma of the slow wind is compressed and a boundary structure forms in the co-rotating interaction region, which can lead to plasma heating and stronger magnetic fields (see Fig. 1.4). A fully developed CIR has a forward and a reverse shock, that can serve as sites of efficient particle acceleration. A discovery by VOYAGER has shown that CIRs can, in turn, catch up to one another, interact and merge forming MIRs (merged interaction regions),

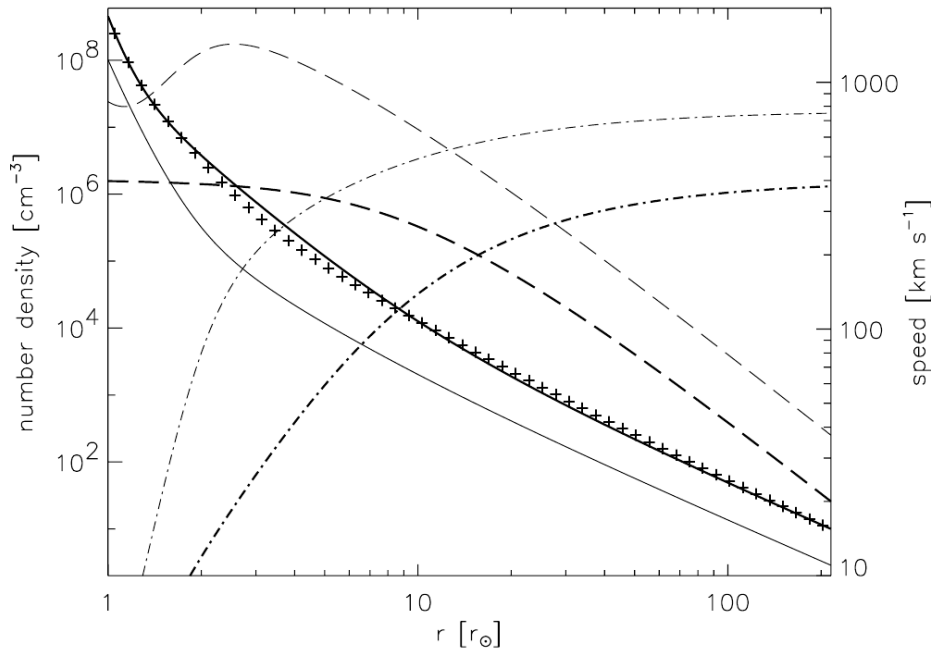


Fig. 1.5: A compilation of fast (thin lines) and slow (thick lines) solar wind properties: number density (solid lines), velocity (dot-dash lines) and Alfvén velocity (dashed lines). Crosses represent observational values. Source: Vainio et al. (2003)

which typically form beyond 10AU and persist to great distances in the heliosphere.

1.1.6 Heliospheric Turbulence

A very good numerical model of the solar wind's properties, especially the velocity profile depending on the distance to the Sun, can be found in Vainio et al. (2003). As can be seen in Fig. 1.5 the slow solar wind is significantly denser than the fast one. This leads to special consequences for the heliospheric turbulence, making it impossible to describe the interplanetary medium as a whole with one turbulence model. It is necessary to differentiate compressible (slow SW) and incompressible (fast SW) regions.

However, both solar wind components do exhibit turbulent behaviour due to the high velocities in a thin medium leading to high magnetic Reynolds numbers. Analogous to the Reynolds number in hydrodynamics this dimensionless quantity is defined as the ratio of inertial forces (induction) to viscous forces (diffusion):

$$R_m = \frac{v \cdot L}{\eta} = \frac{v \cdot L \cdot 4\pi\sigma}{c^2}, \quad (1.1)$$

with the characteristic length L , velocity v and the magnetic diffusivity $\eta = c^2/(4\pi\sigma)$. Reynolds numbers higher than a critical value indicate flows that are sensitive to per-

turbations, lose their laminar behaviour and start to produce eddies, vortices and other instabilities.

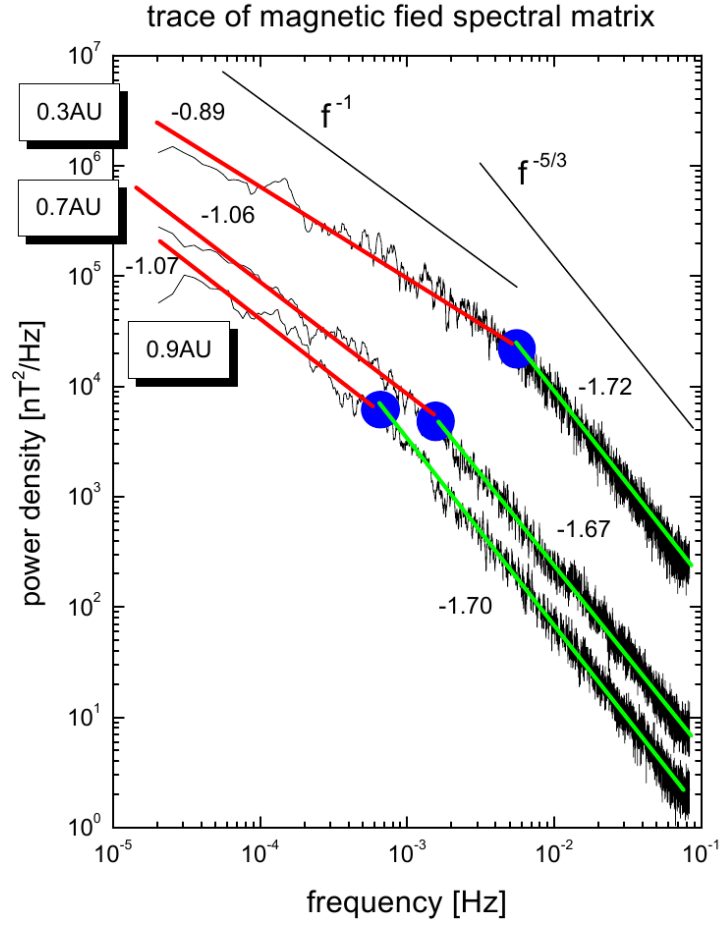


Fig. 1.6: Helios-2 measurements of the magnetic field fluctuations in the solar wind. The data follows a power-law with a spectral break. The inertial range with the spectral index ≈ -1.7 gets larger with growing distance from the Sun. Source: Bruno and Carbone (2005)

Typical values for the Reynolds number in the heliosphere are around 10^{14} (Borovsky and Funsten, 2002; Borovsky and Gary, 2008) which makes the solar wind plasma a highly turbulent medium. In the early 70s the Mariner probes were the first to measure magnetic field fluctuations in situ. The measurement of an energy spectrum of the turbulence was made a few years later by Helios-2. This kind of turbulent spectra follow a characteristic power-law of the form $E(\omega) \propto \omega^\alpha$, with the “spectral index” α , that forms through energy cascades. Fig 1.6 shows those typical spectra, with a spectral break - a change of the spectral index - marked by a blue dot. The measured spectra correspond to different positions of the Helios-2 probe on its way towards the Sun in

the same co-rotating plasma stream. As can be seen, the spectral break moves towards the higher frequencies as the measurement site moves closer to the Sun. This is in good agreement with current understanding of turbulence stating that a typical spectrum is driven by energy injection at low wave numbers, the energy then cascading through the inertial region towards smaller scales forming a steeper power-law in the process. Therefore closer to the Sun the injection range gets bigger and pushes the spectral break to the inertial range towards higher frequencies.

As mentioned before the complex nature of the interplanetary medium prevents a simple explanation with one turbulence model leading to a controversial discussion of all concurrent theories. A particular problem concerning the incompressible solar wind is that most measurements indicate a spectral index of about $\alpha \approx -1.6$ in the inertial range (Tu and Marsch, 1995): a value that lies in between the predictions of the turbulence theory of Kraichnan and Iroshnikov ($\alpha \approx -3/2$) and that of Kolmogorov, Goldreich and Sridhar ($\alpha \approx -5/3$). This problem is still not completely solved, nevertheless the assumption of incompressible turbulence in large parts of the heliosphere appears justified and indicates a dominance of the fast solar wind (Bruno et al., 1985; Tu et al., 1989). The primary wave mode in an incompressible, magnetized plasma is the Alfvén-wave: a transversal dispersionless mode propagating (anti-)parallel to the background magnetic field. Additionally, it has been shown that compressible plasma waves like the fast and slow magnetosonic modes are damped efficiently by Landau (Barnes, 1966) and viscous damping (Spanier, 2005).

1.2 Solar energetic events

Plasma turbulence is by far not the only dynamic and perturbative process in the heliosphere. Other solar phenomena that introduce disturbances to the “ideal” solar wind flow and make the interplanetary medium dynamic and variable shall be discussed in the following sections. A rather significant effect of these events is the introduction of highly energetic particles into the heliosphere.

1.2.1 Solar Flares

A solar flare is a huge release of electromagnetic radiation from the Sun’s upper layers. Even though most of the energy is released in the radio and X-ray range and at the $H\alpha$ wavelength in the optical range, some especially large flares produce enough intensity in the visible light spectrum to be observed with a filterless optical telescope. This led to their relatively early discovery by Richard Carrington in 1859. Discovering the intense $H\alpha$ emission made the flares much easier to observe using passband filters and led to a classification scheme based on the intensity of the $H\alpha$ line. With the dawn of radio astronomy in the 1940s and X-ray satellites ca. 30 years later, solar flare emissions could be observed over their full spectrum, revealing a quite complex emission structure. Whereas the $H\alpha$ emission has a smooth rise in intensity and lasts a few minutes, the emission in γ -rays, hard X-rays, EUV and microwaves exhibits an impulsive phase, characterized by a short bright flash, followed by the main/decay phase lasting from

30 minutes up to one hour. A precursor just before the flash exhibits all properties of thermal radiation at 10^7 K. Nowadays solar flares are monitored regularly by multiple geostationary X-ray satellites and are classified primarily according to the intensity of the X-ray flux on a logarithmic scale.

The general mechanism that causes a flare is rather well understood by now, the details however are still unclear and open to debate. A crucial point is the explanation how magnetic energy is released from an active region. For large scales of 10^5 km this can not happen by simple diffusion because of magnetic diffusion times of the order of hundreds of years. If the gradient scale length is compressed to below a few kilometres, the diffusion times become much more compatible to the observed variability of flares of the order of minutes, leading to the conclusion that thin current sheets, their formation and stability are at the heart of the issue.

The current understanding is that flares occur above coronal loops where the magnetic energy is released by explosive reconnection. Soft X-rays are radiated by the heated loop itself, while hard X-rays and γ -rays are produced by the highly energetic non-thermal particles. Magnetic reconnection is a process of reordering the field line configuration between regions of opposite polarity and as such needs very specific configurations of plasma flows and current sheets to occur. There are multiple theories that try to explain the formation of those structures. A “unified” flare model proposed by Shibata et al. (1995) assumes a twisted flux rope overarching a coronal loop, stretching the field between the two. This leads to a stronger plasma flow towards the reconnection region, compressing the current sheet and effectively boosting the reconnection rate by lowering the diffusion timescale. The energy released in the reconnection heats the plasma and accelerates the particles. In this model the flux rope tears off a plasmoid - a coherent magnetic structure - from the top of the loop accelerates it upwards from the Sun. This would make the release of a Coronal Mass Ejection (CME, see next chapter) an essential part of the model.

Other models (Shiota et al., 2005) assume that the reconnection takes place above solar arcades - a series of magnetic field line loops occurring close to each other. Again, the reconnection is aided by a large flux rope above the arcade that “tears off” the tops of the loops releasing the heated plasma as an eruptive prominence and forming a lower set of arcades. The reconnection energy drives the flare.

It should be mentioned that the microphysics of magnetic reconnection are not completely understood. Especially the geometry of the reconnection region, that is crucial to the onset of reconnection, can be much more complicated in the solar structures than in the quasi two dimensional models. Works like Petschek (1964) have addressed some problems concerning influx timescales and the stretching of the diffusion region, yet other questions remain open, e.g. reconnection in collisionless plasmas, where MHD is not applicable.

On the other hand, there is a large number of different magnetic structures and instabilities in the solar corona, so that a unified flare model might very well be impossible. Furthermore, flares vary widely in sizes (released energy $10^{21} - 10^{25}$ J), even if one excludes microflares (10^{19} J) and nanoflares (10^{16} J), and associated phenomena (e.g. not every flare is accompanied by a CME and vice versa).

1.2.2 Coronal mass ejections

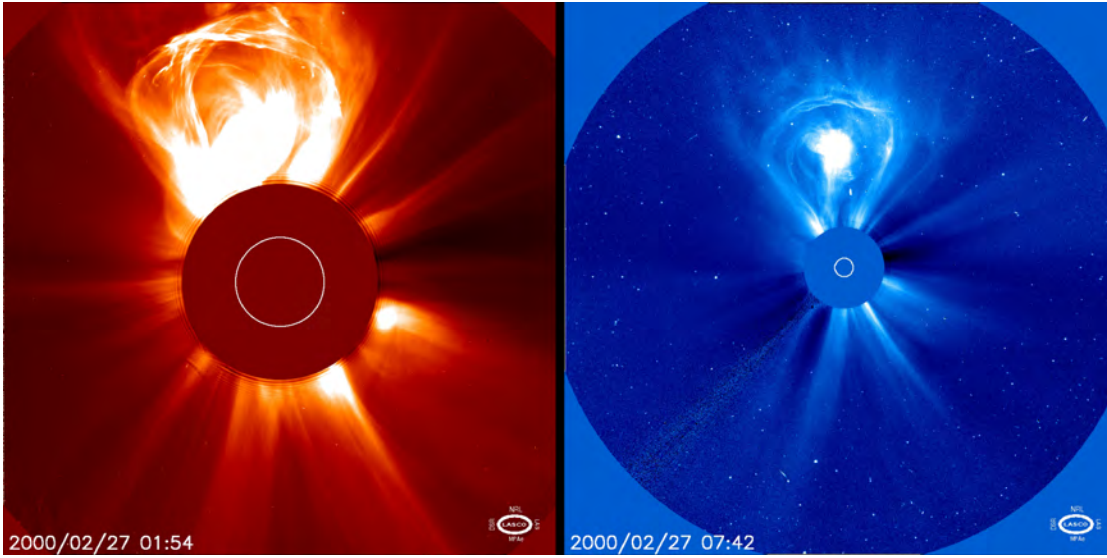


Fig. 1.7: Pictures of a coronal mass ejection taken by the LASCO C2 and C3 instruments on the SOHO spacecraft. The huge light bulb shaped structure carries about 10^{16} g of particles into the heliosphere. Sunlight scattered by the central plasmoid of the CME is clearly visible. Source: NASA/ESA.

Coronal mass ejections (CMEs) are a further type of solar activity that might look very similar to flares at first glance. The typical energy released in a CME is on the order of $10^{24} - 10^{25}$ J, so energy-wise they are comparable to the stronger flares. The crucial difference is that while the flare energy is released as broad-spectrum radiation into a wide angle, CMEs themselves do not radiate. This makes them much harder to observe and explains their relatively late discovery in the early 1970s, the first clear detection being made on December 14, 1971 by Tousey (1973) with the white-light coronagraph on the OSO-7 satellite. A coronagraph occults the Sun disk basically producing an artificial solar eclipse. What is then seen of the CME are in fact visible-light photons from the Sun that are thompson-scattered on the electrons in the ejected plasma cloud. The brightness of the scattered light is proportional to the electron density but not to the temperature, so that the density structure of a CME can be inferred from a coronagraph picture. A typical example of such a picture is shown in Fig. 1.7. The extent of the light bulb shaped CME structure much larger than the Sun (white ring in the middle) can be clearly seen. The bright patch indicates higher-density plasma of the central plasmoid.

CMEs occur with an average rate of about one per day during a solar minimum and 4-6 per day during the maximum, so they are not a seldom phenomenon and are correlated with the sunspot cycle like flares. However, their latitudinal development is different from the sunspots, that appear at mid-latitudes after a minimum and then move towards the equator. The evolution of the CME source region shows an inverted behaviour: in a

minimum they appear in a narrow band around the equator, the band then widens with increasing solar activity and the CME structure becomes more complicated, reflecting the more complicated magnetic field structure.

CMEs are a mechanism to get rid of excess magnetic flux and helicity produced in the tachocline. While the continuous solar wind carries away magnetic flux from regions of open field lines, the ejection of toroidal closed-field structures in CMEs carries away surplus closed flux. For that to happen, the closed-field structure has to be torn off from the Sun, leading to the conclusion that reconnection has to play an important role in the launching of CMEs as it does with flares. A systematic analysis by Munro et al. (1979) has shown that about 40% of CMEs correlate with the appearance of flares, while about 70% appear together with solar prominences. The current understanding is that CMEs are followed by flares, rather than the other way around, and that high-velocity CMEs appear together with H α -flares (Kahler et al., 1989).

A class of CMEs that plays a special role for the space weather at Earth are the so-called Halo CMEs that propagate towards the Earth and consequently appear as a white-light halo in coronagraph pictures. When observed in-situ at 1 AU they are usually called interplanetary CMEs (ICMEs). Since strong ICMEs lead to magnetic storms that can have dramatic impact on electronic equipment at Earth, the forecast of ICMEs is crucial. This includes associating ICMEs with coronal CME observations and estimating the time until arrival at 1 AU. The usual travel time is about 80h (Brueckner et al., 1998), although some CMEs appearing around solar maxima are significantly faster. More accurate models are needed to understand the formation and propagation of CMEs in more detail.

1.2.3 Solar energetic particles

As already mentioned, in both flares and CMEs particles are accelerated to very high energies. On Earth solar particles with energies up to 1 GeV have been detected. Such energies are reached by protons and heavier ions, while electrons are accelerated to keV - 100 MeV regions. Of course, not all particles accelerated at the Sun escape into the heliosphere and become observable at Earth, some produce X-rays, γ -rays and radio waves in their interaction with solar plasma. A first observation of these solar energetic particles (SEP) was made by Forbush (1946). A clear distinction to the ever-present solar wind can be made both by the high energy and much lower density. The typical flux of SEPs caused by a flare is some 5 orders of magnitude lower than the typical solar wind flux density of $5 \cdot 10^{12} \text{ m}^{-2} \text{ s}^{-1}$. The typical flux of galactic cosmic rays at 1 AU is again about 5 orders of magnitude lower.

Measurements of SEPs show two distinct types: impulsive and gradual events with strongly diverging characteristics according to Lang (2009).

Impulsive events are observed over relatively short times up to a few hours, as the name suggests. They are associated with impulsive (minute-long) X-ray flares and Type III and IV radio bursts. Observations with coronagraphs show no correlation with CMEs suggesting that the particles are accelerated in flares or on flare-amplified wave modes. Impulsive SEPs are the more frequent kind with about 1000 events per year

in a solar maximum. Their particle composition is quite peculiar with a large fraction of electrons, $H/He \approx 10$, $Fe/O \approx 1$ and ${}^3He/{}^4He \approx 1$. The equal abundance of ${}^3He$ and ${}^4He$ is especially interesting because only about 0.05% of all Helium in the solar atmosphere is ${}^3He$, indicating that the acceleration mechanism in impulsive SEPs must be very efficient for ${}^3He$. The similar time-dependent intensity profiles for ${}^3He$ and ${}^4He$ lead to the conclusion that the same mechanism may at least partially also accelerate ${}^4He$ ions. A possible candidate would be gyro-resonant wave-particle interaction with Alfvén waves that propagate at frequencies below the proton gyro frequency, as the gyro frequencies of ${}^3He$ and ${}^4He$ are $2/3$ and $1/2$ of the proton gyro frequency, respectively. The acceleration sites of impulsive SEPs seem to be relatively close to the Sun since they are predominantly observed at eastern longitudes. Since charged particles follow the background magnetic field of the Parker spiral, events that are detected near Earth as coming from the east are mostly particles that underwent no significant scattering or acceleration during their propagation.

Gradual SEPs can be observed over much longer periods of several days and are associated with gradual (hour-long) soft X-ray flares and Type II and IV radio bursts. Coronagraph observations show a very strong connection to CMEs, with a CME observed in 96% of all gradual events. This would suggest that particles in gradual SEPs are accelerated at the shock front that the ICME drives in front of itself. The fact that gradual SEPs have no preferred arrival direction with a longitudinal extent of $\sim 180^\circ$ also reinforces the assumption of shock acceleration, since interplanetary shocks disturb the structure of the Parker spiral and can effectively deflect particles from their normal propagation. The composition of gradual events is proton-rich with $H/He \approx 100$, $Fe/O \approx 0.1$ and the normal ${}^3He$ fraction from the solar atmosphere of ${}^3He/{}^4He \approx 0.0005$. They are less frequent with ≈ 100 events per (solar maximum) year and usually accelerate particles to higher energies than impulsive events. This could be due to the longer duration of the acceleration process, although it is more likely that some sort of pre-acceleration happens in the solar corona for particles to reach energies of hundreds of MeVs. The details of shock acceleration in gradual SEP events will be discussed in the following section.

1.2.4 Particle acceleration in gradual SEP events

The long observation time, the longitude invariance and the missing correlation with flares indicate that particles are accelerated at shocks in gradual SEP events. Efficient shock acceleration requires the shock to travel faster than the average solar wind speed; shock speeds above $V_s \approx 7.5 \cdot 10^7$ cm/s (fast solar wind speed) always lead to SEPs (Reames et al., 1997). The size of the CME, on the other hand, seems to be irrelevant.

The mechanism describing the acceleration of particles that are deflected at magnetic mirrors or field irregularities was first proposed by Fermi (1949). It was later refined to describe particle acceleration at shock fronts (Axford et al., 1977; Bell, 1978; Krymskii, 1977). Contrary to Fermi's original proposal, where the energy gain is quadratic in $\beta = V_A/c$ (earning it the name "Fermi II process"), the shock acceleration (or Fermi I) process is much more efficient, due to its linear dependence on $\beta = V_s/c$ and the fact that

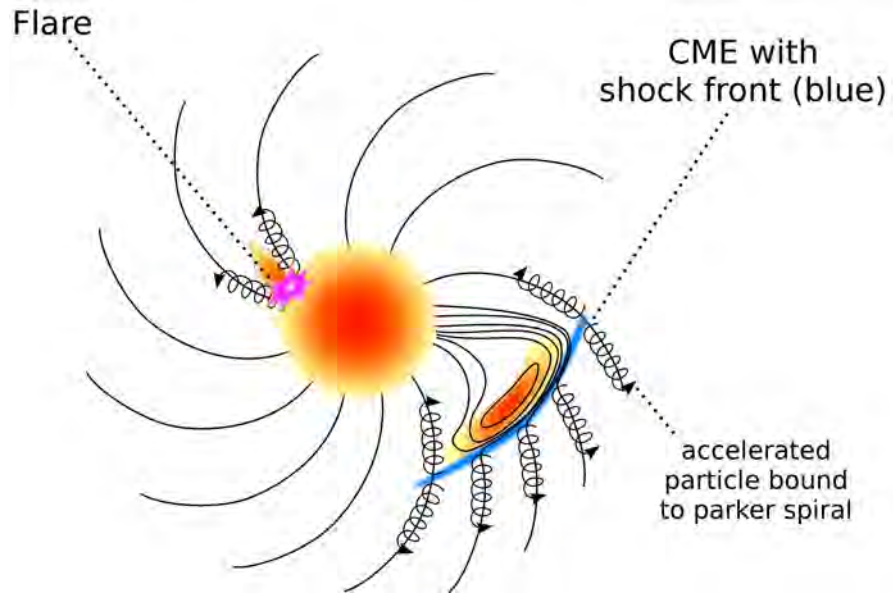


Fig. 1.8: A schematic representation of possible particle acceleration sites in SEP-events. Through frozen-in flux the Parker spiral structure of the magnetic field is strongly modified by a propagating CME. The resulting shock front can accelerate particles to very high energies. Source: Lange (2012)

shock speeds V_s can be much faster than relative Alfvén velocities V_A . The crucial point of Fermi I is that the particles cross the shock multiple times gaining energy with every shock crossing, as shown in Fig 1.9. In the upstream (downstream) reference frame the particle sees the bulk of the plasma from the downstream (upstream) moving towards it with a relative velocity V_p . Each scattering process is elastic in the corresponding reference frame, so the momentum p of the particle does not change, but the change in the particle direction (or more exactly in the pitch angle μ between particle velocity and the magnetic field) leads to a energy gain when the particle crosses the shock and the transformation to the reference frame on the other side is made. Effectively, the particle gains energy in elastic head-on collisions with every shock crossing until it reaches very high energies and escapes the shock region entirely.

For CME shocks there are two relevant mechanisms that cause particles to cross the shock front, and understanding them and the parts of the shock where they occur is crucial for the interpretation of gradual SEP measurements. In regions of the shock further away from the driving CME the magnetic field of the Parker spiral is disturbed only minimally, so that the magnetic field lines are parallel to the shock normal. This is the case at western longitudes (see Fig. 1.8). Here particles follow the field lines and are scattered by magnetic irregularities back and forth across the shock as described earlier. Reflecting the nature of the particle motion the process is called “diffusive shock acceleration”. Near the centre of the shock and at eastern longitudes the magnetic field

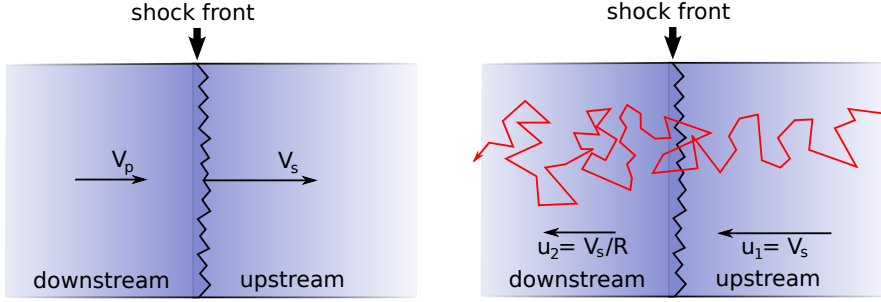


Fig. 1.9: A schematic depiction of the Fermi I acceleration process in the lab frame (left) and in the reference frame of the shock (right) with a particle diffusing from upstream, crossing the shock multiple times and escaping into the downstream.

structure is strongly influenced by the CME to the extent that the magnetic field lines are almost perpendicular to the shock normal. Since charged particles are gyrating around the field lines, they naturally cross the shock front from upstream to downstream and back again with each gyration. The particle crossings are not stochastic as with diffusive shock acceleration, but “forced” by the gyration, leading to faster and stronger particle acceleration. This process is known as “Shock-Drift-Acceleration” (see e.g. Sonnerup, 1969), as particles gyrating around field lines with inhomogeneities feel an additional force perpendicular to the magnetic field, which leads to a drift motion. Measurements with satellites in different parts of the CME structure confirm these distinct acceleration processes. As seen in Fig. 1.10, measurements at western longitudes (Helios 1) show a steep jump in particle intensity followed by a gradual decrease over the observed time. The passage of the shock front is seen as a slight increase in intensity. Measurements in the central region (Helios 2) show a nearly constant, high intensity of SEPs with a gradual decrease after the transition of the shock front. At eastern longitudes (IMP 8) the gradual intensity increase continues some time after the shock passes the satellite, followed by a gradual decrease.

These spectra can be explained as follows: The Parker spiral is almost undisturbed by the CME in the western range. Diffusive shock acceleration produces highly energetic particles that escape quickly from the shock region, leading to the steep intensity increase when the shock has formed, but is still far away from the satellite. After that the number of escaping accelerated particles is nearly constant until the shock passage, at which point it starts to decrease. The measurement near the centre of the shock shows a signature of shock-drift-acceleration with a relatively constant flux of high energy particles that doesn’t change significantly even after the shock transition. This is caused by particles gyrating behind the shock and particles caught in the closed field lines of the CME. The intensity profile in the eastern part can be explained by particles that are scattered back into the downstream from the shock front. These particles propagate from the centre along the bent magnetic field lines towards the western end of the shock, causing an intensity maximum long after the shock front has passed the satellite.

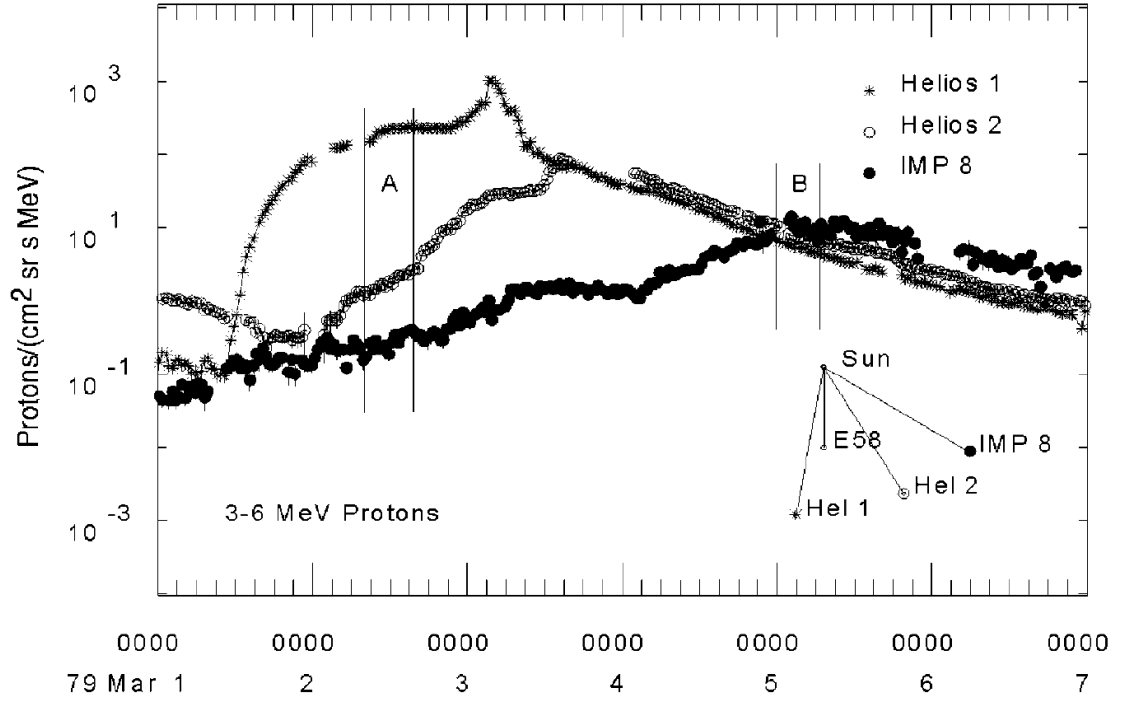


Fig. 1.10: Multi-spacecraft measurements of the intensity profile of a SEP event by Helios 1, Helios 2 and IMP8 positioned from the west to the east end of the shock front, accordingly. Spatial invariance can be seen in region B, as opposed to different spectra in region A. Source: (Reames, 1999)

A further interesting phenomenon that can be seen in Fig 1.10 is the spatial invariance. Spectra from different shock regions coincide at a certain point (Regime B), after starting out as completely different (Regime A). As shown by Reames et al. (1997) this can be explained with particles that are trapped in the magnetic field lines of the CME still connected to the Sun, that form a magnetic bottle due to the divergent magnetic field.

2 Theory

As already exhaustively discussed in the previous chapter, the heliosphere is formed by different kinds of plasma outflows from the sun, with the background magnetic field of the Parker spiral and a sprinkle of highly energetic particles. Section 2.1 of this chapter strives to give a definition of the plasma state and to introduce formal mathematical descriptions that are useful in trying to comprehend the physical properties of and phenomena in a plasma. This leads to the Elsässer plasma description that is used in the numerical model of this work. In section 2.2 the concept of turbulence is introduced and current theories of heliospheric turbulence are presented, especially with regard to turbulence conditions used in the model to study particle transport. The theoretical base for that will be discussed in section 2.3, introducing analytical approximative results as a base for comparison with results obtained numerically in this work.

2.1 Plasma

Plasma is a quasi-neutral medium with unbound charged particles exhibiting collective effects rather than interacting just with the closest neighbour. Plasma is sometimes called the fourth state of matter, because it behaves differently from solids, liquids and gases and can be produced in a natural way by heating a gas to a sufficiently high temperature (about $10^5 - 10^6$ K). However, unlike the phase transitions between the other three states, the transition from gas to plasma is a gradual process showing no abrupt change in properties. In particular, there is no ionisation threshold at which the gas starts to behave like a plasma. An ionisation degree of 0.1% is already sufficient for the gas to show properties of a plasma, at 1% ionisation the conductivity is almost perfect.

Because of the high conductivity there are no large-scale electric fields in a plasma, since they would always be compensated by particle motion and rearrangement. On small scales r particles with a charge q have the electrostatic Coulomb potential

$$\phi = \frac{q}{4\pi\epsilon_0 r} \quad (2.1)$$

(with the vacuum permittivity ϵ_0), that is shielded by neighbouring particles of opposite charge. The measure, how far those electrostatic effects persist, is expressed as an important plasma quantity, the Debye length:

$$\lambda_D = \sqrt{\sum_{\alpha} \frac{\epsilon_0 k_B T_{\alpha}}{n_{\alpha} q_{\alpha}^2}}, \quad (2.2)$$

with the number density n_α and temperature T_α of the particle population and the Boltzmann constant k_B (the derivation of the Debye length assumes a Boltzmann equilibrium of the particle population). Intuitively, the Debye length is the radius of the sphere inside of which electric fields must be taken into account and outside of which the thermal speed of the particles is sufficiently high to escape the Coulomb potential. For plasma behaviour, the sphere must contain enough particles so that the collective properties dominate, i.e.

$$\frac{4\pi}{3}n_0\lambda_D^3 \gg 1. \quad (2.3)$$

This quantity, alongside with $\Lambda = n_0\lambda_D^3$ and its inverse $g = \Lambda^{-1}$ (Boyd and Sanderson, 2003), is called the “plasma parameter”. The other condition is that the volume $V = L^3$ of the plasma must be larger than that of the Debye sphere for it to be quasi-neutral. This constrains the Debye length in a plasma to

$$\frac{1}{\sqrt[3]{n_0}} \ll \lambda_D \ll L. \quad (2.4)$$

Another characteristic quantity of a plasma is the “plasma frequency”. If a plasma is perturbed from the equilibrium, it will start to oscillate with the Coulomb force acting as a restoring force. In the approximation of cold, free electrons ($T_e = 0$) and fixed (or infinitely heavy) background ions the electron plasma frequency is:

$$\omega_{pe} = \sqrt{\frac{n_e e^2}{m_e \epsilon_0}}. \quad (2.5)$$

The ion plasma frequency ω_{pi} is defined in an analogous way with m_e substituted by the ion mass and is a much smaller quantity due to the inverse proportionality to the square root of the mass. The characteristic length scales associated with the plasma frequency are the electron inertial length or skin depth c/ω_{pe} and the ion inertial length c/ω_{pi} . These quantities give the attenuation or damping length scales for fluctuations near the electron and ion plasma frequencies, respectively.

2.1.1 Stochastic description of a plasma

There are multiple ways to derive a mathematical description of a plasma that lead to the same result. The Klimontovich equation is based on the consideration of a single particle i at location $\mathbf{R}_i(t)$ with velocity $\mathbf{V}_i(t)$ in six-dimensional phase space $(\mathbf{r}, \mathbf{v})$, whose trajectory can be expressed as a time-dependent density distribution:

$$N_i(\mathbf{r}, \mathbf{v}, t) = \delta(\mathbf{r} - \mathbf{R}_i(t))\delta(\mathbf{v} - \mathbf{V}_i(t)), \quad (2.6)$$

with Dirac’s delta function δ . A sum over all particles yields the density function for the species $N(\mathbf{r}, \mathbf{v}, t) = \sum_i N_i(\mathbf{r}, \mathbf{v}, t)$. With the equations of motion for each particle in an

electromagnetic field

$$\dot{\mathbf{R}}_i(t) = \mathbf{V}_i(t) \quad (2.7)$$

$$\dot{\mathbf{V}}_i(t) = \frac{q}{m} (\mathbf{E}(\mathbf{R}_i, t) + \mathbf{V}_i \times \mathbf{B}(\mathbf{R}_i, t)) \quad (2.8)$$

the time derivative of the distribution function can be evaluated and yields the Klimontovich equation:

$$\frac{\partial N}{\partial t} + \mathbf{v} \cdot \frac{\partial N}{\partial \mathbf{r}} + \frac{q}{m} (\mathbf{E} + \mathbf{r} \times \mathbf{B}) \cdot \frac{\partial N}{\partial \mathbf{v}} = 0. \quad (2.9)$$

This equation contains the information about all particle trajectories in the form of a sum over δ -functions and is next to impossible to evaluate either analytically or numerically for any relevant astrophysical scenario. Since we are rarely interested in the individual particle orbits, we can take an ensemble average $f(\mathbf{r}, \mathbf{v}, t) = \langle N(\mathbf{r}, \mathbf{v}, t) \rangle$ to get the collisionless Vlasov equation describing the time evolution of the averaged phase-space density f :

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \quad (2.10)$$

A collision term $\partial_t f|_{coll}$ can be added to the right hand side to get the Boltzmann equation with the Lorentz force taking the role of the external force term. The correct treatment of the collision term is not a simple task, since inter-particle collisions can be of variable nature, ranging from elastic to those transferring energy to neutral particles or leading to recombination, ionization and charge exchange. A simple way to approximate the collision term is the Krook model with a constant average collision frequency ν_c and $\partial_t f|_{coll} = -\nu_c(f - f_0)$ with a stable solution f_0 of the Vlasov equation. Fortunately, in most heliospheric and interstellar plasmas collisions can be neglected completely when compared to the effect of the collective interactions.

A different approach leading to the same result employs Liouville's theorem: First we define a probability density function $f(\mathbf{r}, \mathbf{v}, t)$ that expresses the probability of finding a particle in the volume element $d^3r d^3v$ of the six-dimensional phase space $(\mathbf{r}, \mathbf{v})$ at the time t . This is also referred to as the phase space density. Taking its total time derivative and using Liouville's theorem stating that the distribution function is constant along any trajectory in phase space yields:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{r}} \cdot \frac{\partial f}{\partial \mathbf{r}} + \dot{\mathbf{v}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \quad (2.11)$$

The acceleration $\dot{\mathbf{v}}$ is given by the Lorentz force analogous to Eq. 2.8, with two components of the electromagnetic field. One is generated by the collective motion of all particles in the plasma, the other describes the short-range Coulomb potentials of the neighbouring particles located within collision distance. Again, for a large plasma parameter $\Lambda = n_0 \lambda_D^3$ the mean distance between the particles is large and the probability to find a particle within colliding distance is very small. In this case the collision force

term can be neglected and we again get the Vlasov equation 2.10.

The development of the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ is given by Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \frac{\sigma}{\epsilon_0} \quad (2.12)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.13)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.14)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \quad (2.15)$$

with the charge density σ and current density $\mathbf{j}$ as defined in the next section. Together with the Vlasov equation these form the so-called Vlasov-Maxwell system of equations, that contains the complete phenomenology of collisionless plasmas and is the common basis of kinetic plasma physics. To recapitulate: the only assumptions used in the derivation of the equations are the representability of a plasma as a continuous phases space density and the negligibility of collisions for the dynamics of the system.

The Vlasov equation is a non-linear 6-dimensional, hyperbolic partial differential equation which makes finding solutions quite difficult. No general analytical solution is known, only for some linearised special cases closed solutions can be obtained. There are multiple approaches to treat the problem numerically. A direct solution of the Vlasov equation with finite volume methods is possible, however achieving a satisfactory resolution in the 6-dimensional phase space requires large storage and computational resources. Particle in Cell (PiC) methods represent the electromagnetic fields on a numerical grid and calculate the Lorentz force on not grid-bound (macro)-particles, thus lowering the computational requirements at the cost of a stronger dependence on particle statistics. The approach chosen in this work is based on the transition from the kinetic description of a plasma focused on microphysics to a macroscopic fluid description. In the frame of statistical physics one can derive stochastic moments of the Vlasov equation to get the Magneto-Hydro-Dynamic (MHD) system of equations.

2.1.2 Macroscopic variables

The spacial coordinate $\mathbf{r}$ does not designate the position of a particle $\mathbf{r}_i$ any longer, since we don't know where the i -th particle is located after taking the ensemble average. Instead we can define quantities that characterise the entire particle population of the species α by calculating statistic moments of the phase space density $f_\alpha(\mathbf{r}, \mathbf{v}, t)$. The normalisation of f_α is chosen so that the integral over the whole phase space yields the number of all particles of the species:

$$\int f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3x d^3v = N_\alpha. \quad (2.16)$$

The n -th moment of a function $f(v)$ of a variable v is defined as

$$\mu_n = \int_{-\infty}^{\infty} v^n f(v) dv. \quad (2.17)$$

The zeroth order velocity moment of the distribution function yields the particle number density:

$$n_\alpha(\mathbf{r}, t) = \int_{-\infty}^{\infty} f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3v. \quad (2.18)$$

Multiplication with the charge q_α gives the charge density of the species:

$$\rho_\alpha(\mathbf{r}, t) = q_\alpha n_\alpha(\mathbf{r}, t). \quad (2.19)$$

The first order velocity moment is the particle flux:

$$\mathbf{\Gamma}_\alpha(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{v} f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (2.20)$$

multiplying with the charge we get the current density

$$\mathbf{j}_\alpha(\mathbf{r}, t) = q_\alpha \mathbf{\Gamma}_\alpha(\mathbf{r}, t), \quad (2.21)$$

or dividing by the particle density the average velocity

$$\mathbf{u}_\alpha(\mathbf{r}, t) = n_\alpha^{-1} \mathbf{\Gamma}_\alpha(\mathbf{r}, t). \quad (2.22)$$

The second order gives the pressure tensor

$$\mathbf{\Pi}_\alpha(\mathbf{r}, t) = m_\alpha \int_{-\infty}^{\infty} (\mathbf{v} - \mathbf{u}_\alpha) \otimes (\mathbf{v} - \mathbf{u}_\alpha) f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (2.23)$$

which reduces to a scalar pressure that can be used to introduce a temperature definition in a symmetric case:

$$P_\alpha(\mathbf{r}, t) = \frac{m_\alpha}{3} \int_{-\infty}^{\infty} (\mathbf{v} - \mathbf{u}_\alpha)^2 f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3v = n_\alpha k_B T_\alpha. \quad (2.24)$$

For a Maxwellian distribution the temperature definition corresponds to classical thermodynamics. For particle distributions in collisionless plasmas, that can be very different from Maxwellian, temperature is a non-trivial concept.

The scalar particle pressure is used to define the “plasma beta”

$$\beta = \frac{2\mu_0 \sum_\alpha P_\alpha}{B^2}, \quad (2.25)$$

a relation between the particle energy density and the magnetic energy density that

is used to characterise the dominant component of a plasma. If $\beta > 1$, the magnetic field evolution is governed by the collective particle flux. If $\beta \ll 1$, the magnetic field dominates the plasma dynamics.

2.1.3 Macroscopic equations

Analogous to the derivation of the macroscopic variables, one can derive equations that describe the macrophysical behaviour of the plasma by calculating moments of the Vlasov equation 2.10. To get a more complete picture we include the collision term in the derivation, effectively starting from the Boltzmann equation for particle species α :

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \frac{\partial f_\alpha}{\partial \mathbf{r}} + \frac{q_\alpha}{m_\alpha} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} = \left. \frac{\partial f_\alpha}{\partial t} \right|_{coll}. \quad (2.26)$$

The collision effects can be dropped at the macroscopic level when not needed.

The zeroth moment of the Boltzmann equation is obtained by integrating it over velocity space. With the use of Eq. 2.18 the first term yields the temporal change of the particle number density $\partial_t n_\alpha$. The second term is of the first order in $\mathbf{v}$, consequently resulting in the divergence of the particle flux (Eq. 2.20): $\nabla \cdot (n_\alpha \mathbf{u}_\alpha)$. The integral over the force term vanishes for physical particle distributions, since the phase space density has to approach zero for $\mathbf{v} \rightarrow \pm\infty$. Integration of the right hand side results in a source term Q_α that is also zero, if there are no ionizing or recombining collisions nor charge-exchange collisions between ions and neutrals, in short: if the particle number of the charged population does not change. Putting it all together we get the continuity equation:

$$\frac{\partial n_\alpha}{\partial t} + \nabla \cdot (n_\alpha \mathbf{u}_\alpha) = Q_\alpha. \quad (2.27)$$

Multiplying it with q_α and m_α yields the continuity equations for charge and mass densities, respectively. It should be noted that the zeroth order moment equation is connected to the next higher order via the particle flux $n_\alpha \mathbf{u}_\alpha$.

Multiplying Eq. 2.26 with $m_\alpha \mathbf{v}$ and integrating over $\mathbf{v}$ yields:

$$m_\alpha n_\alpha \frac{\partial \mathbf{u}_\alpha}{\partial t} + m_\alpha \int (\mathbf{v} \otimes \mathbf{v}) \cdot \nabla f_\alpha d^3v + q_\alpha \int \mathbf{v} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_\alpha}{\partial \mathbf{v}} d^3v = m_\alpha \int \mathbf{v} \left. \frac{\partial f_\alpha}{\partial t} \right|_{coll} d^3v. \quad (2.28)$$

The second term can be evaluated using Eq. 2.22 and 2.23 resulting in the pressure tensor. The integral over the Lorentz force is solved using integration by parts, effectively substituting the velocity $\mathbf{v}$ with the average particle velocity $\mathbf{u}_\alpha$. The collision term integral vanishes for collisions between the same type of particles. Collisions between different plasma populations transport momentum, therefore resulting in a non-zero

contribution that can for example be approximated with the help of the Krook model:

$$m_\alpha \int \mathbf{v} \frac{\partial f_\alpha}{\partial t} \Big|_{coll} d^3v = - \sum_\beta m_\alpha n_\alpha (\mathbf{u}_\alpha - \mathbf{u}_\beta) \langle \nu_{\alpha\beta} \rangle, \quad (2.29)$$

with the average collision rate $\langle \nu_{\alpha\beta} \rangle$ between particles of the population α and β .

Collecting all terms we get the macroscopic equation of motion or momentum transport equation, since it determines the rate of momentum change per unit volume:

$$m_\alpha n_\alpha (\partial_t + \mathbf{u}_\alpha \cdot \nabla) \mathbf{u}_\alpha + \nabla \cdot \mathbf{\Pi}_\alpha - q_\alpha n_\alpha (\mathbf{E} + \mathbf{u}_\alpha \times \mathbf{B}) = m_\alpha \int \mathbf{v} \frac{\partial f_\alpha}{\partial t} \Big|_{coll} d^3v. \quad (2.30)$$

Again, this first order moment equation involves the second order quantity, namely the pressure tensor $\mathbf{\Pi}$. It is easy to see that every n -th order will involve an $n+1$ order term in $\mathbf{v}$ because of the $\mathbf{v} \cdot \nabla f_\alpha$ term in the Boltzmann/Vlasov equation. Consequently, the second moment yields the heat transport equation or energy conservation law:

$$\frac{3}{2} n_\alpha k_B \left(\frac{\partial T_\alpha}{\partial t} + \mathbf{u}_\alpha \cdot \nabla T_\alpha \right) + P_\alpha \nabla \cdot \mathbf{u}_\alpha + \nabla \cdot \mathbf{H}_\alpha + (\mathbf{\Pi}'_\alpha \cdot \nabla) \cdot \mathbf{u}_\alpha = - \frac{\partial}{\partial t} \left(\frac{n_\alpha m_\alpha \mathbf{u}_\alpha^2}{2} \right)_{coll}, \quad (2.31)$$

with the isotropic (P_α) and anisotropic ($\mathbf{\Pi}'_\alpha$) part of the pressure tensor, temperature T_α and the third-order term $\mathbf{H}_\alpha$ describing the heat flux. An equation for the heat flux is found by taking the third moment, which contains fourth-order terms and so on to infinity. In practice, this series is truncated using a physical argument to find a closing condition, for example by approximating the pressure tensor, so that the continuity equation, the momentum transport equation and Maxwell's equations become a complete description of the plasma. A popular approach is the “cold plasma” approximation in which the pressure tensor vanishes because all particles have the same macroscopic velocity.

2.1.4 Ideal MHD

Assuming that a plasma consists of two particle populations: electrons with the index $\alpha = e$ and ions with $\alpha = i$, the stochastic moments derived in the previous chapter deliver a two-fluid description of the system. The separate fluid components interact via collisions and electromagnetic forces. This is a first step towards a single-fluid model, as it is used in magnetohydrodynamics (MHD). First we can use the assumption of quasi-neutrality, meaning the number densities of the populations must be equal: $n_e = n_i = n$.

Now we can define single-fluid macroscopic variables from the multi-fluid moments:

$$\rho = \rho_e + \rho_i = m_e n_e + m_i n_i = n(m_e + m_i) \quad \text{mass density} \quad (2.32)$$

$$\sigma = en_i - en_e = 0 \quad \text{charge density} \quad (2.33)$$

$$\mathbf{u} = \frac{\rho_e \mathbf{u}_e + \rho_i \mathbf{u}_i}{\rho} = \mathbf{u}_i - \frac{\rho_e}{\rho} (\mathbf{u}_i - \mathbf{u}_e) \quad \text{bulk velocity} \quad (2.34)$$

$$\mathbf{J} = en_i \mathbf{u}_i - en_e \mathbf{u}_e = en(\mathbf{u}_i - \mathbf{u}_e) \quad \text{current density} \quad (2.35)$$

With the help of the last two equations we can express the electron and ion collective velocities as:

$$\mathbf{u}_e = \mathbf{u} - \frac{m_i}{e\rho} \mathbf{J} \quad (2.36)$$

$$\mathbf{u}_i = \mathbf{u} - \frac{m_e}{e\rho} \mathbf{J} \quad (2.37)$$

Plugging these definitions into the moments of the Vlasov equation and taking a sum over the particle species yields the single-fluid description of a plasma. The first moment is the continuity equation of magnetohydrodynamics:

$$\frac{\partial(\rho_e + \rho_i)}{\partial t} + \nabla \cdot (\rho_e \mathbf{u}_e + \rho_i \mathbf{u}_i) = 0 \quad (2.38)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.39)$$

The source term is set to zero under the assumption that no recombination or ionization takes place or is in equilibrium between the particle species. The second moment results in the MHD momentum transport equation:

$$\rho (\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}) = \sigma \mathbf{E} + \mathbf{J} \times \mathbf{B} - \nabla \cdot \sum_{\alpha} \mathbf{\Pi}_{\alpha}, \quad (2.40)$$

which corresponds to the Navier-Stokes equation of hydrodynamics with the viscosity terms hidden in the pressure divergence term. The electric field term can be set to zero with the assumption of quasi-neutrality preventing macroscopic electric fields. The sum over the particle species pressure tensors is defined as the total pressure p . The magnetic part of the Lorentz force $\mathbf{J} \times \mathbf{B}$ can be rewritten using Ampere's law, yielding the ideal MHD momentum transport equation:

$$\rho (\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}) = -\mathbf{B} \times (\nabla \times \mathbf{B}) - \nabla p. \quad (2.41)$$

Plasmas considered in this work are sufficiently conductive quickly negating large-scale electric field fluctuations, so that in the non-relativistic case ($|u| \ll c$) the displacement current in Maxwell's equations can be neglected. External currents are also not considered.

Next, let us consider Ohm's law in the fluid description. While in the particle picture

the plasma current is the sum over all particle motions, in fluid theory the current transport equation is derived from the momentum transport equation by multiplying with the factor q_α/m_α and summing over all populations. We consider the electron equation separately:

$$m_e \frac{d\mathbf{u}_e}{dt} = -e(\mathbf{E} + \mathbf{u}_e \times \mathbf{B}) - \frac{1}{n_e} \nabla(n_e k_B T_e) - \nu_{ie} m_e (\mathbf{u}_e - \mathbf{u}_i), \quad (2.42)$$

with the electron-ion collision term $\nu_{ie} m_e (\mathbf{u}_e - \mathbf{u}_i)$. We consider only low-frequency waves, so that modes around the electron cyclotron frequency Ω_e are not resolved. For electron induced oscillations this means R-modes (or Whistler-modes below $\Omega_e/2$). The consequence is that the temporal variability of the electron collective velocity goes towards zero. Replacing collective velocities with the current density (Eq. 2.35) we get the generalised form of Ohm's law:

$$\mathbf{E} + \mathbf{u}_e \times \mathbf{B} - \frac{1}{en_e} \mathbf{J} \times \mathbf{B} - \frac{1}{en_e} \nabla(n_e k_B T_e) = \eta \mathbf{J}. \quad (2.43)$$

High frequency modes are already excluded here, now the same argument goes for ion wave-modes. If the collision rate between electrons and ions is high, so that the resistivity $\eta \mathbf{J}$ with $\eta = m_e \nu_{ie}/e^2$ is large, the generation of L-modes is suppressed. The so-called Hall-MHD is an extension of the ideal MHD that deals with the L-modes using the Hall term $\mathbf{J} \times \mathbf{B}$, that will be neglected in the frame of this work.

Taking the curl of this equation and using Faraday's law we get the connection to the temporal development of the magnetic field:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \frac{1}{en_e} \nabla n_e \times \nabla k_B T_e - \eta \nabla \times (\nabla \times \mathbf{B}). \quad (2.44)$$

This is the general form of the induction equation in magnetohydrodynamics for low-frequency wave modes. In ideal MHD the resistivity term is set to zero ($\eta = 0$) and the temperature gradient is usually parallel to the density gradient, resulting in a disappearing cross-product between the two. This leads to the induction equation in ideal MHD:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}). \quad (2.45)$$

In the frame of the ideal MHD the current density and correction terms disappear in Eq. 2.43 leaving the simplified form of Ohm's law:

$$\mathbf{E} = -\mathbf{u}_e \times \mathbf{B} \quad (2.46)$$

2.1.5 Resistivity and incompressibility

The ideal MHD neglects macroscopic effects of particle collisions like the magnetic resistivity term in Eq. 2.44 in the previous section. If collisions between electrons and

ions occur, the term $\eta \nabla \times (\nabla \times \mathbf{B})$ cannot be neglected, but we can simplify it with the requirement, that the magnetic field must be solenoidal. In this case the double cross product is reduced and we get the induction equation with a scalar magnetic resistivity ν_B :

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \nu_B \nabla^2 \mathbf{B}. \quad (2.47)$$

It should be noted that magnetohydrodynamics is in principle not collision-free, since the pressure tensor is used in its derivation. The pressure term can only be non-zero if we assume kinetic (or in the classic definition “thermal”) collisions. So the concept of a collisionless plasma has to be seen in two different contexts: first on the microscopic level with collision processes implicitly assumed in the MHD and second in the effect of the MHD equations on specific waves in the fluid. For example a collisionless system would be valid for waves with wave number $k < \lambda_{\text{mfp}}^{-1}$, λ_{mfp} designating the mean free path.

Collisions between particles of the same species, on the other hand, cause another effect: Newtonian friction between fluid layers with different velocities that is described with a resistivity term formally different from the magnetic resistivity. Similar to the induction equation, the momentum transport equation has to be modified with a resistive term:

$$\rho(\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u}) = -\mathbf{B} \times (\nabla \times \mathbf{B}) - \nabla p + \rho \nu_u \nabla^2 \mathbf{u}. \quad (2.48)$$

Analogous to the hydrodynamic case of the Navier-Stokes equation (that this equation transforms to for $\mathbf{B} = 0$) we call ν_u “viscosity”.

Equations 2.47 and 2.48 now both have a dissipation term with a similar magnitude and effect. Therefore it is common in MHD simulations to introduce a general resistivity

$$\nu_n \equiv \nu_B = \nu_u, \quad (2.49)$$

thus limiting oneself to plasmas with a magnetic Prandtl number $Pr_m = \nu_u / \nu_B = 1$, which is the case in wide regions of the heliosphere. Since there are not many measurements of this quantity, the assumption is controversially discussed (Montgomery, 1992), but there are hints that for turbulence dominated by Alfvén waves this parameter is of lower importance (Bershanskii, 2002; Bigot et al., 2008). Additionally, the simulation approach used in this work amplifies the diffusivity term artificially by a power h , so that

$$\nu \nabla^2 \rightarrow \nu_h \nabla^{2h}, \quad (2.50)$$

effectively introducing hyperdiffusivity.

In this work we consider only incompressible plasmas, so that the concept of incompressibility shall be discussed here. According to the continuity equation 2.39, incompressibility is equivalent to $\mathbf{u}$ being solenoidal. On the one hand the time derivative of a constant density is zero, on the other hand the density is not dependent on location and

can be pulled out of the divergence, leaving us with the expression for incompressibility:

$$\nabla \cdot \mathbf{u} = 0. \quad (2.51)$$

In the last steps towards the final form of the equations the magnetic fields have to be adjusted by a factor 4π for the cgs-units, which are used in this work from here on out. The double cross product in the momentum transport equation can be evaluated taking the non-commutativity of the nabla operator and the solenoidality of $\mathbf{B}$ into account. The thermal and magnetic pressure terms can be merged into the total pressure $P = p + \mathbf{B}^2/(8\pi)$ and the magnetic fields renormalised to $\mathbf{b} \equiv \mathbf{B}/\sqrt{4\pi\rho}$. This leads to the set of incompressible MHD equations used throughout this work:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{b} \cdot \nabla \mathbf{b} - \mathbf{u} \cdot \nabla \mathbf{u} - \nabla P + \nu_n \nabla^{2n} \mathbf{u} \quad \text{momentum transport} \quad (2.52)$$

$$\frac{\partial \mathbf{b}}{\partial t} = \mathbf{b} \cdot \nabla \mathbf{u} - \mathbf{u} \cdot \nabla \mathbf{b} + \nu_n \nabla^{2n} \mathbf{b} \quad \text{induction} \quad (2.53)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{incompressibility} \quad (2.54)$$

$$\nabla \cdot \mathbf{b} = 0 \quad \text{solenoidality} \quad (2.55)$$

The incompressibility condition was used to set the density to $\rho = 1$. The normalisation of the magnetic field leads to the same unit of dimension for $\mathbf{b}$ and $\mathbf{u}$, which will be used in the next section together with a background magnetic field to derive a specific wave representation of the equations. The representation of the incompressible MHD equations shown here is comparable to e.g. Maron and Goldreich (2001).

As discussed before the system of equation derived from the stochastic moments is in principle infinite and must be closed (usually using a physical argument). Otherwise we would need to consider the heat transport equation at this point to define the pressure tensor. Instead the system is closed implicitly through the assumption of incompressibility. Taking the divergence of the momentum equation and using the solenoidality of $\mathbf{u}$, we can express the pressure tensor with the magnetic and velocity fields and close the MHD equations:

$$\nabla^2 P = \nabla \mathbf{b} \otimes \nabla \mathbf{b} - \nabla \mathbf{u} \otimes \nabla \mathbf{u}. \quad (2.56)$$

A double integration using Green's function yields the explicit condition for the pressure:

$$P = \int \frac{\nabla \mathbf{u} \otimes \nabla \mathbf{u} - \nabla \mathbf{b} \otimes \nabla \mathbf{b}}{4\pi|\hat{\mathbf{x}} - \mathbf{x}|} d^3 \hat{x}. \quad (2.57)$$

2.1.6 Elsässer notation

In this chapter the physical interpretation of the MHD equation terms shall be discussed. In the context of incompressible magnetised plasma this will lead to an advantageous alternative form of the equations.

The ∇^2 -terms in Eq. 2.52 and 2.53 describe diffusion in real space, which corresponds

to dissipation in fourier space. In connection with the energy dissipation equation, the viscosity ν_n causes dissipation of energy on small scales. The tensor divergence terms describe non-linear interactions that transport the $\mathbf{u}$ and $\mathbf{b}$ fields via advection in real space and diffusion in fourier space¹. Since it's not possible to distinguish these processes and match them to the terms, the nonlinearities of the equations will be referred to as convection terms from now on, as is usual in the literature. As mentioned before, the solenoidality equations 2.54 and 2.55 express the incompressibility of plasma and absence of magnetic monopoles, respectively.

The incompressible MHD equations have two conserved quantities: the total energy

$$E = \frac{1}{2} \int d^3x (\mathbf{u}^2 + \mathbf{b}^2) \quad (2.58)$$

and the cross helicity

$$H = \frac{1}{2} \int d^3x \mathbf{u} \cdot \mathbf{b}. \quad (2.59)$$

As mentioned before, large scale electric fields cannot exist in plasmas with ideal conductivity, which is the case in the heliosphere and the interstellar medium. With the coupling of magnetic fields $\mathbf{B}$ and plasma flow $\mathbf{u}$ in the ideal MHD it can also be shown that ideal conductivity leads to the so-called frozen-in flux. Consider the temporal change of the magnetic flux through an arbitrary surface S :

$$\frac{d\Phi}{dt} = \frac{d}{dt} \int_S d\mathbf{S} \cdot \mathbf{B}. \quad (2.60)$$

Using Gauss's divergence theorem, Stokes' theorem and Faraday's law it can be rewritten in the following form:

$$\frac{d\Phi}{dt} = - \int_S d\mathbf{S} \cdot \nabla \times (\mathbf{E} + \mathbf{u} \times \mathbf{B}). \quad (2.61)$$

The integrand vanishes, if it can be written as a gradient of a scalar field Ψ :

$$\mathbf{E} + \mathbf{u} \times \mathbf{B} = -\nabla\Psi. \quad (2.62)$$

Since in ideal MHD Ψ is a constant, as can be seen in Eq. 2.46, this necessary and sufficient condition to preserve the magnetic flux is fulfilled.

This means that magnetic flux and plasma flow together; the magnetic field lines are “frozen” into the plasma and follow its motion. This concept was first introduced by Hannes Alfvén. Although magnetic field lines are a purely mathematical abstraction and their motion not physical, the concept is still quite useful when interpreted correctly. Neglecting the dissipation term in the induction equation 2.53, we see that the evolution of the magnetic field is determined by the motion of the fluid elements in the convection

¹Note the difference between fourier space diffusion and diffusion in the real space, that corresponds to dissipation in fourier space

terms. In the frame of resistive MHD with $\eta \neq 0$ the conservation of flux in Eq. 2.61 is modified with a diffusion term:

$$\frac{d\Phi}{dt} = - \int_S d\mathbf{S} \cdot \frac{\eta}{4\pi} \nabla^2 \mathbf{B}. \quad (2.63)$$

In this case magnetic fields are approximately frozen-in, if the diffusion time scale is sufficiently long, which is normally the case in astrophysical plasmas.

In contrast to the missing large scale electric fields, the frozen-in flux allows large scale background magnetic fields in plasma, either through large scale plasma dynamics or from external sources, like the solar magnetic field in the heliosphere. The field variables in the MHD equations can be defined as:

$$\mathbf{b} = \delta\mathbf{b} + B\mathbf{e}_z \quad (2.64)$$

$$\mathbf{u} = \delta\mathbf{u} \quad (2.65)$$

with a fluctuating part and a background magnetic field. In the frame of this work the background field is defined along the z -direction and will be denoted as $\mathbf{B}_z$ or $\mathbf{B}_0$. The direction along this global field is parallel, a variable X along this direction is denoted as $X_{\parallel}$, a variable X in the x - y -plane is perpendicular to $\mathbf{B}_0$ and will be written as $X_{\perp}$.

A consequence of a global background magnetic field in an incompressible plasma is the creation of Alfvén waves. When field lines are “pulled along” by the ion motion they react to this perturbation with a restoring force generating oscillations. In the case of Alfvén waves the ion and magnetic field oscillations are transverse to the propagation direction, which is along the background magnetic field. This mode causes shear stress on the magnetic field ($\nabla \cdot (\mathbf{B}\mathbf{B})$) and is therefore often called shear Alfvén wave. Another wave mode is the pseudo Alfvén wave - the incompressible limit of the slow magnetosonic mode. Magnetosonic waves are the equivalent of sound waves in plasma: longitudinal density fluctuations additionally influenced by magnetic pressure. Since the MHD is confined to the low frequency regime below the ion cyclotron frequency, other plasma modes cannot be modelled. For high-frequency waves a kinetic particle description is necessary.

To reiterate: the ideal incompressible MHD contains only two wave modes - shear and pseudo Alfvén waves. Both propagate with the Alfvén velocity

$$v_A = \frac{B_z}{\sqrt{4\pi\rho}}, \quad (2.66)$$

and are dispersionless:

$$\omega^2 = v_A^2 k_{\parallel}^2. \quad (2.67)$$

The difference between the two species lies in their polarisation. Both modes have the

eigenvectors:

$$\hat{\mathbf{u}}_{S/P}(\mathbf{k}, t) = \hat{\mathbf{A}}_{S/P} e^{i(\mathbf{k}\mathbf{x} \pm \omega t)}, \quad (2.68)$$

$$\hat{\mathbf{b}}_{S/P}(\mathbf{k}, t) = \pm \hat{\mathbf{A}}_{S/P} e^{i(\mathbf{k}\mathbf{x} \pm \omega t)}. \quad (2.69)$$

The polarisation vector of shear Alfvén waves is perpendicular to both the wave vector $\mathbf{k}$ and the propagation direction ($\mathbf{e}_z$):

$$\hat{\mathbf{A}}_S = \frac{\mathbf{e}_k \times \mathbf{e}_z}{\sqrt{1 - (\mathbf{e}_k \cdot \mathbf{e}_z)^2}}. \quad (2.70)$$

Pseudo Alfvén waves, on the other hand, have a polarisation vector in the k - z -plane:

$$\hat{\mathbf{A}}_P = \frac{\mathbf{e}_z - (\mathbf{e}_k \times \mathbf{e}_z)\mathbf{e}_k}{\sqrt{1 - (\mathbf{e}_k \cdot \mathbf{e}_z)^2}}. \quad (2.71)$$

For the representation of only those two wave modes an alternative form of the MHD equations is favourable: the so-called “Elsässer notation” (Elsässer, 1950). The Elsässer fields are direct representations of wave packets travelling parallel or antiparallel to the background magnetic field and are defined as follows:

$$\begin{aligned} \mathbf{w}^- &= \mathbf{u} + \mathbf{b} - v_A \mathbf{e}_\parallel \\ \mathbf{w}^+ &= \mathbf{u} - \mathbf{b} + v_A \mathbf{e}_\parallel. \end{aligned} \quad (2.72)$$

With these definitions the MHD Eq. 2.52 - 2.55 can be expressed as:

$$-\frac{\partial}{\partial t} \mathbf{w}^- = \mathbf{w}^+ \cdot \nabla \mathbf{w}^- + \nabla P - v_A \frac{\partial}{\partial z} \mathbf{w}^- - \nu_n \nabla^{2n} \mathbf{w}^- \quad (2.73)$$

$$-\frac{\partial}{\partial t} \mathbf{w}^+ = \mathbf{w}^- \cdot \nabla \mathbf{w}^+ + \nabla P + v_A \frac{\partial}{\partial z} \mathbf{w}^+ - \nu_n \nabla^{2n} \mathbf{w}^+. \quad (2.74)$$

The solenoidality condition of $\mathbf{u}$ and $\mathbf{b}$ applies to both Elsässer fields:

$$\nabla \cdot \mathbf{w}^\pm = 0. \quad (2.75)$$

The closing condition for the pressure tensor transforms to:

$$P = \int \frac{\nabla \mathbf{w}^+ \otimes \nabla \mathbf{w}^-}{4\pi |\hat{\mathbf{x}} - \mathbf{x}|} d^3x. \quad (2.76)$$

The Elsässer notation allows us to draw conclusions about the behaviour of the waves directly from the equations with relative ease. One can see, for example, that an arbitrary wave packet propagating along the background magnetic field ($\mathbf{w}^+ \neq 0, \mathbf{w}^- = 0$) will be moving undisturbed. Same goes for propagation against B_0 , of course. The reason for this is the disappearance of the nonlinear interaction terms in the propagation equations. The pressure tensor, that could also influence the wave motion, is also zero in this case.

Consequently, only waves propagating in opposite directions can collide and interact with each other. These collisions have to conserve energy and cross helicity. Using these conservation laws and multiplying the propagation equations with $\mathbf{w}^\pm$, the pressure and nonlinear terms disappear and we get:

$$\frac{d}{dt} \int d^3x |\mathbf{w}^\pm|^2 = 0. \quad (2.77)$$

From the Elsässer field definitions one can see that the square of the absolute field values corresponds to the energy in the fields:

$$E_\pm = \int d^3x \frac{|\mathbf{w}^\pm|^2}{4}, \quad (2.78)$$

so that Eq. 2.77 can be interpreted as the energy change of the wave packets. Since the change is zero, the collisions between wave packets have to be elastic, there is no energy dissipated in the interaction (assuming viscosity-free equations, of course). The collisions are characterized by the nonlinear terms of the MHD equations. The nonlinear interactions generate perturbations in the plasma that lead to a turbulent energy cascade. The development of turbulence in magnetised plasma is still a subject of active research and shall be introduced in the next chapter.

2.2 Turbulence

“Fluid turbulence remains one of the toughest problems in classical physics and turbulence in collisionless magnetised plasmas is an even harder problem.” - Koskinen (2011).

Fluid turbulence is a very interesting phenomenon and the last topic of classical physics that is still not completely understood. Even today, after more than 70 years after the first turbulence theory was developed by Kolmogorov (1941a), there is an active controversial discussion on its validity. Turbulence, especially in the frame of hydrodynamics, can easily be observed in our everyday life. Even normal movement through air or water results in high Reynolds numbers² ($\approx 10^4$) due to the relatively low viscosity, and the Reynolds numbers grow even further with higher velocities, e.g. in aviation. Magnetised plasma turbulence is a completely different matter, however, since matter in the plasma state is relatively rare on earth or it has less than ideal conductivity leading to low magnetic Reynolds numbers. Considering the baryonic matter in the whole universe, on the other hand, we find most of it in the form of turbulent plasma (at least when considered by volume) that can be described with magnetohydrodynamics.

There are certain properties that characterise turbulence. First, turbulent flows are highly irregular in that they are very sensitive to initial and boundary conditions, making them a subset of chaotic systems (however, not all chaotic flows are turbulent). As a consequence, no closed solutions that are stable at high Reynolds numbers have been found

²See definition in chapter 1.1.6 Eq. 1.1

even for the Navier-Stokes equations (let alone the MHD equations). Thus turbulence problems are treated in the frame of statistical physics rather than deterministically. Second, turbulent flows are rather energetic and the energy contained in turbulence tends to accelerate the homogenisation (or mixing) in the fluid. Turbulent flows are said to have high diffusivity, in that they have increased rates of mass, momentum and energy transport across the flow.

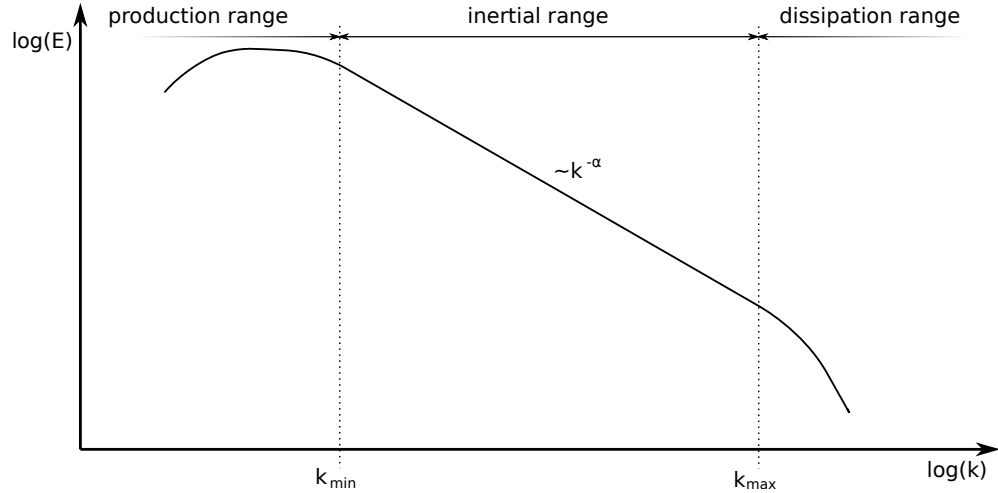


Fig. 2.1: Schematic depiction of a turbulent energy spectrum. Energy is injected into the system in the initial range, cascades to larger wave numbers, forming a power-law over several orders of magnitude in the inertial range, until it is converted to internal energy by viscous stress in the dissipation range. Turbulence theories vary in their predictions of the power-law exponent, or “spectral index”.

The next property of turbulence is the vorticity and it leads to the crucial concept of an energy cascade. In fluid dynamics, turbulent flows can be described as vortices - rotating fluid elements - that are subjected to stretching. Due to conservation of angular momentum the vorticity in the stretching direction increases, conservation of fluid element volumes causes a thinning of vortices perpendicular to the stretching direction. So effectively, large vortices break down into more numerous smaller structures. This is the central mechanic of the energy cascade - the transfer of energy from large to small scales. Turbulent flows need a persistent energy input to be sustainable, because of the rapid energy dissipation due to viscosity at the smallest scales. This energy injection usually happens at large length scales (also called integral length scales), for example heliospheric turbulence is driven by solar events and turbulence in the ISM is most likely driven by supernovae. In the energy spectrum of the turbulence (see Fig. 2.1) this corresponds to the initial or driving range at low wave numbers. Next comes the inertial range where large eddies break down into smaller ones, transporting the energy to ever smaller scales. The inertial range usually covers multiple orders of magnitude in wave number k and the energy spectrum follows a power-law of the form $E(k) \propto k^{-\alpha}$ with a

constant negative exponent or “spectral index”. The cascade continues until the structures created are so small that molecular diffusion becomes important. At this point kinetic energy is converted to internal energy by viscous shear stress and dissipated. This happens in the dissipation range at large wave numbers below the so-called Kolmogorov length scales, where the energy input from non-linear interactions and the energy drain from viscous dissipation are balanced. At these scales the turbulence is highly isotropic and homogeneous.

While in hydrodynamics generation of turbulence is usually described by directly introducing vorticity into the medium (“swirling”), there is a second basic mechanism that generates vortex-like structures from the superposition of waves. The vortices generated through superposition can always be decomposed in single wave modes again. In order to put the system in a turbulent state, the waves have to be far from thermal equilibrium. Zakharov et al. (1992) identifies the basic requirement as stochastic long term behaviour of an ensemble of weakly non-linear dispersive waves. The energy transfer and dissipation usually take place through resonant interactions of wave packets. Just like in vortex turbulence this leads to high Reynolds numbers and an energy cascade in the inertial range with a power-law with a constant spectral index over many orders of magnitude of the wave number. This can be seen from the stochastic wave description where the four-point-correlation function of the wave amplitudes leads to a flux of conserved quantities from the initial range to the dissipation range (Zakharov and Filonenko, 1966). Wave turbulence is the predominant effect in weakly turbulent MHD plasmas, but also in the case of strong incompressible turbulence. This primarily follows from the dispersion behaviour of plasma waves and their non-linear interactions, as discussed in the previous chapter. Since the focus of this work lies in MHD turbulence generated by Alfvén waves, we shall briefly discuss Kolmogorov’s original theory in the next chapter, before moving on to the works of Goldreich and Sridhar, who made ground-breaking progress in the theoretical field describing those systems.

2.2.1 Kolmogorov turbulence theory

Kolmogorov’s theory of hydrodynamic turbulence (Kolmogorov, 1941a) was the first comprehensive explanation of the turbulence’s statistics and cannot be dismissed out of hand to this day. Notably, Kolmogorov arrived at his results with assumptions about scale behaviour and dimensional analysis rather than in a frame of physical equations - a fact that compelled many scientists to find a more elegant description of turbulence. The continued popularity of Kolmogorov’s theory indicates that no complete success has been achieved so far on that front.

Kolmogorov started his considerations by postulating three hypotheses.

First hypothesis: For very high Reynolds numbers, the turbulent motions at small scales are statistically isotropic.

In general, the same cannot be said about flows at large scales (here denoted with L), since they are determined by the geometrical features of the system, i.e. the size and form of the boundaries and the directional information in the large scale energy injection generating the turbulence. Kolmogorov’s postulated that this information is lost in the

energy cascade so that the statistics at small scales are universally the same for all turbulent flows with a sufficiently high Reynolds number. The next step is to find out, which parameters determine these universal statistics. An obvious choice would be the rate of energy transfer from large to small scales that is related to the dissipation rate ϵ at which turbulent kinetic energy is converted into thermal internal energy. Dimension-wise, the scaling of the dissipation rate at large scales can be expressed as:

$$\epsilon \propto \frac{u^3}{l}. \quad (2.79)$$

At small scales the kinematic viscosity ν becomes important. This is the content of Kolmogorov's

Second hypothesis: The statistics of small scales in turbulent flows at high Reynolds numbers are universally and uniquely determined by the kinematic viscosity ν and the energy dissipation rate ϵ .

From dimensional analysis we can now derive the so-called “Kolmogorov microscales” describing the length η , velocity u_η and time scales τ_η of the smallest vortices:

$$\eta = \sqrt[4]{\frac{\nu^3}{\epsilon}} \quad (2.80)$$

$$u_\eta = \sqrt[4]{\nu\epsilon} \quad (2.81)$$

$$\tau_\eta = \sqrt{\frac{\nu}{\epsilon}}. \quad (2.82)$$

These scales define the range of the energy cascade where kinetic energy dissipation takes place. The other end is defined by the energy input into the cascade that comes from the large scales of the order L . Between those two extremes lies a range of scales (each denoted by a characteristic length r) that can span multiple orders of magnitude at high Reynolds numbers. Each of the scales r that is smaller than the system size ($r \ll L$) forms at the expense of the energy of larger scales. Since the structures are still much larger than the Kolmogorov scale ($r \gg \eta$) there is essentially no dissipation due to viscous effects and energy is merely transferred to smaller scales. This range is called the “inertial range” since inertial effects are much larger than viscous effects. This leads to Kolmogorov's final hypotheses.

Third hypothesis: In turbulent flows at high Reynolds numbers the statistics of scales in the range $\eta \ll r \ll L$ are universally and uniquely determined by the scale r and the energy dissipation rate ϵ .

Now we can consider the energy spectrum function $E(k)$ that characterises the way kinetic energy is distributed over the length scales, represented here by the wave number $k = 2\pi/r$. The contribution to kinetic energy from the Fourier modes in the range between k and $k + dk$ is $E(k)dk$, so the mean kinetic energy of the flow can be written

as

$$\frac{1}{2}\langle u_i u_i \rangle = \int_0^\infty E(k) dk. \quad (2.83)$$

The dimension of the left hand side is $L^2 T^{-2}$, the dimension of the integration L^{-1} . Taking the third hypotheses into consideration, $E(k)$ can only depend on r with $[r] = L$ and ϵ with $[\epsilon] = L^2 T^{-3}$, so dimensional analysis yields the only possible form for the energy spectrum:

$$E(k) = C \epsilon^{\frac{2}{3}} k^{-\frac{5}{3}} \quad (2.84)$$

with a universal constant C . This famous result is known as the Kolmogorov -5/3 spectrum and it is supported by a considerable amount of experimental evidence (see e.g. Frisch, 1995).

Nevertheless, Kolmogorov's theory is not without problems. First, consider the statistical moments of velocity fluctuations $\delta \mathbf{u}(\mathbf{r}) = \mathbf{u}(\mathbf{x} + \mathbf{r}) - \mathbf{u}(\mathbf{x})$, that is, the difference in flow velocities between two points separated by vector $\mathbf{r}$. From the assumption of scale invariance and other results of Kolmogorov (1941a) follows the scaling behaviour of n -th order moments of velocity fluctuations (or structure functions):

$$\langle \delta \mathbf{u}(\mathbf{r})^n \rangle = C_n \epsilon^{\frac{n}{3}} r^{\frac{n}{3}}, \quad (2.85)$$

where C_n are, again, universal constants.

However, for higher orders the scaling exponents seem to deviate from the $n/3$ theoretical value, becoming non-linear functions in n , indicating a breakdown of the statistical self-similarity (Mathieu and Scott, 2000). Also, the universality of C_n is debated. Both problems are related with the phenomenon of intermittency in turbulence which is an important subject of contemporary research in the field of turbulence theory.

2.2.2 MHD turbulence

A more general problem with Kolmogorov's theory is that it was derived for pure hydrodynamics without considering the effects of charged or magnetised fluid elements. Turbulence in magnetized plasmas has to fulfil the induction equation 2.53 of the MHD, making the problem more complex but simultaneously allowing for simpler solutions in specific limits. In contrast to hydrodynamics, in MHD turbulence can have different strength characteristics that are determined by the ratio of the magnetic field fluctuations δb to the background field B_z . In general we differentiate two regimes of turbulence: for $(\delta b)^2 / B_z^2 \ll 1$ we speak of weak turbulence, if the ratio is approximately 1, the system is strongly turbulent. In the case of vortex turbulence we can also compare the Alfvén time scale ($\propto (k_{\parallel} B)^{-1}$) to the “eddy turnover time” - the effective lifetime of a vortex. In weak turbulence the Alfvén time scale is much shorter than the turnover time. The advantage of describing MHD turbulence in two different regimes is that we can now find approximations that lead to separate theories for both. In the case of weak incom-

pressible turbulence, a perturbative approach can be made since the system's behaviour is governed by linear (Alfvén) waves. In the weak, compressible case the MHD equations can be reduced to include only two-wave interactions as a linear process and three-wave interactions as the only non-linear contribution. The scope of this work is limited to the incompressible case, however.

There is a controversial discussion about the basic mechanism of turbulence generation in MHD: is it a direct vortex cascade or wave interactions. On the one hand, Alfvén waves are linear solutions of the MHD equations, especially in the incompressible case, but they are non-dispersive with a direct proportionality between frequency and wave number. Another open question for a long time was if a relation for the scaling behaviour of the turbulence structure functions analogous to Kolmogorov's Eq. 2.85 can be found. For strong turbulence Politano and Pouquet (1998) could show that due to the coupling of the magnetic and velocity fields via the frozen-in flux, an analogy to the "4/5 law" (Kolmogorov, 1941b) for the third order structure function can be found. This allows us to express it in Elsässer variables:

$$\langle (\hat{w}_{\parallel}^+ - w_{\parallel}^+) \sum_j (\hat{w}_j^- - w_j^-)^2 \rangle = -\frac{4}{5}\epsilon r, \quad (2.86)$$

and consequently assume interactions between Alfvén waves as the generating mechanism of MHD turbulence.

A first attempt at a theory of incompressible MHD turbulence was made in the independent works of Kraichnan (1965) and Iroshnikov (1964). Both applied an approach analogous to Kolmogorov's original work to Alfvén waves propagating with

$$v_A = \frac{B}{\sqrt{4\pi\rho}} \quad (2.87)$$

through a magnetized plasma. This leads to a change of the characteristic system time scale from the Kolmogorov time scale τ_{η} to the Alfvén time $\tau_A = r_{\parallel}/v_A$, that can be estimated as $\tau_A \sim 1/(kB_z)$. Applying this to the energy dissipation rate estimation yields the structure function of the order p :

$$S_p(r) \sim v^p \sim (\epsilon r B_z)^{\frac{p}{4}}. \quad (2.88)$$

Applying the same procedure as in Kolmogorov's theory (Eq. 2.83), from the Fourier transform $\mathcal{E}(k)$ of the second order structure function we get the energy spectrum by deriving in respect to k :

$$E(k) = \frac{d}{dk} \mathcal{E}(k) \propto (\epsilon B_z)^{\frac{1}{2}} k^{-\frac{3}{2}}. \quad (2.89)$$

The theory of Kraichnan and Iroshnikov predict a power-law energy spectrum with the modified spectral index $-3/2$. The theory assumes that wave packets can collide multiple times without significantly changing their properties, implying weak turbulence. In another implicit assumption, the three-wave interaction is taken to be the main process

as shown by Sridhar and Goldreich (1994). Together with the explicit assumption of isotropy, which cannot be made in a medium with a background magnetic field (Shebalin et al., 1983), this leads to a problem that can be easily shown using momentum and energy conservation. For three-wave interactions these conservation laws can be written as:

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 \quad (2.90)$$

$$\omega_1 + \omega_2 = \omega_3. \quad (2.91)$$

In incompressible magnetised plasmas the interacting waves are Alfvén modes. Using their dispersion relation from Eq. 2.67, we can rewrite the conservation laws as:

$$k_{1,\parallel} + k_{2,\parallel} = k_{3,\parallel} \quad (2.92)$$

$$|k_{1,\parallel}| + |k_{2,\parallel}| = |k_{3,\parallel}|. \quad (2.93)$$

As has been discussed before in section 2.1.6, it is evident from the MHD equations in Elsässer notation that only wave packets propagating in different directions can interact with each other. So in the above interaction the wave numbers of packets 1 and 2 must have opposite signs. To then fulfil both resonance conditions, either of the wave numbers $k_{1,\parallel}$ or $k_{2,\parallel}$ would have to be zero. Consequently, there can be no three-wave interactions under these conditions.

Despite its shortcomings the theory of Kraichnan and Iroshnikov provides valuable insights into the nature of MHD turbulence. The theory progressed the idea that collisions of Alfvén wave packets drive the turbulence, instead of decaying vortices like in hydrodynamics. Kraichnan (1965) also demonstrated that in neutral fluids the energy cascade is only driven by interactions of vortices on the same length scale. The influence of a large scale velocity field on smaller structures disappears after a Galilei transformation. Magnetic fields, on the other hand, persist and consequently influence smaller structures making it an important ingredient of the energy cascade. Since the effect of the magnetic field of a large vortex on a small one is similar to the effect of a background magnetic field on Alfvén waves, Kraichnan concluded that the underlying process of incompressible MHD turbulence are Alfvén wave collisions.

Goldreich-Sridhar theory

The first comprehensive theory of incompressible MHD turbulence was provided by Goldreich and Sridhar (abbreviated as GS below). It is based on wave turbulence rather than vortex turbulence and consists of three parts. In Sridhar and Goldreich (1994) the authors describe the weak turbulence regime, while Goldreich and Sridhar (1995) concentrates on strong turbulence. Since both regimes can occur in the same system on different scales, there has to be a transition between them. This was presented in Goldreich and Sridhar (1997) as a theory of “intermediate” turbulence. We will give a short overview over all three models and show how they lead to Kolmogorov-like spectra, for details the interested reader is referred to the original publications Goldreich and

Sridhar (1995, 1997); Sridhar and Goldreich (1994).

Weak turbulence.

As already discussed, three-wave interactions can be ruled out as the basic mechanism in the frame of incompressible MHD in a weakly turbulent magnetised plasma, since they cannot fulfil the resonance condition. Hence, the GS theory focuses on the next higher order as the predominant process: the four-wave interaction. The resonance condition that has to be fulfilled in this case is:

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4 \quad (2.94)$$

$$\omega_1 + \omega_2 = \omega_3 + \omega_4. \quad (2.95)$$

Again, $\mathbf{k}_1 \cdot \mathbf{e}_z$ and $\mathbf{k}_2 \cdot \mathbf{e}_z$ must have opposite signs so the waves can interact. If we now choose $k_{1,z} > 0$ and $k_{2,z} < 0$ without loss of generality, momentum conservation yields $k_{3,z} > 0$ and $k_{4,z} < 0$, modifying the resonance condition to:

$$k_{1,z} + k_{2,z} = k_{3,z} + k_{4,z} \quad (2.96)$$

$$k_{1,z} - k_{2,z} = k_{3,z} - k_{4,z} \quad (2.97)$$

resulting in

$$k_{1,z} = k_{3,z} \quad (2.98)$$

$$k_{2,z} = k_{4,z}. \quad (2.99)$$

Thus, there can be no energy cascade parallel to the magnetic field based on the four-wave interaction, since the z -component of the waves remains unchanged.

So the turbulence must develop perpendicular to B . Analogous to Kolmogorov's theory we start by estimating the energy dissipation rate as

$$\epsilon \sim \frac{v_\lambda^2}{t_c}, \quad (2.100)$$

with the velocity of the fluid fields v_λ and the cascading time scale t_c that is proportional to the number N of collisions necessary to cancel the correlation to the initial state of a wave packet:

$$t_c \sim \frac{N}{k_z v_A}. \quad (2.101)$$

The velocity of a wave packet changes in a four-wave interaction by

$$|\delta v_\lambda| \sim \left| \frac{d^2 v_\lambda}{dt^2} \frac{1}{(k_z v_A)^2} \right|. \quad (2.102)$$

A dimensional analysis of this expression can be made with the help of the MHD equations, specifically the term $\partial_t \mathbf{v} = -(\mathbf{v} \cdot \nabla) \mathbf{v}$, where an approximation for shear Alfvén waves can be made with $\mathbf{v} \cdot \nabla \sim k_\perp v_\lambda$. The GS model for weak turbulence generally

neglects pseudo Alfvén waves, which can be justified by the fact that they are very effectively damped in weak turbulence by the Barnes mechanism (Barnes, 1966). The result of the dimensional analysis

$$\frac{d^2 v_\lambda}{dt^2} \sim \frac{d}{dt}(k_\perp v_\lambda^2) \sim k_\perp v_\lambda \frac{dv_\lambda}{dt} \sim k_\perp^2 v_\lambda^3 \quad (2.103)$$

can now be used in Eq. 2.102 to estimate the relative velocity change of a wave packet in a collision:

$$\left| \frac{\delta v_\lambda}{v_\lambda} \right| \sim \left(\frac{k_\perp v_\lambda}{k_z v_A} \right)^2, \quad (2.104)$$

which in turn yields the number of collisions necessary for decorrelation:

$$N \sim \left(\frac{k_z v_A}{k_\perp v_\lambda} \right)^4. \quad (2.105)$$

This can now be used to express the relation between the energy dissipation rate ϵ and the velocity fields v_λ :

$$v_\lambda^2 \sim \epsilon t_c \sim \frac{\epsilon N}{k_z v_A} \sim \frac{\epsilon (k_z v_A)^3}{k_\perp^4 v_\lambda^4}. \quad (2.106)$$

Assuming a constant energy rate $\epsilon = \text{const.}$ we get an expression for the square of fluid velocity, which is essentially the second order structure function:

$$v_\lambda^2 \sim \frac{\epsilon^{1/3} (k_z v_A)}{k_\perp^{4/3}}. \quad (2.107)$$

A sum over all scales gives the mean kinetic energy of the flow

$$\sum v_\lambda^2 = \int \frac{d^3 k}{8\pi^3} E(k_\perp, k_z) \quad (2.108)$$

with the three dimensional energy spectrum

$$E(k_\perp, k_z) \propto \frac{\epsilon^{1/3} v_A}{k_\perp^{10/3}}. \quad (2.109)$$

In addition to this power-law in $k_\perp$ we also get an estimate for the scaling of the collision rate. Since v_λ is proportional to $k_\perp^{-2/3}$, the proportionality $N \propto k_\perp^{-4/3}$ follows, meaning that for larger $k_\perp$ fewer collisions are necessary for decorrelation to the initial state. If N is approaching unity, the system is strongly turbulent and the approximations above are no longer valid.

To summarise:

According to GS there can be no three-wave interactions in weak incompressible MHD

turbulence, since the resonance condition cannot be met. Instead, the energy cascade develops through four-wave interactions perpendicular to the background magnetic field. The power-law of the energy spectrum is flatter than Kolmogorov's $-5/3$. The turbulence strength is not a constant, it grows at larger wave numbers. At certain fluctuation strengths and length scales the GS theory for weak turbulence loses its validity, as the system cannot be described by four-wave interactions any more.

A detailed derivation of the above statements using perturbation theory and direct solution of the kinetic four-wave interaction equation can be found in Sridhar and Goldreich (1994).

Strong turbulence.

In the previous section we gained the insight that turbulence strength scales with the wave number. As we go to smaller scales, the collision rate N necessary for decorrelation becomes so small, that the energy cascade cannot continue with four-wave interactions. The GS model for strong turbulence uses the fact, that in a strong fluctuation regime the resonance condition can be violated, which allows three-wave interactions. In order to characterise the transition from weak to strong turbulence, GS introduced the “critical balance”, with a definition basically derived from the collision rate:

$$\zeta \sim \frac{k_{\perp} v_{\lambda}}{k_z v_A}. \quad (2.110)$$

This parameter is a measure of “balance” between the Alfvén time scale of the linear regime and the non-linear time scale of the energy transfer from large to small scales. For $\zeta \ll 1$ the GS model for weak turbulence relies on four-wave interactions. The spectrum for strong turbulence is derived analogous to the previous section with the modified assumption $\zeta \sim 1$. This is equivalent to $N \sim 1$ and to the assumption of a strong perturbation in the velocity field $v_{\lambda} \sim v_A$. Using these assumption, the cascading time scale can be written as

$$t_c \sim \frac{1}{k_z v_A}. \quad (2.111)$$

From the critical balance we get the relation between v_{λ} and v_A :

$$v_{\lambda} = \frac{k_z}{k_{\perp}} v_A. \quad (2.112)$$

Applying both to the energy dissipation rate estimation yields:

$$\epsilon \sim \frac{v_{\lambda}^2}{t_c} \sim \frac{k_z^3 v_A^3}{k_{\perp}^2}. \quad (2.113)$$

From this GS derive two important statements concerning the scaling behaviour of the

velocity and wave numbers in strong turbulence:

$$k_z \sim k_\perp^{2/3} L^{1/3} \quad (2.114)$$

$$v_\lambda \sim v_A (K_\perp L)^{-1/3}. \quad (2.115)$$

The first shows the correlation between the parallel and perpendicular extents of the turbulent structures, that can be seen as vortices in this context. Written as

$$\frac{k_\perp}{k_z} \sim \sqrt[3]{k_\perp L} \quad (2.116)$$

it shows the elongation of the vortices in the direction of the background magnetic field for large $k_\perp$.

The second relation can be used to derive the energy spectrum via the second order structure function:

$$v_\lambda^2 \sim \frac{v_A^2}{k_\perp^{2/3} L^{2/3}} = \frac{v_A^2}{k_\perp^{4/3} L^{1/3}} f\left(\frac{k_z L^{1/3}}{k_\perp^{2/3}}\right). \quad (2.117)$$

Again, this is the k -space integral in $d^2 k_\perp dk_z$ over the three dimensional energy spectrum:

$$E(k_\perp, k_z) \sim \frac{v_A^2}{k_\perp^{10/3} L^{1/3}} f\left(\frac{k_z L^{1/3}}{k_\perp^{2/3}}\right). \quad (2.118)$$

Hence, the one dimensional spectrum in perpendicular direction shows the Kolmogorov spectral index ($\propto k_\perp^{-5/3}$). The derivation is complicated by the inapplicability of perturbation theory due to the strong fluctuations, making a direct solution of the kinetic wave equations necessary. More details can be found in Goldreich and Sridhar (1995).

In summary, their findings are:

Turbulence strength grows with larger wave numbers $k_\perp$ until the number of collisions between wave packet necessary for decorrelation from the initial state approaches unity. The concept of critical balance ζ describes the onset of strong turbulence. At $\zeta \sim 1$ four-wave interaction is no longer the dominant process and non-resonant three-wave interaction contributes to the energy cascade. This effectively limits the inertial range of weak turbulence. The cascade is anisotropic, perpendicular to the background magnetic field and follows a power-law with the Kolmogorov exponent $-5/3$.

Intermediate turbulence.

For the sake of completeness the third theory of Goldreich and Sridhar shall be presented briefly. The previous two models have shown that weak and strong turbulence can occur in the same system simultaneously, as the turbulence strengthens at smaller scales. Consequently, there has to be a transition region, where both model are valid. This region is described by the theory of intermediate turbulence. The main idea is to apply three-wave interactions to the derivation of Kraichnan and Iroshnikov. This

approach modifies the velocity change of the wave packets in a collision (compare Eq. 2.102) to:

$$\frac{\delta v_\lambda}{v_\lambda} \sim \frac{dv_\lambda}{dt} \frac{1}{v_A k_z v_\lambda}. \quad (2.119)$$

This is a result from the kinetic three-wave interaction equations and is used, as previously, to derive the energy spectrum:

$$E(k_\perp, k_z) \sim \frac{v_L^2}{k_\perp^3}. \quad (2.120)$$

The resulting spectrum is significantly flatter than that of GS theories for weak and strong turbulence. The energy cascade is anisotropic, unlike in the models of Kraichnan and Iroshnikov, and develops only perpendicular to the background field. As already implied by the critical balance, the inertial range of intermediate turbulence is strongly limited. Consequently, observing this kind of energy spectrum in astrophysical environments is highly unlikely.

Two conditions have been formulated to help classify which model is applicable in which region:

1. Weak interactions imply, that the velocity change in a single interaction of wave packets is small (see Eq. 2.102).
2. The change has to be small enough for a linearisation in the frame of perturbation theory to be valid.

For the GS theory of weak turbulence to be valid, both conditions must be fulfilled. In the intermediate regime only (1) is valid, but the wave packets change significantly, so that the perturbative approach fails. This is essential, since with the perturbative approach three-wave interactions would be forbidden. Consequently, in strong turbulence both conditions are violated and Goldreich and Sridhar (1995) applies.

In conclusion, the works of Goldreich and Sridhar offer the best foundation for models of incompressible MHD turbulence. They describe the turbulence consistently with different orders of wave interaction processes and explain the observed spectra. In particular, the theory of strong turbulence is widely accepted as it succeeded in deriving the Kolmogorov spectral index $-5/3$ in a magnetised plasma, which has only been possible in a purely hydrodynamical context before. The models of weak and intermediate turbulence are more controversially debated (e.g. Galtier et al., 2000), especially the predominant turbulence driving process is still under discussion, as four-wave interaction seems to be not the most important phenomenon, necessarily. Another issue seems to be the onset of the intermediate regime. Some works favour a mix of the weak and intermediate models, suggesting that they cannot be considered separately and proposing to substitute “intermediate” with “weak” universally (Lithwick and Goldreich, 2003). We abstain from a more detailed consideration of weak turbulence, as strong turbulence is of higher importance for this work.

2.3 Particle transport

The last and most important theoretical cornerstone for this work concerns the transport of charged particles in a magnetised turbulent medium. In principle, this is of course completely covered by the Klimontovich equation 2.9. In practice, this equation describing the trajectory of every particle is cumbersome at best and completely impossible to handle for any type of realistic astrophysical scenario, as has been discussed before. Again, a statistical description is needed and the Vlasov equation 2.10 will serve as a starting point in finding such a description.

Since particle transport in the heliosphere and in the interstellar medium involves highly energetic particles that reach relativistic velocities, a relativistic form of the Vlasov equation is needed, which can be found by using the relativistic definition of velocity $\mathbf{v} = \mathbf{p}/(m\gamma)$, yielding in (cgs-units):

$$\frac{df_T}{dt} = \frac{\partial f_T}{\partial t} + \frac{\mathbf{p}}{m\gamma} \frac{\partial f_T}{\partial \mathbf{x}} + q \left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c} \right) \frac{\partial f_T}{\partial \mathbf{p}} = S_T. \quad (2.121)$$

The equation describes the temporal evolution of the 6-dimensional phase space density $f_T(\mathbf{x}, \mathbf{p}, t)$ under the influence of electromagnetic fields. A source term has been added to the right hand side for the sake of completeness. With a few sensible assumptions and a coordinate transformation the equation's complexity can be reduced until useful statements about particle transport can be made.

First, we will exploit the fact that charged particles with the mass m_α and atomic number Z_α gyrate around the field lines in a magnetic field B with the cyclotron frequency

$$\Omega = \frac{Z_\alpha e B}{\gamma m_\alpha c} \quad (2.122)$$

and the Larmor radius

$$r_L = \frac{v_T}{\Omega}. \quad (2.123)$$

This spiral motion (see Fig. 2.2) lends itself to a transformation from the particle coordinates (x, y, z) to the coordinates of the centre of particle gyration (X, Y, Z) . The position vector of the centre absorbs the gyration:

$$\mathbf{R}(X, Y, Z) = \mathbf{x} + \frac{\mathbf{v} \times \mathbf{e}_z}{\pm \Omega}. \quad (2.124)$$

The sign in the denominator is chosen according to the sign of the particle's charge. It's also advantageous to transform the particle momentum to spheric coordinates, determined by the modulus of the momentum p , the azimuth angle in the x - y -plane ϕ and the pitch angle between the momentum vector and the background magnetic field as shown

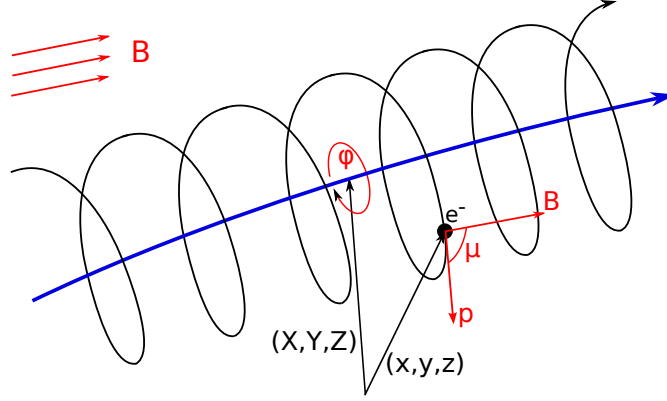


Fig. 2.2: Gyration motion of a charged particle in a magnetic field B showing the coordinates of the gyration centre (X, Y, Z) and the spheric momentum coordinates (p, μ, ϕ)

in Fig. 2.2:

$$\mu = \cos\theta = \frac{\mathbf{v} \cdot \mathbf{B}_0}{|\mathbf{v}| |\mathbf{B}_0|}. \quad (2.125)$$

With the transformations to the new coordinates $X_\sigma = (p, \mu, \phi, X, Y, Z)$

$$\begin{aligned} p_x &= p\sqrt{1-\mu^2}\cos\phi \\ p_y &= p\sqrt{1-\mu^2}\sin\phi \\ p_z &= p\mu \\ x &= X - \frac{c}{Z_\alpha e B_0} p\sqrt{1-\mu^2}\sin\phi \\ y &= Y + \frac{c}{Z_\alpha e B_0} p\sqrt{1-\mu^2}\cos\phi \\ z &= Z. \end{aligned} \quad (2.126)$$

the Vlasov equation can now be written in the form:

$$\frac{\partial f_T}{\partial t} + v\mu \frac{\partial f_T}{\partial Z} - \Omega \frac{\partial f_T}{\partial \phi} + \frac{1}{p^2} \frac{\partial}{\partial X_\sigma} (p^2 g_{X_\sigma} f_T) = S_T. \quad (2.127)$$

The generalised forces g_{X_σ} include the effects of the fluctuating electromagnetic fields, for explicit forms of the terms see Schlickeiser (2002). Solutions of the equation where those terms vanish indicate unperturbed orbits.

2.3.1 Quasilinear theory

The Quasilinear theory (QLT) is a perturbative method based on the assumption that particles (or rather their guiding centres) are propagating along unperturbed trajectories. Consequently, fluctuations are assumed to be small and vanish, when an ensemble average is taken over f_T . This also implies that the electric and magnetic fields can be decomposed in a large scale constant part and random fluctuations that are comparatively small and cancel each other out in an ensemble average:

$$\begin{aligned}\mathbf{E} &= \delta\mathbf{E}(\mathbf{x}, t) \\ \mathbf{B} &= \mathbf{B}_0 + \delta\mathbf{B}(\mathbf{x}, t) \\ \langle \mathbf{E}(\mathbf{x}, t) \rangle &= 0 \\ \langle \mathbf{B}(\mathbf{x}, t) \rangle &= \mathbf{B}_0.\end{aligned}\tag{2.128}$$

This method is based on the argument in Jokipii (1966) that the two-point correlation function describes particle diffusion sufficiently well, meaning that the non-linear fluctuation terms can be neglected making the system quasilinear.

Under these assumptions, the Vlasov equation can be transformed to the Fokker-Planck equation:

$$\frac{\partial F_T}{\partial t} + v\mu \frac{\partial F_T}{\partial Z} - \Omega \frac{\partial F_T}{\partial \phi} = S_T + \frac{1}{p^2} \frac{\partial}{\partial X_\sigma} \left(p^2 D_{X_\sigma X_\nu} \frac{\partial F_T}{\partial X_\nu} \right).\tag{2.129}$$

The distribution function $F_T = \langle f_T \rangle$ is the non-vanishing part of the ensemble averaged phase space density

$$f_T(\mathbf{x}, \mathbf{p}, t) = F_T(\mathbf{x}, \mathbf{p}, t) + \delta f_T(\mathbf{x}, \mathbf{p}, t).\tag{2.130}$$

The quantity $D_{X_\sigma X_\nu}$ in the second-order derivative term is the Fokker-Planck coefficient. It is defined as the integral over the correlation functions of the generalised fluctuating forces along unperturbed particle orbits:

$$D_{X_\sigma X_\nu}(t) = \int_0^t ds \langle \bar{g}_{X_\sigma}(t) \bar{g}_{X_\nu}(s) \rangle.\tag{2.131}$$

For stationary fluctuations the correlation function only depends on the value of the time difference $\xi = t - s$. Another important condition is that we consider time scales much larger than the correlation time ($t \gg t_c$), so that the correlation function becomes negligible, allowing to change the integration boundary to infinity:

$$D_{X_\sigma X_\nu} = \int_0^\infty d\xi \langle \bar{g}_{X_\sigma}(t) \bar{g}_{X_\nu}(t + \xi) \rangle.\tag{2.132}$$

The importance of the Fokker-Planck coefficients becomes evident once we apply a further assumption. The distribution function $F_T(\mathbf{x}, p, \mu, \phi, t)$ is split in an isotropic and

an anisotropic part:

$$F_T(\mathbf{x}, p, \mu, \phi, t) = M_T(\mathbf{x}, p, t) + G_T(\mathbf{x}, p, \mu, \phi, t). \quad (2.133)$$

Substituting this into the Fokker-Planck equation and averaging over μ and ϕ leaves us with only the isotropic part

$$M_T(\mathbf{x}, p, t) = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu F_T(\mathbf{x}, p, \mu, \phi, t), \quad (2.134)$$

thus reducing the number of Fokker-Planck coefficients (or transport parameters) from 25 to 7 and yielding the diffusion-convection transport equation, which is the go-to equation for particle transport in astrophysical settings:

$$\begin{aligned} \frac{\partial M_T}{\partial t} - S_T(\mathbf{r}, \mathbf{p}, t) = & \frac{\partial}{\partial z} \kappa_{zz} \frac{\partial M_T}{\partial z} - \frac{1}{4p^2} \frac{\partial(p^2 v A_1)}{\partial p} \frac{\partial M_T}{\partial z} \\ & + \frac{\partial}{\partial X} \left[\kappa_{XX} \frac{\partial M_T}{\partial X} + \kappa_{XY} \frac{\partial M_T}{\partial Y} \right] \\ & + \frac{\partial}{\partial Y} \left[\kappa_{YX} \frac{\partial M_T}{\partial X} + \kappa_{YY} \frac{\partial M_T}{\partial Y} \right] \\ & + \frac{1}{p^2} \frac{\partial}{\partial p} \left(p^2 A_2 \frac{\partial M_T}{\partial p} \right) + \frac{v}{4} \frac{\partial A_1}{\partial z} \frac{\partial M_T}{\partial p} \end{aligned} \quad (2.135)$$

The parameters in this equation that determine the transport properties are the elements of the spacial diffusion tensor κ_{ij} , the adiabatic deceleration rate A_1 and the momentum diffusion coefficient A_2 . All of these are determined by the remaining 7 Fokker-Planck coefficients:

$$\kappa_{zz} \equiv \frac{v \lambda_{\text{mfp}}}{3} = \frac{v^2}{8} \int_{-1}^1 d\mu \frac{(1 - \mu^2)^2}{D_{\mu\mu}(\mu)}, \quad (2.136)$$

$$\kappa_{XX} = \frac{1}{2} \int_{-1}^1 d\mu D_{XX}(\mu), \quad (2.137)$$

$$\kappa_{XY} = \frac{1}{2} \int_{-1}^1 d\mu D_{XY}(\mu), \quad (2.138)$$

$$\kappa_{YY} = \frac{1}{2} \int_{-1}^1 d\mu D_{YY}(\mu), \quad (2.139)$$

$$\kappa_{YX} = \frac{1}{2} \int_{-1}^1 d\mu D_{YX}(\mu), \quad (2.140)$$

$$A_1 = \int_{-1}^1 d\mu \frac{(1 - \mu^2) D_{\mu p}(\mu)}{D_{\mu\mu}(\mu)}, \quad (2.141)$$

$$A_2 = \frac{1}{2} \int_{-1}^1 d\mu \left[D_{pp}(\mu) - \frac{D_{\mu p}^2(\mu)}{D_{\mu\mu}(\mu)} \right], \quad (2.142)$$

which are, in turn, determined by the random fluctuations of the electromagnetic field components. Consequently, whether the transport parameters are non-zero and appear in the equation at all, depends only on the statistical properties of the plasma turbulence and the background medium.

A glance at the transport parameter equations 2.136-2.142 already foreshadows the significance of the pitch angle diffusion coefficient $D_{\mu\mu}$, as it appears in three of the equations and is directly tied to the mean free path λ_{mfp} . And indeed, especially in the frame of the QLT, pitch angle scattering of particles on magnetic field fluctuations or propagating waves is the most important phenomenon, which warrants a closer look at the coefficient:

$$\begin{aligned} D_{\mu\mu} &= \frac{1}{2\Delta t} \iint_t^{t+\Delta t} dt_1 dt_2 \langle \dot{\mu}(t_1) \dot{\mu}(t_2) \rangle \\ &= \lim_{t \rightarrow \infty} \frac{(\Delta\mu)^2}{2\Delta t} \end{aligned} \quad (2.143)$$

In the frame of the QLT a scattering process is a resonant interaction of a wave with the parallel wave number $k_{\parallel}$ and frequency ω with a particle with a parallel velocity $v_{\parallel} = v\mu$ and gyrofrequency Ω that fulfils the resonance condition

$$k_{\parallel} v\mu - \omega - n\Omega = 0. \quad (2.144)$$

The order of the resonance is determined by $n \in \mathbb{Z}$.

2.3.2 Kinetic wave-particle interaction

Since this work focuses on the effect of these resonant interactions on particles, a kinetic description of the pitch angle change shall be derived here, following Lange (2012). We consider a charged particle interacting with a magnetic field fluctuation δB generated by an Alfvén wave. The resulting pitch angle change can be expressed directly as a change in the parallel momentum $\frac{dp_{\parallel}}{dt} = p \frac{d\mu}{dt}$. The time derivative of the parallel momentum component of a particle with constant total energy E can be calculated from the Lorentz force term $\mathbf{v} \times \delta \mathbf{B}$:

$$\frac{dp_{\parallel}}{dt} = \frac{ec}{2E} \sum_{\pm} (\mp i) p_{\perp} \delta B_{\pm}. \quad (2.145)$$

During an unperturbed motion, the parallel momentum is unchanged while the perpendicular component $p_{\perp} = p_x \pm ip_y = p_{\perp 0} \exp(\mp i\Omega t)$ is modulated by the circular motion, yielding for the temporal change of the pitch angle:

$$\frac{d\mu}{dt} = \frac{ec}{2pE} \sum_{\pm} (\mp i) p_{\perp 0} \exp(\mp i(\phi_0 - \Omega t)) \delta B_{\pm}(\mathbf{r}). \quad (2.146)$$

In order to split the expression into parallel and perpendicular components, the magnetic field is decomposed into Fourier modes

$$\delta B_{\pm}(\mathbf{r}) = \frac{1}{8\pi^3} \int d^3k \delta B_{\pm}(\mathbf{k}) \exp(i\mathbf{k}\mathbf{r}) \quad (2.147)$$

and the exponential function of the parallel components of $\mathbf{k}$ and $\mathbf{r}$ is expressed via the generating Bessel functions J_n :

$$\exp(iz \sin \theta) = \sum_{n=-\infty}^{\infty} J_n(z) \exp(in\theta). \quad (2.148)$$

Substituting this into Eq. 2.146 yields the sought-after decomposition:

$$\begin{aligned} \frac{d\mu}{dt} = & \frac{ec}{2pE} \sum_{\pm} (\mp i) p_{\perp 0} \exp(\mp i\phi_0) \\ & \cdot \frac{1}{8\pi^3} \int d^3k \delta B_{\pm}(\mathbf{k}) \sum_n J_n \left(\frac{k_{\perp} v_{\perp}}{\Omega} \right) \\ & \cdot \exp(-in(\Psi - \phi_0) + it(k_{\parallel} v_{\parallel 0} \pm \Omega - n\Omega)), \end{aligned} \quad (2.149)$$

where Ψ is the azimuth angle in $\mathbf{k}$ -space so that $\mathbf{k} = (k_{\perp} \cos \Psi, k_{\perp} \sin \Psi, k_{\parallel})$. The expression can be simplified by using the property of Alfvén waves, that their fluctuations are perpendicular to both the magnetic field and the wave vector, i.e. in the direction $\mathbf{k} \times \mathbf{B}_z$ and can be written as $\delta B_{\pm}(\mathbf{k}) = \delta B(\mathbf{k})(\mp i) \exp(\pm i\Psi)$. Finally, a time integral yields the change of the pitch angle from the initial value:

$$\begin{aligned} \Delta\mu = \mu - \mu_0 = & -\frac{ecp_{\perp 0}}{2pE8\pi^3} \int d^3k \delta B(\mathbf{k}) \sum_n \frac{2n\Omega}{k_{\perp} v_{\perp}} J_n \left(\frac{k_{\perp} v_{\perp}}{\Omega} \right) \exp(-in(\Psi - \phi_0)) \\ & \times \int_0^t dt' \exp(it'(k_{\parallel} v_{\parallel 0} - n\Omega)). \end{aligned} \quad (2.150)$$

This is an important result, as it can relatively easily be used to compare numerical results with theory. It can also be used in the definition 2.143 to find an expression for the pitch angle diffusion coefficient $D_{\mu\mu}$ by calculating $\langle \Delta\mu \Delta\mu \rangle$. With $\langle \delta B(\mathbf{k}) \delta B^*(\mathbf{k}') \rangle = (2\pi)^6 P(\mathbf{k}) \delta(\mathbf{k} - \mathbf{k}')$ we get:

$$\begin{aligned} \langle \Delta\mu \Delta\mu \rangle = & \frac{e^2 c^2 p_{\perp 0}^2}{p^2 E^2} \int d^3k P(\mathbf{k}) \sum_n \frac{n^2 \Omega^2}{k_{\perp}^2 v_{\perp}^2} J_n^2 \left(\frac{k_{\perp} v_{\perp}}{\Omega} \right) \\ & \times \int_0^t dt' \int_0^t dt'' \exp(i(t' - t'')(k_{\parallel} v_{\parallel 0} - n\Omega)). \end{aligned} \quad (2.151)$$

The time integral can be evaluated approximatively for $t \rightarrow \infty$

$$\int_0^t dt' \int_0^t dt'' \exp(i(t' - t'')(k_{\parallel} v_{\parallel 0} - n\Omega)) \simeq 2\pi t \delta(k_{\parallel} v_{\parallel 0} - n\Omega) \quad (2.152)$$

yielding the coefficient

$$D_{\mu\mu} = \frac{\pi e^2 c^2}{E^2} (1 - \mu^2) \int d^3k P(\mathbf{k}) \cdot \sum_n \frac{n^2 \Omega^2}{k_{\perp}^2 v_{\perp}^2} J_n^2 \left(\frac{k_{\perp} v_{\perp}}{\Omega} \right) \delta(k_{\parallel} v_{\parallel 0} - n\Omega). \quad (2.153)$$

The physical interpretation of this equation is that a particle interacts with a wave mode in discrete intervals n resonantly. For $n = 0$ there is no scattering on the shear modes, as Alfvén waves cannot interact with particles moving along the magnetic field. This is the so-called Cherenkov resonance that can only be generated by pseudo Alfvén waves. For $n = \pm 1$ the parallel component of the wave dominates and the resonance condition Eq. 2.144 can be applied. For $|n| > 1$ the contributions from the perpendicular component of the wave become significant and the Bessel functions J_n cannot be neglected. The derivation has been made in the limit of time scales much larger than the correlation time. At small time scales the wave will interact with arbitrary particles in the ballistic regime, but only those particles near the resonance condition will be scattered effectively. After a longer development time, when multiple wave-particle collisions occurred, the shape of the coefficient, which is effectively a distribution function of the scattering processes, will approach that of a δ -function. Particles that don't fulfil the resonance condition are effectively not scattered on average.

2.3.3 Wave generation with plasma instabilities

In the previous section, wave-particle interactions were discussed with the focus on their effects on the particle. The other side is, of course, to consider which effects the interactions have on the waves (or the plasma). For a complete description of this topic we would need to introduce the theory of plasma instabilities that describes the development of a perturbation that will grow, if it is not dampened effectively (stable plasma), and generate waves. A plasma instability is a non-linear effect that changes the dispersion relation of the plasma waves. Due to the large number of different instabilities and wave modes they can induce, we will limit the discussion here to the relevant phenomenon for this work, the streaming instability. A good overview of the remaining plasma instabilities can be found in Baumjohann and Treumann (1997); Koskinen (2011); Melrose (1989). We already encountered the streaming instability implicitly in the first chapter, as a disturbance mechanism of the solar wind plasma through streams of high-energy particles e.g. from flares and CMEs. Depending on the energy of the particles, non-linear generation of Alfvén waves can occur.

In a resonant particle-wave interaction that fulfils the resonance condition 2.144 the

energy density of the wave W changes by

$$\frac{dW}{dt} = -\Gamma(\mathbf{k})W(\mathbf{k}). \quad (2.154)$$

The change can be positive or negative depending on the sign of Γ and the particle will be decelerated or accelerated in the process, respectively. According to Vainio et al. (2003), the energy growth rate depends on the distribution function $f(\mathbf{r}, \mathbf{p}, t)$ of the interacting particles and can be written in the QLT approximation as:

$$\Gamma(k_{\parallel}) = \frac{1}{W} \frac{dW(\mathbf{k})}{dt} = \frac{\pi\Omega}{2nv_A} \int d^3p v(1 - \mu^2)|k|\delta\left(k_{\parallel} + \frac{\Omega}{v\mu}\right) \frac{\partial f}{\partial \mu}. \quad (2.155)$$

The resonance condition in the form of a δ -function can be converted to a μ -independent form according to Skilling (1975) and Bell (1978). The energy growth rate then simplifies to:

$$\begin{aligned} \Gamma(\mathbf{r}, p, t) &= \frac{\pi}{2} \frac{\Omega p}{nv_A} 2\pi \int_{-1}^1 d\mu v \mu p^2 f \\ &= \frac{\pi}{2} \frac{\Omega p}{nv_A} S(\mathbf{r}, p, t). \end{aligned} \quad (2.156)$$

The function $S(\mathbf{r}, p, t)$ is the injection rate of the in-streaming particles - the number of particles of unit momentum streaming through a unit area per unit time. This equation shows the direct connection between an influx of a particle population, for example SEPs, and the generation of wave modes in a plasma.

This generation and growth of wave modes corresponds to the explanation of the Fermi process at the shock of a CME event, as described in 1.2.4. The shock region is composed of self-generated waves for example from a coronal mass ejection. The species of waves especially relevant for SEP events are the already mentioned low frequency Alfvén waves. They are generated preferably through SEPs and, in turn, interact resonantly with protons by accelerating them. It's easy to picture a flux of particles with a high momentum streaming through the shock region and partially transferring it's energy to the Alfvén modes. This reduces the mean free path for the following particles, since the scattering probability is higher with the stronger wave mode. Higher scattering probability leads to more shock front crossings and, in turn, to a more efficient acceleration of particles. This effect is known as streaming-limited intensity of SEPs (Reames and Ng, 1998). Measurements of gradual SEP events show an upper limit in intensity, depending on particle energy. Another strong indication of this mechanism of self-generated waves is the strong rise in intensity, when the shock passes the probe, explained by particles captured in wave-particle-interactions near the shock front.

3 Methodology and Numerics

This chapter reflects the two-fold nature of the work done in the scope of the thesis at hand. On the one hand, a large share of the work consisted of improving and using the incompressible MHD code GISMO (Lange, 2012) to obtain realistic turbulence scenarios in which particle transport was studied. In order to understand the range of scenarios in which the particle transport results are applicable, a description of the methods, assumptions and limitations in the code is necessary and will be given in 3.1. Furthermore, some results of numerical testing and validation beyond the already extensive work in Lange (2012) will be presented in chapter 4.

On the other hand, the large amounts of raw data is of little use without careful statistical analysis and interpretation. The relevant statistical basics as well as the new analysis methods developed in the frame of this work will be presented in 3.2.

A brief excursus in 3.3 will be made to give an overview of other approaches to particle transport, such as stochastic differential equations and Perlin noise simulations, which were mainly used for comparison and validation purposes.

3.1 Plasma simulation

3.1.1 Pseudo-spectral description

The foundation of the work at hand is the representation of plasma turbulence through a numerical solution of the ideal MHD equations as they are presented in 2.1.6, specifically in their Elsässer variables notation (Eq. 2.73-2.76). In general, the equations cannot be solved analytically, and a transition to a numerical description on a discretised grid gives rise to a Riemann problem. Several approaches to tackle Riemann problems are available - so-called Riemann solvers - like piecewise linearisation in the Godunov method (Zrake and MacFadyen, 2012), or weighting of the spacial derivatives in the CWENO-scheme (Levy, Doron et al., 1999). They share the disadvantage of high numerical dissipation that is intrinsic to the scheme and cannot be controlled. In a simulation of turbulence this can lead to a severely limited inertial range of the spectrum, somewhat diminishing its usefulness.

A completely different approach was chosen here in bypassing the spatial derivatives by using spectral methods. The idea is to disassemble the solution of the PDE into “basis functions”, specifically as a Fourier series with a Fast-Fourier-Transform algorithm. Since the nonlinear terms in the MHD equations still have to be treated in real space, the method employed here is categorised as pseudo-spectral. The Fourier transform of a

function $f(\mathbf{x})$ in real space can be defined as

$$FT(f(\mathbf{x})) = F(\mathbf{k}) = c \cdot \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} f(\mathbf{x}) \quad (3.1)$$

and the inverse transform

$$FT^{-1}(f(\mathbf{k})) = F(\mathbf{x}) = c \cdot \int d^3k e^{-i\mathbf{k} \cdot \mathbf{x}} f(\mathbf{k}). \quad (3.2)$$

The normalisation c must be chosen so that a consecutive forward and inverse transform of a function deliver the function itself. Assuming commutativity of differentiation and integration, which is always given for a Fourier integral, we can now transform the spacial derivatives

$$\mathbf{f}(\mathbf{x}) = \nabla u(\mathbf{x}) \quad (3.3)$$

into their Fourier space form

$$F(\mathbf{k}) = c \cdot \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} \nabla u(\mathbf{x}) = i\mathbf{k} \tilde{u}(\mathbf{k}), \quad (3.4)$$

where the tilde-notation is used to indicate quantities in the Fourier space.

Analogous transformation of the MHD equations 2.52-2.55 yields:

$$\frac{\partial \tilde{\mathbf{u}}}{\partial t} = \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} (\mathbf{b} \cdot \nabla \mathbf{b}) - \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} (\mathbf{u} \cdot \nabla \mathbf{u}) - i\mathbf{k} \tilde{P} - \nu_n \mathbf{k}^{2n} \tilde{\mathbf{u}} \quad (3.5)$$

$$\frac{\partial \tilde{\mathbf{b}}}{\partial t} = \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} (\mathbf{b} \cdot \nabla \mathbf{u}) - \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} (\mathbf{u} \cdot \nabla \mathbf{b}) - \nu_n \mathbf{k}^{2n} \tilde{\mathbf{b}} \quad (3.6)$$

$$\mathbf{k} \cdot \tilde{\mathbf{u}} = 0 \quad (3.7)$$

$$\mathbf{k} \cdot \tilde{\mathbf{b}} = 0. \quad (3.8)$$

It can be easily seen why a fully spectral approach for MHD is impossible. The non-commutativity of the nonlinear terms and the closing condition 2.76 prevent an easy computation of the terms in Fourier space, since e.g. $\tilde{\mathbf{u}} \cdot \tilde{\mathbf{u}} \neq \widetilde{\mathbf{u} \cdot \mathbf{u}}$. A closed representation can be found in a vector component notation. The components are denoted by α, β, γ and the Einstein summation convention is used. The pressure term in Fourier space can then be written as:

$$\tilde{P} = -\frac{k_\beta k_\gamma}{k^2} \left(\widetilde{u_\beta u_\gamma} - \widetilde{b_\beta b_\gamma} \right). \quad (3.9)$$

This corresponds to the closing condition for the MHD equations in real space:

$$\Delta P = \nabla \mathbf{b} \otimes \nabla \mathbf{b} - \nabla \mathbf{v} \otimes \nabla \mathbf{v}. \quad (3.10)$$

The same can be done for the remaining nonlinear terms, yielding the MHD equation

system in Fourier space in component notation:

$$\frac{\partial \tilde{u}_\alpha}{\partial t} = -ik_\gamma \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) \left(\widetilde{u_\beta u_\gamma} - \widetilde{b_\beta b_\gamma} \right) - \nu k^2 \tilde{u}_\alpha \quad (3.11)$$

$$\frac{\partial \tilde{b}_\alpha}{\partial t} = -ik_\beta \left(\widetilde{u_\beta b_\alpha} - \widetilde{b_\beta u_\alpha} \right) - \nu k^2 \tilde{b}_\alpha \quad (3.12)$$

$$k_\alpha \tilde{u}_\alpha = 0 \quad (3.13)$$

$$k_\alpha \tilde{b}_\alpha = 0. \quad (3.14)$$

Using the definition of the Elsässer variables 2.72 and the fact that summations over β and γ are independent of the α -summation yields the final form of the equations with Elsässer notation in Fourier space:

$$(\partial_t - v_A k_z) \tilde{w}_\alpha^- = i \frac{k_\alpha k_\beta k_\gamma}{k^2} \widetilde{w_\beta^+ w_\gamma^-} - i k_\beta \widetilde{w_\alpha^- w_\beta^+} - \frac{\nu}{2} k^{2h} \tilde{w}_\alpha^- \quad (3.15)$$

$$(\partial_t + v_A k_z) \tilde{w}_\alpha^+ = i \frac{k_\alpha k_\beta k_\gamma}{k^2} \widetilde{w_\beta^+ w_\gamma^-} - i k_\beta \widetilde{w_\alpha^+ w_\beta^-} - \frac{\nu}{2} k^{2h} \tilde{w}_\alpha^+ \quad (3.16)$$

$$k_\alpha \tilde{w}_\alpha^\pm = 0. \quad (3.17)$$

It is easy to see on these equations, why it is necessary to artificially add hyperdiffusivity via the parameter h . As the pseudo-spectral approach replaces all spacial derivatives with simple multiplications, there is no intrinsic numerical dissipation at all, which is normally a big advantage and a crucial reason to use spectral methods. In turbulence, however, dissipation is a crucial element of the energy cascade, as has been discussed in section 2.2. So depending on the problem at hand, hyperdiffusivity is used to amplify the energy dissipation. The potency of k-space dissipation is strengthened and wave modes at high wave numbers are damped more efficiently. As a rule of thumb, a parameter of $h = 2$ is enough for sufficient dissipation.

3.1.2 Fourier transform implementation

For a numerical approach the Fourier integral in Eq. 3.1 and 3.2 must be discretised, turning it into a sum in every dimension:

$$X_k = c \sum_{n=0}^{N-1} x_n \exp\left(\frac{2\pi i}{N} nk\right) \quad (3.18)$$

$$x_n = c \sum_{k=0}^{N-1} X_k \exp\left(-\frac{2\pi i}{N} nk\right). \quad (3.19)$$

This form is called the “discrete Fourier transform” or DFT. Specifically, in three dimensions we have three nested sums over variables $X(k_1, k_2, k_3)$ and $x(n_1, n_2, n_3)$ that are represented on a discretised grid of $K_1 \times K_2 \times K_3$ and $N_1 \times N_2 \times N_3$ cells, respectively. The numbers $n \in \{0; N-1\}$ represent the grid points in a simulation box of size L and k

defines the corresponding wave number via $k = \{-N/2 \cdots 0 \cdots N/2\} \cdot 2\pi/L$. The usual convention in numerical simulations is to shift the summation index so that $k \in \mathbb{N}$. The negative wave numbers from $-N/2 \cdot 2\pi/L$ to $-1 \cdot 2\pi/L$ are shifted by N so that they are represented by the summation index from $N/2$ to N . The positive wave numbers stay unchanged from 0 to $N/2$.

As the values on the spacial grid are real numbers, their complex Fourier transform has to be hermitian symmetric due to information conservation:

$$X_k = X_{N-k}^*. \quad (3.20)$$

This property is used by the numeric DFT algorithm to reduce the number of calculations, since it's not necessary to calculate the three-dimensional grid in Fourier space completely. Instead, the size of the last dimension is halved, as a variable in a spacial grid of the size $N_1 \times N_2 \times N_3$ is completely determined in a wave number grid of the size $K_1 \times K_2 \times K_3/2$. In an actual numeric implementation the last dimension is usually $K_3/2 + 1$, since the wave numbers 0 and $N/2$ don't have symmetrical partners.

The simple discretisation approach in Eqs. 3.18 and 3.19 is rarely used for larger field transformations, since its runtime scales quadratically with the number of grid points, as can easily be seen from the equations. Instead, a divide-and-conquer approach was suggested by Cooley and Tukey (1965)¹ that splits the sums recursively in odd and even indices, effectively reducing the number of complex operations and improving the scaling for the run time from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$. This approach is known as the fast Fourier transform or FFT, and has the advantages of being widely used and having a large variety of implementations in different programming languages available.

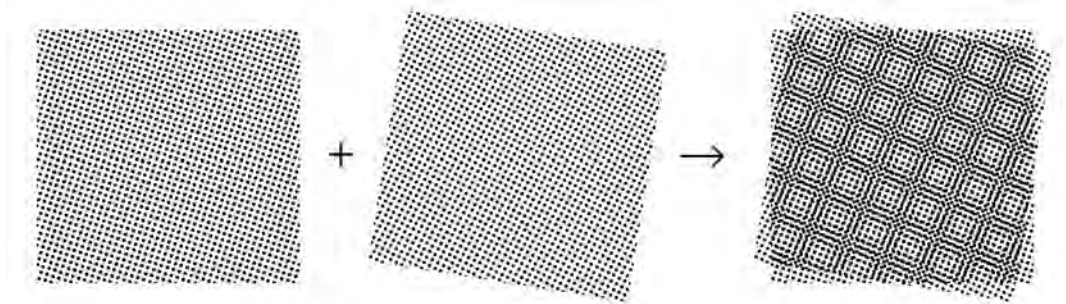


Fig. 3.1: A moiré pattern generated by the superposition of two grids of points, rotated by 15° relative to each other. Source: Wikimedia Commons

The variables are discretised on the grid by sampling at grid points. According to the Nyquist-Shannon sampling theorem, signals that vary faster than the Nyquist frequency, which is the sampling frequency halved, are not reproduced correctly. Instead they are distorted and interpreted as lower frequencies due to undersampling. This undesirable

¹A form of the algorithm was known to Carl Friedrich Gauss around 1805, and was subsequently rediscovered in several other works

effect is called aliasing and is encountered in every day life for example in the wagon-wheel effect or in Moiré patterns. The first is a temporal effect, where the wheels of a moving car or the propellers on a plane seem to stand still or move backwards in a film, because the sample rate of 24-30 frames per second is not enough to reproduce the fast motion. A Moiré pattern is the spacial variation of the aliasing effect that can be seen when two regular grid-like structures, that are very similar but not completely identical, overlap. In order for the patterns to not be identical, a small displacement or rotation is sufficient. An example of a Moiré pattern is shown in Fig. 3.1. In general, less serious aliasing effects can appear even below the Nyquist frequency, resulting from the discretisation of the Fourier integral 3.1. According to Canuto (1988) aliasing errors can be defined as the difference between the continuous Fourier transformation (Eq. 3.1) and the DFT (Eq. 3.18). To avoid these errors, antialiasing methods are employed. The MHD code in this work uses zero-padding according to Orszag (1971), where all wave modes above a certain wave number $k > oN$ are set to zero. The fraction o lies between the best case scenario $o = 1/2$ that corresponds to the Nyquist theory and the worst case with $o = 1/3$ that emerges on a chequered grid, where neighbouring cells alternate between two values. The wave modes in the lower third are never affected by the padding and always alias-free, and are called active modes. Consequently, to get a simulation with a chosen number of active modes, the grid resolution must be tripled.

The discrete Fourier transform is the central part of the spectral MHD code. We use the popular and mature implementation of the FFT that was developed by Matteo Frigo and Steven G. Johnson at the Massachusetts Institute of Technology: the Fastest Fourier Transform in the West, or FFTW. In order to simulate systems with grid sizes larger than 128^3 , a parallel version of the FFTW was chosen: the Parallel 3-Dimensional Fast Fourier Transform (P3DFFT), developed by Dmitry Pekurovsky at the San Diego Supercomputer Center. It uses FFTW3 as a foundation and parallelises it with MPI (Message Passing Interface) in Fortran. It exploits the hermitian symmetry of the Fourier transform by splitting the last dimension in half to save computation time and storage space.

3.1.3 Particle pusher

The test particle simulations are based on solving the relativistic Lorentz force equation

$$\frac{d}{dt}\gamma\mathbf{v} = \frac{q}{mc} (c\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (3.21)$$

which is an ordinary differential equation in time t and can easily be discretised with the Euler-method:

$$\frac{\gamma(\mathbf{v}^{t+\Delta t} - \mathbf{v}^t)}{\Delta t} = \frac{q}{mc} (c\mathbf{E} + \mathbf{v}^t \times \mathbf{B}). \quad (3.22)$$

Since Euler is a first-order method, this would lead to large errors in the resolution of the small scale particle gyration motion, especially for larger time steps Δt . For a stable solution of the gyration orbit the time step would have to be very small. Alternatively,

an implicit method suggested by Boris (1970) can be used. The idea is to divide the iteration in three parts. First the particle is accelerated by the electric field over half a time step, then it is rotated by the magnetic field and finally accelerated by the E-field again. This translates directly to the discretised equations for a Boris-push time step:

$$\mathbf{v}^{t-\Delta t/2} = \mathbf{v}^- - \frac{q\mathbf{E}}{2m}\Delta t \quad (3.23)$$

$$\frac{\mathbf{v}^+ - \mathbf{v}^-}{\Delta t} = \frac{q}{2\gamma m} (\mathbf{v}^+ - \mathbf{v}^-) \times \mathbf{B} \quad (3.24)$$

$$\mathbf{v}^{t+\Delta t/2} = \mathbf{v}^+ + \frac{q\mathbf{E}}{2m}\Delta t. \quad (3.25)$$

The biggest advantage of this implicit method is its numerical stability. A big problem with explicit approaches like Eq. 3.22 is that at larger time steps particles deviate from the correct circular motion onto spiral trajectories. The Boris-push merely deviates from the analytically correct Larmor radius, while still maintaining “correct” gyration. Even for $\Delta t \ll 1/\Omega$ the circular motion is stable, converging towards the adiabatic drift. A small issue is that the electric field is shifted by half a time step relative to the magnetic field in the equations. This can be neglected, however, since the evolution time scales of the fields are much larger than the gyration time scales of particles.

To calculate the Boris-push time step, Eq. 3.24 must be solved for $\mathbf{v}^+$. This can be done by introducing auxiliary vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ following bir:

$$\mathbf{h}_1 = \frac{q\mathbf{B}}{2\gamma mc}\Delta t \quad (3.26)$$

$$\mathbf{v}' = \mathbf{v}^- + \mathbf{v}^- \times \mathbf{h} \quad (3.27)$$

$$\mathbf{h}_2 = \frac{2\mathbf{h}_1}{1 + \mathbf{h}_1 \cdot \mathbf{h}_1} \quad (3.28)$$

$$\mathbf{v}^+ = \mathbf{v}^- + \mathbf{v}' \times \mathbf{h}_2. \quad (3.29)$$

As discussed by Vay (2008) the Boris push method fails for ultra-relativistic particle velocities. In this case the energy conservation is violated, as ideal Ohm’s law is no longer valid and particles experience pseudo forces. The method can still be applied to mildly relativistic particles, however, as significant deviations appear for Lorentz factors larger than 10^3 .

3.1.4 Code structure

In order to study particle motion in realistic MHD turbulence, the C++ hybrid code GISMO (Genuine Incompressible Spectral Magnetohydrodynamic Odyssey), developed by Lange (2012), was used in the frame of this work. This chapter gives an overview of code structure as well as some details of the implementation of the techniques presented in the previous sections. GISMO consists of two independent parts: GISMO-MHD, which simulates the evolution of the turbulent plasma through the MHD equations, and GISMO-Particles, which calculates the particle motion from the electromagnetic fields provided

by GISMO-MHD.

Gismo-MHD

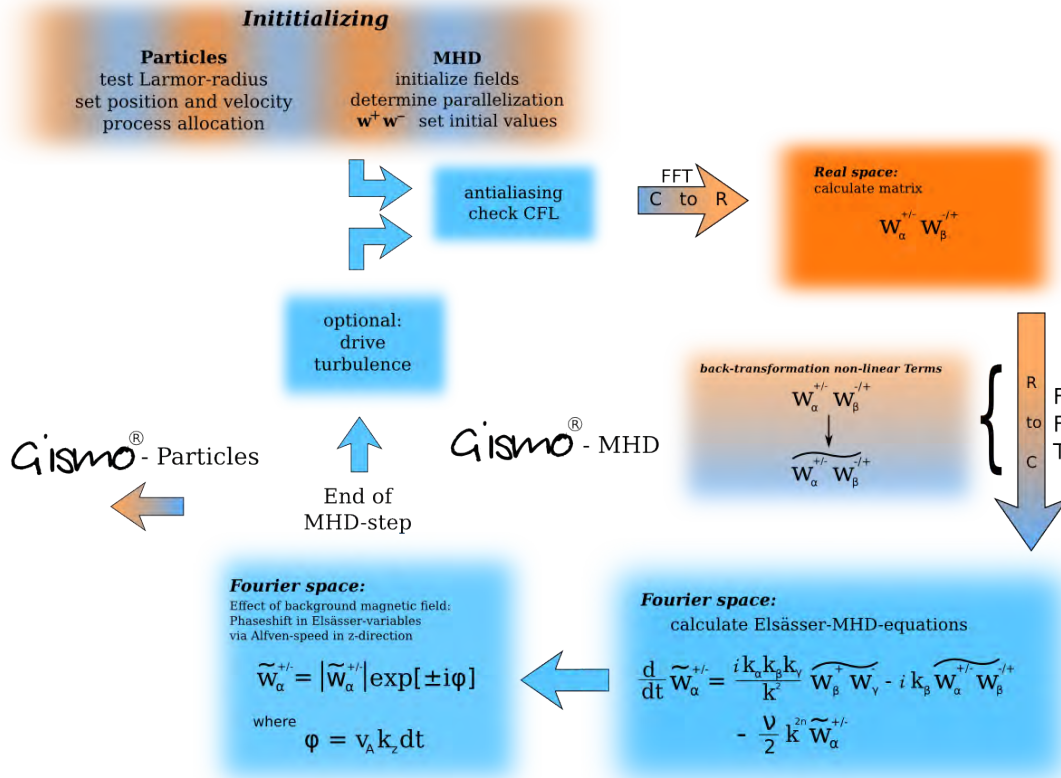


Fig. 3.2: Structure chart of the MHD calculations by GISMO-MHD. Steps with a blue background are calculated in Fourier space, orange background indicates calculations in real space. At the end of each MHD time step the fields can be passed to the particle pusher. Source: modification of Lange (2012)

As can be seen from the diagram 3.2, the major part of the calculations takes place in the Fourier space. After setting the simulation parameters like system size, grid and time resolution, Alfvén velocity, resistivity and so on, the Elsässer fields $\mathbf{w}^\pm$ are set to their initial values. This can be done by initialising a single propagating wave or multiple wave packets that generate turbulence through non-linear interactions. The advantage of the Elsässer notation becomes apparent at this point, as an Alfvén wave is fully defined by setting the amplitude at a chosen wave vector k in one component of the Elsässer fields $\mathbf{w}^\pm$. For example setting

$$\tilde{w}_y^+(k_x = 1, k_y = 0, k_z = 0) = A + i \cdot 0 \quad (3.30)$$

would result in a cosine wave of the lowest possible frequency along x ($k_x = 1$) with an amplitude A in y propagating in the positive direction along the background field. The polarisation of the wave can be changed from linear to circular by initialising the x and y components of the fields $\mathbf{w}^\pm$ adhering to the condition $iE_x/E_y = \pm 1$. The sign indicates the left/right handedness of the polarisation. For example, a right handed circularly polarised wave in Elsässer notation would be generated by

$$\tilde{w}_x^\pm(k_x, k_y, k_z) = 0 + i \cdot \frac{A}{2} \quad (3.31)$$

$$\tilde{w}_y^\pm(k_x, k_y, k_z) = -\frac{A}{2} + i \cdot 0. \quad (3.32)$$

To initialise a turbulence development in the plasma, the system can be seeded with arbitrary wave packets, which then interact with each other. It is advisable to initialise modes at low wave numbers, since in a turbulent system the energy cascades from larger to smaller scales until it dissipates, as discussed in 2.2. GISMO offers the possibility to inject arbitrary amplitude distributions with random phases either isotropic at $|\mathbf{k}| < k_{\text{init}}$ or anisotropic at $\mathbf{k} < \mathbf{k}_{\text{init}}$, for example to allow a faster development of intermediate turbulence in the parallel direction. Pure slab-turbulence can also be simulated by only setting field values along the background field direction. This can and should be done over a broad range of wave numbers, since pure slab turbulence has no non-linear development according to the Eq. 3.15-3.17. After initialisation, the $\mathbf{w}^\pm$ fields are checked for solenoidality with a divergence cleaner, which enforces $\nabla \cdot \mathbf{w}^\pm = 0$ with a projection operator. The fields are then scaled to reflect the physical problem at hand with the energy injection parameter that can be chosen to represent the source and type of injection, for example energy from solar processes like flares and CMEs or supernovae. The energy can be injected continuously throughout the simulation for driven turbulence or only once in the beginning for decaying turbulence.

Once the initialisation is finished, the iteration of the MHD equations begins. First, antialiasing via zero-padding is carried out to prepare for the coming FFT. Next, the CFL condition is checked by calculating $c = u_{\text{max}}\Delta t/\Delta x$ (Courant et al., 1928). The CFL number indicates how far an element can move during a numeric time step and it must be smaller than a problem-dependent constant to guarantee stability. In general the constant is determined by the maximal characteristic of the problem, e.g. the Alfvén velocity in the case of subalfvénic MHD. For pseudo-spectral methods this cannot be adopted directly, because the spacial coordinates are discretised, but the discretisation does not enter the calculations directly. Still, there is an upper boundary on how far a fluid element can propagate during a time step to preserve numerical stability, that is influenced by the time step size and the convection and diffusion strength Gottlieb and Tadmor (1991).

The next step is the Fourier transform of the Elsässer fields into real space (C2R, complex to real). Two important details should be taken into account here. The first is the normalisation of the FFT. The normalisation of the forward and reverse transforms are not fixed, but chosen by convention. The only condition is that the product of

the normalisation factors is $1/N$. In Gismo, a factor $1/\sqrt{N}$ is multiplied to the fields after the C2R transform. After calculating the quadratic interaction terms, they have a resulting normalisation factor $1/N$ and don't need to be normalised in the reverse transform. One should keep in mind, that the real space fields differ from the Fourier space fields by a factor of $\sqrt{N}$. The second point to consider is the conservation of the hermitian symmetry during the Fourier transform. When transforming from real to complex (R2C) the P3DFFT ensures this condition automatically. In the reverse C2R transform this has to be done manually. The Fourier space fields are already reduced by the hermitian symmetric half, so we only need to treat the border regions. Since the summation in the FFT always runs over both hermitian symmetric partner values, we can weigh one of the values with a factor 2 and set the other to zero. This produces correct real space fields, that can be then transformed with R2C to hermitian symmetric k-space fields, that have to be normalised with the factor $\sqrt{N}$.

The transformation to the real space serves the purpose of calculating the non-linear MHD terms - the only calculation step not done in the Fourier space. This is done by calculating the tensor

$$\mathbf{w}^- \otimes \mathbf{w}^+ = \begin{pmatrix} w_x^- w_x^+ & w_x^- w_y^+ & w_x^- w_z^+ \\ w_y^- w_x^+ & w_y^- w_y^+ & w_y^- w_z^+ \\ w_z^- w_x^+ & w_z^- w_y^+ & w_z^- w_z^+ \end{pmatrix} \quad (3.33)$$

as it contains all the interaction terms appearing in Eq. 3.15-3.17.

After the non-linear terms have been transformed back to k-space, the iteration of a time step of the MHD equations can be performed. This is done without the effect of the background magnetic field term $v_A k_z \tilde{w}_\alpha^\pm$ at first. Subsequently, the effect is taken into account through a phase-shift of the Elsässer-fields. Since this can be done semi-analytically by calculating

$$\tilde{w}_\alpha^\pm(t + dt) = |\tilde{w}_\alpha^\pm(t)| e^{i\phi} = |\tilde{w}_\alpha^\pm(t)| e^{iv_A k_z dt}, \quad (3.34)$$

the approach has the advantage of better stability and precision as compared to the discretisation of the term.

This step completes the iteration cycle of the MHD equations. Depending on the chosen turbulence scenario, energy can be injected into the system at this point to drive the turbulence. If test particles are simulated, the fields will be passed to the Gismo-Particles sub-routine. After that the iteration cycle begins anew.

Gismo-Particles

The test particle simulation is a sub-routine of the Gismo-MHD program, so particle parameters are set within the general initialisation routine, if a flag indicates that particles should be simulated. The particle species is specified by the mass and charge. The specified number of particles is distributed randomly and homogeneously in the three dimensional simulation volume. The momentum modulus is usually set to a constant value, the azimuth angle Φ in the p_x - p_y -plane is random as well. The pitch-angle μ can

be distributed randomly between -1 and +1 (Type I), set to a constant value (Type II), or set to follow a pre-determined distribution function, for example in the form of a half-parabola that is necessary for the application of the new $D_{\mu\mu}$ calculation method (Type III). After setting the components of the particle momenta in spherical coordinates, a transformation to Cartesian coordinates is made.

A further intialisation step is required in case of a particle simulation. Since the kinetic particle model operates on much smaller time scales than the magnetohydrodynamic turbulence model, the time resolution needs to be adjusted. The particle integration is always stable due to the use of the implicit Boris-push method, but insufficient time resolution will introduce an error to the gyration radius. In order to set the time step, a single particle is injected into the background field. After a short transient phase, multiple measurements of its gyration radius are taken and the relative deviation from the analytical value $r_L = v_T/\Omega$ is calculated. The time step Δt is then reduced successively until either the desired precision or the maximal sensible number of steps per gyration is reached.

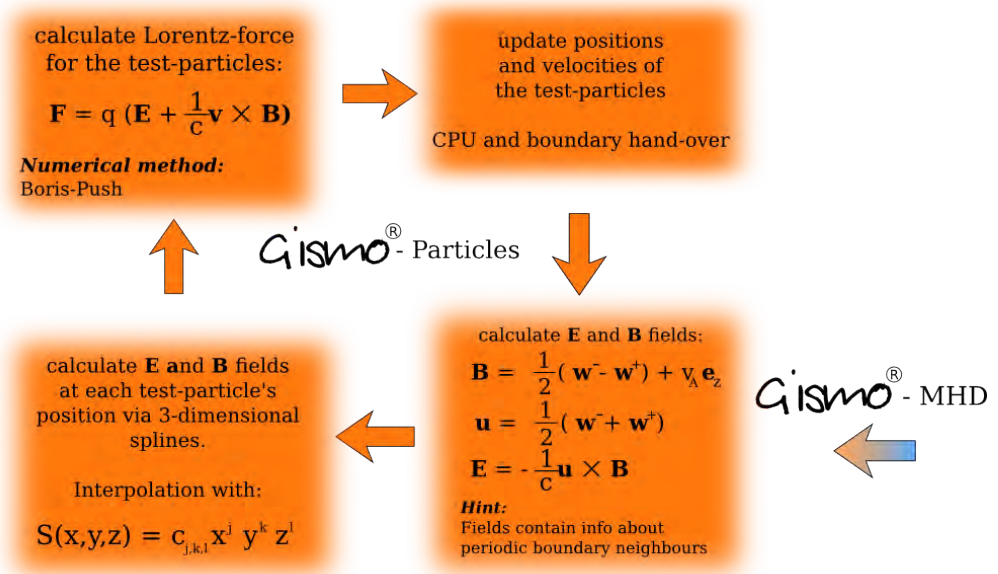


Fig. 3.3: Schematic flow chart of the test particle simulation. GISMO-Particles is a sub-routine of GISMO-MHD and is initialised and controlled from the main program. All particle calculations are done in the real space. Source: modification of Lange (2012)

After the Elsässer fields are provided by GISMO-MHD, the particle iteration begins as illustrated in Fig. 3.3. Three significant differences of the Elsässer fields to the electric and magnetic fields used in the particle simulation must be considered. First, the fields must be transformed from the forward and backward propagating waves $\mathbf{w}^\pm$ to $\mathbf{B}$ - and $\mathbf{u}$ -fields according to Eq. 2.72. The ideal Ohm's law $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$ can then be used

to calculate the electric fields. Secondly, the Elsässer fields are implicitly calculated with periodic boundaries, as their evolution is computed in Fourier space. For the real space fields, the periodic boundaries and the CPU boundaries have to be handled explicitly. This is done by adding margins with the width of two cells to every boundary and filling the margin cells with field values from the opposite side. The third point arises from the different representation of the spacial variables in the MHD and particle codes. The Elsässer fields in the MHD simulation are discretised on a grid with the resolution $N = N_x \times N_y \times N_z$, whereas the particles are described with continuous spacial coordinates in the simulation volume. This is alleviated by spline interpolation of the fields on the grid to the position of the particle. A good compromise of accuracy and computational effort is achieved by interpolation with a third-order polynomial:

$$S(x) = c_0 + c_1x + c_2x^2 + c_3x^3. \quad (3.35)$$

In three dimensions, using Einstein's summation convention this yields in the matrix equation:

$$S(x, y, z) = c_{ijk}x^iy^jz^k. \quad (3.36)$$

The coefficient matrix c_{ijk} needs to be calculated only once, as it is independent of the grid resolution. The result of the calculation in Weiss (2011) is used in GISMO. The interpolation of the fields from two neighbouring cells in each direction is the reason why the periodic boundaries must be added manually to insure a correct calculation of the fields for particles near a boundary.

After the fields have been interpolated to the position of the particle, the Boris-push (Eq. 3.26-3.29) can be applied to calculate the new particle velocity. The position of the particle can then be updated with $\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{v}^+dt$. Finally, the check is made, if the new position coordinates are still inside the simulation box. If the particle crossed the boundary, it is transferred to the opposite side to enforce periodic boundary conditions. In the frame of the parallelisation of the code it must also be checked, if the particle left its local coordinates, and transferred to the next CPU where necessary.

For further details on the implementation and parallelisation of GISMO the reader is referred to Lange (2012). This chapter shall be concluded with a short assessment of the applicability and the limits of the model. The goal of this work is to investigate particle transport in plasma turbulence on a microphysical level, specifically in the frame of the heliosphere. As has been laid out in the previous chapters, heliospheric plasma is a rather complex subject that is near to impossible to describe with a single method in all of its totality. So the question of applicability becomes a question of scales. For length scales larger than the mean free path of the particles, the interplanetary medium can be described as a collision-dominated plasma, which invites a magnetohydrodynamic approach as long as kinetic effects above the ion cyclotron frequency can be neglected. The next limitation comes with the decision between compressible and incompressible forms of the MHD equations. The first is a more general description, but it is more limited in its implementation. The inevitably numeric dissipation strongly limits the

inertial range of the energy spectrum as compared to observations. The incompressible method offers a realistic depiction of plasma turbulence down to the smallest scales. As has been discussed in chapter 1 many phenomena lie in the incompressible regime and wide ranges of the heliosphere are dominated by Alfvén waves due to the strong damping of compressible modes. In particular, the solar wind can largely be considered to be incompressible. Care should be taken when choosing the extent of the simulation volume, however, as it should be small enough to prevent negative effects. For example, the model is not applicable on scales where the curvature of the Parker spiral becomes significant or in special regions like CIRs where compressibility is non-negligible. A further advantage of incompressible MHD is the availability of established theoretical descriptions (Goldreich and Sridhar, 1995; Sridhar and Goldreich, 1994), which make predictions that can be compared with the numerical results. A comprehensive theory of compressible turbulence is harder to come by.

With GISMO, a wide range of scenarios in incompressible MHD turbulence can be investigated. The driving of the turbulence can be tuned to fit the desired scenario. It's possible to drive specific wave modes, make the energy injection isotropic or anisotropic, continuous or singular (to study the decay of turbulence). The parallelisation of the code allows to use it on supercomputers to achieve a large spectral and temporal resolution.

The simulation of the particle transport provides a very exact and realistic description due to the choice of the methods. The influence of the electromagnetic fields through the Lorentz force makes very little approximations, so that particles move on their actual trajectories influenced by field fluctuations. By using the Boris-push method combined with an automatic adjustment of the time step, very high stability and accuracy is guaranteed. A limit of the Boris-push solver lies in the simulation of ultra-relativistic particles with gamma-factors above 10^3 . It should also be kept in mind that by choosing a test-particle approach, any back-reaction of the particles on the plasma is prevented. This is a consequence of the strict separation of the kinetic and the hydrodynamic regimes in GISMO. This is a minor problem as long as one is not interested in the specific scenarios where influence of the particles on the turbulence is important, for example Landau damping. Finally, the parallelisation of GISMO should be stressed again at this point, since it allows to simulate large particle numbers (on the order of 10^6) to obtain very good statistics, which helps to distinguish physical effects from numeric noise.

3.2 Statistical analysis

A single simulation with GISMO-Particles usually produces multiple gigabytes of data for thousands of time steps and millions of particles. On the one hand, this is a big advantage of a numerical simulation, as in general we know everything about every single particle and can trace their exact trajectories through the whole simulation volume for the complete duration of the run. On the other hand, the overwhelming amount of data does not automatically translate to any gain of insight into the underlying mechanisms and properties of particle propagation. Statistical analysis needs to be employed to obtain propagation parameters that characterise the effect of plasma turbulence on the

particles. The most useful parameters are, arguably, the coefficients of the diffusion convection equation 2.135 and the Fokker-Planck coefficients (Eq. 2.131) that they are derived from.

As has been argued before, the pitch angle diffusion coefficient $D_{\mu\mu}$ is of major importance among the Fokker-Planck coefficients, as it reflects the predominant process of pitch angle scattering on magnetic field irregularities and has a direct connection to the mean free path of particles in a turbulent plasma. Two established methods - the running diffusion coefficient and the Kubo formalism - to derive $D_{\mu\mu}$ shall be introduced first and their advantages and drawbacks discussed, before we present the new method developed in the frame of this work. Calculation of spacial diffusion coefficients shall be discussed afterwards.

3.2.1 Running diffusion coefficient

The running diffusion method simply uses the definition of the diffusion coefficient (see Eq. 2.143) to calculate it. The mean square change of the pitch angle is defined as

$$\langle (\Delta\mu)^2 \rangle = \langle (\mu(t) - \mu(0))^2 \rangle, \quad (3.37)$$

where the ensemble averaging $\langle \dots \rangle$ is done over all particles with the same initial value of $\mu(t=0)^2$. The temporal behaviour of the mean square angle change is usually assumed to be a power-law:

$$\langle (\Delta\mu)^2 \rangle \sim t^\alpha, \quad (3.38)$$

where the parameter α can be used to characterise the type of diffusion:

$$0 < \alpha < 1 : \quad \text{subdiffusion.} \quad (3.39)$$

$$\alpha = 1 : \quad \text{regular Markovian diffusion.} \quad (3.40)$$

$$1 < \alpha < 2 : \quad \text{superdiffusion.} \quad (3.41)$$

$$\alpha = 2 : \quad \text{ballistic motion.} \quad (3.42)$$

As the name suggests, the predominant case is $\alpha = 1$ that is also known as “true diffusion”. It results in a time-independent diffusion coefficient which makes the diffusion process only dependent on the current configuration and not on past time steps, or “memoryless”. This kind of process is called “Markovian”. Subdiffusion and superdiffusion, on the other hand, are classified as anomalous diffusion, as they are seldomly observed in particle transport. In fact, to even have a chance to observe anomalous diffusion in GISMO would incur a high computational cost, since large simulation volume sizes would be needed for superdiffusion and very long running times for subdiffusion. Therefore we assume normal diffusion, unless stated otherwise. The diffusion coefficient

²Or more precisely, in the case of numeric evaluation, all particles in the same $[\mu(t=0), \mu(t=0) + \Delta\mu]$ -bin

can then be defined as:

$$D_{\mu\mu} = \lim_{t \rightarrow \infty} \frac{\langle (\Delta\mu)^2 \rangle}{2\Delta t} \stackrel{t \gg t_c}{\approx} \frac{\langle (\Delta\mu)^2 \rangle}{2\Delta t}. \quad (3.43)$$

In practice it is not possible to consider infinite time differences due to the finite nature of the simulation and the condition $t \rightarrow \infty$ has to be relaxed to $t \gg t_c$, where t_c is a characteristic time scale that the particle needs to develop resonant interactions and show diffusive behaviour. Together with the transition from $\Delta t = t - t_0$ to t (by setting $t_0 = 0$) this yields a definition for a time-dependent or “running” diffusion coefficient:

$$D_{\mu\mu}(t) = \frac{\langle (\Delta\mu)^2 \rangle}{2t}. \quad (3.44)$$

The clear advantage of the running diffusion method is its ease of use in numeric simulations. The pitch angle of every particle at time t is calculated and the difference to the initial value is squared. All that needs to be done now is to divide the μ -range in a number of bins and assign the particles to them according to their pitch angle value, after which the ensemble average is taken. A significant conceptual problem with the method is that every particle is assigned a pitch angle value over its whole trajectory, although the pitch angle changes over the (potentially long) run time of the simulation by definition. This is partially an artefact of the quasi-linear theory, where the integration over the generalised forces yielding the Fokker-Planck coefficients is performed along the unperturbed particle trajectories. This assumption works reasonably well in weak turbulence when the pitch angle changes are minimal and particles mostly stay in their initial μ -bins, in strong turbulence, however, this condition cannot be met any longer, especially at long time scales.

Another effect of the ambiguous assignment of the particle’s pitch angle is a more challenging interpretation of the results. Fig. 3.4 shall serve as a demonstration. After 30 gyration periods a quasi-stationary state in pitch-angle scattering is reached and the clearly discernible double-peak structure could arguably be interpreted as showing two values of $\mu \approx -0.15$ and $\mu \approx 0.15$ that fulfil the resonance condition. However, the resonance condition is actually fulfilled for $\mu \approx 0$ and the double peak structure is an effect caused by the point symmetry of the corresponding $\Delta\mu(\mu_0)$ scatter plot that is shown in Fig. 3.5. As can be seen in this plot, particles with an initial μ around the resonance preferably change the pitch angle towards the resonance and even apparently cross it, thus producing a tilt in the resonance peak structure. When $\Delta\mu$ is squared, the negative half of the peak is effectively folded upwards, forming the double-peak structure. This effect can be alleviated somewhat by choosing a different μ -definition for the abscissae. For example plotting the pitch angle change over the final value μ_f will cause a tilt in the opposite direction and plotting over an average $\bar{\mu} = 0.5(\mu_0 + \mu_f)$ will produce a symmetrical peak structure. While this can improve the appearance of the plot, it does not change its fundamental meaning.

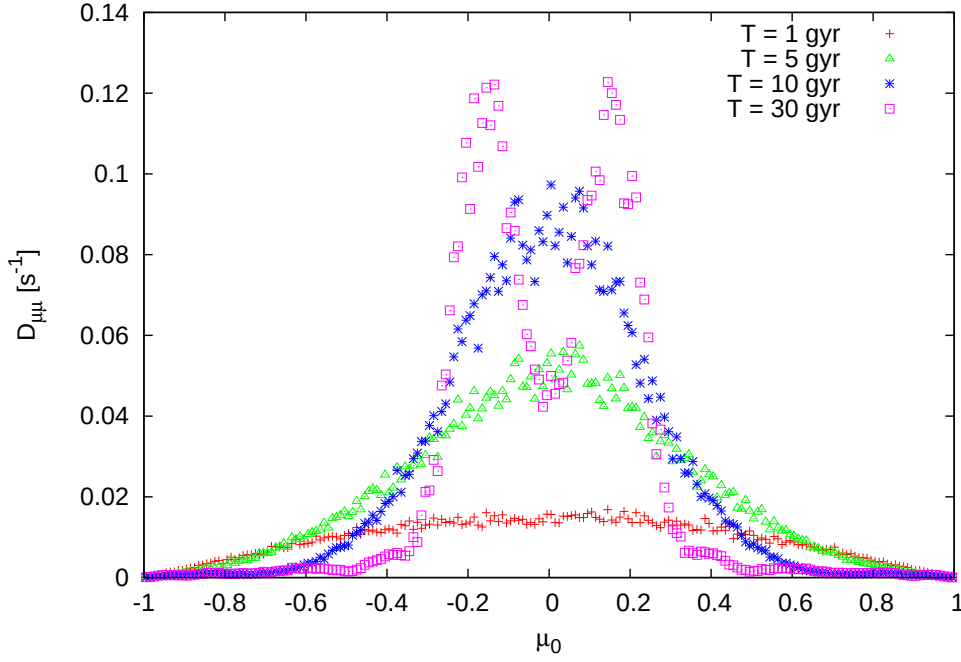


Fig. 3.4: Time evolution of the pitch angle scattering coefficient $D_{\mu\mu}$ calculated with the running diffusion method. A clear resonant structure develops between 10 and 30 gyration periods. By the comparison to the scatter plot Fig. 3.5 the maxima can be connected to the Cherenkov resonance.

3.2.2 Kubo formalism

The Kubo formalism or the Taylor-Green-Kubo formulation (Green, 1951; Kubo, 1957; Taylor, 1922) is another very popular method to calculate diffusion coefficients. It is primarily useful in analytic calculations, but can also be employed numerically. The main difference to the running diffusion method lies in taking the integration approach for the mean square pitch angle change:

$$\langle (\Delta\mu)^2 \rangle (t) = \left\langle \left(\int_0^t d\tau \dot{\mu}(\tau) \right)^2 \right\rangle = \int_0^t d\tau \int_0^t d\xi \langle \dot{\mu}(\tau) \dot{\mu}(\xi) \rangle. \quad (3.45)$$

Assuming homogeneity in time the $\dot{\mu}$ correlation functions depend not on the specific time but only on the time difference. This can be used to transform the integrals and solve them with partial integration (Shalchi, 2009):

$$\langle (\Delta\mu)^2 \rangle (t) = 2 \int_0^t d\tau (t - \tau) \langle \dot{\mu}(\tau) \dot{\mu}(0) \rangle. \quad (3.46)$$

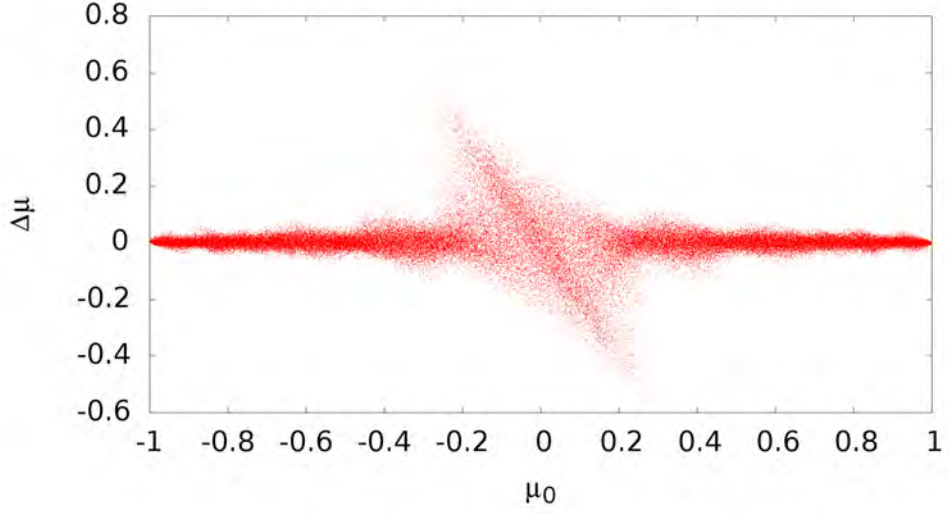


Fig. 3.5: Change of the pitch angle $\Delta\mu$ in dependence of the initial value μ_0 for 10^5 particles within a turbulent plasma after 30 gyrations. This *scatter plot* is a valuable tool to investigate resonant interactions. The dominant structure is the Cherenkov resonance $n = 0$.

From the result, the diffusion coefficient can be derived:

$$D_{\mu\mu}(t) = \frac{1}{2} \frac{d}{dt} \langle (\Delta\mu)^2 \rangle (t) = \int_0^t d\tau \langle \dot{\mu}(\tau) \dot{\mu}(0) \rangle. \quad (3.47)$$

In the case of true diffusion the coefficient $D_{\mu\mu}(t)$ will be constant by definition. The upper limit of the integral can then be extended to infinity yielding the Kubo formula for the diffusion coefficient:

$$D_{\mu\mu} = \int_0^\infty d\tau \langle \dot{\mu}(\tau) \dot{\mu}(0) \rangle. \quad (3.48)$$

This is basically the same expression that has been derived for Fokker-Planck coefficients in the frame of the quasi-linear theory (compare Eq. 2.132).

The Kubo formalism is a very useful tool in analytical calculations, provided one can find analytical expressions for the pitch angle time derivatives, which can be done for a limited number of field configurations (e.g. a single interaction with an Alfvén wave as derived in 2.3.2). The numeric application of the method, however, is rather tedious when done correctly. The integral in Eq. 3.48 can be discretised, yielding

$$D_{\mu\mu} = \sum_{\tau=0}^t \Delta t \sum_P \frac{1}{N_P} \dot{\mu}(\tau) \dot{\mu}(0), \quad (3.49)$$

where the second sum over all particles in a $[\mu, \mu + \Delta\mu]$ -bin represents the ensemble averaging. It should be noted, that the Kubo method can only partially address the drawbacks of the running diffusion method in that it still traces particle trajectories and assigns a pitch-angle value to each one, even though the pitch angle can change significantly over the simulation run. This can be somewhat accounted for with Kubo by calculating the ensemble average individually for each time step of the integration, although that requires some computational effort, which is intrinsically heavier for Kubo as it is, since the correlation functions are computed along the entire particle trajectory instead of taking the initial and final values as in the running diffusion method. The time derivatives of the pitch angle present another problem: they amplify the stochastic noise of the $\mu(t)$ trajectories. Even though the integration smoothes the results again, if the derivatives are done in a numerically naive way, the results will be dominated by stochastic noise. This means that a more sophisticated method to calculate time derivatives is needed, e.g. Savitzky-Golay, which again increases the computational cost.

3.2.3 Diffusion equation approach

To address the weaknesses of the described methods a new method to determine diffusion coefficients was developed in the frame of this work with a focus on increased applicability especially for particle transport scenarios that are not covered by QLT, e.g. pitch angle scattering in strong turbulence, where the initial pitch angle of the particle changes significantly over the simulation run time. The starting point is the Fokker-Planck equation 2.129:

$$\frac{\partial F_T}{\partial t} + v\mu \frac{\partial F_T}{\partial Z} - \Omega \frac{\partial F_T}{\partial \phi} = S_T + \frac{1}{p^2} \frac{\partial}{\partial X_\sigma} \left(p^2 D_{X_\sigma X_\nu} \frac{\partial F_T}{\partial X_\nu} \right). \quad (3.50)$$

The basic assumption of the method is that pitch angle diffusion is the predominant diffusion process, in particular, momentum diffusion and the off-diagonal mixed components of the diffusion tensor must be negligible in comparison, i.e. $D_{pp} \ll D_{\mu\mu}$ and $D_{\mu p} \ll D_{\mu\mu}$. This assumption is valid in the case of low-frequency waves with non-relativistic phase speeds ($v_{ph}/c \ll 1$), as the Fokker-Planck coefficients $D_{\mu p}$ and D_{pp} scale with the factors v_{ph}/c and $(v_{ph}/c)^2$, respectively Schlickeiser (1994). In this case the Fokker-Planck equation can be simplified to:

$$\frac{\partial F_T}{\partial t} + v\mu \frac{\partial F_T}{\partial Z} - \Omega \frac{\partial F_T}{\partial \phi} = \frac{\partial}{\partial \mu} \left(D_{\mu\mu} \frac{\partial F_T}{\partial \mu} \right). \quad (3.51)$$

An isotropic homogeneous particle distribution allows us to average the equation over the spacial coordinates and the phase angle ϕ , yielding the well known Fick's diffusion equation:

$$\frac{\partial F_T}{\partial t} - \frac{\partial}{\partial \mu} \left(D_{\mu\mu} \frac{\partial F_T}{\partial \mu} \right) = 0. \quad (3.52)$$

With knowledge of the particle distribution $F_T(\mu, t)$ and its derivatives this equation can easily be solved numerically. Assuming true diffusion the coefficient $D_{\mu\mu}$ will be a time-independent function of μ only, in which case applying the product rule yields the form:

$$\frac{\partial F_T(\mu, t)}{\partial t} = \left(\frac{\partial}{\partial \mu} D_{\mu\mu}(\mu) \right) \frac{\partial F_T(\mu, t)}{\partial \mu} + D_{\mu\mu}(\mu) \cdot \frac{\partial^2 F_T(\mu, t)}{\partial \mu^2}. \quad (3.53)$$

The simulations provide us with discrete distributions $f^n \equiv F_T(\mu^n)$ that express the number of particles in a bin of width $\Delta\mu$ centred around $\mu^n = -1 + n \cdot \Delta\mu$ with $n \in \mathbb{N}$, or μ -histograms. The diffusion equation is discretised accordingly in $D_{\mu\mu}^n \equiv D_{\mu\mu}(\mu^n)$:

$$\partial_t f^n = \frac{D_{\mu\mu}^{n+1} - D_{\mu\mu}^{n-1}}{2 \cdot \Delta\mu} \cdot \partial_\mu f^n + D_{\mu\mu}^n \partial_{\mu\mu} f^n. \quad (3.54)$$

As can be easily seen, this corresponds to a matrix equation with a tridiagonal matrix:

$$\begin{pmatrix} \partial_{\mu\mu} f^0 & \frac{\partial_\mu f^0}{2\Delta\mu} & 0 & 0 \\ -\frac{\partial_\mu f^1}{2\Delta\mu} & \partial_{\mu\mu} f^1 & \ddots & 0 \\ 0 & \ddots & \ddots & \frac{\partial_\mu f^{N-1}}{2\Delta\mu} \\ 0 & 0 & -\frac{\partial_\mu f^N}{2\Delta\mu} & \partial_{\mu\mu} f^N \end{pmatrix} \cdot \begin{pmatrix} D_{\mu\mu}^0 \\ D_{\mu\mu}^1 \\ \vdots \\ D_{\mu\mu}^N \end{pmatrix} = \begin{pmatrix} \partial_t f^0 \\ \partial_t f^1 \\ \vdots \\ \partial_t f^N \end{pmatrix}. \quad (3.55)$$

This linear equation system can easily be solved with standard matrix inversion methods (as long as the matrix is invertible of course), but a problem arises when one considers the input quantities of the method. Since the distribution function is composed from data of a finite number of test particles, the sampling of the phase space will be imperfect, which results in stochastic noise superposing the μ -histogram. While this can be alleviated somewhat with suitable binning, taking the first and second μ -derivatives that are also necessary for the method makes the stochastic noise dominate the signal-to-noise ratio rather quickly, as demonstrated in Fig. 3.6. There are several approaches to handle this problem: the correct way to increase the signal-to-noise ratio would be by improving the statistics either by increasing the particle number in a simulation or by making multiple runs and taking the ensemble average over them. The feasibility of this approach is limited for numerical simulations as it would mean a huge computational effort and non-existent for real measurements, since they cannot be generated at will. The other way to handle this issue is to apply a more sophisticated method to obtain usable derivatives, for example by applying smoothing kernels or fitting analytic functions to the data.

Savitzky-Golay smoothing

While many smoothing methods produce acceptable results, we found that fitting subsets of the distribution function with low-degree polynomials as described by Savitzky and Golay (1964) yields the best representation of the main features of the distribution function while reducing the stochastic noise to a manageable level. Additionally,

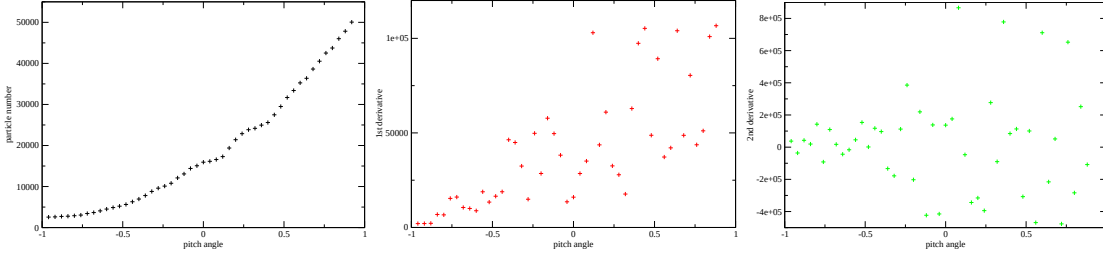


Fig. 3.6: A typical particle distribution function from a GISMO simulation and its first and second derivative in μ . When taken in a naive way, the stochastic noise will dominate any useful signal.

the Savitzky-Golay method offers the advantage of easily obtainable derivatives. The method works by fitting a polynomial with linear least squares to a subset of m (called the window) adjacent data points separated by intervals h_i . For equally spaced data points the intervals are identical $h_i \equiv h$ and the least squares equations can be solved analytically. After transforming the independent variable x to

$$z = \frac{x - \bar{x}}{h} \quad (3.56)$$

where $\bar{x}$ is the central point in the considered window, the fitting polynomial of degree k takes on the form

$$Y = \sum_{i=0}^k a_i z^i \quad (3.57)$$

where the z -values are integers running from $(1 - m)/2$ to $(m - 1)/2$. The coefficients a_i of the polynomials are the solutions to the normal equations

$$\mathbf{a} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{y}, \quad (3.58)$$

where the matrix $\mathbf{J}$ is composed of first k powers of z or, in element notation:

$$\mathbf{J}_{ij} = \left(i - \frac{m+1}{2} \right)^{j-1}. \quad (3.59)$$

Although the equations can be solved to obtain the fitting polynomials, usually it is sufficient to calculate the convolution coefficients, which are the elements of the matrix

$$\mathbf{C} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T. \quad (3.60)$$

With the convolution coefficients the smoothed values Y_j can be calculated from the actual data values y_{j+i} in the respective window, where the index j runs to within half

the width of the window $(m + 1)/2$ to the borders of the data set:

$$Y_j = (\mathbf{C} \otimes \mathbf{y})_j = \sum_{i=-\frac{m-1}{2}}^{\frac{m-1}{2}} C_i y_{j+i}. \quad (3.61)$$

The first and last $(m - 1)/2$ data points have to be treated differently, especially with non-periodic boundary conditions for which the data set can not be simply extended with values from the opposite boundary. In this case, instead of using only the central value of the polynomial, the complete left half of the first fitting polynomial and the complete right half of the last one can be used. Derivatives of the data set are equally easy to obtain from the analytical derivative of the polynomials. Due to the widespread use of the Savitzky-Golay method, the convolution coefficients are implemented in many common programming languages and algebra environments, e.g. Mathematica. Even if tables of convolution coefficients are unavailable, closed formulations can be derived and easily implemented. For our purposes the coefficients with polynomials up to the third degree for up to the second derivative completely suffice:

$$\begin{aligned} C_i &= \frac{3(3m^2 - 7 - 20i^2)}{4m(m^2 - 4)} && \text{smoothing, polynomial degree 2,3} \\ C_i &= \frac{12i}{m(m^2 - 1)} && \text{1st derivative, polynomial degree 1,2} \\ C_i &= \frac{75(3m^4 - 18m^2 + 31)i - 420(3m^2 - 7)i^3}{m(m^2 - 1)(3m^4 - 39m^2 + 108)} && \text{1st derivative, polynomial degree 3,4} \\ C_i &= \frac{360mi^2 - 30m(m^2 - 1)}{m^2(m^2 - 1)(m^2 - 4)} && \text{2nd derivative, polynomial degree 2,3} \end{aligned}$$

Integration

A fundamentally different approach to solve Fick's diffusion equation 3.52 for the diffusion coefficient is to integrate it over μ :

$$\int_{-1}^{\mu} d\mu' \frac{\partial f(\mu', t)}{\partial t} = D_{\mu\mu}(\mu) \frac{\partial f(\mu, t)}{\partial \mu} = -j_{\mu}(\mu). \quad (3.62)$$

Note that the integration of the right hand side assumes $D_{\mu\mu} \partial_{\mu} f|_{\mu=-1} = 0$. The same assumption is made implicitly when formulating the discretised matrix equation; it is justified by the choice of spherical momentum coordinates, which adds an intrinsic factor $(1 - \mu^2)$ to the diffusion coefficient causing it to approach zero at the borders of the μ -interval. The integration of the simulation data is done numerically, transforming the integral on the left hand side into a sum:

$$\int_{-1}^{\mu} d\mu' \frac{\partial f(\mu', t)}{\partial t} \rightarrow \sum_{i=0}^n \Delta\mu \frac{\partial f(\mu^i, t)}{\partial t}. \quad (3.63)$$

This integration yields the effective pitch angle current j_μ^n from which the diffusion coefficient $D_{\mu\mu}^n$ is calculated through division by $\partial_\mu f^n$. The clear advantages of this approach are the elimination of the second derivative in μ from the equation and the smoothing of the time derivative of the distribution function by the integration. Because of this the integration method retains its validity, though somewhat limited, even without application of any smoothing techniques, which makes it a valuable tool to investigate the influence of the smoothing on the resulting diffusion coefficient. Naturally, any smoothing method used for the matrix solver method can also be applied to the integration method, in particular Savitzky-Golay, for a better signal-to-noise ratio.

3.3 Perlin noise turbulence

Despite all the simplifying assumptions incorporated into the incompressible MHD equation system, the simulation of realistic turbulence development through them still takes a significant computational effort, especially for large spacial resolutions ($> 512^3$) and long run times that are necessary for the complete development of the inertial range power-law. In order to make investigation of particle transport in turbulence more convenient, a method was developed (Kilian et al., 2016) that makes a reasonable compromise between a realistic turbulence representation and a low computational effort. The pitch angle computation methods developed in this work were used, amongst others, to validate the resulting turbulence with respect to realistic wave-particle scattering.

The new turbulence simulation approach is based on Perlin noise - a method to generate stochastic noise named after its developer Ken Perlin (Perlin, 1985) with the original purpose of generating more realistic-looking computer graphics, for which it was widely used³. However even though this may be the most widespread use of the Perlin noise function, it is by far not the only one, as even in the original paper the possibility of realistic turbulence generation based on Perlin noise was mentioned. One important property that makes Perlin noise so popular for computer graphics is the existence of a characteristic correlation length, as it causes all of the visual features to be at the same size while maintaining a random appearance. In other words, the noise is correlated on scales smaller than the characteristic correlation length and looks smooth. On larger scales it is uncorrelated and looks random. The statistical properties of the noise, e.g. amplitude and correlation length, are invariant under translations, making the output of the noise function homogeneous. In multiple dimensions a small improvement as detailed in Perlin (2002) can be implemented to obtain isotropic output with statistical properties that are invariant under rotations. In general, it is very straightforward to extend the Perlin noise function to multiple dimensions, although the implementation used in this work is limited to a single one. For a large number of dimensions a more computationally efficient alternative exists in the form of Simplex noise (Perlin, 2001).

Naive noise generation is usually done with a “value noise” function by using coor-

³So far as to win a Technical Achievement Award from the Academy of Motion Picture Arts and Sciences in 1997 “to Ken Perlin for the development of Perlin Noise, a technique used to produce natural appearing textures on computer generated surfaces for motion picture visual effects”

ordinates in spacial grid and assigning computed pseudo-random output values directly at the grid locations. Perlin (and Simplex) noise take a different approach by assigning pseudo-random gradients to the grid locations. The values of the noise function at arbitrary spacial coordinates are then computed by interpolation between the grid location using smooth functions that match the assigned gradients and the boundary conditions. This kind of noise function falls in the category of “gradient noise”.

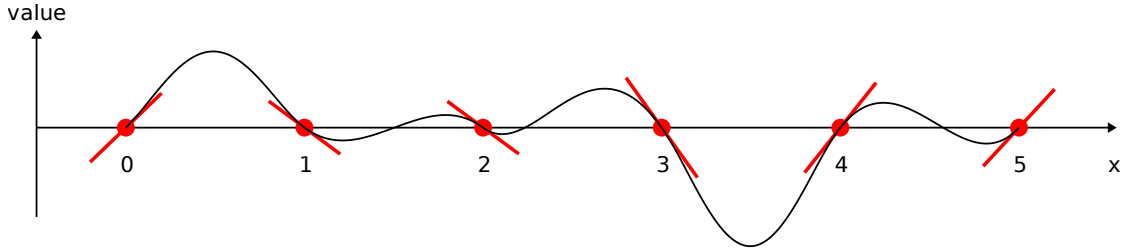


Fig. 3.7: Perlin noise generation in 1D: random gradients (indicated by red lines) are assigned at all grid points (red dots), then a smooth function (indicated by the black line) that matches the gradients at the grid points is used to interpolate the value of the noise function at all desired locations. Source: Kilian et al. (2016).

An illustration of the noise generation principle in one dimension is shown in Fig. 3.7. First, each supporting grid point is assigned a uniform random gradient from the interval $[-1; 1]$ and the value of the noise function is set to zero. This is a slight deviation from the multi-dimensional case, where each grid point is assigned a gradient vector of unit length. The values for the gradients can be picked from a pre-generated set to improve performance and offer the possibility to reproduce a particular noise function output when necessary without storing it. For our purposes these features are not necessary allowing us to use a random number generator to produce the gradient values. The value of the noise function at arbitrary positions between support nodes can now be calculated using an interpolating function. This function is typically a polynomial constructed to vanish at $x = 0$ and $x = 1$, as x is rescaled to $[0; 1]$ between adjacent nodes, and to match the gradients at the end points. In this work the chosen polynomial is

$$f(x) = 2(a - b)x^4 - (3a - 5b)x^3 - 3bx^2 + ax, \quad (3.64)$$

constructed to satisfy the requirements $f(0) = f(1) = 0$, $f'(0) = a$ and $f'(1) = b$.

A power spectrum generated in this manner is shown in Fig. 3.8. A property of Perlin noise is clearly visible here: the characteristic length scale, at which all features in the curve are concentrated. As indicated in the figure, this characteristic wavelength corresponds exactly to the separation between the supporting grid points. The power in the noise drops fairly quickly in both directions from this characteristic wave number. Towards smaller k the drop roughly follows a power law. The drop towards larger k is even quicker due to the use of smooth interpolation functions. In general, even though the exact spectral shape can be influenced by the choice of the interpolation function,

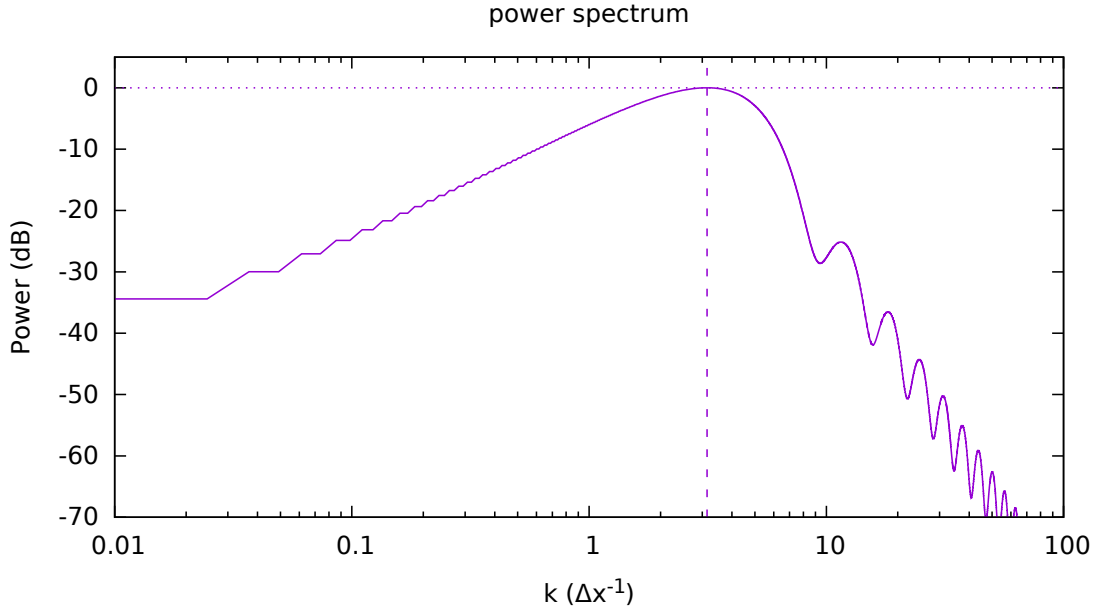


Fig. 3.8: An example of a Perlin noise power spectrum. The wavelength of the peak position matches the distance between the points at which gradients are chosen (indicated by the vertical dashed line in the plot). The power spectrum was normalized to the peak value indicated by the dotted line. Source: Kilian et al. (2016).

the power drop from the characteristic length towards smaller scales over an order of magnitude will typically be large enough to reach the round-off level of floating point numbers.

The output of one Perlin noise function does obviously not meet our expectations of a realistic turbulent power spectrum. The solution to the problem lies in the superposition of the outputs from multiple noise functions with different characteristic lengths. Since the power spectrum is strongly peaked at the characteristic length and falls off quickly and smoothly in both directions, the desired power spectrum can be constructed by adding the output of multiple iterations of Perlin noise. At each successive iteration the distance between the grid support points is reduced. In our case specifically, the distance is halved and the iterations are called octaves. Each octave of noise can be scaled separately by an amplitude factor called persistence that can be changed to determine the shape of the superposed spectrum. Effectively, the turbulent energy cascade is constructed by modelling single length scales with Perlin noise, setting the amplitude ratio between the scales to obtain the desired spectral shape and superposing them.

In the code this is done to directly generate the Elsässer fields $w_{x,y}^{\pm}$. The fields are the propagated semi-analytically in the Fourier space by multiplying the initial spectrum

with a phase factor

$$\tilde{w}_{x,y}^{\pm}(k, t) = \tilde{w}_{x,y}^{\pm}(k, 0) \exp(\pm i v_A k_z t) \quad (3.65)$$

and then transformed back to the real space fields $w_{x,y}^{\pm}(x, t)$. From the definition of Elsässer variables $\mathbf{u}$ and $\mathbf{b}$ can be then computed, which are in turned used to calculate $\mathbf{E} = -\mathbf{u} \times \mathbf{b}$. The fields are then interpolated to the spacial coordinates of the test particles and used to update their positions and velocities via the Boris push (see chapter 3.1.3).

4 Results

The results obtained with the new methods developed in this dissertation will be presented in this chapter along with comparisons to established methods where applicable. Additionally to the extensive validation and testing of the Gismo code done in Lange (2012), we present validation of some special cases relevant to this work. The newly developed Perlin noise turbulence method will be compared both with real MHD turbulence and analytical test cases focusing on particle transport.

The last part will present results obtained for particle transport parameters, specifically pitch angle and spacial diffusion coefficients, and discuss their relevance to particle transport in the heliosphere.

4.1 Method validation

Validation is a crucial point when dealing with numerical simulation results. Construction of test case scenarios that are analytically computable and physically comprehensible is an important part of validation and helps to gain a deeper insight into the physics of the model. For Gismo extensive validation and testing has been done by Lange (2012). As this work focuses primarily on particle transport, we present an improved test of the wave-particle interactions comparing the results to analytical calculations in the frame of the QLT.

4.1.1 Wave-particle interaction

As has been demonstrated in chapter 2.3.2, the pitch angle change of a particle interacting with a single Alfvén wave can be calculated when using the quasi-linear approximation (see Eq. 2.150). This can be used for a relatively complex test case to verify that a large number of particles with random initial values interact correctly with a wave propagating through the simulation volume and show resonances at correct pitch angles as predicted by the resonance condition 2.144:

$$\mu = \frac{\omega - n\Omega}{k_{\parallel}v} = \frac{v_A k_z - n\Omega}{k_z v}. \quad (4.1)$$

The choice of the magnetic field configuration for this test case is a single circular polarised Alfvén wave with the wave vector $\mathbf{k} = \pm k \cdot \mathbf{e}_z$. In the reference frame of the wave this yields:

$$\mathbf{B} = B_0 \mathbf{e}_z + \delta B (\mathbf{e}_x \cos(\pm k z + \psi) + \mathbf{e}_y \sin(\pm k z + \psi)). \quad (4.2)$$

The time derivative of μ is proportional to the time derivative of the parallel momentum, which can be calculated from the z -component of the Lorentz force:

$$\dot{\mu} = \frac{1}{p} \dot{p}_{\parallel} = \frac{q}{\gamma p m c} (p_x B_y - p_y B_x). \quad (4.3)$$

Keeping in mind that the perpendicular momentum component is modulated by the gyration of the particle with $p_x = p_{\perp} \cos(\Omega t)$ and $p_y = p_{\perp} \sin(\Omega t)$ and using trigonometric identities, the pitch angle time derivative can be written as:

$$\dot{\mu} = \frac{\Omega}{\gamma} \frac{\delta B}{B_0} \sqrt{1 - \mu^2} \sin((\pm k v \mu - \Omega)t + \psi). \quad (4.4)$$

This differential equation can be integrated in an approximate manner by neglecting the implicit time-dependence of the μ -terms on the right hand side. This is generally equivalent to the approach of the QLT to integrate along unperturbed particle orbits and only valid under the assumption of small pitch angle changes, implying weak turbulence ($\delta B \ll B_0$). The integration yields an expression for the pitch angle change after the time t :

$$\Delta\mu = \mu(t) - \mu(0) = \frac{\Omega}{\gamma} \frac{\delta B}{B_0} \sqrt{1 - \mu^2} \frac{\cos(\psi) - \cos((\pm k v \mu - \Omega)t + \psi)}{\pm k v \mu - \Omega}. \quad (4.5)$$

Since the randomly initialised particles cover the complete range of phases ψ , for comparison purposes we are interested in the extremal values of $\Delta\mu$. Those are reached for phase values of

$$\psi_{\max} = \arctan\left(\frac{\sin(\pm k v \mu - \Omega)\Delta t}{1 - \cos(\pm k v \mu - \Omega)\Delta t}\right) \quad (4.6)$$

and at the identical branch of the arctangent, shifted by π , $\psi_{\max} + \pi$.

The above formulae 4.5 and 4.6 give the maximum pitch angle change in the reference frame of the wave. In order to compare it to simulation results, it needs to be transformed to the lab frame. As noted by Schreiner et al. (2017) the transformation used in Lange (2012)

$$\mu' = \mu \pm \frac{v_A}{v} \quad (4.7)$$

is in fact not correct as it does not cover the complete range of μ and produces nonsensical μ -values with $|\mu| > 1$, although it yields reasonable results in the case $v_A/v \ll 1$. The correct transformations are adopted from Schreiner et al. (2017) with the simplifying assumption of non-relativistic wave velocities, which is fulfilled in our case. The remaining transformations for the primed quantities in the wave frame as functions of

the unprimed quantities in the lab frame read as follows:

$$\mu' = \frac{v\mu - v_A}{\sqrt{v^2 - 2v\mu v_A + v_A^2 - v^2 v_A(1 - \mu^2)/c^2}} \quad (4.8)$$

$$v' = \frac{\sqrt{v^2 - 2v\mu v_A + v_A^2 - v^2 v_A(1 - \mu^2)/c^2}}{1 - vv_A\mu/c^2} \quad (4.9)$$

Applying these transformations to Eq. 4.5 and 4.6 yields the scattering amplitude $\Delta\mu'(\bar{\mu}, t)$ in the wave frame expressed with quantities in the lab frame. The free variable is the arithmetic average of the initial and final pitch angle of the particle: $\bar{\mu} \equiv (\mu(t) - \mu(0))/2$. In order to transform $\Delta\mu$ itself into the lab frame, we need the inverse transformation for the pitch angle:

$$\mu = \frac{v'\mu' + v_A}{\sqrt{v'^2 + 2v'\mu'v_A + v_A^2 - v'^2 v_A(1 - \mu'^2)/c^2}}. \quad (4.10)$$

This transformation is then applied to $\mu'(t) = \Delta\mu'/2 + \bar{\mu}'$ to get $\mu(t)$ in the lab frame, which can then be used to calculate the scattering amplitude in the lab frame:

$$\Delta\mu(\bar{\mu}, t) = \mu(t) - \mu(0) = 2(\mu(t) - \bar{\mu}) \quad (4.11)$$

For the first test a circularly polarised Alfvén wave propagating in positive z -direction was initialised in Gismo. The negative solutions of equations 4.5 and 4.6, transformed into the lab frame as described above, were used as comparison. The amplitude of the wave was set to satisfy $\delta B/B_0 = 0.05$ with a background magnetic field of $B_0 = 4.34 \cdot 10^{-4}$ G, resulting in a proton gyration frequency of $\Omega = 4.16 \text{ s}^{-1}$. In order to ensure a good statical coverage of the phase space, 500000 particles with a fixed velocity $v = 2 \cdot 10^8$ cm/s and randomly and homogeneously distributed spacial coordinates, pitch angles μ and azimuth angles Φ were injected into the simulation. The initial particle speed lies above the Alfvén speed $v_A = 10^7$ cm/s, ensuring resonant interactions.

The resulting scatter plot of pitch angle changes over the average pitch angle for every single particle after one gyration ($t = 1.5$ s) can be seen in 4.1, together with the theoretical prediction for the maximal scattering amplitude. At this low level of magnetic field perturbation $\delta B/B_0 = 0.05$ theory and simulation are in very good agreement. The position and amplitude of the main resonance as well as the secondary ballistic peaks are matched perfectly.

Fig. 4.2 shows the same simulation with the running time advanced to $t = 3$ s (2 gyrations) and $t = 4.5$ s (3 gyrations). Again, theory and simulation are in very good agreement. With advancing time the secondary ballistic peaks become more numerous and at the same time smaller in amplitude, while the resonant peak becomes both narrower and higher. This can be understood as indication that the notion of the sharp δ -function shaped resonances in wave-particle interactions is only valid in an infinite time limit, as particles in resonance will always be scattered stronger and dominate the

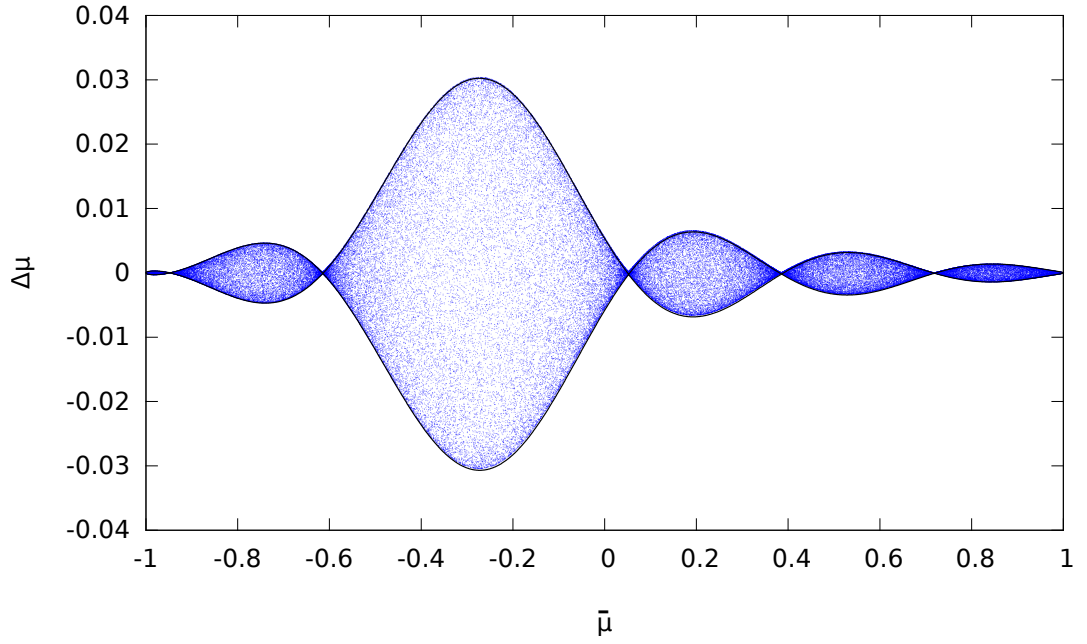


Fig. 4.1: Scatter plot of pitch angle changes of individual particles from a Gismo simulation (blue dots) compared with the theoretical prediction of the maximal scattering amplitude (black lines) after 1 gyration period ($t = 1.5s$). The particles are interacting resonantly with a circularly polarised Alfvén wave with $\delta B/B_0 = 0.05$.

secondary scatterings. This leads to a δ -shaped resonance after very long times, while real resonances after finite times will always be broadened.

The same test configuration was also simulated with the Perlin noise code. The comparison of numerical results with analytical prediction is shown in Fig. 4.3. It can be seen that the results here are also in fairly good agreement, although the resonant peak seems to be slightly underestimated by the simulation.

4.1.2 Monte-Carlo validation

To verify that the newly developed pitch angle calculation methods yield the correct $D_{\mu\mu}$, they were applied to the output of a Monte-Carlo particle transport code that uses a predetermined diffusion coefficient as input. For a detailed description of the Monte-Carlo approach the reader is directed to Agueda and Vainio (2013); Agueda et al. (2008). The diffusion coefficient that was used has the form:

$$D_{\mu\mu}(\mu) = \frac{\nu}{2}(1 - \mu^2)$$

$$\nu(\mu) = \nu_0 \left(\frac{|\mu|}{1 + |\mu|} + \epsilon \right), \quad (4.12)$$

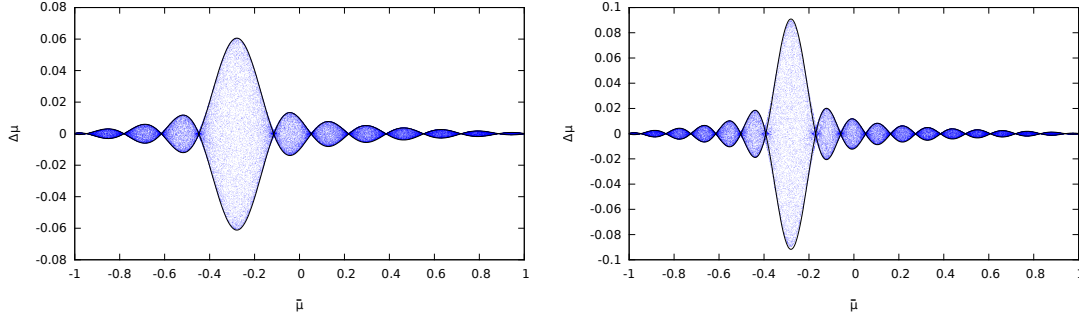


Fig. 4.2: Comparison of pitch angle scattering amplitudes from a Gismo particle simulation (blue dots) and theoretical prediction (black lines) at $t = 3$ s (2 gyrations, left) and $t = 4.5$ s (3 gyrations, right) in the same configuration as Fig. 4.1.

where the parameter $\epsilon = 0.01$ and ν_0 is tied to the time step Δt by $2\nu_0\Delta t = 0.01$. The initial particle distribution $f(\mu)$ was chosen to fulfil the condition of non-vanishing derivatives in μ , specifically $\partial_\mu f(\mu) \neq 0$, $\partial_{\mu\mu} f(\mu) \neq 0$. This is fulfilled by a distribution with the shape of a shifted half-parabola

$$f(\mu) = a \cdot (\mu \pm 1)^2 + c, \quad (4.13)$$

which was used in this simulation. The initial distribution and its change through particle diffusion can be seen in Fig. 4.4 (left).

The pitch angle diffusion coefficients calculated by solving the diffusion equation with matrix inversion (MIIa) and numerical integration (MIIb) are shown in Fig. 4.4 on the right. The dashed line shows the actual form of $D_{\mu\mu}$ that was used in the simulation for comparison. In general, both methods are able to reconstruct the form and amplitude of the coefficient. The resonant maxima are underestimated somewhat, especially by the integration method MIIb. This can be explained with a look at the particle distribution and its change at these μ -values. As seen in the left figure, the number of particle remains almost constant while only the slope changes significantly. This leads to an increase in the relative statistical error when calculating the time derivative of the distribution, $\partial_t f(\mu)$, causing in turn the deviation in $D_{\mu\mu}$ at these points. It is also worth noting that the discontinuous slope at $\mu = 0$ cannot be reproduced, as the employed Savitzky-Golay smoothing method uses continuous polynomials.

4.1.3 Running diffusion comparison

In order to further verify the diffusion equation approach methods MIIa and MIIb and point out their benefits and limitations, additional comparisons were made with $D_{\mu\mu}$ obtained with the running diffusion coefficient method (hereafter referred to as MI) in different turbulence scenarios. The simulation setup is as described in Ivaschenko et al. (2016) and attempts to approximate the conditions in the solar corona at three solar radii

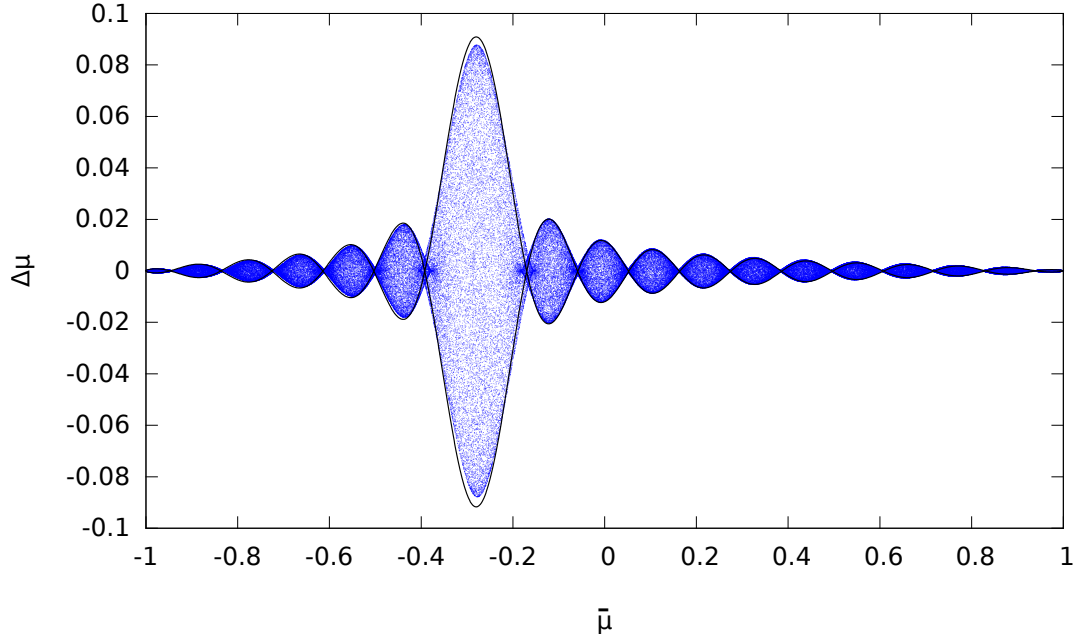


Fig. 4.3: A comparison of resonant wave-particle interactions in a Perlin noise simulation (blue dots) with theoretical predictions based on QLT (black lines). The pitch angle scattering amplitude is shown after 3 gyration periods at $t = 4.5$ s.

(Vainio et al., 2003) with a background magnetic field of $B_0 = 0.174$ G and a particle density $\rho = 10^5 \text{ cm}^{-3}$, yielding an Alfvén speed of $v_A = 1.2 \cdot 10^8 \text{ cm/s}$. The simulation length scale is $L = 3.4 \cdot 10^8 \text{ cm}$, resolved on a 256^3 grid. Several different turbulence setups were used following Lange and Spanier (2012), as shall be described below.

Saturated background turbulence

The first setup was achieved using an anisotropic driver at small wave numbers until a saturated turbulent stage develops through an energy cascade. In order to enhance the development of the parallel cascade in this relatively weak turbulence, the energy injection in the parallel direction was performed up to higher wave numbers compared to the perpendicular injection. After a steady-state in the turbulence development was reached, 10^6 protons with $v = 1.21 \cdot 10^{10} \text{ cm/s}$ were injected and their propagation simulated for 30 gyration periods. The proton velocity was chosen to fulfil the resonance condition 2.144 within the μ -range $[-1, 1]$ for wave numbers k within the inertial range of the spectrum. For the initial μ -distribution of the protons, again a shifted half-parabola shape (Eq. 4.13) was chosen to ensure applicability of the new methods MIIa and MIIb. The running diffusion coefficient method MI does not depend on the pitch-angle distribution as long as the number of particle is sufficiently high to ensure good statistics.

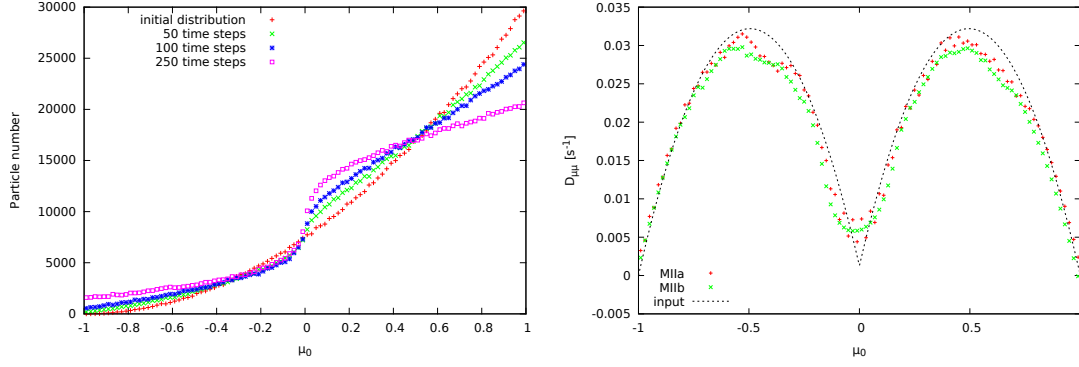


Fig. 4.4: Left: Initial particle distribution in the Monte-Carlo simulation and its evolution after n time steps under the influence of pitch angle diffusion. Right: A comparison of the $D_{\mu\mu}$ used as input for the Monte-Carlo propagation model (dashed line) with the results obtained from the output of the model by solving the diffusion equation with a matrix solver (MIIa) and with numerical integration (MIIb)

Figure 4.5 shows the results obtained with the various methods. The left figure shows the evolution of a resonant structure calculated with MI as the simulation progresses. This is caused by an increasing amount of particles in the simulation volume interacting resonantly with the wave modes running through it and should not be confused with the development of the turbulence itself. The evolution as seen in Fig. 4.5 is a purely statistical effect of wave-particle interactions in a completely developed turbulence. A second notable effect is the apparent splitting of the resonant peak at later times. As has been discussed in detail in chapter 3.2.1 this is a purely visual effect caused by the choice of the independent variable (in this case the initial pitch angle μ_0). As can be seen in the figure on the right, the diffusion equation solutions from MIIa and MIIb do not suffer from this problem and show otherwise rather well comparable results with a central Cherenkov resonance with an amplitude of $\approx 0.1 \text{ s}^{-1}$ and much smaller secondary resonances at $|\mu| \approx 0.4$ that are just becoming visible with MI at 30 gyration periods. This sums up the advantages of the new methods rather nicely: they are capable of producing more exact results from a shorter simulation run time without the need for an interpretive help in the form of a scatter plot.

A note on the Kubo formalism as described in chapter 3.2.2: Since the particle simulations show an evolution phase in the beginning as shown in Fig. 4.5, during which the pitch angle coefficient is not constant due to an increasing number of wave-particle interactions, the time-dependent version of the Kubo formula was applied. Specifically, we give up the assumption that the correlation integrals in Eq. 3.45 depend only on the time difference (or lag) and not on the specific time itself. This yields a time dependent pitch angle diffusion coefficient $D_{\mu\mu}(t)$ as shown in Fig. 4.6. It should be noted that application of the Kubo formalism in this way to computational data involves a

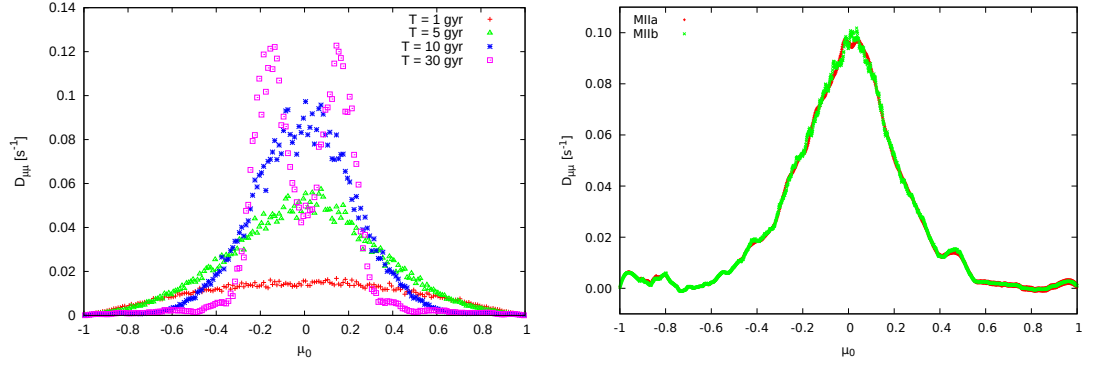


Fig. 4.5: Comparison of pitch angle diffusion coefficients from the saturated background turbulence simulations. Left: time evolution of $D_{\mu\mu}$ calculated with the running diffusion method with apparent splitting of the Cherenkov resonance at 30 gyration periods. Right: $D_{\mu\mu}$ obtained with MIIa and MIIb after 10 gyration periods. Source: Ivascenko et al. (2016).

lot of effort, since it means tracing every particle trajectory for the whole simulation run and calculating the pitch angle time derivative $\dot{\mu}$ at every time step. Again, the Savitzky-Golay method was employed to obtain useful numerical derivatives. Next, the autocorrelation time integral was performed and, finally, the ensemble average over the particle trajectories in the same μ -bin was calculated at every time step. For all this effort, the results in Fig. 4.6 roughly reproduce the Cherenkov resonance shape at $\mu = 0$ and the scheme shows oscillatory behaviour in both time and pitch angle with a distinct coupling between the both (wave pattern), thus leading to the conclusion that the Kubo formalism is better suited for analytic calculations. It will be discarded from further comparisons.

Amplified wave mode

The second simulation setup is mostly identical to the first, with an important addition. Together with the energy injection at low wave numbers, a single wave mode at the numerical wave number $k'_{\parallel} = 2\pi \cdot 24$ (corresponding to a physical wave number $k = 4.44 \cdot 10^{-7} \text{ cm}^{-1}$) was driven at small amplitudes compatible with the QLT. The energy at the driven wave mode was injected at regular time intervals throughout the particle simulation. A comparison of the diffusion coefficients calculated with the methods MI, MIIa and MIIb are shown in figure 4.7. The amplified wave mode creates additional resonances, at which particles are scattered. The running diffusion method shows again a split in the maxima, so that the calculated resonance positions lie almost exactly between them. This is of course explained with the point symmetry of resonance structures in the scatter plot. The diffusion equation solutions MIIa and MIIb show a very good agreement with the pre-calculated resonances. The absolute values of $D_{\mu\mu}$ from all

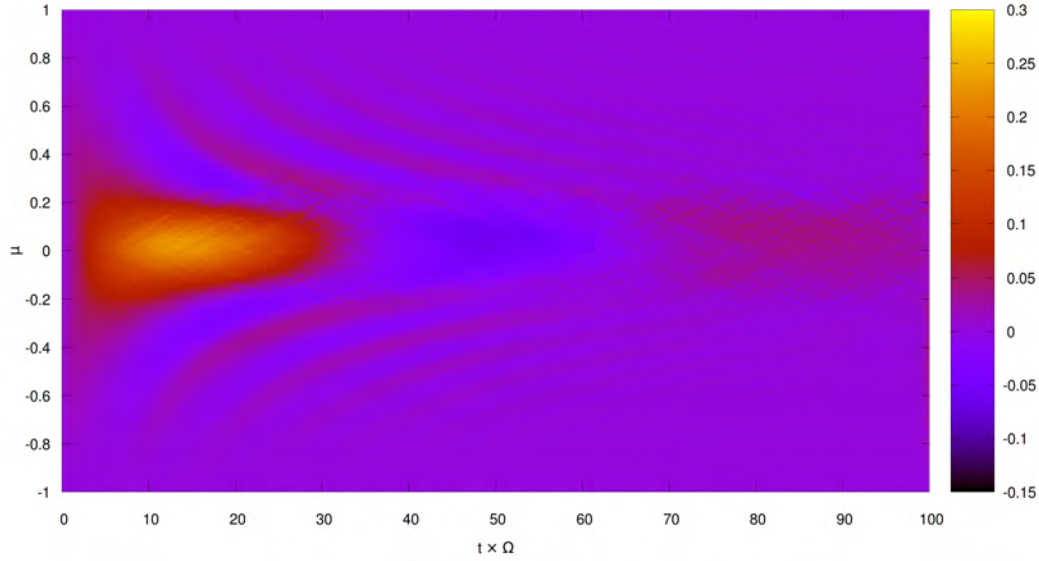


Fig. 4.6: Time dependent pitch angle diffusion coefficient results obtained with the Kubo formalism. The central structure of the Cherenkov resonance is reproduced and the $D_{\mu\mu}$ -evolution up to 30 gyration periods is comparable with the running diffusion results. The method shows strong oscillatory behaviour.

methods are comparable, especially MIIa and MIIb are in very good agreement. A slight disadvantage of these methods can be seen at the right edge of the μ -range: Since the half-parabola distribution required by these methods has its minimum there, two effects come into play. The total particle number at higher μ -values is relatively low causing worse statistics and the μ -derivatives of the distribution approach zero, making the calculations more susceptible to numerical fluctuations. This causes deviations as seen at $\mu > 0.5$, where the diffusion coefficient partially becomes negative. This problem can be addressed with more sophisticated distribution function shapes or by additionally performing every simulation with a mirrored half-parabola distribution and taking an average over the results.

Decaying wave mode

This simulation setup differs only in a small but important detail compared to the previous one. Instead of a continuous turbulence driving at the peak mode, the energy is injected into the mode only once followed by a decay stage. During the decay the energy of the peak mode is transported across perpendicular wave numbers via convection and diffusion. Higher order resonances become relevant leading to a more complex scattering

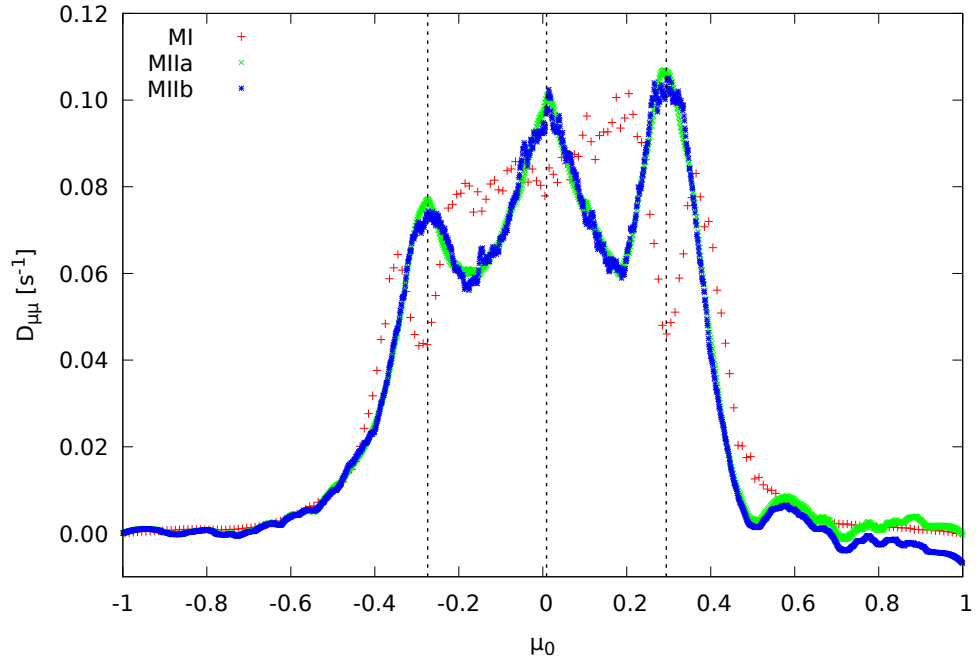


Fig. 4.7: $D_{\mu\mu}$ obtained with MI, MIIa and MIIb for the saturated background turbulence simulation with a single amplified wave mode after 10 gyration periods. The vertical dashed lines indicate resonant pitch angles according to Eq. 2.144. Source: Ivascenko et al. (2016).

pattern. The calculated $D_{\mu\mu}$ are shown in Fig. 4.8. The summary here is more or less the same as in the previous case. All methods produce comparable shapes of the coefficients, when we account for the apparent splitting of the maxima of the running coefficient method, which fall between the calculated resonances, again, due to point-symmetrical resonant scattering. It is apparent from the weaker scattering at negative pitch-angles that the left-hand polarised wave has not developed fully. The right-hand and central resonances were reproduced with MIIa and MIIb very nicely. For this simulation, the half-parabola shape of the μ -distribution was mirrored, identifiable by the numerical instabilities at the left edge of the μ -interval, that do not correspond to resonance peaks. Apart from that all methods show comparable amplitudes, with MIIa and MIIb in excellent agreement. The higher amplitudes compared to the previous cases indicate a stronger turbulence.

Strong turbulence

For the last simulation setup the strength of the turbulence driving was significantly increased for the single wave mode at $k'_{\parallel} = 2\pi \cdot 24$ by increasing the energy injection. This

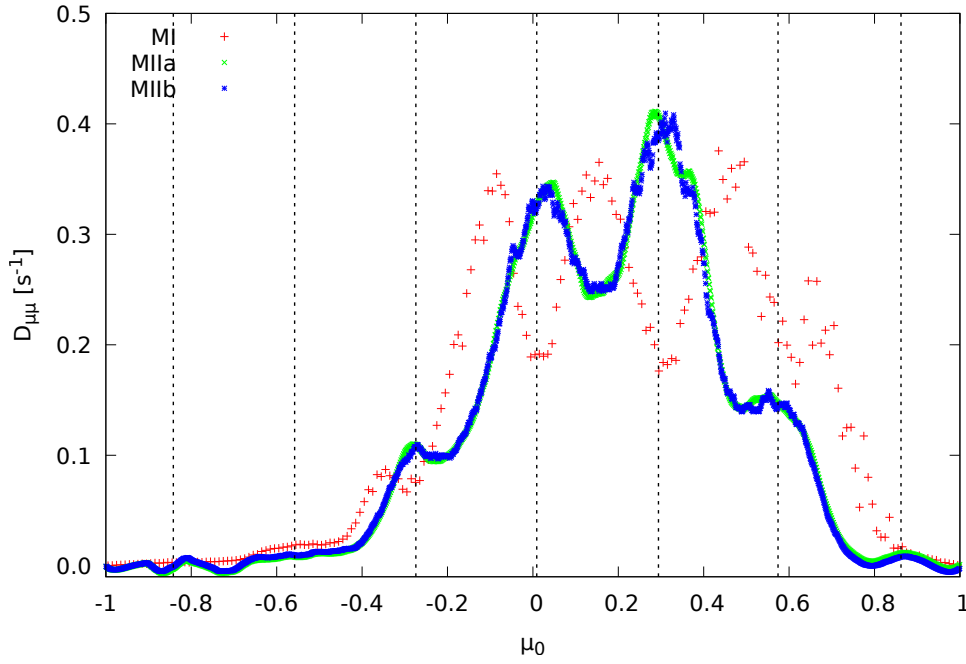


Fig. 4.8: $D_{\mu\mu}$ obtained with MI, MIIa and MIIb for the decaying wave mode simulation after 10 gyration periods. The vertical dashed lines indicate resonant pitch angles according to Eq. 2.144. Source: Ivascenko et al. (2016).

leads to local field amplitudes reaching values up to $\delta B/B_0 \approx 1$, which is far outside the validity regime of the quasi-linear theory. The calculated $D_{\mu\mu}$ for this scenario are shown in Fig. 4.9. As expected, the running diffusion coefficient method MI fails completely due to its strong reliance on the QLT. The diffusion coefficient shape from MI does not match the expected resonance structure at all, even the match of the resonance at $\mu = -0.56$ is purely coincidental. The scatter plot $\Delta\mu(\mu_0)$ does not help to interpret the result in this case, either, since the pitch angle changes from the initial values are too large, making the resonant structures completely indiscernible. Methods MIIa and MIIb yield a similar order of magnitude in the $D_{\mu\mu}$ -amplitude, but a much more structured and realistic general shape. Atop the background scattering, at least the major resonance peaks can be distinguished. The rightmost and two leftmost resonances are too weak to be seen under the background noise. It should be noted that with a strong turbulence like this, results can be achieved easier (and will in most cases be better) with a shorter running time than the 10 gyration periods used here. Due to the strong scattering, the particle distribution function changes significantly at earlier times, which can then be sufficient to calculate the diffusion coefficients with MIIa and MIIb.

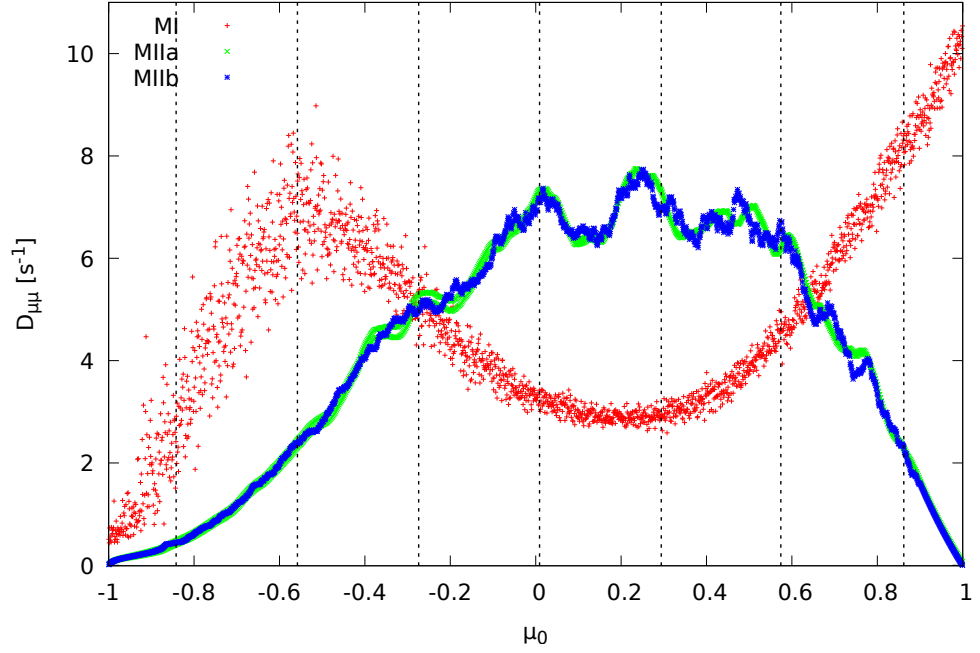


Fig. 4.9: $D_{\mu\mu}$ obtained with MI, MIIa and MIIb for the single wave mode with strong turbulence simulation after 10 gyration periods. The vertical dashed lines indicate resonant pitch angles according to Eq. 2.144. Source: Ivascenko et al. (2016).

4.2 Pitch angle diffusion coefficients

This chapter presents pitch angle diffusion coefficients calculated with the new methods MIIa and MIIb from a range of different turbulence scenarios, simulated with the Gismo MHD-code. The common parameters of all simulations are the outer scale of the simulation volume $L = 3.4 \cdot 10^8$ cm and the particle density $\rho = 10^5$ cm $^{-3}$. The turbulence was driven with continuous energy injection at $k'_\perp \leq 4$ and $k'_\parallel \leq 14$ with a power law scaling $E(k) \propto k^{-2}$. The main reason for the anisotropic turbulence driving is the limited development of the parallel energy cascade, especially in the case of weak turbulence, as has been discussed in chapter 2.2.2. An isotropic energy injection, even in a saturated state, would not lead to a fully developed slab-turbulence. To match heliospheric observations that detect turbulent modes at high parallel wave numbers $k_\parallel$, we artificially strengthen the parallel cascade development by extending the driving region.

4.2.1 Weak turbulence

The first simulation SI was selected as a relatively faithful recreation of the conditions in the solar corona at three solar radii. The magnetic field was set to $B_0 = 0.174$ G, together

with the proton mass density $m_p \cdot \rho = 1.67 \cdot 10^{-19} \text{ g/cm}^3$ resulting in an Alfvén velocity $v_A = 1.2 \cdot 10^8 \text{ cm/s}$. Numerically, the simulation was performed with a grid resolution of 512^3 , which corresponds to a wave number grid of 256^3 modes, of which only 85 are active due to antialiasing. To maximise the inertial region of the turbulence spectrum a hyperdiffusivity parameter of $h = 2$ and a resistivity of $\nu = 1$ were chosen. The hyperdiffusivity approach has the advantage of strengthened dissipation at large wave numbers, so that the steady-state is achieved faster, while simultaneously preventing numerical artefacts caused by wave modes at the edge of the antialiasing region.

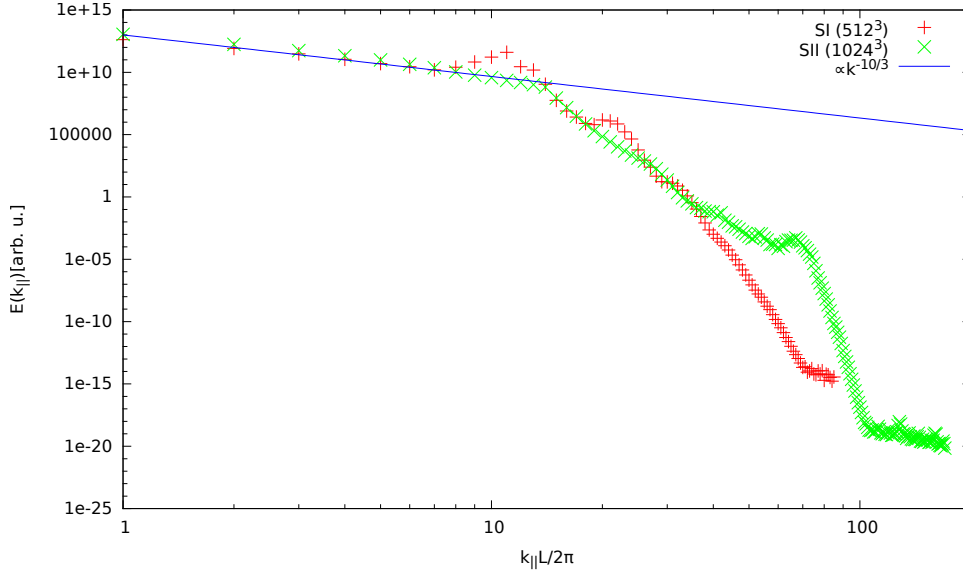


Fig. 4.10: Slab energy spectrum in the saturated state of simulations SI and SII. The magnetic field energy is plotted over the numeric parallel wave number $k'_{\parallel}$ (integrated over $k_{\perp}$). The solid line indicates a power-law spectrum with a spectral index of $-10/3$.

A crucial physical parameter is the turbulence strength $\delta B/B_0$, that is influenced by the energy injection and dissipation rate and, most directly, by the strength of the background magnetic field B_0 . As has been extensively discussed in chapter 2.2, the turbulence strength influences the development of the energy cascade by determining the possible wave-interaction processes. The critical balance parameter $\zeta \propto (k_{\perp} v_{\perp})/(k_{\parallel} v_A)$ depends directly on the turbulence strength and influences the direction of the turbulence development parallel or perpendicular to B_0 .

In the first run we simulate relatively weak turbulence with $\delta B/B_0 \approx 0.15$. Consequently, the parallel cascade development is suppressed but not completely disabled. As can be observed in Fig. 4.10 a small pile-up develops at $k'_{\parallel} = 11$ and energy is transported in parallel direction mainly by exciting higher-order harmonics at $k'_{\parallel} = 22, 33, 44$. This behaviour is consistent with findings of Lange (2012) and results in a rather short

inertial range where a $-5/3$ power-law is reproduced, followed by a much steeper cascade that is mainly defined by the shape of the harmonic peaks.

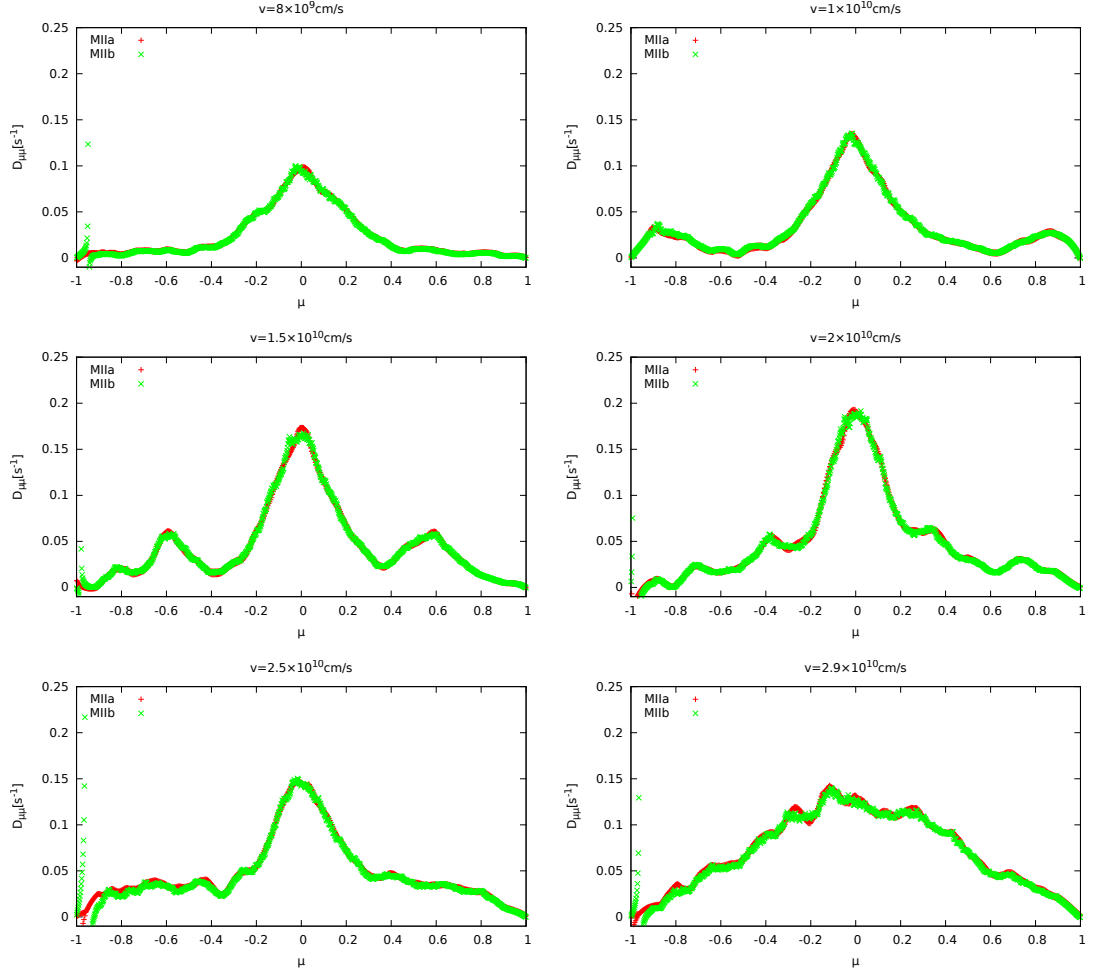


Fig. 4.11: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated with methods MIIa and MIIb for 6 different test-particle energies in the weak turbulence simulation SI with $\delta B/B_0 \approx 0.15$. The calculations were performed after 10 gyration periods. Test particle speed is indicated above the respective plot.

The shape of the spectrum is also reflected in the pitch angle coefficients, shown in Fig. 4.11 for multiple different test-particle speeds. The lowest particle speed $v = 8 \cdot 10^9$ cm/s was selected to have no $|n| \geq 1$ resonances for $k'_{\parallel} < 12$ according to the resonance condition 2.144. Thus the main feature of the pitch angle coefficient is the Cherenkov ($n = 0$) resonance at $\mu \approx 0$, which is quite prominent with an amplitude of 0.1 s^{-1} .

As the particle speed increases to $v = 10^{10}$ cm/s, the more energetic waves at $k'_{\parallel} = 11$ can interact resonantly with particles moving at extremely acute angles to the magnetic

field, specifically at $\mu = -0.81$ and $\mu = 0.83$. Consequently, secondary maxima appear in the $D_{\mu\mu}$ -shape just at the borders of the μ -interval. There is also a notable increase in the amplitude of the Cherenkov resonance peak to 0.14 s^{-1} .

The amplitude continues to grow as the particle speed is increased to $v = 1.5 \cdot 10^{10} \text{ cm/s}$ and $v = 2 \cdot 10^{10} \text{ cm/s}$ with $D_{\mu\mu}(0) = 0.17 \text{ s}^{-1}$ and $D_{\mu\mu}(0) = 0.19 \text{ s}^{-1}$, respectively. As expected from the resonance condition, the secondary peaks move closer to the central one with increasing particle speed over $|\mu| \approx 0.55$ to $|\mu| \approx 0.4$.

As we go to even higher particle energies ($v = 2.5 \cdot 10^{10} \text{ cm/s}$), the diffusion coefficient starts to lose its structure, as the amplitude of the Cherenkov resonance decreases to 0.14 s^{-1} again while simultaneously becoming broader and less prominent. The background scattering over the whole μ -range increases, making the secondary peaks at $|\mu| \approx 0.33$ practically invisible.

This process continues at extreme particle energies. As can be seen in the bottom right panel of Fig. 4.11, for particles with $v = 2.9 \cdot 10^{10} \text{ cm/s}$ the structure of the pitch angle coefficient disappears almost completely as the background scattering increases significantly at all pitch angles. The secondary peaks can be barely distinguished above the background noise at $|\mu| \approx 0.29$. The most probable explanation for the increased scattering lies in the relativistic modification of the resonance condition. In the relativistic case, the gyrofrequency is modified by a factor γ^{-1} . When taken into account in the resonance condition Eq. 2.144, this lowers the resonant wave number for a given μ

$$k_{\parallel} = \frac{n\Omega}{\gamma(\mu v - v_A)}. \quad (4.14)$$

In other words, the possible pitch-angle range of resonant interactions with every wave-mode $k_{\parallel}$ is broadened by the factor γ , effectively leading to more particle scattering, especially at low wave numbers that contain most of the turbulent energy.

4.2.2 Weak turbulence with higher resolution

In addition to the simulation SI discussed in the previous chapter, a very similar simulation SII with the exact same physical parameters but a higher resolution of 1024^3 was performed. The slab spectrum of the turbulence is shown in Fig. 4.10 and has not reached a saturated steady-state yet. The obvious distinction to SI is the missing pile-up peak at $k'_{\parallel} = 11$, instead the spectrum shows a strong break to a much steeper power-law after $k'_{\parallel} = 14$. The total turbulence strength of the magnetic field with $\delta B/B_0 \approx 0.06$ is lower than in SI by a factor of ~ 2.5 , though from a comparison of the spectra it is clear that the excess energy in SI must almost exclusively be located in the pile-up peak around $k'_{\parallel} = 11$, since the energy in the first 7 wave modes is slightly higher in SII and the contribution from modes after the spectral break becomes negligible rather quickly.

Figure 4.12 shows the pitch angle coefficients for different test-particle energies. The most notable difference is that despite very similar turbulence spectra, the Cherenkov resonance peaks are more prominent and consistently higher by a factor of ≈ 1.5 compared to SI. This inevitably leads to the conclusion that the $n = 0$ resonance is caused by

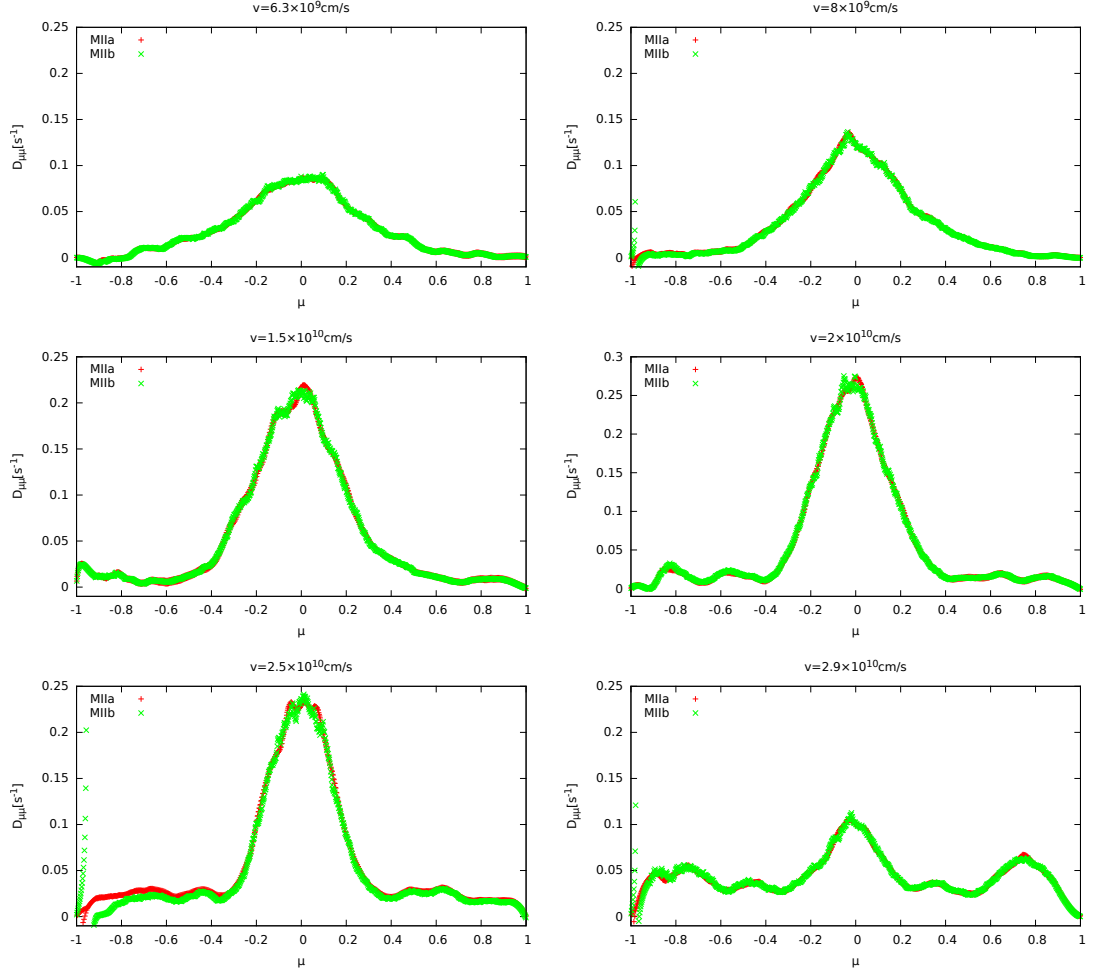


Fig. 4.12: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated with methods MIIa and MIIb for 6 different test-particle energies in the weak turbulence simulation SII with $\delta B/B_0 \approx 0.06$. The calculations were performed after 10 gyration periods. Test particle speed is indicated above the respective plot.

scattering at large wavelengths, since the energy in small wave numbers in SII is larger by a factor ≈ 2 , while the overall turbulence is weaker.

A rather obvious difference is the absence of the secondary maxima, especially for $v = 1.5 \cdot 10^{10}$ cm/s and $v = 2 \cdot 10^{10}$ cm/s particles, in SI caused by scattering with the wave modes around $k'_{\parallel} = 11$, which are not amplified in SII. Missing interactions with these wave modes also explain the overall lower noise level, which is especially notable for the highest particle energy. With a Lorentz factor of $\gamma = 3.9$ the relativistic correction becomes significant, allowing more interactions at lower pitch angles and wave numbers. The peaks at $|\mu| \approx 0.8$ in the lower right panel of Fig. 4.12 are caused by the

$n = 1$ resonance with the most energetic wave mode at $k'_{\parallel} = 1$, which makes them rather prominent. The corresponding plot in Fig. 4.11 shows much higher scattering amplitudes at small pitch angles, since resonances with higher wave numbers play a bigger role there.

The lowest particle energy with $v = 6.3 \cdot 10^9$ cm/s was chosen to have only resonant interactions above the spectral break at $k'_{\parallel} = 14$. Consequently the overall scattering amplitude is low and the pitch angle coefficient shows no prominent features.

4.2.3 Strong turbulence

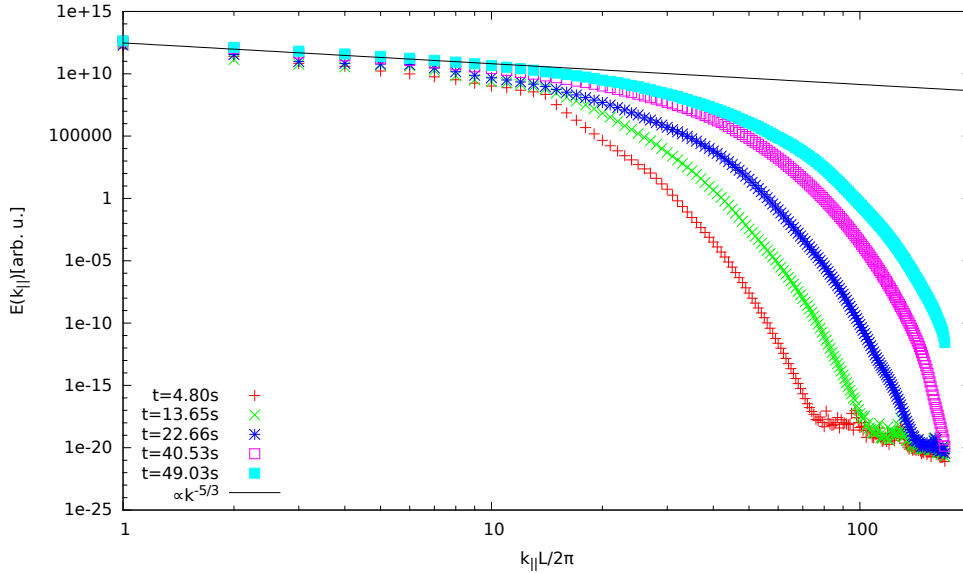


Fig. 4.13: Selected time steps of development and the final saturated state of the slab energy spectrum for the simulation setup SIII. The magnetic field energy is plotted over the numeric parallel wave number $k'_{\parallel}$ (integrated over $k_{\perp}$). The solid line indicates a power-law spectrum with a spectral index of $-5/3$.

Next, we investigate pitch angle scattering in strong turbulence. The simulation setup SIII chosen for this purpose was achieved by reducing the background magnetic field from the previous setups by a factor 10, yielding $B_0 = 0.0174$ G. Since the proton mass density was kept constant, the Alfvén velocity also declined by a factor 10 to $v_A = 1.2 \cdot 10^7$ cm/s, as it is coupled directly to the background magnetic field. In order to achieve the saturated steady-state of the energy cascade faster, the hyperdiffusivity was strengthened to $h = 3$. Accordingly, the resistivity parameter had to be reduced to $\nu = 5 \cdot 10^{-7}$, to prevent energy transport across the grid with unphysical velocities. After the slab spectrum has reached a saturated state, that is shown in figure 4.13 at $t = 49.03$ s, the overall turbulence strength amounted to $\delta B/B_0 \approx 0.45$, which is

classified as relatively strong turbulence.

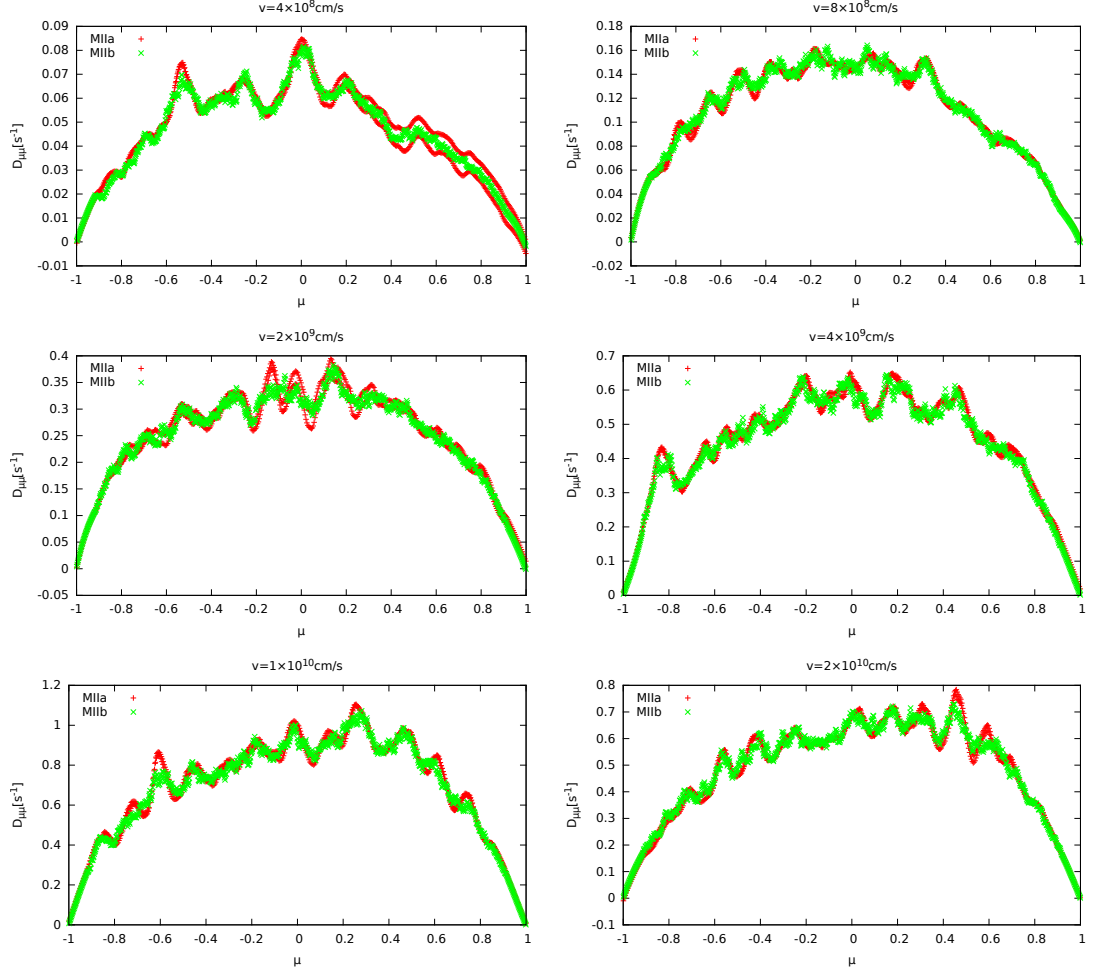


Fig. 4.14: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated with methods MIIa and MIIb for 6 different test-particle energies in the strong turbulence simulation SIII with $\delta B/B_0 \approx 0.45$ after 1 gyration period. The pitch angles definition in this calculation uses the local magnetic field at each particle's position. Test particle speed is indicated above the respective plot.

Since the perturbations in SIII lie in the same order of magnitude as the background magnetic field, it warrants the discussion to what extent the definition of a background field is still meaningful. The primary significance of the magnetic field definition in our particular case is its role in the calculation of the pitch angle. As described in chapter 2.3 the transformation of the particle momentum to spherical coordinates and, consequently, the definition of the pitch angle μ (see Eq. 2.125) stem from the assumptions of QLT, where the background magnetic field is always well-defined and the turbulence can

be seen as a small perturbation. The validity of this assumption becomes hard to argue in this case with $\delta B/B_0 \approx 0.45$, as the concept of “unperturbed particle trajectories” loses its relevance. An alternative definition of the pitch angle can be introduced, in which the local magnetic field at the position of the particle is used to calculate μ . Since the magnetic field from the grid needs to be interpolated to every particle’s position to calculate the velocity and coordinate update in every time step, the local field information is already available. In order to analyse the effect of the pitch angle definition, the coefficients $D_{\mu\mu}$ have been calculated for 6 particle energies in SIII, first using the local field definition of μ (see figure 4.14) and second, with μ calculated against the background field B_0 (see figure 4.15).

As can be expected from strong turbulence, the scattering amplitudes are rather high across the whole pitch angle range (going to zero at $|\mu| = 1$ due to the implicit $(1 - \mu^2)$ term in the definition). There is no prominent central Cherenkov resonance, though the maximum is mostly near $\mu = 0$, again, most likely due to the $(1 - \mu^2)$ term. The peaks overlaying the general structure are most likely numerical fluctuations for the biggest part, although due to the long extent of the inertial range of the spectrum there are more wave modes available for resonant interactions.

A notable feature of the pitch angle coefficients calculated with the local fields is the strong dependence of the amplitude on the particle energy. As can be seen in Fig. 4.14, there is an almost perfect direct proportionality between the particle speed and the maximal scattering amplitude up to $v = 10^{10}$ cm/s. At $v = 2 \cdot 10^{10}$ cm/s the scattering amplitude decreases again, which is also generally consistent with our findings for weak turbulence.

In direct comparison, the scattering coefficients calculated with the background magnetic field in Fig. 4.15 show qualitatively similar results with strong scattering across the μ -range with fluctuations overlaying the general structure. The slant in the plots is most likely owing to the depletion of the particle distribution at large pitch angles due to the strong scattering. The μ -histogram reaches a steady state with a level slope relatively fast, after which the shape remains constant, underestimating the time derivative of the distribution $\partial_t f(\mu)$ and, consequently, underestimating $D_{\mu\mu}$.

The most significant differences between the coefficients based on the two different pitch angle definitions lie in their absolute values. The $D_{\mu\mu}$ in Fig. 4.15 show significantly bigger scattering amplitudes with only a very weak dependency on the particle energy up to the $v = 10^{10}$ cm/s particles. At relativistic particle speeds the scattering strength decreases noticeably (seen in the lower right panel), which is consistent with the behaviour observed in all previous simulations. The comparison of the amplitude scaling behaviours is shown in figure 4.16. Power-law curves were fitted to the data points to indicate the linear increase of the local μ amplitudes and the slight decrease of the background μ results with the particle velocity v . Also visible is the strong amplitude drop at relativistic particle velocities.

The strong pitch angle changes with respect to the background field are interpreted as a sign that particles in strong turbulence are almost completely decorrelated from the magnetic field lines and their behaviour is governed predominantly by the local field fluctuations. The concept of unperturbed particle trajectories becomes meaningless in

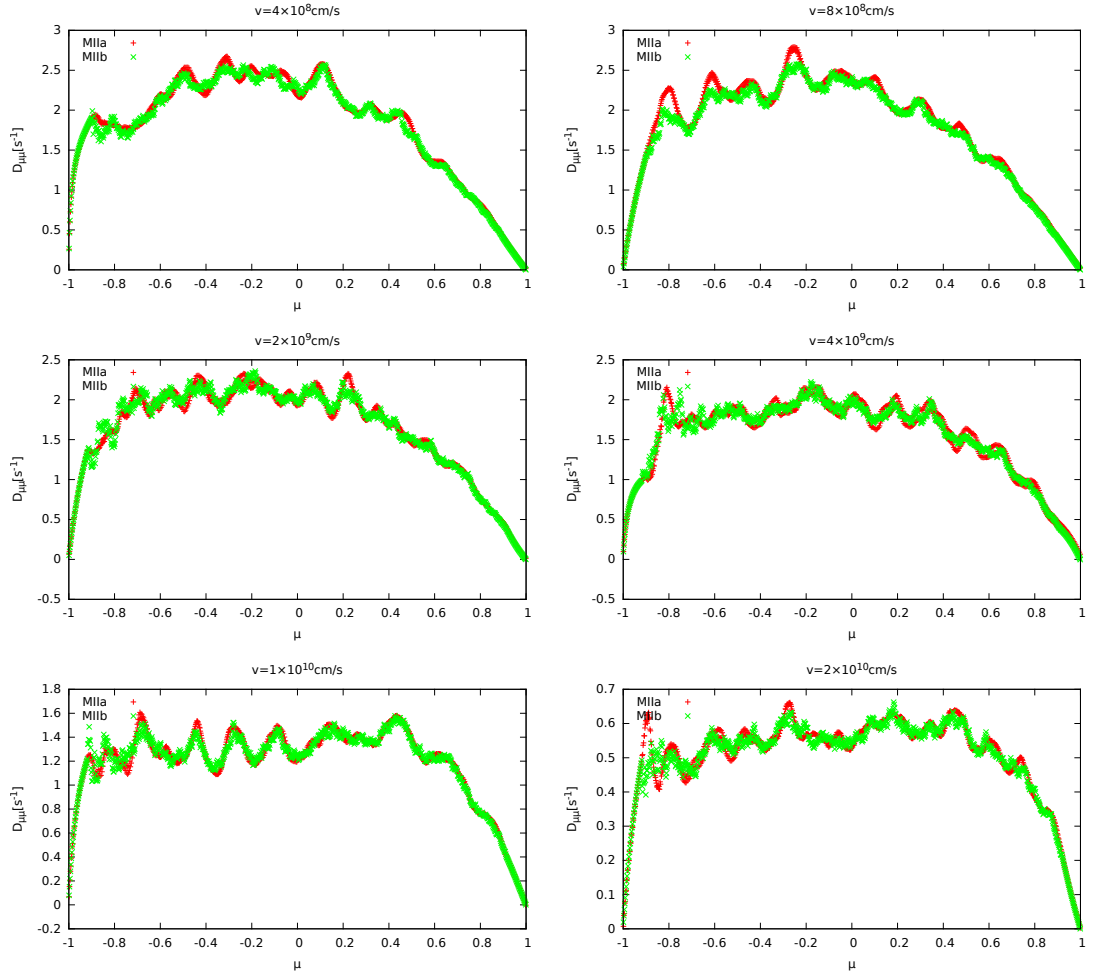


Fig. 4.15: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated with methods MIIa and MIIb for 6 different test-particle energies in the strong turbulence simulation SIII with $\delta B/B_0 \approx 0.45$ after 1 gyration period. In this calculation the background field B_0 was used in the pitch angle definition. Test particle speed is indicated above the respective plot.

this context as does the classical definition of the pitch angle diffusion coefficient, as the pitch angle of the particles changes significantly over the course of the calculation. Considering the total particle distribution, however, can still yield valid results, as long as care is taken to calculate the immediate local μ of the particles and not to let the distribution function reach a dynamic equilibrium.

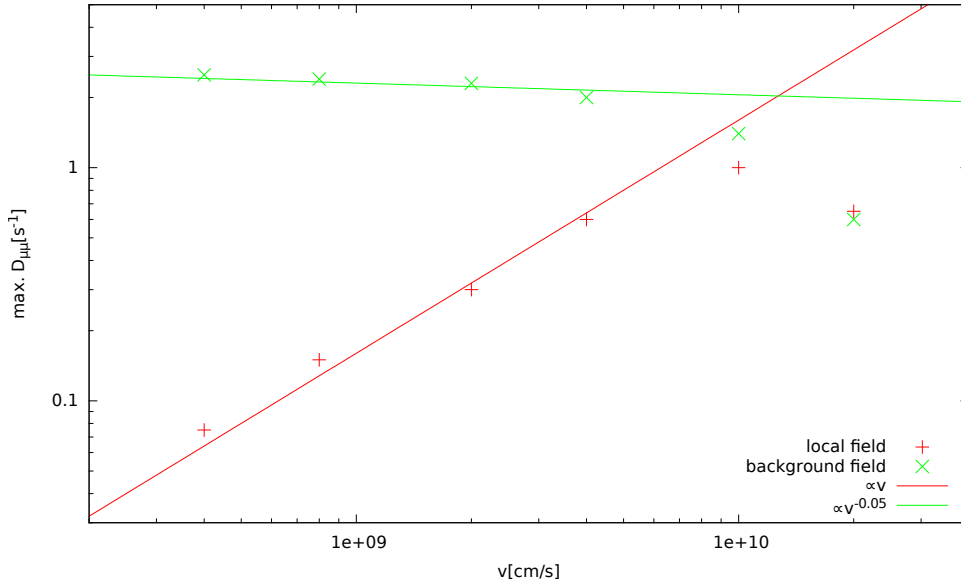


Fig. 4.16: A comparison of maximal $D_{\mu\mu}$ amplitudes in SIII with the pitch angle calculated with respect to the global background field and to the local fields at the particles' positions.

4.2.4 Perlin noise results

The last step of the section is to investigate the properties of the Perlin noise turbulence by calculating diffusion coefficients. For this purpose the simulation setup SIII, particularly the slab turbulence spectrum, was recreated with Perlin noise as far as possible. Subsequently the turbulence amplitude was adjusted to analyse particle scattering at different turbulence strengths. The first simulation set SIV aims to produce weak turbulence with $\delta B/B_0 \approx 0.1$. The resulting $D_{\mu\mu}$ for 6 test-particle energies are shown in Fig. 4.17. A distinct feature of the coefficients is immediately apparent: the local minimum at $\mu \approx 0$. Since the Perlin noise is generated in one dimension and applied only perpendicularly to the background magnetic field, the condition $\delta \mathbf{B} \cdot \mathbf{B} \neq 0$ for the Cherenkov ($n = 0$) resonance can never be fulfilled. In other words, there are no pseudo-Alfvén waves for resonant interactions with particles at $\mu = 0$. The absence of the Cherenkov peak is especially noticeable at low overall interaction energies: low particle velocities and low turbulence strengths. At higher energies the overall scattering amplitudes grow so that particles are scattered significantly at and around $\mu = 0$, thus smearing out the minimum. Additionally, in stronger turbulence the definition of the pitch angle becomes more ambiguous, as discussed in the previous section. Using the local magnetic field at the particles' positions increases the prominence of the missing Cherenkov resonance slightly, except for the two lowest particle energies.

In general, the diffusion coefficients in Perlin noise are much less sensitive to the

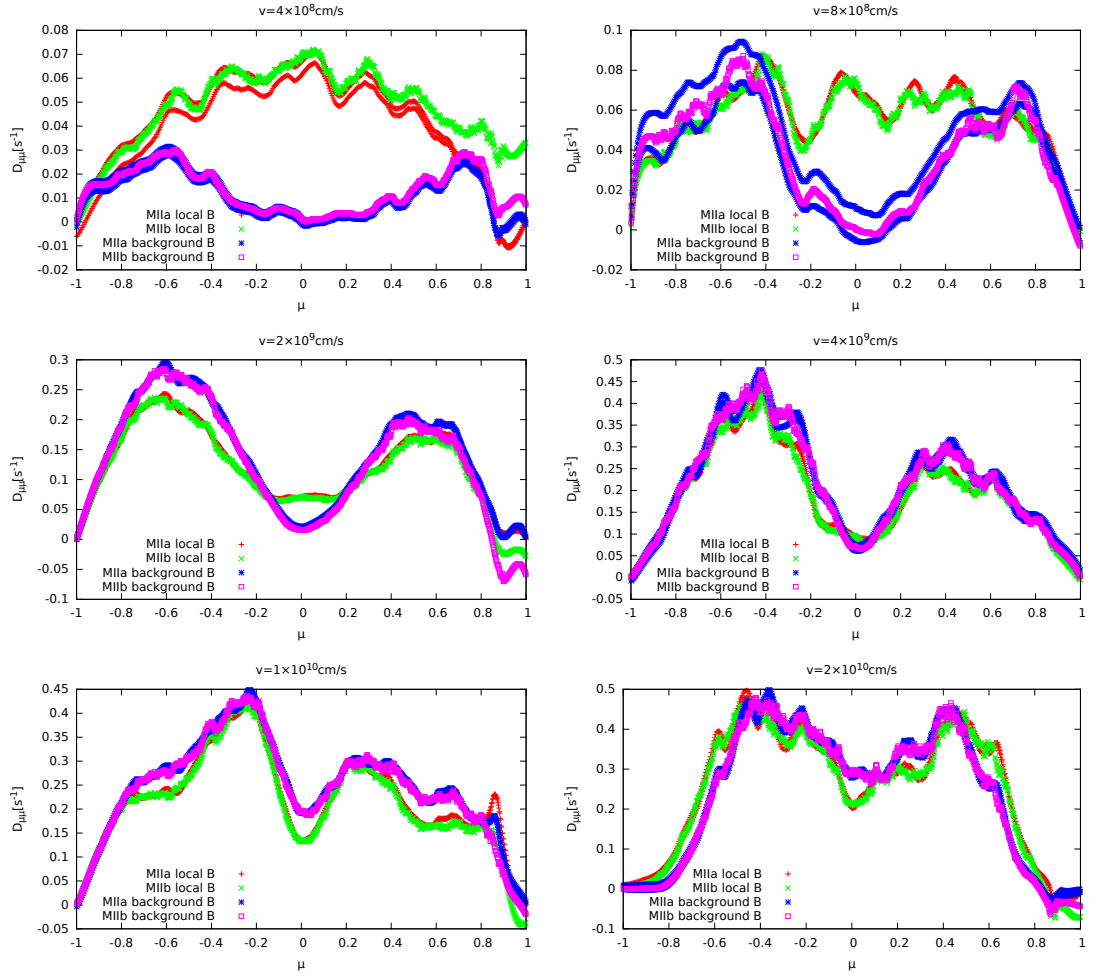


Fig. 4.17: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated from the Perlin noise simulation SIV with $\delta B/B_0 \approx 0.1$ after 1 gyration period. Test particle speed is indicated above the respective plot.

magnetic field choice used in the calculation of the pitch angles compared to the GISMO simulations. A reasonable explanation of the cause can be found in the dimensionality of the turbulence. It is apparent that the three-dimensional magnetic field fluctuations in GISMO are much more chaotic, since they have an additional degree of freedom over the two-dimensional turbulence of the Perlin noise. This leads to an overall stronger perturbation and “twisting” of the field lines in MHD turbulence.

The scattering amplitudes in Perlin noise are consistently higher compared to GISMO simulation of similar turbulence strength. The reason lies in the scaling method that is used in the generation of the Perlin noise spectrum, i.e. the fluctuation strength was chosen so that the total slab spectra $E(k_{\parallel})$ are equal. In GISMO this fluctuation energy

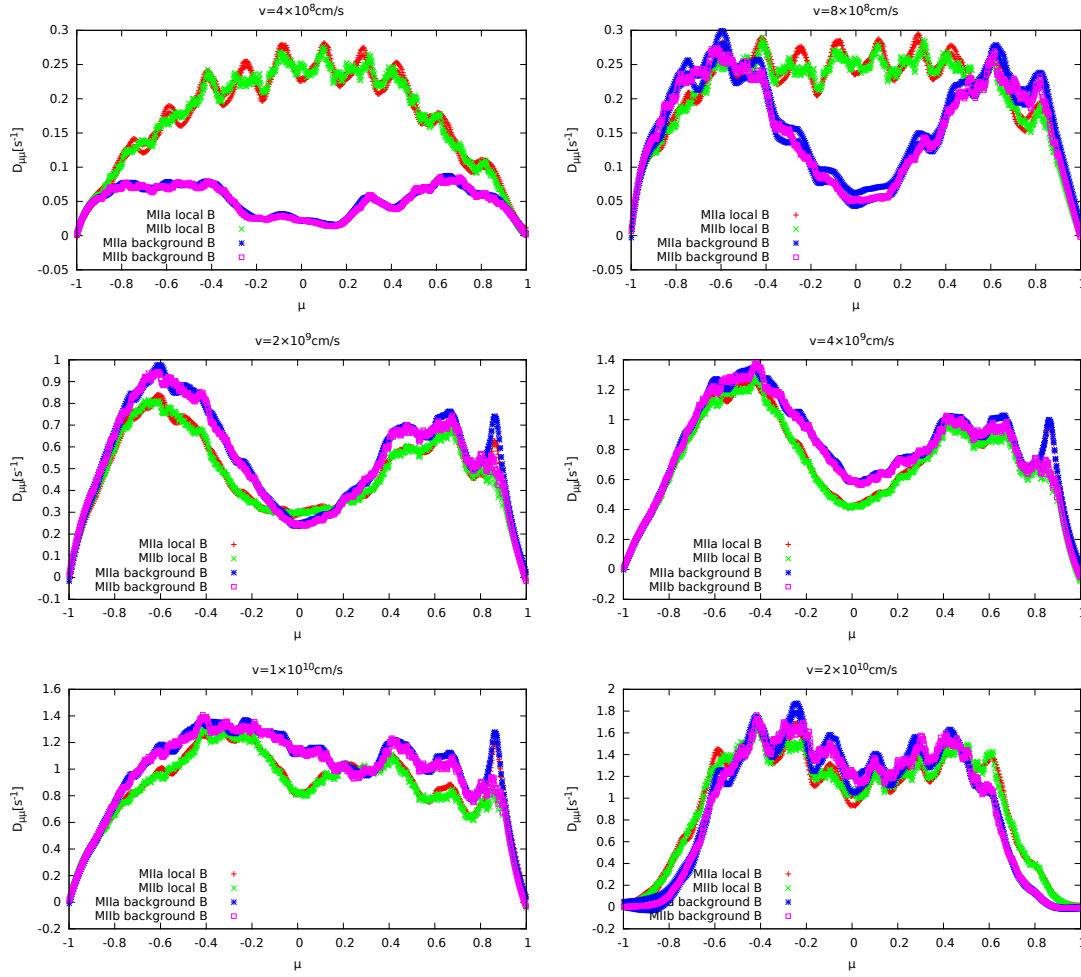


Fig. 4.18: Pitch angle diffusion coefficients $D_{\mu\mu}$ calculated from the Perlin noise simulation SV with $\delta B/B_0 \approx 0.2$ after 1 gyration period. Test particle speed is indicated above the respective plot.

$|\delta B|^2$ is distributed over three all dimensions, while in Perlin it is confined to the plane perpendicular to the background magnetic field to conserve solenoidality. This leads to a higher effective amplitude of the shear Alfvén waves propagating parallel to B_0 in Perlin, thus leading to stronger scattering.

In order to account for this effect, in the Perlin simulation SVI (see Fig. 4.19) a correction factor of $\sqrt{2/3}$ was applied to the turbulence strength to reproduce the results of SIII (Fig. 4.14). The factor $\sqrt{2/3}$ is based on the assumption of an isotropic turbulence, which is only fulfilled approximately in this case. However, for this relatively strong turbulence, the assumption can be made.

The scattering amplitudes, especially when calculated against the background mag-

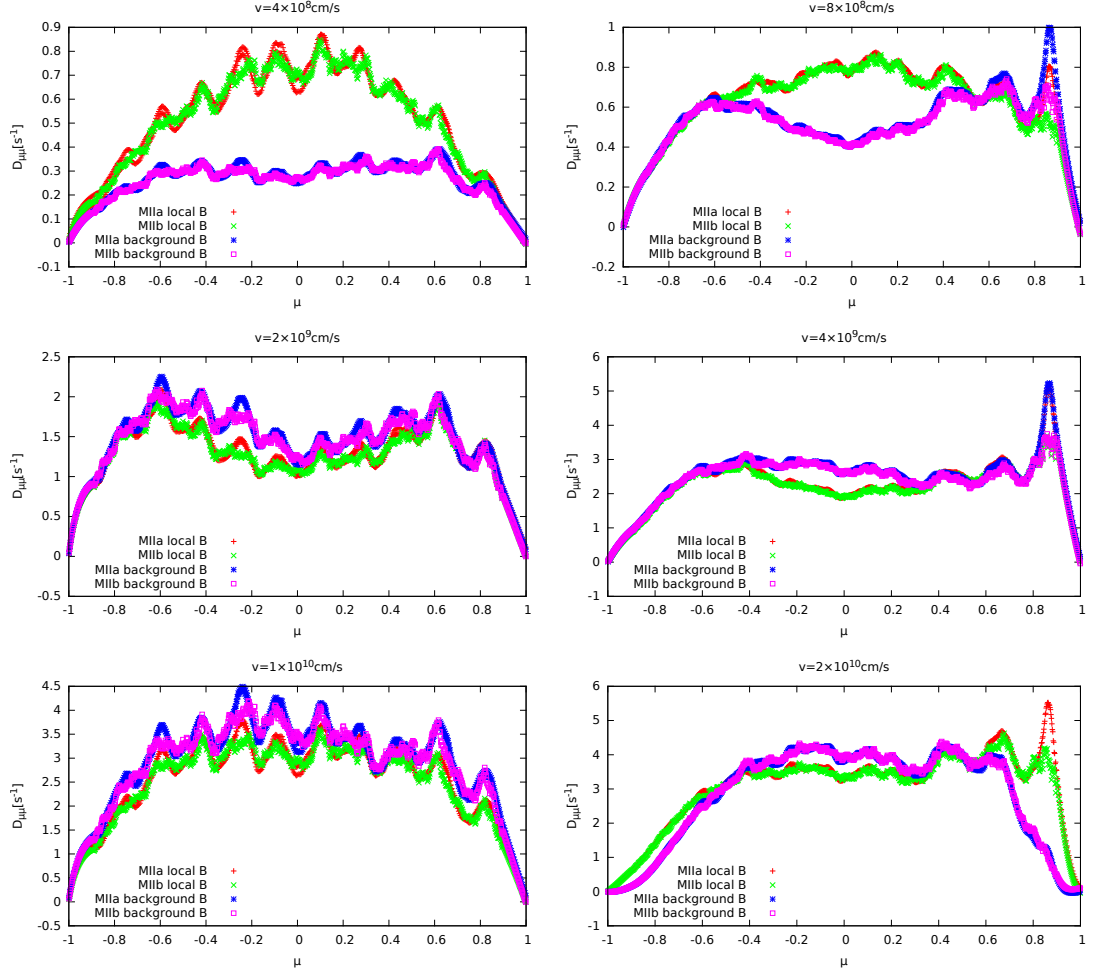


Fig. 4.19: Pitch angle diffusion coefficient $D_{\mu\mu}$ calculated from the Perlin noise simulation SVI with $\delta B/B_0 \approx 0.365$ after 1 gyration period. Test particle speed is indicated above the respective plot.

netic field, are still consistently higher in the Perlin noise by a factor of ≈ 5 , even with the correction. The form and amplitude scaling are, however, in relatively good agreement. Due to the overall strong scattering, the absence of the Cherenkov resonance peak in Perlin is less prominent (though still noticeable), so that the form of the coefficients is determined by strong scattering at all pitch angles. The amplitude scaling with test particle energy is, again, almost linear at low energies with a noticeable down-turn at relativistic particle velocities $v = 1 \cdot 10^{10}$ cm/s and $v = 2 \cdot 10^{10}$ cm/s, as can be seen in Fig. 4.20.

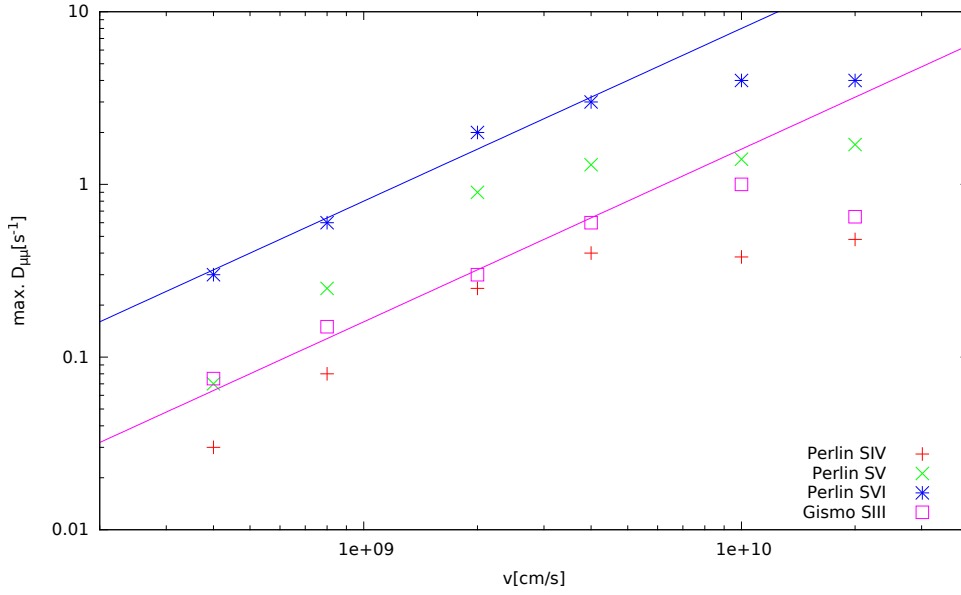


Fig. 4.20: A comparison of maximum scattering amplitudes from the Perlin noise simulations SIV-SVI and the strong turbulence GISMO simulation SIII at different particle energies. The pitch angle was calculated with respect to the global background field. Solid lines indicate the linear scaling of $D_{\mu\mu}$ with particle energy.

4.3 Perpendicular diffusion coefficients

In order to demonstrate the wide applicability of the concept used in calculating the pitch angle diffusion coefficients and to examine a further important particle transport property, in the final section the same principal shall be applied to calculation of the perpendicular spacial diffusion coefficient $D_{\perp}$. Using the diffusion-convection transport equation 2.135 as the starting point, we choose a particle distribution function that is monoenergetic with isotropic phase Ψ and uniformly distributed pitch-angle μ and two of the spacial coordinates y and z . For the third spacial coordinate x , normalised to values in $[0, 1]$, a rectangular distribution is chosen as defined by:

$$f(x) = \begin{cases} 0, & 0 \leq x < 0.4, \\ f_0, & 0.4 \leq x \leq 0.6, \\ 0, & 0.6 < x \leq 1, \end{cases} \quad (4.15)$$

where f_0 is a constant particle line density, defined by the total particle number divided by the width of the rectangular distribution (see Fig. 4.21). Since the perpendicular turbulence is homogeneous and isotropic, we can assume the diffusion coefficient to have no spacial dependence. Together with the choice of the distribution function, the

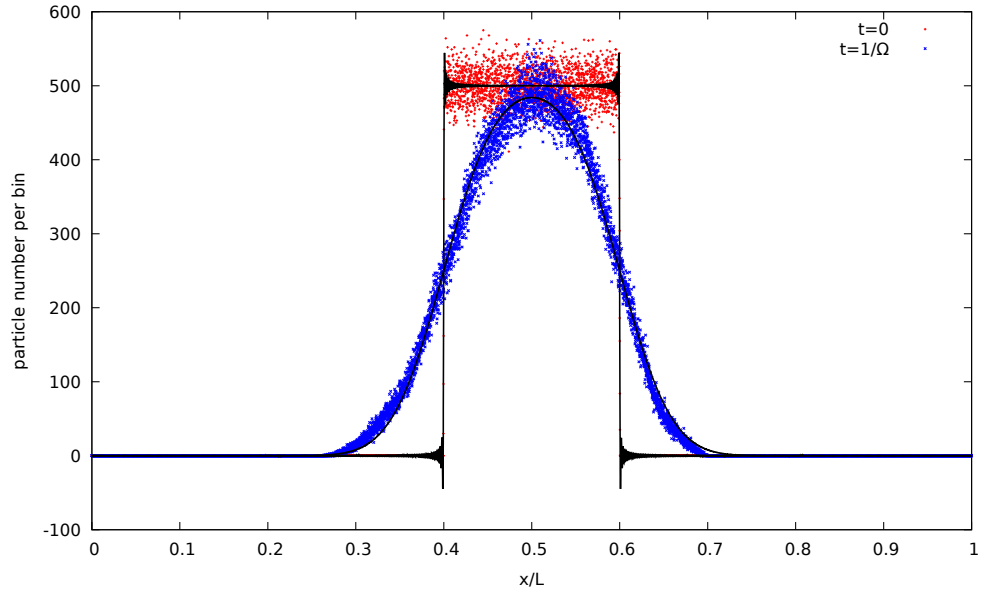


Fig. 4.21: Initial rectangular particle distribution $f(x)$ and broadening after a run time of one gyration period for simulation setup SI with test particle velocity $v = 2 \cdot 10^{10}$ cm/s. Black lines indicate the semi-analytical solution.

transport equation is thus simplified to:

$$\frac{\partial f(x)}{\partial t} = D_{\perp} \frac{\partial^2 f(x)}{\partial x^2}. \quad (4.16)$$

This differential equation can be solved analytically using separation of variables. Applying the boundary conditions of the initial distribution results in a Fourier series decomposition for the spacial part and an exponential time dependency for each Fourier mode:

$$f(x, t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos(2\pi n x) + B_n \sin(2\pi n x)) \cdot \exp(-(2\pi n)^2 D_{\perp} t), \quad (4.17)$$

with the Fourier coefficients

$$A_n = 2 \int_0^1 dx f(x) \cos(2\pi n x), \quad (4.18)$$

$$B_n = 2 \int_0^1 dx f(x) \sin(2\pi n x). \quad (4.19)$$

The infinite sum converges quickly, allowing us to evaluate it numerically for the first N elements, where N was chosen to be equal to the spacial grid resolution in the considered

simulation, resulting in a relative error that is below machine precision.

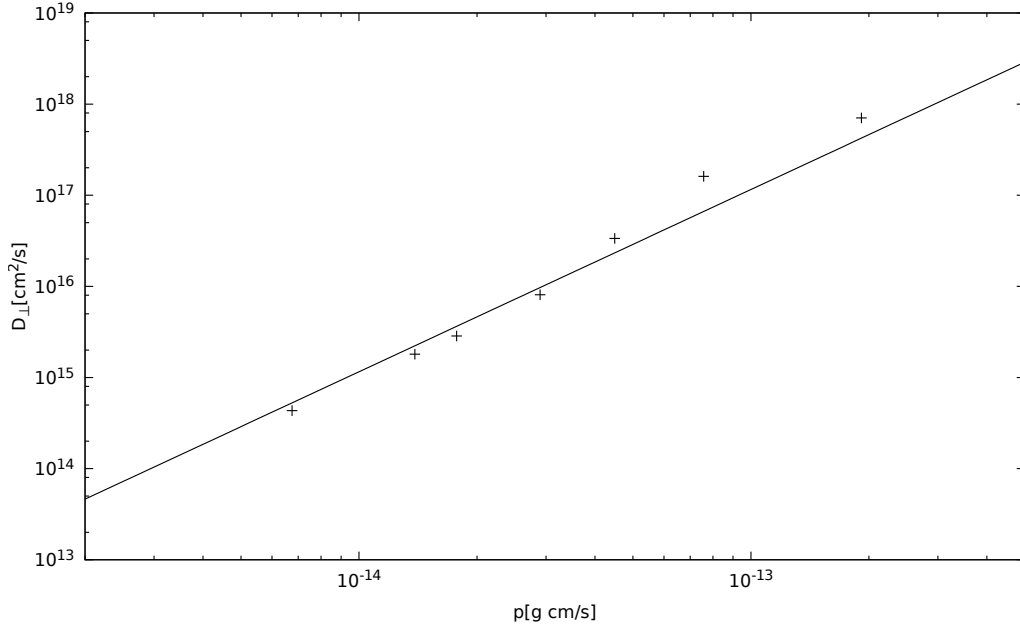


Fig. 4.22: Perpendicular diffusion coefficients plotted over particle momentum. A least-square fit indicates a power-law with a spectral index of 2.3, indicated by the black line.

In order to calculate the diffusion coefficient, the simulation is evaluated after a run time t , when the particle distribution is significantly broadened by the diffusion as can be seen in Fig. 4.21. A characteristic property of the distribution is chosen and compared to the analytical solution, specifically, we evaluate the number of particles $N_{\text{in}}(t)$ that are still within the original boundaries $0.4 \leq x \leq 0.6$ and numerically solve the equation

$$N_{\text{in}}(t) = \int_{0.4}^{0.6} dx f(x, t) \quad (4.20)$$

for $D_{\perp}$.

The calculations of the perpendicular diffusion coefficients were done using turbulence setup SI, as described in section 4.2.1, which is a relatively weak turbulence with $\delta B/B_0 \approx 0.15$ in a magnetic field of $B_0 = 0.174$ G. Test particle distributions with the described properties and seven different particle energies were injected and the number of particles in the x -interval $[0.4, 0.6]$ after one gyration period was counted. The resulting $D_{\perp}$ are summarised in Table 4.1 and plotted against the particle momenta in Fig. 4.22.

A least-squares fit of $D_{\perp}(p) \propto p^{\alpha}$ indicates a spectral index of $\alpha = 2.3 \pm 0.1$ or, if plotted over the particle energy, $\alpha = 1.33 \pm 0.07$. When compared with Giacalone and Jokipii (1999), the diffusion coefficients are roughly of the same order of magnitude,

4.3 Perpendicular diffusion coefficients

$v[\text{cm/s}]$	$4 \cdot 10^9$	$8 \cdot 10^9$	10^{10}	$1.5 \cdot 10^{10}$	$2 \cdot 10^{10}$	$2.5 \cdot 10^{10}$	$2.9 \cdot 10^{10}$
$D_{\perp}[\text{cm}^2/\text{s}]$	$4.3 \cdot 10^{14}$	$1.8 \cdot 10^{15}$	$2.9 \cdot 10^{15}$	$8.1 \cdot 10^{15}$	$3.4 \cdot 10^{16}$	$1.6 \cdot 10^{17}$	$7.1 \cdot 10^{17}$

Table 4.1: Perpendicular diffusion coefficient $D_{\perp}$ calculated in MHD simulation SI for seven test particle energies.

although the power-law in this work is significantly steeper.

5 Summary

Particle transport in astrophysical plasmas, be it the heliosphere, interstellar or intergalactic medium, is a wide field and our analytical models of it are rather limited, and yet it is crucial to our comprehension of many astrophysical phenomena. This work can only hope to make a small step on the long way towards a more complete picture by showing limits of some analytical approaches and possibilities to pass those limits with new numerical and statistical methods.

By combining a realistic model of incompressible heliospheric MHD turbulence delivered by the self-consistent spectral code GISMO with an apt choice of the particle distribution function, we were able to devise a method to calculate transport parameters in situations where other methods fail. In particular, statistical analysis of the distribution function delivers reliable results in strong turbulence scenarios, where other diffusion coefficient calculation methods, like running diffusion and the Kubo formalism, that were primarily designed for analytic calculations and with the assumptions of the quasi linear theory in mind, reach their limits. Additionally, the newly developed methods are very easy to apply to simulation data and could even be used on satellite measurements of angular particle distributions to calculate pitch angle coefficients from heliospheric data.

The new methods have their limitations and areas of potential improvement that have been discussed both in the “Methodology” and “Results” sections. The sensitivity to statistical noise could largely be negated by using the Savitzky-Golay smoothing method, but the impact of the initial distribution choice remains. In simulations, where the particle distribution can be chosen practically arbitrarily, this is a minor nuisance, for experimental data this could potentially be a bigger obstacle, if the slope of the distribution is insufficient (although one could argue that in the interesting cases an anisotropy of the particle distribution would produce a non-vanishing slope).

In the end, these implementation problems are only of minor importance compared to the still open questions, some of them of a rather fundamental nature. For example, as has been shown in this work, leaving the realm of validity of the QLT and going to stronger turbulence brings some interesting implications for the classical particle transport picture. When the magnetic field fluctuations become too large to be seen as small perturbations of an otherwise homogeneous background field, the definition of the pitch angle becomes at least ambiguous, if not completely misleading. So it is not surprising that diffusion coefficients with pitch angles calculated against the background magnetic field are deviating from those calculated with the local field that acts on the particle, and that this deviation grows stronger with increasing turbulence strength. Considering this, the question should be asked, if the definition of a pitch angle is useful at all in strong turbulence when $\delta B/B_0 \approx 1$. Can we still assume a gyration motion around field

lines if the field lines themselves are lost in stochastic noise? And if not, where is the transition region between a well defined background field with applicability of the QLT and random scattering, and what happens in that region?

Fundamental questions aside, the result and methods of this work can be used as a foundation for further research. A potentially important topic would be the application of the presented methods to observational data and a stronger comparison of simulations and measurements. The EU project “SEPServer” that was concluded in 2013 can be used to comfortably gain access to satellite data. A further interesting topic of research, that has only been mentioned briefly in Ivascenko et al. (2016), would be the investigation of anomalous diffusion. There are two possible ways to approach this subject with the diffusion equation methods. One way would be to change the integer derivatives to a fractional form and (numerically) solve the Laplace integrals yielded by their inversion. This could potentially destroy the simple and elegant form of the solution and make it more sensitive to noise. A simpler approach would be to use the method as is on very short subsequent time steps, which was shown to be feasible, as long as the distribution function change is not vanishing. If the particle transport exhibits super- or subdiffusion, the fractional order can then be determined from the time dependence of the diffusion coefficient. Arguably, large simulation sizes would be needed to observe superdiffusion and long run times to observe subdiffusion, which would at some point confront us with the problem of limited computational resources. A way to at least delay the onset of that problem, that is worth looking into, would be to take a domain decomposition approach to the spectral MHD code, following Vay et al. (2013).

Another possibility to lower the computational cost is the use of Perlin noise turbulence instead of the self-consistently generated MHD turbulence. As the analysis in this work has shown, it can yield very plausible results at a small fraction of the computational cost, especially in the case of strong turbulence. An extension of the one-dimensional Perlin code to three dimensions is a possibility to be considered, since it would solve some of the scaling and structural problems of the spectrum (like the missing Cherenkov resonance). At the very least, the dimensionality of the model should be taken into account, when interpreting the results.

The new analysis methods were used in this work to derive pitch angle diffusion coefficients $D_{\mu\mu}$ and perpendicular diffusion coefficients $D_{\perp}$, since these parameters are deemed to be the most interesting ones and because the focus lay mainly on the implementation and validation of the methods. But as the derivation of the $D_{\perp}$ has shown, the underlying principle can be used to derive other transport parameters as well, although some modification might be needed.

Finally, while this work focused on turbulence and particle transport in the heliosphere, this is only one of the possible applications. The simulation setup and the analysis methods are, in principle, transferable to other astrophysical environments and phenomena, allowing systematic parameter studies for many different topics. Hence, this thesis can be used as a starting point for further projects.

Bibliography

- Abdo et al. Measurement of the Cosmic Ray $e^+ + e^-$ Spectrum from 20GeV to 1TeV with the Fermi Large Area Telescope. *Physical Review Letters*, 102(18):181101–+, May 2009. DOI: 10.1103/PhysRevLett.102.181101.
- Ackermann et al. Fermi LAT observations of cosmic-ray electrons from 7 GeV to 1 TeV. *Physical Review D*, 82(9):092004–+, Nov. 2010. DOI: 10.1103/PhysRevD.82.092004.
- Adriani et al. An anomalous positron abundance in cosmic rays with energies 1.5–100 GeV. *Nature*, 458:607–609, Apr. 2009. DOI: 10.1038/nature07942.
- Adriani et al. A statistical procedure for the identification of positrons in the PAMELA experiment. *Astroparticle Physics*, 34:1–11, Aug. 2010. DOI: 10.1016/j.astropartphys.2010.04.007.
- Agueda, N. and Vainio, R. On the parametrization of the energetic-particle pitch-angle diffusion coefficient. *Journal of Space Weather and Space Climate*, 3(27):A10, Mar. 2013. DOI: 10.1051/swsc/2013034.
- Agueda, N., Vainio, R., Lario, D., and Sanahuja, B. Injection and Interplanetary Transport of Near-Relativistic Electrons: Modeling the Impulsive Event on 2000 May 1. *ApJ*, 675:1601–1613, Mar. 2008. DOI: 10.1086/527527.
- Aguilar, M., Aisa, D., Alvino, A., Ambrosi, G., Andeen, K., Arruda, L., Attig, N., Azzarello, P., Bachlechner, A., Barao, F., and et al. Electron and Positron Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station. *Physical Review Letters*, 113(12):121102, Sept. 2014. DOI: 10.1103/PhysRevLett.113.121102.
- Allahverdi, R., Dutta, B., Richardson-McDaniel, K., and Santoso, Y. A Supersymmetric B-L Dark Matter Model and the Observed Anomalies in the Cosmic Rays. *ArXiv e-prints*, Dec. 2008.
- Axford, W. I., Leer, E., and Skadron, G. The Acceleration of Cosmic Rays by Shock Waves. In *International Cosmic Ray Conference*, volume 11 of *International Cosmic Ray Conference*, pages 132+, 1977.
- Barnes, A. Collisionless Damping of Hydromagnetic Waves. *Physics of Fluids*, 9:1483, Apr. 1966. DOI: 10.1063/1.1761882.
- Baumjohann, W. and Treumann, R. A. *Advanced space plasma physics*. Imperial Coll., London, 1997. URL <https://cds.cern.ch/record/431397>.

- Bell, A. R. The acceleration of cosmic rays in shock fronts. II. *Monthly Not. Royal Astron. Soc.*, 182:443–455, Feb. 1978.
- Bershadskii, A. Anomalous three-dimensional symmetries of solar-wind plasma. *Phys. Rev. E*, 66:046410, Oct 2002. DOI: 10.1103/PhysRevE.66.046410. URL <http://link.aps.org/doi/10.1103/PhysRevE.66.046410>.
- Biermann, L. Kometenschweife und solare Korpuskularstrahlung. *ZAp*, 29:274, 1951.
- Biermann, L. Solar corpuscular radiation and the interplanetary gas. *The Observatory*, 77:109–110, June 1957.
- Bigot, B., Galtier, S., and Politano, H. An anisotropic turbulent model for solar coronal heating. *A&A*, 490:325–337, Oct. 2008. DOI: 10.1051/0004-6361:20079227.
- Birdsall, C. K. and Langdon, A. B. *Plasma Physics via Computer Simulation*. 1991.
- Boris, J. P. Relativistic plasma simulation-optimization of a hybrid code. *Proceedings of the Fourth Conference on Numerical Simulation Plasmas*, pages 3–67, 1970.
- Borovsky, J. E. and Funsten, H. O. The role of solar-wind turbulence in the coupling of the solar wind to the Earth’s magnetosphere. *AGU Spring Meeting Abstracts*, page B1, May 2002.
- Borovsky, J. E. and Gary, S. P. The Viscosity of the Collisionless Solar Wind and the Reynolds Number of Its MHD Turbulence. *AGU Fall Meeting Abstracts*, page A2, Dec. 2008.
- Boyd, T. and Sanderson, J. *The Physics of Plasmas*. Cambridge University Press, 2003. ISBN 9780521459129. URL <https://books.google.de/books?id=bAmqvUGTJ4C>.
- Brueckner, G. E., Delaboudiniere, J.-P., Howard, R. A., Paswaters, S. E., St. Cyr, O. C., Schwenn, R., Lamy, P., Simnett, G. M., Thompson, B., and Wang, D. Geomagnetic storms caused by coronal mass ejections (CMEs): March 1996 through June 1997. *Geophys. Res. Lett.*, 25:3019–3022, 1998. DOI: 10.1029/98GL00704.
- Bruno, R. and Carbone, V. The Solar Wind as a Turbulence Laboratory. *Living Reviews in Solar Physics*, 2:4, Sept. 2005.
- Bruno, R., Bavassano, B., and Villante, U. Evidence for long period Alfvén waves in the inner solar system. *J. Geophys. Res.*, 90:4373–4377, May 1985. DOI: 10.1029/JA090iA05p04373.
- Büsching, I., de Jager, O. C., Potgieter, M. S., and Venter, C. A Cosmic-Ray Positron Anisotropy due to Two Middle-Aged, Nearby Pulsars? *The Astrophysical Journal*, 678:L39–L42, May 2008. DOI: 10.1086/588465.

- Bzowski, M., Mäkinen, T., Kyrölä, E., Summanen, T., and Quémerais, E. Latitudinal structure and north-south asymmetry of the solar wind from Lyman-alpha remote sensing by SWAN. *A&A*, 408:1165–1177, Sept. 2003. DOI: 10.1051/0004-6361:20031022.
- Canuto, C. *Spectral Methods in Fluid Dynamics*. Springer, Berlin Heidelberg New York, 1988. ISBN 3-540-52205-0.
- Chang et al. An excess of cosmic ray electrons at energies of 300-800GeV. *Nature*, 456: 362–365, Nov. 2008. DOI: 10.1038/nature07477.
- Cooley, J. W. and Tukey, J. W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.*, 19:297+, 1965.
- Courant, R., Friedrichs, K., and Lewy, H. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100:32–74, 1928.
- Elsasser, W. M. The Hydromagnetic Equations. *Physical Review*, 79:183–183, July 1950. DOI: 10.1103/PhysRev.79.183.
- Erdélyi, R. and Ballai, I. Heating of the solar and stellar coronae: a review. *Astronomische Nachrichten*, 328(8):726–733, 2007. ISSN 1521-3994. DOI: 10.1002/asna.200710803. URL <http://dx.doi.org/10.1002/asna.200710803>.
- Fermi, E. On the Origin of the Cosmic Radiation. *Physical Review*, 75:1169–1174, Apr. 1949. DOI: 10.1103/PhysRev.75.1169.
- Forbush, S. E. Three Unusual Cosmic-Ray Increases Possibly Due to Charged Particles from the Sun. *Physical Review*, 70:771–772, Nov. 1946. DOI: 10.1103/PhysRev.70.771.
- Frisch, U. *Turbulence*. Nov. 1995.
- Galtier, S., Nazarenko, S. V., Newell, A. C., and Pouquet, A. A weak turbulence theory for incompressible magnetohydrodynamics. *Journal of Plasma Physics*, 63:447–488, June 2000. DOI: 10.1017/S0022377899008284.
- Giacalone, J. and Jokipii, J. R. The Transport of Cosmic Rays across a Turbulent Magnetic Field. *ApJ*, 520:204–214, July 1999. DOI: 10.1086/307452.
- Goldreich, P. and Sridhar, S. Toward a theory of interstellar turbulence. 2: Strong alfvenic turbulence. *ApJ*, 438:763–775, Jan. 1995. DOI: 10.1086/175121.
- Goldreich, P. and Sridhar, S. Magnetohydrodynamic Turbulence Revisited. *ApJ*, 485: 680, Aug. 1997. DOI: 10.1086/304442.
- Gottlieb, D. and Tadmor, E. The CFL condition for spectral approximations to hyperbolic initial-boundary value problems. *Mathematics of Computation*, vol. 56, no. 194, pp. 565–588, 56:565–588, Apr. 1991. DOI: 10.1090/S0025-5718-1991-1066833-9.

- Green, M. S. Brownian Motion in a Gas of Noninteracting Molecules. *J. Chem. Phys.*, 19:1036–1046, Aug. 1951. DOI: 10.1063/1.1748449.
- Grimani, C. Experimental clues on pulsar energy losses and the role of circumpulsar supernova fallback disks. *Classical and Quantum Gravity*, 26(23):235009–+, Dec. 2009. DOI: 10.1088/0264-9381/26/23/235009.
- Iroshnikov, P. S. Turbulence of a Conducting Fluid in a Strong Magnetic Field. *Soviet Ast.*, 7:566, Feb. 1964.
- Ivascenko, A., Lange, S., Spanier, F., and Vainio, R. Determining pitch-angle diffusion coefficients from test particle simulation. *The Astrophysical Journal Supplement Series*, 2016. accepted.
- Jokipii, J. R. Cosmic-Ray Propagation. I. Charged Particles in a Random Magnetic Field. *ApJ*, 146:480, Nov. 1966. DOI: 10.1086/148912.
- Jokipii, J. R. and Thomas, B. Effects of drift on the transport of cosmic rays. IV - Modulation by a wavy interplanetary current sheet. *ApJ*, 243:1115–1122, Feb. 1981. DOI: 10.1086/158675.
- Kahler, S. W., Sheeley, Jr., N. R., and Liggett, M. Coronal mass ejections and associated X-ray flare durations. *ApJ*, 344:1026–1033, Sept. 1989. DOI: 10.1086/167869.
- Kilian, P., Ganse, U., Ivascenko, A., Spanier, F., and Vainio, R. Using Perlin Noise to study charged particle diffusion in magnetic turbulence. *in prep*, 2016.
- Kolmogorov, A. The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds’ Numbers. *Akademiia Nauk SSSR Doklady*, 30:299–303, 1941a.
- Kolmogorov, A. The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds’ Numbers. *Akademiia Nauk SSSR Doklady*, 30:301–305, 1941b.
- Koskinen, H. E. J. *Physics of Space Storms*. Springer-Verlag Berlin Heidelberg, 2011.
- Kraichnan, R. H. Inertial-range spectrum of hydromagnetic turbulence. *Phys. Fluids*, 8:1385–1387, 1965.
- Krymskii, G. F. A regular mechanism for accelerating charged particles at the shock front. *Akademiia Nauk SSSR Doklady*, 234:1306–1308, June 1977.
- Kubo, R. Statistical-Mechanical Theory of Irreversible Processes. I. *Journal of the Physical Society of Japan*, 12:570–+, June 1957.
- Lang, K. R. *The Sun from space*. Springer-Verlag Berlin Heidelberg, 2009.
- Lange, S. *Dissertation: Turbulenz und Teilchentransport in der Heliosphäre*. Institut für theoretische Physik und Astrophysik, Julius-Maximilians-Universität Würzburg, 2012.

- Lange, S. and Spanier, F. Evolution of plasma turbulence excited with particle beams. *A&A*, 546:A51, Oct. 2012. DOI: 10.1051/0004-6361/201219579.
- Levy, Doron, Puppo, Gabriella, and Russo, Giovanni. Central weno schemes for hyperbolic systems of conservation laws. *ESAIM: M2AN*, 33(3):547–571, 1999. DOI: 10.1051/m2an:1999152. URL <http://dx.doi.org/10.1051/m2an:1999152>.
- Lithwick, Y. and Goldreich, P. Imbalanced Weak Magnetohydrodynamic Turbulence. *ApJ*, 582:1220–1240, Jan. 2003. DOI: 10.1086/344676.
- Maron, J. and Goldreich, P. Simulations of Incompressible Magnetohydrodynamic Turbulence. *ApJ*, 554:1175–1196, June 2001. DOI: 10.1086/321413.
- Mathieu, J. and Scott, J. *An Introduction to Turbulent Flow*. June 2000.
- McIntosh, S. W., de Pontieu, B., Carlsson, M., Hansteen, V., Boerner, P., and Goossens, M. Alfvénic waves with sufficient energy to power the quiet solar corona and fast solar wind. *Nature*, 475:477–480, July 2011. DOI: 10.1038/nature10235.
- Melrose, D. B. *Instabilities in space and laboratory plasmas*. Cambridge University Press, 1989.
- Montgomery, D. Modifications of magnetohydrodynamics as applied to the solar wind. *J. Geophys. Res.*, 97:4309, Apr. 1992. DOI: 10.1029/92JA00144.
- Munro, R. H., Gosling, J. T., Hildner, E., MacQueen, R. M., Poland, A. I., and Ross, C. L. The association of coronal mass ejection transients with other forms of solar activity. *Sol. Phys.*, 61:201–215, Feb. 1979. DOI: 10.1007/BF00155456.
- Orszag, S. A. Elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *Atmospheric Sciences*, 28:1074+, 1971.
- Parker, E. N. Dynamics of the Interplanetary Gas and Magnetic Fields. *ApJ*, 128:664, Nov. 1958. DOI: 10.1086/146579.
- Perlin, K. An image synthesizer. *SIGGRAPH Comput. Graph.*, 19(3):287–296, July 1985. ISSN 0097-8930. DOI: 10.1145/325165.325247. URL <http://doi.acm.org/10.1145/325165.325247>.
- Perlin, K. noise hardware. In Olano, M., editor, *Real-Time Shading Languages*, number 36. siggraph course notes, 2001.
- Perlin, K. Improving noise. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '02, pages 681–682, New York, NY, USA, 2002. ACM. ISBN 1-58113-521-1. DOI: 10.1145/566570.566636. URL <http://doi.acm.org/10.1145/566570.566636>.
- Petschek, H. E. Magnetic Field Annihilation. *NASA Special Publication*, 50:425, 1964.

- Politano, H. and Pouquet, A. von Kármán-Howarth equation for magnetohydrodynamics and its consequences on third-order longitudinal structure and correlation functions. *Phys. Rev. E*, 57:21, Jan. 1998. DOI: 10.1103/PhysRevE.57.R21.
- Reames, D. V. Particle acceleration at the Sun and in the heliosphere. *Space Sci. Rev.*, 90:413–491, Oct. 1999. DOI: 10.1023/A:1005105831781.
- Reames, D. V. and Ng, C. K. Streaming-limited Intensities of Solar Energetic Particles. *ApJ*, 504:1002, Sept. 1998. DOI: 10.1086/306124.
- Reames, D. V., Kahler, S. W., and Ng, C. K. Spatial and Temporal Invariance in the Spectra of Energetic Particles in Gradual Solar Events. *ApJ*, 491:414, Dec. 1997. DOI: 10.1086/304939.
- Savitzky, A. and Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36:1627–1639, 1964.
- Schlickeiser, R. Cosmic-ray transport and acceleration. *ApJS*, 90:929–936, Feb. 1994. DOI: 10.1086/191927.
- Schlickeiser, R. *Cosmic Ray Astrophysics*. 2002. DOI: 10.1007/978-3-662-04814-6.
- Schreiner, C., Kilian, P., and Spanier, F. Particle Scattering off of Right-handed Dispersive Waves. *ApJ*, 834:161, Jan. 2017. DOI: 10.3847/1538-4357/834/2/161.
- Shalchi, A., editor. *Nonlinear Cosmic Ray Diffusion Theories*, volume 362 of *Astrophysics and Space Science Library*, 2009. DOI: 10.1007/978-3-642-00309-7.
- Shebalin, J. V., Matthaeus, W. H., and Montgomery, D. Anisotropy in MHD turbulence due to a mean magnetic field. *Journal of Plasma Physics*, 29:525–547, June 1983. DOI: 10.1017/S0022377800000933.
- Shibata, K., Masuda, S., Shimojo, M., Hara, H., Yokoyama, T., Tsuneta, S., Kosugi, T., and Ogawara, Y. Hot-Plasma Ejections Associated with Compact-Loop Solar Flares. *ApJ*, 451:L83, Oct. 1995. DOI: 10.1086/309688.
- Shiota, D., Isobe, H., Chen, P. F., Yamamoto, T. T., Sakajiri, T., and Shibata, K. Self-Consistent Magnetohydrodynamic Modeling of a Coronal Mass Ejection, Coronal Dimming, and a Giant Cusp-shaped Arcade Formation. *ApJ*, 634:663–678, Nov. 2005. DOI: 10.1086/496943.
- Skilling, J. Cosmic ray streaming. III - Self-consistent solutions. *MNRAS*, 173:255–269, Nov. 1975.
- Sonnerup, B. U. Ö. Acceleration of particles reflected at a shock front. *J. Geophys. Res.*, 74:1301, 1969. DOI: 10.1029/JA074i005p01301.
- Spanier, F. *Dissertation: Turbulenz im ISM: Einschränkung der Turbulenzparameter durch Berücksichtigung von Plasmawellenheizung*. Institut für theoretische Physik IV Weltraum- und Astrophysik, Ruhr-Universität Bochum, 2005.

- Sridhar, S. and Goldreich, P. Toward a theory of interstellar turbulence. 1: Weak Alfvénic turbulence. *ApJ*, 432:612–621, Sept. 1994. DOI: 10.1086/174600.
- Taylor, G. I. Diffusion by continuous movements. *Proceedings of the London Mathematical Society*, s2-20(1):196–212, 1922. DOI: 10.1112/plms/s2-20.1.196. URL <http://plms.oxfordjournals.org/content/s2-20/1/196.short>.
- Tousey, R. The solar corona. In Rycroft, M. J. and Runcorn, S. K., editors, *Space Research Conference*, volume 2 of *Space Research Conference*, pages 713–730, 1973.
- Tu, C.-Y. and Marsch, E. MHD structures, waves and turbulence in the solar wind: Observations and theories. *Space Sci. Rev.*, 73:1–210, July 1995. DOI: 10.1007/BF00748891.
- Tu, C.-Y., Roberts, D. A., and Goldstein, M. L. Spectral evolution and cascade constant of solar wind Alfvénic turbulence. *J. Geophys. Res.*, 94:13575–13578, Oct. 1989. DOI: 10.1029/JA094iA10p13575.
- Vainio, R., Laitinen, T., and Fichtner, H. A simple analytical expression for the power spectrum of cascading Alfvén waves in the solar wind. *A&A*, 407:713–723, Aug. 2003. DOI: 10.1051/0004-6361:20030914.
- Vay, J. Simulation of beams or plasmas crossing at relativistic velocity. *Physics of Plasmas*, 15(5):056701–+, May 2008. DOI: 10.1063/1.2837054.
- Vay, J.-L., Haber, I., and Godfrey, B. B. A domain decomposition method for pseudo-spectral electromagnetic simulations of plasmas. *Journal of Computational Physics*, 243:260–268, June 2013. DOI: 10.1016/j.jcp.2013.03.010.
- Weiss, M. *Dissertation: Numerische Untersuchung von Turbulenz und Teilchentransport in der Heliosphäre*. Institut für theoretische Physik und Astrophysik, Julius-Maximilians-Universität Würzburg, 2011.
- Zakharov, V. E. and Filonenko, N. N. The energy spectrum for stochastic oscillations of a fluid surface. *Akademiia Nauk SSSR Doklady*, 170:1292–1295, 1966.
- Zakharov, V. E., L’vov, V., and Falkovich, G. E. *Kolmogorov Spectra of Turbulence I: Wave Turbulence*. Springer-Verlag, Berlin, 1992.
- Zrake, J. and MacFadyen, A. I. Numerical Simulations of Driven Relativistic Magnetohydrodynamic Turbulence. *ApJ*, 744:32, Jan. 2012. DOI: 10.1088/0004-637X/744/1/32.

List of Publications

Peer-reviewed Publications

- **Ivascenko, A.** and Spanier, F.: Semi-analytical model of cosmic ray electron transport. *Astrophysics and Space Sciences Transactions*, 7 (2011), 265–269, DOI: 10.5194/astra-7-265-2011.
- **Ivascenko, A.**, Lange, S., Spanier, F. and Vainio, R.: Determining pitch-angle diffusion coefficients from test particle simulations. *The Astrophysical Journal Supplement Series*, (2016, accepted), Preprint: <https://arxiv.org/abs/1610.03240>.

Other Publications

- Spanier, F., **Ivascenko, A.**, Lange, S. and Schreiner, C. : Charged Particle Transport in Turbulent Media. *8th International Conference of Numerical Modeling of Space Plasma Flows (ASTRONUM 2013)*, 2014, Vol 288, p. 35, ADSURL: <http://adsabs.harvard.edu/abs/2014ASPC..488...35S>
- **Ivascenko, A.**, Spanier, F. and Ganse, U.: New Method for Determination of Diffusion Coefficients in Turbulent Plasmas. In *Proceedings of the ICRC 2015* (2015), PoS(ICRC2015)208.
- additionally Co-Author on 23 H.E.S.S. collaboration papers

Acknowledgements

If this feels a little rushed, it's because it is...

First of all I have to thank my long-time supervisor Prof. Felix Spanier, who has pulled me through from the very beginning of my diploma thesis to the finalisation of this work with his boundless optimism, long, sometimes recurring, but nevertheless always helpful discussions and support also in personal matters, which is not at all par for the course in a supervisor. I would also like to thank him, Prof. Markus Böttcher and the rest of the Centre for Space Research for the unique opportunity of conducting my PhD in South Africa. It was certainly quite an experience that I will not forget quickly!

My deepest thanks go to Petro Sieberhagen. I am not exaggerating when I say that without her personal and professional dedication and support this PhD project would not have been possible. Her ability to stay on top of all the paperwork and to overcome all the bureaucratic hurdles is amazing and her skills on the phone will forever be a source of admiration to me.

Next, I would like to thank all of my colleagues from Würzburg, Potchefstroom and Bochum. Special thanks go to Patrick Kilian, whom I have known since the time we both begun our studies at Würzburg University. Thank you for all the interesting discussions, science related or not, for the help and support and, of course, for proof-reading this work.

And of course I have to return the mention of my “companion in misfortune” and fellow PhD student Cedric Schreiner and thank him for all the help, discussions and companionship on road trips and conferences.

I will not mention every other colleague by name, but you know who you are, and know that I am very grateful for everything you've done for me.

Last but not least, I would like to thank Prof. Reinhard Schlickeiser for sheltering me at the TPIV at Bochum University in the last stage of my thesis.