

Master's thesis

# Time-dependent Modeling of Diffusive Shock Acceleration Using Stochastic Differential Equations

Patrick Günther



**ASTRO WÜRZBURG**

LEHRSTUHL FÜR ASTRONOMIE  
FAKULTÄT FÜR PHYSIK UND ASTRONOMIE  
JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG



ASTRO WÜRZBURG

LEHRSTUHL FÜR ASTRONOMIE  
FAKULTÄT FÜR PHYSIK UND ASTRONOMIE  
JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG

Master's thesis

# Time-dependent Modeling of Diffusive Shock Acceleration Using Stochastic Differential Equations

Patrick Günther

May 2022



Supervisors: Prof. Dr. Karl Mannheim and Sarah Wagner

# Abstract

While it is expected that in the center of every galaxy there is a black hole, in some galaxies the region around the black hole shows significant activity. Those objects are called active galactic nuclei. Some active galactic nuclei form a highly energetic outflow of matter, so called *jets*, that are tightly collimated and can reach far beyond the host galaxy. There are still few details of the jet physics known with certainty. Jets are assumed to consist of a highly relativistic plasma and are powered by the accretion of matter onto the black hole or even the black hole itself. The radiation spectrum of jets is strongly non-thermal and energetic emission is observed which can only be explained by the presence of ultrarelativistic particles with suprathreshold energies. The acceleration site of those is expected to be far from the black hole and a variety of acceleration mechanisms have been proposed. It is an open question which plasma conditions and radiation mechanisms dominate and extensive modeling is done and simulations are conducted to better understand the nature of these complex objects.

In this work, details of the jet structure and composition are reviewed before discussing the framework of equations that is used to describe the therein contained plasma. In the test-particle limit these results in the cosmic-ray transport equation describing the propagation of charged particles. Several mechanisms capable of accelerating particles in the jet to high energies are considered, focusing on diffusive shock acceleration (also known as first-order Fermi acceleration). A shock is a steep gradient in the plasma bulk velocity which can cause particles to accelerate when crossing the shock several times due to diffusion. A set of particles undergoing diffusive shock acceleration will eventually show a power-law spectrum in energy. A framework for the numerical solution of the cosmic-ray transport equations is discussed, based on a Monte-Carlo method using stochastic differential equations. This method is tested on diffusive shock acceleration. It is shown that the algorithm is in principle capable of replicating the behavior expected from analytic solutions, while care must be taken for the choice of integration step sizes. Acceleration is found to lose efficiency and show steeper power laws if the advection and diffusion is not thoroughly balanced and if the numerically required shock smoothing is too strong or weak. Furthermore, the choice of the integration scheme is shown to impact the numerical accuracy significantly. The basic Cauchy-Euler scheme fails to accurately predict the solution and a semi-implicit second-order scheme is shown to reproduce expected results significantly better. Emission mechanisms describing the electromagnetic radiation produced by the charged particles in the jet plasma are discussed. Corresponding losses are incorporated in the transport equation and the resulting particle distribution is used to predict spectral energy densities. It is demonstrated that the synthetic spectrum can be compared to observations.

# Zusammenfassung

Man geht davon aus, dass sich im Zentrum jeder Galaxie ein schwarzes Loch befindet. In einem Teil der beobachteten Galaxien zeigt die Region um das zentrale schwarze Loch starke Aktivität. Diese Regionen nennt man aktive Galaxienkerne. Manche aktiven Galaxienkerne emittieren einen hochenergetischen Materiestrom, der als *Jet* bezeichnet wird und weit über die Galaxie seines Ursprungs hinaus kollimiert ist. Obwohl bereits 1918 zum ersten Mal ein Jet beobachtet wurde, sind immer noch wenige sichere Details über seine Physik bekannt. Es wird angenommen, dass Jets aus einem hochrelativistischen Plasma bestehen und die Akkretion von Materie des schwarzen Loches oder das schwarze Loch selber die mit dem Plasma transportierte Energie bereitstellt. Das elektromagnetische Spektrum der Jets ist kein thermisches Spektrum und Emission wird auch weit entfernt vom schwarzen Loch beobachtet. Diese Phänomene lassen sich nur durch ultrarelativistische Teilchen und deren Beschleunigung in der Nähe der beobachteten Emissionsregionen erklären. Welche der möglichen Mechanismen für Teilchenbeschleunigung und Emission von Strahlung tatsächlich ablaufen ist eine offene Frage, die durch Modellierung und entsprechende Simulationen geklärt werden soll.

In dieser Arbeit werden die Details des Jet-Aufbaus und der Jet-Zusammensetzungen gezeigt und die grundlegenden Gleichungen zur Beschreibung des Plasmas diskutiert. Weiterhin werden verschiedene Mechanismen der Beschleunigung geladener Teilchen veranschaulicht, vornehmlich die Fermi-Beschleunigung erster Ordnung. Die Fermi-Beschleunigung erster Ordnung tritt an Schock-Fronten auf, welche steile Gradienten in Parametern des Plasmas darstellen. Wenn Teilchen effizient an Schocks beschleunigt werden, konvergiert ihr Impulsspektrum zu einem Potenzgesetz. Ein Modell basierend auf einer Monte-Carlo-Methode mit stochastischen Differentialgleichungen zur numerischen Lösung der Transportgleichung kosmischer Strahlen wird diskutiert und am Beispiel der Fermi-Beschleunigung erster Ordnung getestet. Es wird gezeigt, dass der Algorithmus das zu erwartende, physikalische Verhalten aufweist, wenn eine sinnvolle Wahl der Größe der Integrationschritte getroffen wird. Die Beschleunigung ist weniger effizient und zeigt daher steilere Potenzgesetze wenn Advektion und Diffusion nicht in einem sinnvollen Verhältnis zueinander stehen oder wenn das von der Numerik benötigte Glätten des Schocks unpassend ist. Des Weiteren ist die Wahl des Integrationsverfahrens von zentraler Relevanz, wobei das grundlegende Cauchy-Euler-Verfahren nicht in der Lage ist, die richtigen Ergebnisse exakt vorherzusagen. Im Gegenzug wird ein semi-implizites Verfahren zweiter Ordnung gezeigt, das deutlich exaktere Vorhersagen trifft. Darüber hinaus werden Emissionsmechanismen der beschleunigten, geladenen Teilchen diskutiert. Die daraus resultierenden Verluste können zur Transportgleichung hinzugefügt werden und aus den Impulsspektren der numerischen Lösung können spektrale Energiedichten berechnet werden. Diese können mit Beobachtungen verglichen werden.

# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Active galactic nuclei</b>	<b>3</b>
2.1. Observational classification	3
2.2. Jets	6
2.2.1. Launching and power source	6
2.2.2. Structure and morphology	7
2.2.3. Spectral footprint	8
2.2.4. Particle acceleration	10
<b>3. Theoretical description of the jet plasma</b>	<b>12</b>
3.1. Kinetic plasma physics	13
3.1.1. Vlasov equations	13
3.1.2. Quasilinear approximation	14
3.1.3. Diffusion approximation	16
3.1.4. Plasma waves	18
3.2. Particle acceleration mechanisms	19
3.2.1. Acceleration mechanisms at shocks	19
3.2.2. Diffusive shock acceleration	19
3.2.3. Second-order Fermi acceleration	22
<b>4. Stochastic differential equations</b>	<b>24</b>
4.1. Random variables	24
4.2. Stochastic processes	25
4.3. Markov chains	26
4.4. Diffusion processes	27
4.5. Itô calculus and stochastic differential equations	28
4.6. Sampling the stochastic process	29
<b>5. Numerical solution of the transport equation</b>	<b>33</b>
5.1. Explicit Euler scheme	34
5.1.1. Numerical steps	36
5.1.2. Shock width	41
5.1.3. Asymptotic behavior	41
5.1.4. Momentum diffusion	41
5.2. Higher-order and implicit numerical schemes	43
5.3. Technologies used for numerical integration and code structure	47
<b>6. Radiative transfer</b>	<b>50</b>
6.1. Elements of radiation physics	50
6.2. Synchrotron radiation	52
6.3. Synchrotron self-Compton radiation	54
6.4. Computation of synthetic SEDs	55
<b>7. Conclusion</b>	<b>56</b>

<b>A. Pitch angle- and gyrophase-averaged distribution function</b>	<b>57</b>
<b>Bibliography</b>	<b>59</b>
<b>Acknowledgements</b>	<b>67</b>
<b>Declaration</b>	<b>68</b>

# 1. Introduction

Active galactic nuclei are on many scales amongst the most extreme objects that can be found in our universe. Located in the center of galaxies (hence their name) they emit vast amounts of energy in form of electromagnetic radiation and are the most luminous, persistent objects observed. In the center of active galactic nuclei themselves there is a black hole of very high mass which, while naturally not radiating itself, is the ultimate source of energy of an active galactic nucleus through its gravitational potential. Typical masses of the central black hole reach from  $1 \cdot 10^6$  to  $1 \cdot 10^{10}$  solar masses which is approximately  $2 \cdot 10^{36}$  to  $2 \cdot 10^{40}$  kg and the active galactic nucleus region typically emits  $1 \cdot 10^{43}$  to  $1 \cdot 10^{47}$  erg s<sup>-1</sup> (Woo and Urry, 2002). Because of these properties and the resulting unique physical conditions in the environment of active galactic nuclei they offer unmatched opportunities for research on scales not achieved anywhere else. In contrast to the possibilities, research on active galactic nuclei requires huge effort in building telescopes and detector systems since their large distance from Earth makes them hard to observe despite their high radiative power.

Many active galactic nuclei are emitting a so called *jet* which is a large-scale, highly collimated matter outflow which can extend far beyond the host galaxy. The first observation of a jet is attributed to Curtis (1918) who observed a “curious straight ray”, as he described what we now call the jet of the active galactic nucleus M87. The radiation detected from jets ranges over the whole electromagnetic spectrum and shows complex non-thermal behavior which hints at a complex set of prevalent physical conditions. This is different from what astronomers expected for an accumulation of usual stars, which apart from absorption lines show a thermal spectrum, and left them both puzzled and intrigued for over a century up until today. Jets are one of the main sources of radiation observed in radio galaxies, quasars and blazars and emit radiation from the radio band extending through the optical and X-ray band up to highly energetic  $\gamma$ -rays. An open question to this day is the composition of the matter in the jets and the conditions therein that lead to the emission across the whole electromagnetic spectrum. It is clear from observations that non-thermal emission is dominating since smooth power-law spectra extending over many orders of magnitude are found. Remarkably, the flux spectral index is observed universally at around  $-0.5$  to  $0.5$  in the radio band in many different sources (Abdo et al., 2010) so a common mechanism is probable. The radio emission is due to synchrotron radiation which results from relativistic charged particles orbiting under the influence of a magnetic field. For the observed synchrotron spectrum it is necessary that the emitting particles have ultrarelativistic energies. Processes that are expected to be involved in the acceleration of those are the first- and second-order Fermi processes, also called diffusive shock acceleration and momentum diffusion. Diffusive shock acceleration occurs at shock fronts in the plasma, which can be described as infinitely thin gradients in pressure and bulk velocity. If particles with momentum much higher than the background plasma cross a shock front multiple times due to diffusion they gain energy in each crossing. In contrast, momentum diffusion happens when particles randomly scatter at plasma waves which also leads to acceleration since net gains in energy are more likely than losses. Both mechanisms are explained in more detail in section 3.2.

Since *in situ* measurements of jet conditions are impossible astronomers must rely on models predicting the radiative output and the comparison of those with actual observations. Equations governing the plasma conditions can analytically only be solved under very simplified conditions so that numerical methods were used to predict observations, since modern computers made it possible. The large-scale structure of a magnetized plasma can be described by the equations of magnetohydrodynamics (MHD) which treat the plasma as a fluid subjected to Maxwell’s equations.

They are commonly simulated using MHD codes such as PLUTO (Mignone, Bodo, Massaglia, et al., 2007) which are able to resolve and reproduce the observed large-scale structure of the jet and its launching from disk winds (Porth and Fendt, 2010; Pudritz et al., 2007). In addition to that it is necessary to find small-scale models which can describe the acceleration of particles to ultrarelativistic energies as required by the observations. This regime is governed by a Fokker-Planck transport equation when the system is described in test-particle approximation, *i.e.* the acceleration of particles has no substantial effect on the large-scale structure. The coupling of both scales may not be negligibly small since some amount of energy is extracted from the plasma by the accelerated particles. The large-scale MHD and small-scale cosmic-ray transport can be simultaneously solved in both the relativistic regime using particle-in-cell simulations (Mignone, Bodo, Vaidya, et al., 2018) or Lagrangian tracer particles (Vaidya et al., 2018) and in the nonrelativistic regime (Winner et al., 2019). An alternative method to solving the small-scale propagation and acceleration of cosmic rays is the use of stochastic differential equations, first proposed for cosmic-ray transport by Achterberg and Krülls (1992). For an equation of Fokker-Planck type an equivalent stochastic differential equation can be found which can be solved efficiently using a Monte-Carlo method. Its advantages are the absence of the need to impose a discrete grid on the phase space and its very straightforward extension to multi-threaded processors. In contrast to finite-difference methods for solving differential equations, using stochastic differential equations also stays numerically stable when extending to more dimensions (Kopp et al., 2012). It has been successfully applied in simulating the transport and acceleration of cosmic rays in AGN jets (Achterberg, Gallant, et al., 2001; Krülls and Achterberg, 1994; Zhang, 2000) and the transport of energetic particles in the heliosphere (Dröge et al., 2010) but convergence issues caused by the infinitely thin shock front have also been pointed out (Achterberg and Schure, 2011). In this work the method of solving cosmic-ray transport equations using stochastic differential equations is studied in the environment of jets from active galactic nuclei. This is done to estimate if this method is a useful tool for efficient calculation of synthetic observational signatures from jets in the context of large-scale simulations. These can possibly include the production of synthetic spectral energy densities or synthetic lightcurves of blazars.

## 2. Active galactic nuclei

The structural complexity of active galactic nuclei (AGN) combined with the sparsity of observational detail made astronomers believe for a long time they found many different types of objects, not realizing they were looking at only a single phenomenon. This led to the rise of a lot of different terms for classes of observations, for example radio galaxy, Seyfert galaxy and quasar. Eventually, the *unified model* was introduced, which explains a variety of observations as single type of source seen from different angles, and the term “AGN” in its modern sense was proposed. After decades of use, the observational classes were well established and are still used even after the general consensus about the unified model of AGN. The commonly used picture of an AGN is shown in Figure 2.1 which consists of a supermassive black hole in the center, surrounded by an accretion disk and a dust torus both in a common plane. Despite being small in comparison to the whole structure, the black hole is the central motor of the AGN providing the energy driving the rest of the AGN through its gravitational potential. Matter from surrounding stars and the interstellar medium spirals into the black hole at an accretion rate  $\dot{M}$ , forming an accretion disk, and heats up due to dissipation which in turn leads to radiation. To explain the luminosities observed in a bright quasar approximately  $100 M_{\odot}$  have to be accreted per year (Courvoisier, 2013). The best resolved insights of the inner region of an AGN are from the sources M87\* and SgrA\*, observed by the Event Horizon Telescope (Event Horizon Telescope Collaboration et al., 2019, 2022). The mass of the black hole in the center of an AGN can be estimated using the Eddington luminosity which for a black hole with mass  $M$  is the luminosity that is required so that the radiation pressure overweighs the gravitational force. In the case of spherical-symmetric accretion it is given by

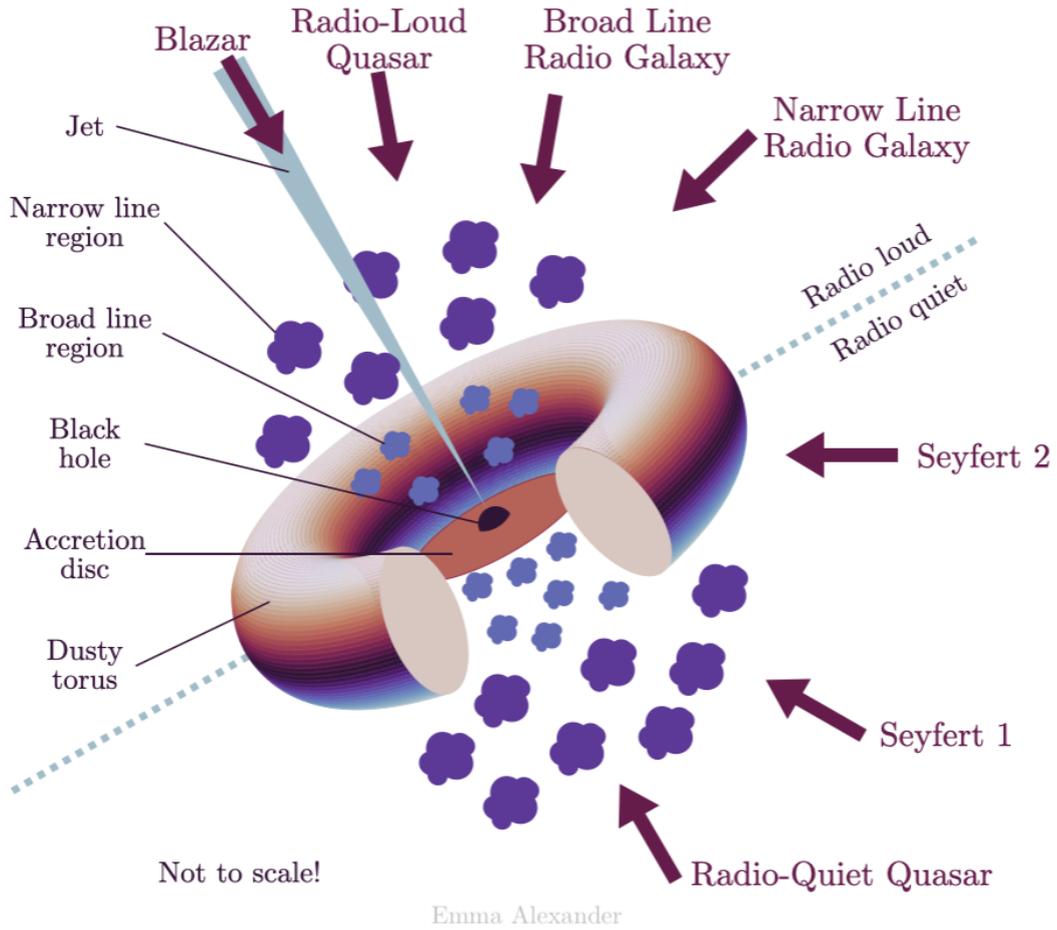
$$L_{\text{Edd}} = \frac{4\pi GMm_p c}{\sigma_T} \approx \frac{M}{M_{\odot}} \cdot 1.3 \cdot 10^{38} \text{ erg s}^{-1} . \quad (2.1)$$

Measuring the accretion flow luminosity imposes a lower limit on the black hole mass since if the mass would be any lower the accretion disk would be blown away from the black hole and could not be observed. Therefore if the black hole mass is higher, the accretion rate and in turn the accretion flow luminosity will decrease again, reaching equilibrium.

The presence of the black hole and the properties of the magnetized accretion disk can lead to the launching of jets approximately perpendicular to the accretion disk. The jets are large-scale structures extending far from the galaxy itself, forming *hotspots* of high emission and terminating in *lobes*. They consist of an outflow of plasma which is propagating at highly relativistic velocities as can be inferred from subsequent radio observations (e.g. Lister et al., 2009). They are the central objects under scrutiny in this work. The core region is surrounded by clouds of ionized gas forming the broad- and narrow-line regions, further described below, and a dust torus. The latter is a region which is assumed to lie around the core and accretion disk hindering the radiation from passing through because of its optical thickness. In the following an overview of the properties of AGN and especially their jets is given.

### 2.1. Observational classification

Mainly depending on the angle between the line of sight from Earth and the orientation of the accretion disc, surrounding ionized clouds and the jet in which the AGN system is seen from Earth, but also on the presence or absence of a jet, the variety of seemingly unconnected sources can be



**Figure 2.1.:** The standard model of active galactic nuclei. Different phenomena observed are unified into a single typical anisotropic structure which is randomly seen from different angles from the perspective of Earth (Urry and Padovani, 1995). Note that jets can be launched from both sides and not at all, too. Illustration by Emma Alexander, licensed under the Creative Commons licence (CC-BY). [https://emmaalexander.github.io/images/unified\\_agn.png](https://emmaalexander.github.io/images/unified_agn.png)

explained. The jets, if existing in a particular source, are most likely bipolar, *i.e.* one jet is emitted on both sides of the AGN.

### Seyfert galaxies

Seyfert galaxies, named after Carl Seyfert, first caught the attention of astronomers due to their unusual broad emission lines observed in the optical range. They are further split into two categories, namely Seyfert I galaxies and Seyfert II galaxies, depending on whether broad emission lines are present. The AGN contains clouds of ionized gas which are excited by the emission from the accretion disk and therefore radiate emission lines themselves. Emission lines are Doppler-broadened because the orbital velocity of the gas increases the closer it is to the black hole. Therefore clouds close to the core are in the so-called *broad line region* and clouds further from the core are in the *narrow line regions*. Depending on the viewing angle different line widths are observed. Additionally the density of the galactic medium decreases as  $\rho \propto r^{-\delta}$ ,  $\delta = 1 - 2$  with the distance from the core (Bicknell, Dopita and O’Dea, 1997). This allows for the emission of *forbidden lines* further away from the core which are emission lines only observed in very low densities and therefore not on Earth. In Seyfert galaxies not only the active galactic nuclei can be seen but also the surrounding galaxy which indicates that they are both closer to Earth and less luminous than quasars, which outshine their host galaxy. They are also characterized by typically not showing any jet structures. The emission-line features observed in Seyfert galaxies are found across most types of AGNs, the prominent exception being BL Lac objects.

### Radio galaxies

Another type of AGN are radio galaxies whose prominent feature is the intense emission in the radio band, orders of magnitude higher than the radio emission of our own galaxy. Alfvén and Herlofson (1950) proposed first that this emission is due to synchrotron radiation of charged particles because of the non-thermal power-law spectrum and the detected polarisation. First only observed as point sources and misnamed “radio stars” high-resolution radio images resolved the presence of jets in those galaxies which are the distinguishing factor for most radio galaxies. Similar to Seyfert galaxies, radio galaxies can also show narrow and broad emission lines depending on where gas clouds are on the line of sight from Earth or obscured by the torus. One distinguishes between radio-loud and radio-quiet AGN. The main difference is that in radio-loud AGN a jet emits a substantial part of the total emitted radiation while in the radio-quiet case no such jet is observed. They also differ in the type of galaxy they live in: radio-loud and radio-quiet AGNs seem to exist preferentially in elliptical galaxies and spiral galaxies, respectively (Wilson and Colbert, 1995).

### Fanaroff-Riley classification

Fanaroff and Riley (1974) categorized radio sources in terms of the position of their brightest spots. Sources with brighter spots closer to the center of the source are called FR class I objects and sources with brighter spots further out are called FR class II objects after the initials of the authors. They observed that FR-I sources typically have lower luminosities than their FR-II counterparts. In the high-luminosity regions the spectra are generally flatter than in low-luminosity regions which indicates that the age of radiating particles increases as luminosity decreases.

### Quasars

When the line of sight is approximately parallel to the jet axis (usually  $< 10^\circ$ ) the source in question is a quasar. The major contribution to the emission from these sources is considered to originate from the relativistic jet, the angle between the jet and the line of sight being small (Blandford and Königl, 1979). In this case the extent of the source in the plane of the sky is relatively small which is why they were named *quasi-stellar radio source*, later abbreviated to

quasar. It is also common to talk about radio-loudness or quietness of quasars, similar to radio-loud and radio-quiet galaxies. In the case of radio-quiet quasars the source of the radiation is a more intriguing question since no jets are observed. Liao et al. (2022) correlate radio emission intensity of these sources with optical/UV variability, which is an indicator for disk activity, and find that the radio intensity is not correlated with the disk activity, which may point to a weak jet that is not directly observed. Radio-loud quasars also show high variability in their intensity, sometimes over orders of magnitude, in between days in the case of  $\gamma$ -rays and weeks in the case of radio observations.

The most compact quasars typically emit a nearly flat radio spectrum which is thought to originate from the superposition of many emission components. These sources are called flat-spectrum radio quasars (FSRQ) and typically show superluminal motion. The term BL Lac refers to a source which shows only weak absorption and emission features as detected in the optical (Urry and Padovani, 1995). They are named after BL Lacertae which was the first object with those characteristics that was observed. Similar to FSRQ they show strong variability in their luminosity. Both FSRQ and BL Lac are commonly known as blazars. The blazar sequence (Prandini and Ghisellini, 2022) links the multitude of different observations of blazars to physical conditions. It shows that the spectral properties of blazars are linked to the mass of the central black hole and the accretion rate, however, the correlation is rather weak.

## 2.2. Jets

### 2.2.1. Launching and power source

The total radiative power of the jet ranges from  $1 \cdot 10^{43}$  to  $1 \cdot 10^{47}$  erg s<sup>-1</sup> and correlates with the luminosity of the accretion disk (Ghisellini, Tavecchio, et al., 2014). AGN with more massive core black holes are generally found to be more luminous in the radio and X-ray band (Merloni, Heinz and di Matteo, 2003). Launching and powering mechanisms providing this vast amount of energy can be put into two categories: accretion disk powered and black hole powered. The powering by rotating black holes has been proposed by Blandford and Znajek (1977) and is known as Blandford-Znajek mechanism after the authors. A potentially large part of the energy of a rotating black hole can be extracted, how much depends on the angular momentum of the black hole. This is possible if it is surrounded by a magnetic field to which angular momentum can be transferred which leads to an outwards-directed Poynting flux (Romero and Vila, 2014).

Jets can also be powered by the accretion disk, which contains magnetic fields, as proposed by and named after Blandford and Payne (1982). The fields are dragged along with the rotational movement of the accretion disk matter when the plasma's conductivity is high as one would expect. This can lead to the winding up of the field lines to form collimated magnetic towers rising perpendicular to the accretion disk (Lynden-Bell, 2003). The highly collimated structure of the jet can not be explained by self-collimation without taking external pressure into account (Begelman, 1995). However, in recent years magnetohydrodynamic simulations can reproduce the jet formation and collimation behavior by the accretion disk over large scales (Ramsey and Clarke, 2019) and in general relativistic MHD studies both the jet launching by the Blandford-Znajek and the Blandford-Payne mechanism could be shown under different conditions (Dihingia, Vaidya and Fendt, 2021). Correlations between the activity of the accretion disk and the emergence of knots visible in the radio band could be found for the source 3C 111 (Chatterjee et al., 2011) supporting the idea of disk-powered jets.

The mechanism providing energy for the jet is also expected to influence its composition. Black-hole powered jets are made up of a very tenuous leptonic plasma consisting of electrons and positrons and are described as a collimated electromagnetic wave (Romero and Vila, 2014). In contrast, if the jet power originates from the accretion disk, a hadronic plasma is expected. It is still an open question which of the two mechanisms are, or in which ratio they are, responsible for

the widely observed emergence of jets and why only some black holes in the center of galaxies form jets.

### 2.2.2. Structure and morphology

Jets themselves show varying intensity and spectra along their extent which can be best observed in radio galaxies when viewing the jet edge on. For observing their structure and their development the radio band is generally the most useful since with interferometry-based telescope arrays such as the VLBA (Very Long Baseline Array) or the EHT (Event Horizon Telescope) the highest resolution images to this day can be made. In radio images jets generally consist of bright spots, called *blobs* or *knots* which are travelling outwards from the core (for example Kadler, Ros, et al., 2008). Their propagation can either be ballistic, *i.e.* the knot travels independently from its environment, or advective with the background jet plasma. Most of the observed knots seem to be advected with the bulk motion of the jet and many move with apparently superluminal velocities ranging from several times up to multiple tens of the speed of light (Lister et al., 2009). For example in the jet of M87 the apparent velocities are estimated to be within  $4 - 6c$  (Biretta, Sparks and Macchetto, 1999). Closer to the core the jets are well collimated while further outside they bend in some sources until the collimation of the jet decreases and it forms looser *plumes* that decrease in intensity.

Also the spectrum changes across the jet, where flatter spectra are observed closer to the core and higher frequency parts decay further outside the central region. Explanations for this behavior include the increase of the magnetic field strength, leading to a shorter synchrotron loss timescale. Therefore, particles cannot accelerate to energies as high as closer to the core (Sambruna and Harris, 2012). Another theory is that the bulk Lorentz factor  $\Gamma = (1 - \beta^2)^{-1/2}$  decreases and therefore the apparent intensity in the plasma reference frame of external photon fields decreases. In this case fewer cosmic rays can scatter external photons to higher energies via the inverse-Compton effect.

#### Relativistic effects

The velocity of the plasma in the jet is close to the speed of light, therefore relativistic effects have to be discussed (Romero, Böttcher, et al., 2017). As already mentioned, superluminal motion of jet features are observed for many AGN jets. The apparent velocity of a feature in the jet is related to its physical velocity through the angle between the velocity vector and the line of sight to Earth  $\theta$  as (Biretta, Zhou and Owen, 1995)

$$\beta_a = \frac{\beta \sin \theta}{1 - \beta \cos \theta}. \quad (2.2)$$

This imposes a lower limit on  $\beta$  and an upper limit on the angle  $\theta$  and suggests highly relativistic motion for the observed apparent motions. Examples are for  $\beta_a = 1$  at least  $\beta \approx 0.71, \theta \approx 45^\circ$  and for  $\beta_a = 5$  at least  $\beta \approx 0.98, \theta \approx 11^\circ$ . VLBA observations enabled the estimation of the proper velocity of the jet features. They arrived at Lorentz factors in the range of 5 to 40 (Jorstad et al., 2005).

If radiating particles are at high Lorentz factors relative to the observer's frame of reference, the measured radiation's frequency is Doppler-shifted and the measured flux is Doppler-boosted (see section 6.1). The relativistic Doppler factor is given by

$$\delta_D = \frac{1}{\Gamma(1 - \beta \cos \theta)} \quad (2.3)$$

where  $\Gamma$  is the bulk Lorentz factor. The observed intensity from a radiating blob in the jet is then boosted by a factor  $\propto \delta_D^4$ . This matches the observation that in the majority of sources only a single jet is observed. The counter jet is rarely observed because the luminosity of the jet travelling

towards Earth is relativistically boosted and the luminosity of the jet travelling away from Earth is de-boosted strongly. This also imposes a bias on the jets we are able to observe: If the jet's matter propagates with higher velocities its radiation is boosted to higher luminosities and can be therefore observed with the available telescopes while slower features remain hidden.

### 2.2.3. Spectral footprint

The spectrum of a blazar extends over the whole electromagnetic spectrum up to  $\gamma$ -rays and typically forms two peaks which is commonly known as the double hump spectrum. After radio observations could provide images with a resolution high enough to resolve the jet as the source of AGN emission, the development of the CHANDRA X-ray telescope, for instance, made it possible to associate also X-ray emission from AGN with jets detected in the radio band (Kadler, Kerp, et al., 2004). In the radio and optical band jet radiation is usually polarized which was first discovered in the jet of M87 by Baade (1956) and quickly lead to the conclusion of a synchrotron process (Burbidge, 1956). Angel and Stockman (1980) later proposed that blazars are AGNs and again showed polarized emission in the optical and infrared region in those sources.

The lower-frequency hump of the double-hump spectrum is mainly accounted to synchrotron emission while the high-frequency part of the spectrum is most often assumed to be caused by inverse-Compton scattering. Apart from inverse-Compton scattering hadronic and nuclear processes may be at work which indirectly lead to the emission of very high energetic photons and other particles such as neutrinos (Mannheim, 1993).

#### Variability

The assumption of a single-blob emission model does not hold in practice as can be seen for example at the fast variability of the radio emission polarisation angle. It shows that most likely multiple emission regions with differently oriented magnetic fields are at play (Agudo et al., 2018). Violent variability in luminosity is also observed in some types of AGN such as BL Lac objects and optically violent variable quasars. The timescales of the variability limit the extent of the radiating regions to the corresponding light-travel distances, which is a hint at the existence of local particle acceleration and radiation mechanisms. In this case energy is transported with the bulk motion until it reaches the region of radiation, where it is radiated away.

Variability is observed in all regions of the spectrum and correlation between different parts of the spectrum can be observed (e.g. Sbarrato et al., 2012). The X-ray variability correlates with the mass of the central black hole and the luminosity, with brighter and more massive sources exhibiting less variability (Lanzuisi et al., 2014).

#### Emission processes

One possible radiation process is the free-free emission or bremsstrahlung which happens if particles electrostatically scatter at each others potentials in a relatively dense plasma and therefore radiate because of the experienced acceleration. Thermal bremsstrahlung is not expected to be a relevant emission mechanism in jets because it would require densities that are in conflict with estimates, although sometimes models could in principle be fitted to observed data (Schwartz et al., 2000). It is also ruled out as the main driver of jet emission because it emits unpolarized radiation contrary to observations.

Synchrotron emission forms one of the main parts of the spectrum and is observed universally in jets, mainly in the radio band. Its popularity originates in its ability to explain the polarized radiation and power laws observed over many orders of magnitude. It is produced by relativistic, charged particles that gyrate due to external magnetic fields and radiate because of their continued acceleration on the orbit. If magnetic fields and particle energies are high enough, synchrotron emission from electrons can be observed extending to the optical and even the X-ray band

(Sambruna, Gambill, et al., 2004). Electrons are more efficient radiators of synchrotron radiation than protons because of their lower mass and are therefore usually considered the source of the observed radiation. Nevertheless, synchrotron radiation can naturally also be emitted by protons if they are accelerated to very high energies  $> 1 \cdot 10^{18}$  eV and are a possible explanation of X-ray emission (Aharonian, 2002).

Synchrotron-emitted photons can rescatter with the population of electrons emitting synchrotron radiation and therefore not escape the source region and reach the observer. This is called synchrotron self-absorption and happens in optically thick conditions, *i.e.* if the mean free path of reabsorption scattering is smaller than the source region. In this case the synchrotron spectrum is no longer dependent on the electron momentum spectrum but has a constant spectral index of 5/2 (Rybicki and Lightman, 1986). In the optically thick region the electron population therefore also loses less energy. Synchrotron self-absorption is observed in radio spectra as a break of the power laws at the low-frequency end of the spectrum.

Synchrotron radiation, especially as an explanation for optical and X-ray emission, requires the local acceleration of charged particles and excludes the possibility for radiating particles originating in the core region. The lifetimes of particles decreases with increasing frequency as  $\tau \sim 1/\sqrt{\nu}$  and is on the order of a few hundred years in the optical band while emission is detected multiple kpc from the core (Perlman, 2012). Another argument that even extends to radio sources is that for electrons having high energies at the emission sites they must have had unrealistically high energies at the nucleus if not accelerated near the emission site (Begelman, Blandford and Rees, 1984). In contrast, Gopal-Krishna et al. (2001) show that plausible parameters can be found that explain optical synchrotron emission in many sources without the need of local acceleration mechanisms.

Excess radiation in the optical and X-ray bands above extrapolated radio spectra is observed for example in 3C 273 (Jester et al., 2002) among other sources and requires another contribution to the radiation apart from synchrotron radiation. The additional contribution is assumed to be inverse-Compton upscattering of some photon field which can also explain  $\gamma$ -ray emission. External Compton scattering occurs if relativistic electrons scatter at photons, while in contrast to the classical Compton scattering process, the electrons lose energy and photons with higher energy are emitted. This can be attributed to a single electron population producing synchrotron and inverse-Compton radiation which explains the spectra of multi-band observations (Dermer and Atoyan, 2002), but multiple populations are also possible.

Inverse Compton scattering can occur with a variety of different photon fields that may be present at the emission zone. First the accelerated particles can upscatter the synchrotron photons produced by themselves which is called synchrotron self Compton (SSC) emission. It is considered unlikely to be the only explanation of the observed X-ray luminosity (Finke, Dermer and Böttcher, 2008) and IC scattering at other targets must be taken into account. Scattering of the infrared emission of hot dust heated by the emission from the accretion disk has also been proposed as a source of X-ray emission (Błażejowski et al., 2000) as well as the direct comptonization of the accretion disk's radiation itself for radiation originating near the core. This was proposed by Dermer and Schlickeiser (1993) who explain the emission of  $\gamma$ -rays among other spectral features with electrons accelerated near the core upscattering accretion disk photons. Another popular approach is to assume inverse-Compton scattering of the cosmic microwave background (CMB) radiation. This is thought to dominate over SSC especially in high-redshifted sources since the apparent density of the CMB photon field is boosted from the viewpoint of the relativistic plasma and its energy density can exceed that of the magnetic field (Ghisellini, Celotti, et al., 2014). High redshifts also suggest that emission occurred when the universe was younger and the CMB radiation was more intense.

Another mechanism is the photo-pair production which attenuates the emission in the  $\gamma$ -ray range by the production of electron-positron pairs.

### Jet composition

The composition of the matter in the jet naturally influences the radiation emitted by the jet. It is assumed that it consists of a collisionless plasma, which is an ionized gas where the free charges influence its properties considerably. Since collisions are sparse (the collision length being longer than the size of the system) the plasma has no means of relaxing to a thermal (Maxwellian) distribution and therefore can be far from thermodynamic equilibrium. This is called *non-thermal plasma* and is likely present in jets. The plasma must be quasi-neutral, that is the positive and negative charges balance each other, because confinement as observed in jets would not be possible in charged plasmas. For the magnetic field strength, estimates vary greatly, but limits can be derived from the peak intensity of the synchrotron hump and the inverse-Compton luminosity. In the jet of M87, it has been estimated to be at least 230  $\mu\text{G}$  (Snios et al., 2019) while for other sources estimates range from 0.007 to 612 G in different radiation models (Sambruna, Maraschi and Urry, 1996).

There are two common models for the charges making up the plasma. The first is that of a *leptonic* jet consisting of an electron-positron plasma and the second is a *hadronic* jet which consists of protons and electrons. It is assumed that the composition depends on the mechanism of jet launching as discussed in section 2.2.1. A model assuming that a mixed population of both plasma types is present in the jet is proposed by Sol, Pelletier and Asseo (1989) and matches the expectation that both launching mechanisms play a role. The leptonic plasma is expected to move much faster than the hadronic plasma which could explain the detection of both superluminal and subluminal motion in the same jet.

Hadronic contents in the jet are also probable as they can explain emission of x- and  $\gamma$ -ray photons when protons are accelerated to very high energies (Mannheim, 1993). Protons are also accelerated by diffusive shock acceleration and can interact with the synchrotron photons emitted by relativistic electrons to trigger the production of energetic pions. These pions then decay into pairs of photons which start a cascade of electron-positron pair production and synchrotron emission (Mannheim, Krüß and Biermann, 1991).

#### 2.2.4. Particle acceleration

As discussed above, the cooling times of the radiating particles are generally much shorter than the light travel time from the core to the radiating lobes and knots. This requires the presence of acceleration mechanisms at or near the hotspots of radiation (e.g. Romero, Böttcher, et al., 2017). Some of those are summarized here and discussed in section 3.2. The most common mechanism is first-order Fermi acceleration, also called diffusive shock acceleration. At a shock front, that is a steep gradient in pressure, density and velocity of the bulk plasma, particles are accelerated by crossing the shock many times while undergoing diffusive motion. If the acceleration region is large enough and particles are scattered sufficiently so that they are able to cross many times before escaping, a substantial amount of energy can be gained. The gain of energy per crossing is of first order in the shock's velocity difference, hence the name. In an AGN jet, there are plasma conditions, notably the highly relativistic bulk motion, that question the applicability of this mechanism. Furthermore, diffusive shock acceleration is only efficient in re-accelerating particles that already are much faster than the background medium. This is known as the *injection problem* and details on diffusive shock acceleration are discussed in section 3.2.2.

Another mechanism that is possible at shock fronts is the shock drift acceleration, in which particles cross the shock only a single time (Begelman and Kirk, 1990). In superluminal shocks (in the sense that the intersection of field lines with the shock front travels superluminally), which are expected in a plasma with highly relativistic bulk motion, the particle travels along electric field gradients which accelerate it (Marcowith et al., 2016). The possible energy gain, however, is

relatively limited, and is at maximum

$$\frac{\gamma_1}{\gamma_2} = \sqrt{\frac{B_2}{B_1}} \quad (2.4)$$

where the ratio of magnetic field strengths is at maximum 4 when assuming the plasma is an ideal gas (Kirk, 1994).

Without the need of a shock front, second-order Fermi acceleration takes place if particles randomly scatter at so-called *magnetic mirrors*, that are most commonly assumed to be plasma waves. In principle, the particles can equally gain or loose energy in those collisions, but Fermi (1949) reasoned that head-on collisions in which particles gain energy are more likely to happen than lossy collisions and therefore a net energy gain is achieved.

### 3. Theoretical description of the jet plasma

The description of plasmas is the starting point for describing the physics in the jets of AGN. A plasma is a gas with substantial ionization of the atoms or molecules, which substantially changes its characteristics. In a plasma consisting of a population of positive and negative charges the particles mainly interact according to their collective Lorentz forces instead of their collisions. A fundamental quantity of a plasma is the plasma frequency  $\omega_p$

$$\omega_p = \sqrt{\frac{4\pi e^2 n}{m}} \quad (3.1)$$

which can be given for any particle species in the plasma with  $n$  as the (electron) density,  $m$  as the (electron) mass and  $e$  the elementary charge, written in Gaussian units. It is the oscillation frequency of a disturbance of charge density equilibrium, *i.e.* a infinitesimal charge separation and gives the fundamental timescale of plasmas. The particles surrounding a test particle shield the electrostatic forces of all other particles beyond a distance  $\lambda_D$  which is called the Debye length. It is given by the distance a typical particle with thermal velocity of  $v_t = \sqrt{k_B T/m}$  travels during a period of the plasma oscillations at a temperature  $T$  with the Boltzmann constant  $k_B$ :

$$\lambda_D = \frac{v_t}{\omega_p} = \sqrt{\frac{k_B T}{4\pi n e^2}} \quad (3.2)$$

Describing some medium as a plasma is reasonable if the size of the considered region is much larger than this length. The number of particles in the Debye sphere is given by

$$\Lambda = \frac{4\pi \lambda_D^3}{3} n \quad (3.3)$$

and is called the *plasma parameter*. The collisions are dominated by the Coulomb interaction of the particles and a collision frequency  $\nu = \nu n \sigma$  is introduced. To estimate the collision cross section  $\sigma$  one can estimate the distance of closest approach  $r$  by the distance where the average kinetic and potential energies equal each other (Hasegawa, 1975):

$$E_{\text{kin}} = \frac{\bar{m} v_t^2}{2} = \frac{e^2}{r} = E_{\text{pot}} \Rightarrow r = \frac{2e^2}{\bar{m} v_t^2} \quad (3.4)$$

Therefore the collision frequency in the case of electron-electron collisions with the reduced mass  $\bar{m} \approx m$  and cross section  $\sigma = \pi r^2$  reads

$$\nu = \frac{4\pi n e^4}{m^2 v_t^3} = \frac{\omega_p}{4\pi n \lambda_D^3} = \frac{\omega_p}{3\Lambda} \quad (3.5)$$

where the plasma frequency and the number of particles in the Debye sphere can be identified. For the shielding of electromagnetic forces to be effective, many particles have to be in the Debye sphere, *i.e.*  $\Lambda \gg 1$ . In this case of the so-called weakly coupled plasma  $\nu \ll \omega_p$  collisions occur rarely in comparison to the timescale of the plasma oscillations and therefore can barely influence the latter. Based on the same argument it also follows that the mean free path  $\lambda_{\text{mfp}} = v_t/\nu$  is much

larger than the Debye sphere so collisions play a minor role in those plasmas. For this reason they are also called *collisionless* especially compared to neutral gases where collisions play a significant role. Collisionless plasmas are typically relatively hot and tenuous since  $\Lambda \propto \sqrt{T^3/n}$  and are an appropriate description for most plasmas that can be found in space (Fitzpatrick, 2014) and likely also in AGN jets.

While in neutral gases collisions between molecules quickly establish thermodynamic equilibrium, *i.e.* the particles' velocities are always distributed closely to the thermal Maxwell-Boltzmann-distribution, the weak particle interactions in plasmas allow the plasma to be driven far from thermal equilibrium. It must be noted that this is only an approximation because since the Coulomb interaction is long-ranged particles are also scattered when crossing much farther away than  $r$ , although by a smaller magnitude. Fitzpatrick (2014) shows that a correction of order  $\ln \Lambda$  must be introduced but this does not conflict with the general argument.

### 3.1. Kinetic plasma physics

#### 3.1.1. Vlasov equations

To describe the evolution from a known initial state of a plasma one starts from the exact distribution function  $f(\vec{x}, \vec{p}, t)$  in phase space  $(\vec{x}, \vec{p})$ . The distribution function must be conserved in phase space, *i.e.*

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} + \sum_i \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} = 0. \quad (3.6)$$

The distribution function is essentially a sum of  $\delta$ -distributions, one for each particle, and leads to a many-body problem which is not solvable for a realistic number of particles. (Hazeltine and Waelbroeck, 2004) To overcome this problem an ensemble average  $\bar{f} = \langle f \rangle$  is taken. Because of the coupling of the electromagnetic field to the exact particle trajectories the force  $dp/dt$  is correlated with  $f$  and when averaging this term in equation (3.6) a correlation term  $C(f)$  has to be introduced (Hazeltine and Waelbroeck, 2004):

$$\left\langle \frac{dp_i}{dt} \frac{\partial f}{\partial p_i} \right\rangle = \left\langle \frac{dp_i}{dt} \right\rangle \frac{\partial \bar{f}}{\partial p_i} - C_i(\bar{f}). \quad (3.7)$$

The correlation term is dominated by close interactions of particles and is therefore also called collision operator. Taking the ensemble average of equation (3.6) and identifying the velocity  $v_i = dx_i/dt$  and the force  $F_i = dp_i/dt$  leads to the Boltzmann equation which reads

$$\frac{\partial \bar{f}}{\partial t} + \sum_i v_i \frac{\partial \bar{f}}{\partial x_i} + \sum_i F_i \frac{\partial \bar{f}}{\partial p_i} = \sum_i C_i(\bar{f}) \quad (3.8)$$

The averaged distribution function  $\bar{f}(\vec{x}, \vec{p}, t)$  now gives the probability density for finding a particle at the given phase space point and time. Vlasov (1968) started from this equation and proposed two ideas when applying the Boltzmann equation to a weakly coupled plasma:

- The force term is expressed by the Lorentz force caused by an electromagnetic field  $\vec{E}, \vec{B}$  that itself is determined by Maxwell's equations. Those depend on the charge and current densities which can be given in terms of the distribution function  $\bar{f}$ .
- The collision term is set to zero since in the case of a weakly coupled plasma the Coulomb interactions become negligible. This is equivalent to assuming that the particle trajectories are not correlated as described above and corresponds to the collisionless approximation.

Replacing  $\bar{f}$  with  $f$  for simplicity one arrives at a system of coupled equations that is commonly known as the Vlasov equations, where the kinetic equation (3.9) is known as *the Vlasov equation* which reads

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla f - e(\vec{E} + \frac{1}{c} \vec{v} \times \vec{B}) \cdot \nabla_p f = S \quad (3.9)$$

$$\rho = -e \int f d^3 p \quad (3.10)$$

$$\vec{j} = -e \int \vec{v} f d^3 p \quad (3.11)$$

$$\nabla \cdot \vec{E} = 4\pi\rho$$

$$\nabla \cdot \vec{B} = 0$$

$$\nabla \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \quad (3.12)$$

$$\nabla \times \vec{B} = \frac{1}{c} \left( 4\pi\vec{j} + \frac{\partial \vec{E}}{\partial t} \right)$$

introducing a particle source and sink term  $S$ . The above and especially the equations (3.9) to (3.11) are given for electrons but hold for each species of particles separately and charge and current densities  $\vec{j}, \rho$  have to be summed up replacing  $e$  by the respective charges. The Vlasov equations also hold in the relativistic case when using the relativistic momentum  $p = \gamma mv$  (Schlickeiser, 2002).

The coupling of this system of equations can be partially resolved in two fundamental ways. On the one hand the *test-particle approach* can be used where the fields  $\vec{E}, \vec{B}$  are assumed to be given and the propagation of the test particles described by  $f$  is studied. On the other hand there is the *test-field approach* which prescribes the particle distribution  $f$  and solves for the resulting fields  $\vec{E}, \vec{B}$ . Both of these are limits where parts of the physical coupling of fields and particles are ignored and therefore are only valid if the influence of one component on the other can be neglected.

### 3.1.2. Quasilinear approximation

One major assumption for arriving at a transport equation for test particles in plasma is that the electromagnetic fields deviate from some given field by only small turbulences. Since the plasma is assumed to be collisionless its conductivity is nearly infinite. Electric fields are therefore short-circuited so there will be no macroscopic electric fields. Without loss of generality the magnetic field can be chosen to align with the  $z$ -axis and the fields can therefore be written as

$$\vec{B} = B_0 \hat{z} + \delta\vec{B}, \quad \vec{E} = \delta\vec{E} \quad (3.13)$$

where  $\delta\vec{E}, \delta\vec{B}$  denote the turbulent deviations. Every particle will then perform a gyroscopic trajectory in the magnetic field with a gyroradius of

$$r_g = \frac{p_\perp}{|q|B_0} \quad (3.14)$$

and a gyrofrequency

$$\omega_g = \frac{|q|B_0}{m}. \quad (3.15)$$

Particles therefore move in helical orbits and it is useful to introduce a set of coordinates that separates the gyroscopic motion from the motion of the gyrorotation centers (Schlickeiser, 2002):

$$\vec{x}' = \vec{x} + \frac{\vec{p} \times \hat{z}}{\epsilon m \omega_g} \quad (3.16)$$

where  $\epsilon = |q|/q$ . It is helpful to use spherical coordinates for the momentum space which are

$$\vec{p}' = p \begin{pmatrix} \cos \phi \sqrt{1 - \mu^2} \\ \sin \phi \sqrt{1 - \mu^2} \\ \mu \end{pmatrix} \quad (3.17)$$

where  $\phi$  is the gyrophase and  $\mu = \cos \theta$  is the pitch-angle, *i.e.* the angle between the momentum and the magnetic field direction. The spatial coordinates then read

$$x' = x + \sin \phi \frac{p \sqrt{1 - \mu^2}}{\epsilon m \omega_g}, \quad y' = y - \cos \phi \frac{p \sqrt{1 - \mu^2}}{\epsilon m \omega_g}, \quad z' = z. \quad (3.18)$$

Transforming the Vlasov equation (3.9) into this coordinate set leads to (Achatz, Steinacker and Schlickeiser, 1991)

$$\frac{\partial f}{\partial t} + \mathcal{G}(f) + \frac{1}{p^2} \frac{\partial}{\partial x_\alpha} (p^2 g_\alpha f) = S(\vec{x}, \vec{p}, t) \quad (3.19)$$

with  $\mathcal{G}(f)$  giving the influence of the unperturbed field

$$\mathcal{G}(f) = v \mu \frac{\partial f}{\partial z'} - \epsilon \omega_p \frac{\partial f}{\partial \phi} \quad (3.20)$$

and the effect of the field fluctuations are separated into a set of generalized forces  $g_\alpha$ . A single stochastic particle orbit which results from a specific history of the fluctuating fields is not of particular interest; instead one considers, again, an ensemble average which this time runs over phase space trajectories of  $f$  for a set of random fluctuations. The ensemble average of the fluctuations is zero, *i.e.*  $\langle \delta B \rangle = \langle \delta E \rangle = 0$  and one is interested in an equation for  $F = \langle f \rangle$

$$\frac{\partial F}{\partial t} + \mathcal{G}(F) + \frac{1}{p^2} \frac{\partial}{\partial x_\alpha} (p^2 g_\alpha \delta f) = S(\vec{x}, \vec{p}, t) \quad (3.21)$$

where  $\delta f = f - F$  and the usual summation convention over identical indices is implied. Subtracting the averaged equation (3.21) from equation (3.19) gives an equation in terms of the fluctuations

$$\frac{\partial \delta f}{\partial t} + \mathcal{G}(\delta f) + g_\alpha \frac{\partial F}{\partial x_\alpha} + g_\alpha \frac{\partial \delta f}{\partial x_\alpha} - \left\langle g_\alpha \frac{\partial \delta f}{\partial x_\alpha} \right\rangle = 0. \quad (3.22)$$

If the timescale  $t_g$  of the effects of the fluctuating forces  $g_\alpha$  is long in comparison to the observed timescale the last two terms can be neglected in comparison to the other terms. Achatz, Steinacker and Schlickeiser (1991) solve this equation for  $\delta f$  using the method of characteristics and arrive at

$$\delta f(t) = \delta f(t_0) - \int_{t_0}^t dt' g_\alpha \frac{\partial F}{\partial x_\alpha} \quad (3.23)$$

where the integrand has to be evaluated along the unperturbed orbits of the particle, *i.e.* only the contribution of  $\mathcal{G}(F)$  to the propagation. It is necessary to assume that at the time  $t_0$  the particle fluctuations are completely uncorrelated to the turbulences. Then, plugging equation (3.23) into equation (3.19) only the integral term remains and one arrives at

$$\frac{\partial f}{\partial t} + \mathcal{G}(f) + \frac{1}{p^2} \frac{\partial}{\partial x_\alpha} \left( p^2 \int_{t_0}^t dt' \langle g_\alpha(t) g_\beta(t') \rangle \frac{\partial F(t')}{\partial x_\beta} \right) = S(\vec{x}, \vec{p}, t). \quad (3.24)$$

If there is a correlation time  $t_c$  beyond which the fluctuations  $g_\alpha$  and  $g_\beta$  are uncorrelated the lower limit of integration can be replaced by 0 since there are no contributions to the integral's value

before  $t - t_c$ . Assuming that the variations in  $\partial F/\partial x_\beta$  are small during the time in which the fluctuations are correlated the equation reduces to a diffusion equation when pulling  $\partial F/\partial x_\beta$  out of the integral. This is the result of the quasilinear approximation.

$$\frac{\partial F}{\partial t} + v\mu \frac{\partial F}{\partial z'} - \epsilon\omega_p \frac{\partial F}{\partial \phi} - \frac{1}{p^2} \frac{\partial}{\partial x_\alpha} \left( p^2 D_{\alpha\beta} \frac{\partial F}{\partial x_\beta} \right) = S(\vec{x}, \vec{p}, t) \quad (3.25)$$

where  $D_{\alpha\beta}$  are the Fokker-Planck coefficients which represent the turbulent properties of the plasma and are given by the correlation functions of the generalized forces  $g_\alpha$  integrated along unperturbed orbits:

$$D_{\alpha\beta} = \int_0^t dt' \langle g_\alpha(t) g_\beta(t') \rangle \quad (3.26)$$

These coefficients are lengthy to calculate and resulting equations cannot be solved analytically in most cases so one naturally searches for ways to avoid calculating some of them. One possibility is the diffusion approximation described in the next section.

The quasilinear approximation naturally only holds for small deviations from some given large-scale electromagnetic fields. In AGN jets a significant portion of the energy of the plasma bulk is assumed to accelerate high-energy particles whose high-intensity radiation can be observed on Earth. This of course breaks the assumptions of the perturbation method since the plasma properties (like magnetic fields) itself will be influenced by the particles and challenges the applicability of the test-particle regime.

### 3.1.3. Diffusion approximation

Most of the diffusion terms of the Fokker-Planck equation (3.25) are of the order of  $v_A/v$ ,  $r_g/r$  or of second order of these terms where  $v_A$  is the plasma wave speed and  $r$  the length scale of variations of  $F$  while only  $D_{\mu\mu}$ ,  $D_{\mu\phi}$ ,  $D_{\phi\mu}$  and  $D_{\phi\phi}$  are of lower order (Schlickeiser, 2002). Assuming that the gyroradius  $r_g$  is much smaller than the observed length scales  $r$  and the particles have a velocity  $v$  that is much faster than the wave phase speed  $v_A$  diffusion of pitch- and gyrophase angle are the fastest processes in the system. In other words the time and length scales on which the density varies must be large compared to the pitch angle relaxation time  $\tau$  and the typical distance travelled in this time  $v\tau$  respectively so that particles can reach the isotropic local equilibrium. In this case the particle distribution can be assumed to stay close to isotropy at the timescales one is interested in and the anisotropic part of the distribution can be split off:

$$F = \underbrace{\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu F}_{\text{Isotropic part } F_i} + \underbrace{F - \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu F}_{\text{Anisotropic part } F_a} = F_i(\vec{x}, p, t) + F_a(\vec{x}, p, \mu, \phi, t) \quad (3.27)$$

where

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu F_a = 0. \quad (3.28)$$

and in the diffusion approximation  $F_a \ll F_i$ . Taking the average in  $\phi$  and  $\mu$  of equation (3.25) leads to

$$\begin{aligned} S &= \frac{\partial F_i}{\partial t} + \frac{v}{4\pi} \frac{\partial}{\partial z'} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \mu F_a \\ &- \sum_{\alpha, \beta = x', y', p} \frac{1}{4\pi p^2} \frac{\partial}{\partial x_\alpha} p^2 \frac{\partial F_i}{\partial x_\beta} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu D_{\alpha\beta} \\ &- \sum_{\substack{\alpha = x', y', p \\ \beta \neq z', \phi}} \frac{1}{4\pi p^2} \frac{\partial}{\partial x_\alpha} p^2 \int_0^{2\pi} d\phi \int_{-1}^1 d\mu D_{\alpha\beta} \frac{\partial F_a}{\partial x_\beta} \end{aligned} \quad (3.29)$$

Details of this calculation are shown in appendix A. To achieve the goal of a transport equation only depending on the isotropic part of the density  $F_i$  we subtract this equation from the unaveraged equation (3.25) giving an expression relating the anisotropic part  $F_a$  to the isotropic part  $F_i$  (Schlickeiser, 2002):

$$\begin{aligned} & \left( \epsilon \omega_p + \frac{\partial D_{\mu\phi}}{\partial \mu} \right) \frac{\partial F_a}{\partial \phi} + \frac{\partial}{\partial \phi} \left( D_{\phi\phi} \frac{\partial F_a}{\partial \phi} \right) + \frac{\partial}{\partial \mu} \left( D_{\mu\mu} \frac{\partial F_a}{\partial \mu} \right) \\ & = v\mu \frac{\partial F_i}{\partial z'} - \frac{\partial D_{\mu p}}{\partial \mu} \frac{\partial F_i}{\partial p} - \sum_{\alpha, \beta} \frac{\partial D_{\alpha\beta}}{\partial x_\alpha} \frac{\partial F_i}{\partial x_\beta} \end{aligned} \quad (3.30)$$

with  $\alpha \in \{\mu, \phi\}$ ,  $\beta \in \{x', y'\}$

The anisotropic part can be expressed as a Fourier series in gyrophase  $\phi$  with coefficients  $F_a^k$  that of course are not dependent of  $\phi$ :

$$F_a = \sum_k F_a^k e^{ik\phi} \quad (3.31)$$

Terms with  $|k| > 1$  will drop out in the gyrophase averaging and  $k = 1$  is of second order in  $(\delta B/B_0)^2$ . One arrives at the following equation (Schlickeiser, 2002)

$$\frac{\partial}{\partial \mu} \left( D_{\mu\mu} \frac{\partial F_a}{\partial \mu} + D_{\mu p} \frac{\partial F_i}{\partial p} \right) = v\mu \frac{\partial F_i}{\partial z'} \quad (3.32)$$

which can be integrated to obtain an expression for  $\partial_\mu F_a$  and therefore for term  $\int_{-1}^1 d\mu D_{\mu p} \partial_\mu F_a$  in equation (3.29):

$$\frac{\partial F_a}{\partial \mu} = -\frac{1-\mu^2}{2D_{\mu\mu}} v \frac{\partial F_i}{\partial z'} - \frac{D_{\mu p}}{D_{\mu\mu}} \frac{\partial F_i}{\partial p} \quad (3.33)$$

where the constant of integration is  $-v\partial_z F_i/2$  since  $D_{\mu\alpha}(\mu = \pm 1) = 0$  because there is no pitch-angle scattering when the particle moves parallel to the magnetic field. Integrating again over  $\mu$  leaves us with an expression for the anisotropy in terms of the isotropic distribution:

$$F_a = C - \frac{v}{2} \frac{\partial F_i}{\partial z'} \int_{-1}^{\mu} d\mu' \frac{1-\mu'^2}{D_{\mu\mu}} - \frac{\partial F_i}{\partial p} \int_{-1}^{\mu} d\mu' \frac{D_{\mu p}}{D_{\mu\mu}} \quad (3.34)$$

The anisotropy therefore arises from spatial and momentum gradients in the isotropic distribution. This equation can be used to find an expression for the second term in equation (3.29) in terms of  $F_i$ :

$$\frac{v}{4\pi} \frac{\partial}{\partial z'} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \mu F_a = -\frac{v}{4} \frac{\partial F_i}{\partial z'} \int_{-1}^1 d\mu \frac{(1-\mu^2)^2}{D_{\mu\mu}} - \frac{1}{2} \frac{\partial F_i}{\partial p} \int_{-1}^1 d\mu (1-\mu^2) \frac{D_{\mu p}}{D_{\mu\mu}} \quad (3.35)$$

The integrals have been partially integrated similar to

$$\int_{-1}^1 d\mu \mu \int_{-1}^{\mu} d\mu' \mathcal{F}(\mu') = \left[ \frac{\mu^2}{2} \int_{-1}^{\mu} d\mu' \mathcal{F}(\mu') \right]_{\mu=-1}^{\mu=1} - \int_{-1}^1 d\mu \frac{\mu^2}{2} \mathcal{F}(\mu) = \frac{1}{2} \int_{-1}^1 d\mu (1-\mu^2) \mathcal{F}(\mu) \quad (3.36)$$

and the integration constant (being independent of  $\mu$ ) vanishes because  $\int_{-1}^1 d\mu \mu = 0$ . With equation (3.33) one can also evaluate the term

$$\int_{-1}^1 d\mu D_{\mu p} \frac{\partial F_a}{\partial \mu} = -\frac{v}{2} \frac{\partial F_i}{\partial z'} \int_{-1}^1 d\mu (1-\mu^2) \frac{D_{\mu p}}{D_{\mu\mu}} - \frac{\partial F_i}{\partial p} \int_{-1}^1 d\mu \frac{D_{\mu p}^2}{D_{\mu\mu}} \quad (3.37)$$

using  $D_{p\mu} = D_{\mu p}$ . One can now write down the diffusion-convection equation for the isotropic part of the density by plugging in equations (3.35) and (3.37) into equation (3.29) and reordering the terms.

$$\begin{aligned}
 \frac{\partial F_i}{\partial t} = & \frac{\partial}{\partial z} \left( \underbrace{\frac{v^2}{8} \int_{-1}^1 d\mu \frac{(1-\mu^2)^2}{D_{\mu\mu}}}_{\kappa_{zz}} \frac{\partial F_i}{\partial z} \right) + \sum_{i,j \in \{x',y'\}} \frac{\partial}{\partial x_i} \left( \underbrace{\frac{1}{2} \int_{-1}^1 d\mu D_{ij}}_{\kappa_{ij}} \frac{\partial F_i}{\partial x_j} \right) \\
 & - \frac{1}{p^2} \frac{\partial}{\partial p} \left( \underbrace{p^2 \frac{v}{4} \int_{-1}^1 d\mu (1-\mu^2) \frac{D_{\mu p}}{D_{\mu\mu}}}_{a_1} \frac{\partial F_i}{\partial z} + \frac{\partial}{\partial z} \left( \underbrace{\frac{v}{4} \int_{-1}^1 d\mu (1-\mu^2) \frac{D_{\mu p}}{D_{\mu\mu}}}_{a_1} \right) \frac{\partial F_i}{\partial p} \right) \\
 & + \frac{1}{p^2} \frac{\partial}{\partial p} \left( \underbrace{p^2 \frac{1}{2} \int_{-1}^1 d\mu \left( D_{pp} - \frac{D_{\mu p}^2}{D_{\mu\mu}} \right)}_{a_2} \frac{\partial F_i}{\partial p} \right) + S
 \end{aligned} \tag{3.38}$$

The remaining Fokker-Planck coefficients are  $D_{pp}$  giving momentum change through fluctuating electric fields and  $D_{\mu\mu}$  for pitch-angle scattering. In this equation 7 transport parameters can be identified that are given in terms of pitch-angle averages of  $D_{\alpha\beta}$ .  $D_{\mu p}$  arises from correlations between magnetic and electric field turbulence. The coefficients  $\kappa_{ij}$  are the components of the spatial diffusion coefficient and the momentum diffusion is parametrized by  $a_1, a_2$ . When considering Alfvén waves propagating parallel to the magnetic field and the damping of the waves can be considered small the diffusion coefficients for perpendicular directions  $D_{x'x'} = D_{y'y'} = D_{x'y'} = D_{y'x'} = 0$  and therefore  $\kappa_{ij} = 0$  except for  $\kappa_{zz}$  (Schlickeiser, 2002). They are therefore neglected in the further discussion.

To consider a moving background plasma it is useful to write the transport equation in a mixed comoving coordinate system in which the spatial part is measured in the laboratory frame and the particle's momentum is measured in the rest frame of the plasma's bulk motion. Kirk, Schlickeiser and Schneider (1988) give an expression of equation (3.38) transformed to this coordinate system using a Lorentz transform. In the limit of nonrelativistic background plasma velocities it reads with  $\kappa = \kappa_{zz}$  and renaming  $z := z'$

$$\begin{aligned}
 \frac{\partial F}{\partial t} = & -u \frac{\partial F}{\partial z} + \frac{\partial}{\partial z} \left( \kappa \frac{\partial F}{\partial z} \right) - \frac{1}{p^2} \frac{\partial (p^2 a_1)}{\partial p} \frac{\partial F}{\partial z} \\
 & + \left( \frac{p}{3} \frac{du}{dz} + \frac{\partial a_1}{\partial z} \right) \frac{\partial F}{\partial p} + \frac{1}{p^2} \frac{\partial}{\partial p} \left( p^2 a_2 \frac{\partial F}{\partial p} \right) + S
 \end{aligned} \tag{3.39}$$

where  $p$  is now to be read as comoving with the plasma bulk motion parametrized as  $u(z)$ . The effect of processes leading to energy losses  $\dot{p}$  can be included into this equation by adding additional terms of the form

$$\frac{1}{p^2} \frac{\partial}{\partial p} (p^2 \dot{p} F) \tag{3.40}$$

Equation (3.39) is the basic equation used in the rest of this work.

### 3.1.4. Plasma waves

Alfvén waves are a wave of oscillating ions in the plasma in response to the magnetic field line tension. The group velocity of an Alfvén wave is the Alfvén velocity that is given by

$$v_A = \frac{B_0}{\sqrt{4\pi n_e (m_p + m_e)}} \tag{3.41}$$

The gyroresonant instability (proposed by Kulsrud and Pearce, 1969) enables particles in the plasma to excite Alfvén waves which in turn are able to pitch-angle scatter the particles. The

Alfvén waves are excited by anisotropies in the distribution function which arise from particles travelling along the magnetic field lines of force (Thomas and Pfrommer, 2019).

## 3.2. Particle acceleration mechanisms

As discussed in section 2.2.4 the radiation observed from jets seemingly requires acceleration of particles *in situ* because the lifetime of particles accelerated close to the core is much smaller than the light travel time from the core to the outer regions of the jet, where radiation is observed (Perlman, 2012). Many popular mechanisms used for explaining the acceleration up to extremely relativistic energies assume the presence of a shock.

### 3.2.1. Acceleration mechanisms at shocks

It is assumed that shock fronts in the plasma play a central role in the acceleration of electrons. A shock is essentially a sudden change in pressure and velocity of the plasma bulk medium where a faster travelling part of the plasma hits a slower one. Astrophysical plasmas and therefore the shocks can be considered collisionless because of the long mean free path of Coulomb collisions as opposed to the observed lengthscales. While particle-particle interactions are therefore not significant, wave-particle interactions are and plasma waves scattering particles are of central importance for the acceleration mechanisms. Possible mechanisms for acceleration are discussed in this chapter.

### 3.2.2. Diffusive shock acceleration

Diffusive shock acceleration, also known as first-order Fermi acceleration, happens when particles are able to cross the shock front multiple times due to diffusion and gain energy at each crossing. Consider a shock perpendicular to the  $z$ -axis with bulk velocities  $u_1$  and  $u_2$  for upstream (1) and downstream (2) respectively. If a particle with pitch-angle  $\mu$  and momentum  $p$  crosses the shock front its momentum is not inherently changed (before any scattering occurs) but since in the mixed comoving coordinates used the particle's momentum is measured in the bulk plasma rest frame and the value of the momentum changes. We can use a Lorentz boost to find the particles energy and momentum after crossing the front. The momentum component parallel to the shock normal is denoted by  $p_{\parallel} = \mu p$  and the Lorentz factor of the boost is  $\gamma = 1/\sqrt{1 - \beta^2}$  with  $\beta = (u_1 - u_2)/c$ . The transformation for a crossing in downstream direction (from upstream to downstream) then reads

$$\Lambda_d \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} \overset{\gamma \sim 1}{\approx} \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} \quad (3.42)$$

and the inverse transformation holds for travelling in upstream direction

$$\Lambda_u \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} \overset{\gamma \sim 1}{\approx} \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} E/c \\ p_{\parallel} \end{pmatrix} \quad (3.43)$$

The energy gains of the particle can now be determined by applying both of those transformations and incorporating the fact that the particle's momentum is scattered back in the direction of the shock front between crossings. For a particle travelling downstream this results in

$$\begin{pmatrix} E_d/c \\ p_{\parallel,d} \end{pmatrix} = \Lambda_d \begin{pmatrix} E_u/c \\ p_{\parallel,u} \end{pmatrix} = \begin{pmatrix} E_u/c + \beta p_{\parallel} \\ \beta E_u/c + p_{\parallel} \end{pmatrix} \quad (3.44)$$

where downstream quantities are denoted by an index  $d$  and upstream quantities by an index  $u$ . Therefore the energy gain ratio of the particle crossing  $u \rightarrow d$  is

$$\epsilon_{u \rightarrow d} = \frac{\Delta E_{u \rightarrow d}}{E} = \frac{c\beta p_{\parallel}}{cp} = \beta\mu \quad (3.45)$$

This value has to be averaged over all possible pitch angles  $1 > \mu > -u_u/v \approx -u_u/c \approx 0$  assuming highly relativistic particles with  $v \sim c$  and a non-relativistic bulk velocity. The probability for a particle with some  $\mu$  crossing the shock without being scattered on its way is inversely proportional to the time  $t$  spent in the scattering medium if pitch-angle scattering happens with some collision frequency  $\nu_c$ :

$$P_c \propto \frac{1}{t} = \frac{v + u_u/\mu}{d} = \frac{v\mu + u_u}{d_{\parallel}} \propto v\mu + u_u \quad (3.46)$$

where  $d$  is the distance from the particle's position to the shock in the direction of its motion. The last proportionality holds because particles are equally likely to be at any distance from the shock. The average energy gain ratio weighted by this probability then reads, again assuming  $v \sim c$  and neglecting terms of order  $\beta^2$  and higher:

$$\bar{\epsilon}_{u \rightarrow d} = \int_0^1 d\mu P_c(\mu) \epsilon_{u \rightarrow d} / \int_0^1 d\mu P_c(\mu) = \int_0^1 d\mu (v\mu + u_u) \epsilon_{u \rightarrow d} / \int_0^1 d\mu (v\mu + u_u) \approx \frac{2}{3}\beta. \quad (3.47)$$

Similar calculations also hold for the traversal back to the upstream medium after the particle has been scattered in pitch angle to a value  $-1 < \mu \lesssim 0$ . Here the fundamental assumption is made that the particle distribution is always isotropic on both sides of the shock so that every pitch angle is equally likely. The pitch angle scattering coefficient  $D_{\mu\mu}$  therefore has to be the most significant effect for the diffusion approximation to apply. Using the inverse transformation  $\Lambda_u$  leads similarly to  $\epsilon_{d \rightarrow u} = -\beta\mu$  only with the opposite sign. Integrating over  $\mu$  over the inverse integration range leads to the same result of  $\bar{\epsilon}_{d \rightarrow u} = 2/3\beta$  and therefore the total energy gain ratio per cycle is

$$\epsilon_{u \rightarrow u} = \frac{E'_u - E_u}{E_u} = \frac{E_u(1 + \epsilon_{u \rightarrow d})(1 + \epsilon_{d \rightarrow u}) - E_u}{E_u} \approx \epsilon_{u \rightarrow d} + \epsilon_{d \rightarrow u} \quad (3.48)$$

up to the first order in  $\beta$ . The  $\mu$ -integrated energy gain ratios therefore also sum up and one arrives at

$$\epsilon_{u \rightarrow u} = \frac{4}{3}\beta \quad (3.49)$$

for a full cycle of the particle. The energy gain is of first order which gave this mechanism the commonly known name *first-order Fermi acceleration*.

After the particle has crossed the shock  $N$  full cycles its momentum is

$$p(N) = p_0 (1 + \bar{\epsilon}_{u \rightarrow u})^N \quad (3.50)$$

but there is a nonzero probability  $P_e$  that the particle escapes the acceleration region. This probability is given by the ratio of particle flux in upstream and downstream direction for particles in the downstream medium (Jones, 1994). The downstream particle flux is

$$\mathcal{F}_{\rightarrow d} = \int_{-u_d/v}^1 d\mu P_c(\mu) / \int_{-1}^1 d\mu = \frac{(u_d + v)^2}{4v} \quad (3.51)$$

and the upstream particle flux is

$$\mathcal{F}_{\rightarrow u} = \int_{-1}^{-u_d/v} d\mu P_c(\mu) / \int_{-1}^1 d\mu = \frac{(u_d - v)^2}{4v}. \quad (3.52)$$

The escape probability for one cycle is therefore up to first order in  $u_d/v$  for fast accelerated particles

$$P_e = 1 - \frac{\mathcal{F}_{\rightarrow u}}{\mathcal{F}_{\rightarrow d}} = 4 \frac{u_d}{v} \approx 4 \frac{u_d}{c} . \quad (3.53)$$

Since the per-cycle momentum gain ratio as well as the per-cycle escape probability is independent from the particle momentum and therefore the same in each cycle a power-law spectrum of particles arises. To show this one is interested in finding an expression for the probability  $f$  of finding a particle after  $N$  cycles depending on the momentum gain after those  $N$  cycles. Such an expression is found when considering the following (Jones, 1994):

$$q := -\frac{\log P(N)}{\log(p(N)/p_0)} = -\frac{\log(1 - P_e)^N}{\log(1 + \bar{\epsilon}_{u \rightarrow u})^N} \approx \frac{P_e}{\bar{\epsilon}_{u \rightarrow u}} = -3 \frac{u_d}{u_u - u_d} \quad (3.54)$$

which is a constant in momentum. The count of cycles  $N$  cancels in this equation because the relative gains and losses are the same for each cycle. The probability of finding a particle with momentum  $< p$  is therefore given by

$$P(p) = \left( \frac{p}{p_0} \right)^q \quad (3.55)$$

where  $q$  can also be given in terms of the compression ratio  $r = u_u/u_d$ :

$$q = -\frac{3}{r - 1} \quad (3.56)$$

To stay consistent with the use of the probability density one can write

$$f(p) = \frac{dP}{dp} = \left( \frac{p}{p_0} \right)^s \quad (3.57)$$

with a modified power-law index of

$$s := q - 1 = -\frac{r + 2}{r - 1} . \quad (3.58)$$

The same results can also be derived by solving the Fokker-Planck equation for pitch-angle diffusion only at two sides of the shock (for example Bell, 1978; Kirk, 1994). Since the energy gain of the first-order Fermi process is proportional to  $u/c$  and the acceleration timescale  $\tau_a$  is short it is expected to be efficient. The acceleration timescale is given by (Bednarz and Ostrowski, 1996)

$$\tau_a = \frac{3}{u_u - u_d} \left( \frac{\kappa_u}{u_u} + \frac{\kappa_d}{u_d} \right) . \quad (3.59)$$

Therefore is a popular candidate for the mechanism of high-energy acceleration in a jet. This process is believed to be very efficient in converting a significant portion of the energy of the plasma into the cosmic rays (Drury, Markiewicz and Voelk, 1989). The accelerated cosmic rays are also able to induce plasma instabilities and therefore plasma waves which in turn scatter the cosmic rays themselves and increase the probability of re-crossing the shock front to gain more energy (Bykov et al., 2013). This self-consistent creation of turbulence has successfully been simulated using particle-in-cell simulations (Crumley et al., 2019).

The diffusion approximation used in the derivation of diffusive shock acceleration only holds if the plasma bulk speed is non-relativistic, *i.e.* if  $u \ll c$ . In contrast observations of AGN jets make it seem likely that highly relativistic bulk motion of the plasma is common (Biretta, Zhou and Owen, 1995) which leads to some difficulties in describing the acceleration mechanisms. In the relativistic case the pitch-angle scattering cannot necessarily isotropize the particle distribution on

short enough timescales that the distribution can be assumed to always stay near the equilibrium (Kirk, 1994). Therefore and because the particle's velocity is now comparable to the bulk velocity particles may not be able to cross the shock multiple times which is essential for diffusive shock acceleration. On top of that oblique and highly relativistic shocks are in many cases superluminal in the sense that the magnetic field lines travel along the shock front with velocities  $> c$  (Begelman and Kirk, 1990). Since unscattered particle trajectories follow the magnetic field lines strong scattering across magnetic field lines must be prevalent for diffusive shock acceleration to work in this environment (Kirk, 1994). On the other hand Achterberg, Gallant, et al. (2001) show that for ultrarelativistic shocks these differences lead to a slight steepening in the resulting power laws and Fraix-Burnet (1990) show that in the oblique case Fermi acceleration works as long as the orientation of the magnetic field fulfills certain conditions.

Another problem with diffusive shock acceleration is that it is only effective for highly non-thermal particles and the source of these is unknown. If the gyroradii of the particles is small, the shock transition region is not infinitely small on the scales the particle travels and it cannot gain energy in transitioning it (Balogh and Treumann, 2013). This is known as the *injection problem* and many ideas for the injection of suprathermal particles have been proposed. Many find their origin in the interaction with plasma waves and instabilities (Matthews, Bell and Blundell, 2020) but magnetic reconnection events where magnetic field lines reconnect are also discussed as injection sources or as a direct acceleration mechanism Kagan et al., 2015. Another candidate for highly relativistic flows is shock drift acceleration which is able to accelerate electrons to significant non-thermal energies (Marcowith et al., 2016).

### 3.2.3. Second-order Fermi acceleration

Second-order Fermi acceleration, also called stochastic acceleration, occurs in turbulent plasma conditions when particles undergo scattering in every direction. It was proposed by Fermi (1949) who assumed randomly moving “magnetic mirrors” with a typical velocity  $u$  scattering the particles. Today the random mirrors are usually assumed to be plasma (*i.e.* magnetohydrodynamic) waves. Each scattering event increases or decreases the particle's momentum but head-on collisions increasing momentum are more probable so there is a net gain in momentum. Following Longair (2011) one finds that the energy gain of a particle with velocity  $v$  in one collision is to second order in  $\beta = u/c$

$$\frac{\Delta E}{E} = 2\beta \cos \theta \frac{v}{c} + 2\beta^2 \quad (3.60)$$

with the incident angle  $\theta$ . Assuming relativistic particles, *i.e.*  $v \rightarrow c$  the probability of a collision is proportional to  $1 + \beta \cos \theta$  and averaging over  $\theta$  in similar fashion as in equation (3.47) then gives

$$\left\langle \frac{\Delta E}{E} \right\rangle = 2\beta + 2\beta \int_0^\pi d \cos \theta (1 + \beta \cos \theta) \cos \theta \Big/ \int_0^\pi d \cos \theta (1 + \beta \cos \theta) = \frac{8}{3}\beta^2 . \quad (3.61)$$

This process is happening slower compared to diffusive shock acceleration and is described by the momentum diffusion coefficient  $\kappa_{pp}$  in equation (3.39). With the mean free path between collision  $L$  the acceleration timescale  $\tau_a$  can be defined as

$$\tau_a = \frac{3}{4} \frac{L}{c\beta^2} \quad (3.62)$$

and with the escape timescale  $\tau_e$  the equilibrium state of the mechanism has an energy distribution (Blandford and Eichler, 1987)

$$\frac{\partial N}{\partial E} \propto E^{-1-\tau_a/\tau_e} . \quad (3.63)$$

To explain the relatively narrow range of observed power-law indices with second-order Fermi acceleration it, unrealistically, has to be assumed that very similar plasma conditions are prevalent

over a broad range of sources. Therefore it is unlikely that it is the only reason for particle acceleration. The stochastic nature of second order Fermi acceleration can be described in terms of a diffusion process in momentum space by incorporating a momentum diffusion term into the Fokker-Planck equation (3.39) by (Blandford and Eichler, 1987):

$$a_2 = \frac{\langle V^2 \rangle}{3vL} \quad (3.64)$$

While this mechanism is relatively inefficient in general and works efficient only under strong assumptions it is proposed as “pre-heating” mechanism for particles in the downstream region of a shock (Petrosian, 2012). After they gained enough energy to cross the shock front they can participate in more efficient shock acceleration mechanisms, which may solve the injection problem.

## 4. Stochastic differential equations

To solve the diffusion-convection equation (3.39) numerical methods are widely used. Since the equation is of Fokker-Planck type its solution is apart from describing the physical particle density also the probability density of some random process being the solution of a stochastic differential equation. This equivalency can be exploited and the stochastic differential equation can be solved which is a relatively straightforward task and has numerical advantages. In this section some basics of probability theory are established that will lead to the formal equivalence of the stochastic differential equation's solution and the solution of Fokker-Planck-type equations.

Every time a random experiment is executed exactly one of several outcomes  $\omega$  is produced. The set of possible outcomes of a random experiment is called the sample space  $\Omega = \{\omega\}$ . Probably the most famous example of randomness is the roll of a dice. In this case, the sample space consists of the six possible numbers that the dice can show:  $\Omega = \{\omega_1, \dots, \omega_6\}$ . If we run the experiment  $N$  times and track the frequency of each outcome  $N_\omega$  we can define a probability  $p_\omega = N_\omega/N$  in the limit of infinitely many repetitions, *i.e.*  $N \rightarrow \infty$ . This interpretation of probability is called the *frequentist* interpretation since it is based on the concept of frequency. It stands in contrast to the *Bayesian* interpretation of probability which relates the concepts of *information* or *knowledge* to the probability assigned to an event. To define stochastic differential equations the first interpretation is the most sensible and will be used.

A *probability space* is a triplet  $(\Omega, \mathcal{A}, P)$  with the sample space  $\Omega$  that has already been defined.  $\mathcal{A}$  is the set of *events*, an event being in itself a set of outcomes. Obviously the elements  $A_i$  of  $\mathcal{A}$  are subsets of the sample space  $\Omega$ .  $\Omega$  itself and the empty set  $\emptyset$  can be events where the first always occurs and the latter never. The last item in the triplet is the *probability function*  $P(\omega) : \Omega \rightarrow [0, 1]$ ,  $P(\omega) = p_\omega$  which assigns each possible outcome the respective probability as introduced above. In the following section random variables and stochastic processes are introduced, leading to the definition of stochastic differential equations and the Itô integral.

### 4.1. Random variables

To connect possible observable outcomes  $\omega \in \Omega$  to quantities we define the *random variable*  $X : \Omega \rightarrow \mathbb{R}$  assigning a real value to every  $\omega$ . The probability for a random variable  $X$  taking a value in some Borel subset  $B$  of  $\mathbb{R}$  is  $P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\})$  with the event  $\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ . A Borel subset is any interval of the form  $(a, b)$ ,  $(-\infty, a)$  or  $(a, \infty)$  with  $a, b \in \mathbb{R}$ . Any complement, union or intersection of a finite or countable number of such intervals is also a Borel subset of  $\mathbb{R}$ . This definition gives rise to a cumulative probability density  $F_X(a) = P_X((-\infty, a])$  being a function  $F_X : \mathbb{R} \rightarrow \mathbb{R}$ . The random variable and its probability function  $P_X$  form a new probability space  $(\mathbb{R}, \mathcal{B}, P_X)$  with  $\Omega = \mathbb{R}$  and the set of events being the  $\sigma$ -algebra  $\mathcal{B} = \{B\}$ .

It is instructive to differentiate between two types of random variables. A discrete random variable is a random variable which takes a discrete set of values  $x_i \in \mathbb{R}$ . For those a probability  $P_X(X = x_i) = p(x_i) = p_i$  can be defined since for every  $x_i$  a Borel subset  $B \in \mathcal{B}$  containing only  $x_i$  can be found. This probability may be nonzero for some or all  $x_i$ . In contrast for a continuous random variable the probability hitting *exactly*  $x_i$  is  $P_X(X = x_i) = 0$  for every  $x_i$  since there are infinitely many possible outcomes for  $X$ . Despite the probability being zero it is not in principle impossible to hit a specified value  $x_i$  exactly so we say such an event happens *almost never*. In spite of this for continuous random variables it is possible to define a *probability density function*

(p.d.f.) which is defined in terms of the differential of  $F_X$  as  $p(x)dx = dF_X(x)$ . To retrieve the probability of finding the random variable  $X$  inside the Borel subset  $B$  we integrate the probability density:

$$P_X(B) = \int_B dx p(x) \quad (4.1)$$

### Moments

Many interesting properties of random variables can be quantified by using the notion of moments. Moments can be understood as an abstraction of the concept of an average of a function. The  $n$ -th moment, centered around  $c$  of a function  $f(x)$  is given by

$$\mu_n = \int dx (x - c)^n f(x) \quad (4.2)$$

If not stated otherwise,  $c = 0$  is assumed. In physics moments are often used to describe properties of densities (for example mass or charge density) in space, *i.e.*  $\rho(r), r \in \mathbb{R}^3$ . For example, the zero-th moment of the mass density is obviously the total mass, the first moment of the mass density is proportional to the center of mass and the second moment is the rotational inertia (when rotating around the center point  $c$ ). Similarly for a charge density, the first moment is the electric dipole moment. These examples make it clear that moments carry information about the *shape* of a distribution.

This concept can also be readily applied to a probability density function  $p(x)$  of a random variable  $X$ . In this case, the first moment directly gives the *expectation value*  $E(X) = \langle X \rangle$  of the random variable which can be interpreted as the value that we expect to see after evaluating the random variable for many ( $N \rightarrow \infty$ ) outcomes and averaging the random variable's values. The second moment, centered around the expectation value ( $c = E(X)$ ), gives the *variance*  $\text{Var}(X)$  of the random variable. It is a measure for how much, on average, the values of  $X$  deviate from the expectation value. Higher moments (also centered around the expectation value) are called *skewness* and *kurtosis* but will not be further explored here.

## 4.2. Stochastic processes

A stochastic process  $\{X_t\}$  is commonly known as a set of random variables  $X_t$  indexed by a variable  $t \in T$  all defined on the same probability space  $(\Omega, \mathcal{A}, P)$ . Most of the times  $t$  is interpreted as a time variable so that  $X_t$  represents the time evolution of a quantity. Depending on the choice of  $T$  we differentiate *discrete time* and *continuous time* stochastic processes. The stochastic process can be written as a function  $X(t, \omega) : T \times \Omega \rightarrow \mathbb{R}$  where different  $X(\omega)$  are called different realisations or sample paths of the stochastic process. The joint distributions of the stochastic process  $F_{t_0, \dots, t_n}(x_0, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$  give the probability of a subset of random variables  $X_{t_0} \dots X_{t_n}$  at increasing time instants  $t_0 < t_1 < \dots < t_n$  each evaluating to the corresponding  $x_0, \dots, x_n$ , *i.e.* a particular sample path. When analysing stochastic processes it is useful to look at the conditional probability

$$P(X_n = x_n | X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}) = \frac{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)}{P(X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1})} \quad (4.3)$$

denoting the probability for  $X_n = x_n$  if the random variables  $X_1 \dots X_{n-1}$  take the given values  $x_1, \dots, x_{n-1}$ . The expression  $X_n$  is a shorthand form for  $X_{t_n}$  which is used from now on if the set of time instants  $\{t_n\}$  under consideration is clear. In terms of events the probability  $P(X_n = x_n)$  is given by

$$P(X_n = x_n) = P(\{\omega \in \Omega : X_n(\omega) = x_n\}) \quad (4.4)$$

and the probability for a sample path until the time  $t$  is therefore

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P\left(\bigcup_{i=1}^n \{\omega \in \Omega : X_i(\omega) = x_i\}\right) . \quad (4.5)$$

This notation makes it possible to omit mentioning the probability space on which the random variables are defined but only talk about their joint distributions or probabilities.

Apart from their joint distribution and their probabilities it is also reasonable to talk about expectation values and variance of a stochastic process. These quantities depend on time and are directly written as

$$\mu(t) = \langle X_t \rangle \quad \text{and} \quad \sigma(t)^2 = \text{Var}(t) = \text{Var}(X_t) . \quad (4.6)$$

The time dimension makes it possible to introduce the covariance which is a measure of joint variability or correlation between the values at different time instants  $t_0, t_1$ :

$$\text{Cov}(t_0, t_1) = \langle (X_{t_0} - \langle X_{t_0} \rangle)(X_{t_1} - \langle X_{t_1} \rangle) \rangle \quad (4.7)$$

A basic example of a stochastic process is the random walk of a particle observed for example in the case of Brownian motion where the random variables  $X_t$  represents the position of the tracked particle at time  $t$ . Since for Brownian motion we assume a symmetric model in the sense that there are no preferred directions for the randomly walking particles the expectation value  $\mu(t) = \langle X_t \rangle = 0$ . Intuitively it is also clear that the longer particle propagates the more do we expect to find it randomly at locations further away from its starting point. This is manifested in the variance of the stochastic process increasing with increasing time difference between two observations:  $\text{Var}(t) = t$  or more generally  $\text{Var}(X_{t_1} - X_{t_0}) = t_1 - t_0$ . For Brownian motion the increments  $X_{t_1} - X_{t_0}$  do not depend on the history of the particles trajectory and are distributed according to a Gaussian distribution. The stochastic process fulfilling those requirements is called *Wiener process* and will be of central importance later on.

### 4.3. Markov chains

An important subset of stochastic processes are the Markov processes, also called Markov chains for reasons that will become clear soon. A Markov process is a process in which the probability for  $X_n = x_n$  only depends on the value of the random variable  $X_{n-1} = x_{n-1}$  and not on random variables at time instants  $t_m$ ,  $m < n - 1$ . In conditional probabilities this can be written as

$$P(X_n = x_n | X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1}) . \quad (4.8)$$

This property is called *Markov property* or *memoryless* and allows talking about the process as a “chain” by imagining the subsequent random variables forming a chain of states (*i.e.* their values  $x_n$ ). Similar to a chain in which each link is only attached to exactly one other link in each direction the values of the random variables only depend on their precursor and only influence their successor. Keeping up the metaphor each realisation (sample path) of the stochastic corresponds to one particular chain. The entire stochastic process can be represented in the form of a directed graph with weighted edges, each node corresponding to a value  $x \in \mathbb{R}$ . The edges’ weights then the conditional probabilities defined above with the weight of the edge from node  $x_i$  to node  $x_j$  being  $P(X_{n+1} = x_j | X_n = x_i)$ . This probability is in general time dependent which is expressed by the free parameter  $n$ .

Interpreting the nodes as states of a system we can call those conditional probabilities the *transition probabilities* of the Markov process. One can then define a *transition matrix* with the elements

$$p^{ij}(n) = P(X_{n+1} = x_j | X_n = x_i) \quad (4.9)$$

which gives the probability for transitioning from  $x_i$  to  $x_j$  at the time  $t_n$ . In the same manner a probability vector for each random variable  $X_n$  with components  $p^i$  can be defined with each element being the probability of the random variable taking the value  $x_i$ . Transitioning to the next state can then be written as

$$p^i(n+1) = \sum_j p^{ij}(n)p^j(n) \quad (4.10)$$

which is a simple matrix multiplication.

These ideas can only be applied to discrete-time and discrete-state Markov processes. Generalising them for continuous-time while preserving the discrete state space takes some additional effort. The transition probabilities now depend on two time instants  $t_0, t_1$  and give the probability for finding  $X_{t_1}$  in some state  $x_1$  after having found  $X_{t_0}$  in some (other) state  $x_0$ :

$$P(X_1 = x_1 | X_0 = x_0) \quad (4.11)$$

The Markov property is fulfilled when this probability does not depend on values of the stochastic process for times  $t < t_0$  which is implied in the formulation. Similar to the definitions above a transition matrix  $p^{ij}(t_0, t_1) = P(X_1 = x_j | X_0 = x_i)$  and a probability vector  $p^i(t) = P(X_t = x_i)$  can be defined. The propagation of state from time  $t_0$  to  $t_1$  is then written as

$$p^i(t_1) = \sum_j p^{ij}(t_0, t_1)p^j(t_0) \quad (4.12)$$

#### 4.4. Diffusion processes

To include processes whose states are continuous, *i.e.*  $x \in \mathbb{R}$ , we need to modify the definition of the transition probability and introduce a density similar to the case of continuous random variables before. The probability of finding the stochastic process at time  $t_n$  in the Borel subset  $B$  (that is in any closed or open interval that is  $\subset \mathbb{R}$ ) for a given sample path  $X_i = x_i$ ,  $t_i < t_n$  is

$$P(X_n \in B | X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = \frac{\int_B dx_n p(t_1, x_i; \dots; t_n, x_n)}{\int_{\mathbb{R}} dx_n p(t_1, x_i; \dots; t_n, x_n)} \quad (4.13)$$

in terms of the probability density  $p$  for a general stochastic process. For Markov processes this shortens to  $P(X_n \in B | X_{n-1} = x_{n-1})$  and the corresponding density is defined by integration:

$$P(X_n \in B | X_{n-1} = x_{n-1}) = \int_B dx_n p(t_{n-1}, x_{n-1}; t_n, x_n) \quad (4.14)$$

This transition density takes the place of the transition matrix in the case of discrete-state Markov process. A diffusion process is an important special case of Markov processes which has no instantaneous jumps and is therefore continuous in some sense and can be assigned a *drift coefficient*  $a(t, x)$  and a *diffusion coefficient*  $b(t, x)$ . The coefficients can be interpreted as “local” moments of the probability distribution in the limit of small time increments  $\delta t$  reading

$$a(t, x) = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int_{y \in B_\epsilon(x)} dy (y - x) p(t, x; t + \delta t, y) \quad , \quad (4.15)$$

$$b(t, x)^2 = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \int_{y \in B_\epsilon(x)} dy (y - x)^2 p(t, x; t + \delta t, y) \quad . \quad (4.16)$$

The drift coefficient therefore represents the infinitesimal change in the mean value of the stochastic process over time and the diffusion represents the infinitesimal change in variance. The condition of having not instantaneous jumps is written as

$$\lim_{\delta t \rightarrow 0} \int_{|y-x| > \epsilon} dy p(t, x; t + \delta t, y) = 0 \quad \forall \epsilon \in \mathbb{R} \quad (4.17)$$

Again the most popular example of a diffusion process is the Wiener process for which the transition density  $p(s, x; t, y)$  is a normal distribution centered around  $x$  with the variance being  $t - s$ . It fulfills the conditions mentioned and its probability density  $p(s, x; t, y)$  fulfills a set of differential equations containing the drift and diffusion coefficient. These were introduced by Kolmogorov (1931) from the viewpoint of probability theory and named after him.

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y} (a(t, y)p) + \frac{1}{2} \frac{\partial^2}{\partial y^2} (b(t, y)^2 p) \quad (\text{Forward equation}) \quad (4.18)$$

$$\frac{\partial p}{\partial s} = -a(s, x) \frac{\partial p}{\partial x} + \frac{1}{2} b(s, x)^2 \frac{\partial^2 p}{\partial x^2} \quad (\text{Backward equation}) \quad (4.19)$$

The forward equation gives the forward time evolution of  $p(t, y)$  for a known set of initial conditions  $(s, x)$  and the backward equation gives the backward time evolution of  $p(s, x)$  for a known set of final state  $(t, y)$ . The forward equation was found before by Fokker (1914) and is therefore also known as Fokker-Planck equation after him and Planck (1917). In other fields it is also called Smoluchowski equation.

The Wiener process will be the central process in the following analysis of stochastic differential equations and their applications. While other stochastic processes can also take the place of the Wiener process in the following chapters, this work limits itself to the Wiener process most of the time.

## 4.5. Itô calculus and stochastic differential equations

Going back at the practical example of Brownian motion it makes intuitive sense to write the stochastic process as differential  $dX_t = dW_t$  where  $W_t$  is a Wiener process. This differential does not exist in a strict sense since the Wiener process is continuous but nowhere differentiable so we need to find a useful extension to the notion of a differential and an integral respectively. To motivate stochastic integrals we could ask the question of how the test particle will propagate if we modify the Brownian process such that the average “intensity” of the random motion depends on the position of the particle rather than being a constant. Calling this intensity  $b(x, t)$  this problem can be intuitively formulated in terms of an integral

$$X_t = \int_0^t b(x, t) dW_t \quad (4.20)$$

For a constant  $b$  the sensible solution and therefore the stochastic process representing the test particle’s position is  $X_t = bW_t$ . To generalize the integral for all possible  $b(\omega, t)$  a limiting procedure similar to the procedure used in standard Riemann integrals is used (Kloeden and Platen, 1992). The function  $b^{(n)}(\omega, t)$  is a discretized variant of  $b$  which is a step function for a  $n$  time steps between the integration boundaries which are set to  $[0, t]$  without losing generality, *i.e.*  $b^{(n)}(\omega, t) = b(\omega, \tau_i)$  with  $\tau_i$  being some time instant between  $t_i \leq \tau_i < t_{i+1}$ . This step function as integrand lets us split the integral into  $n$  parts with constant integrand

$$X_t = \int_0^t b^{(n)}(\omega, t) dW_t = \sum_i^n \int_{t_i}^{t_{i+1}} b(\omega, \tau_i) dW_t = \sum_i^n b(\omega, \tau_i) (W_{t_{i+1}} - W_{t_i}) \quad (4.21)$$

Now taking the limit  $n \rightarrow \infty$  gives us the so-called *Itô stochastic integral* which can be seen as a generalisation of Riemann-Stieltjes integral with integrands and integrators being stochastic processes instead of deterministic functions. It has been introduced by Itô (1944). We saw that in general  $b$  can also be a stochastic process itself which is why  $b(\omega, t)$  was written in the above equation. The integral inherits the stochastic properties of the stochastic processes involved. This means the integral itself is again a stochastic process living on the same probability space as these processes.

A central example is the Itô process which is the solution to the stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad (4.22)$$

Apart from the stochastic term  $b(X_t)dW_t$  the evolution of  $X_t$  also consists of a deterministic term  $a(X_t)dt$  which,  $a$  and  $b$  being functions  $\mathbb{R} \rightarrow \mathbb{R}$  can be evaluated as Riemann integral. It can be shown (Kloeden and Platen, 1992) that solutions of this equation are a diffusion process itself if  $a$  and  $b$  are sufficiently well-behaved. Therefore they fulfill the Kolmogorov equations (4.18) and (4.19) with exactly the  $a$  and  $b$  in the differential equation (4.22). The concepts introduced in the preceding sections can be extended to vector spaces with  $n$  dimensions. We then have a multidimensional Itô process which is essentially a set of  $n$  coupled Itô processes

$$d\vec{X}_t = \vec{A}(t, \vec{X}_t)dt + B(t, \vec{X}_t)d\vec{W}_t \quad (4.23)$$

or written in terms of indices using Einstein summation convention

$$dX_{i,t} = A_i(t, \vec{X}_t)dt + B_{ij}(t, \vec{X}_t)dW_{j,t} \quad (4.24)$$

where  $\vec{X}_t$  is now a vector-valued stochastic process. This can also be written in integral form

$$X_{i,t} = X_{i,0} + \int_{t_0}^t dt A_i(t, \vec{X}_t) + \int_{t_0}^t dW_{j,t} B_{ij}(t, \vec{X}_t) \quad (4.25)$$

where the first integral is an ordinary integral and the second is an Itô integral. The drift coefficient  $\vec{A}$  then has to be a vector-valued function and the diffusion coefficient  $B$  is a  $n \times n$ -matrix valued function. The Wiener process also has to be replaced by a multidimensional Wiener process, that is a set of  $n$  independent Wiener processes forming a vector. The Kolmogorov forward equation, referred to as Fokker-Planck equation from now on, then reads

$$\frac{\partial p(t, x)}{\partial t} = -\frac{\partial}{\partial x_i}(A_i(t, x)p(t, x)) - \frac{\partial^2}{\partial x_i \partial x_k}(B_{ij}(t, x)B_{jk}(t, x)p(t, x)) \quad (4.26)$$

for a chosen initial condition  $p(t_0, x_0) = p_0$ , again summing over every index occurring twice. Of course, the inverse statement holds too: if a function  $p$  solves a Fokker-Planck equation with coefficients  $a$  and  $b^2$  (or their vector and matrix valued counterparts) it is the density of an Itô process if the coefficients are sufficiently smooth and bounded. In general more than one stochastic process can be found for a given Fokker-Planck equation because it is only determined up to  $b^2$  (or  $B^T B$  in the matrix case) which does not necessarily imply a fixed choice of  $b$  or  $B$ . This equivalency between a Fokker-Planck equation given in terms of its drift and diffusion coefficients and a Itô process stochastic differential equation is of central importance in this work.

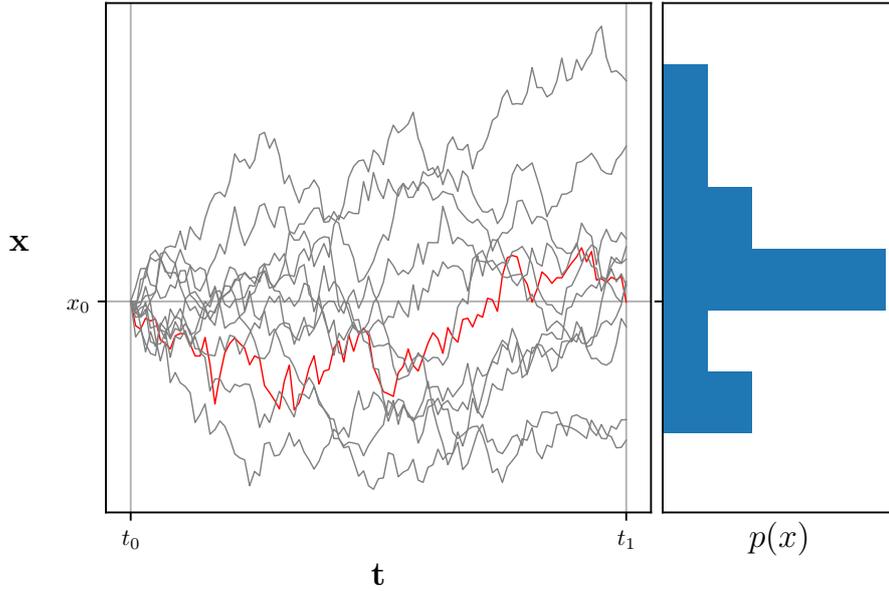
## 4.6. Sampling the stochastic process

Fokker-Planck equations originating from physical models can be solved by writing down an equivalent stochastic differential equation and estimating its probability distribution numerically. To that end we sample the random process and create a histogram of the outputs. Suppose one bin is delimited by  $[x_0, x_1]$ . The expected amount of samples  $N$  in this bin is then given in terms of the probability density  $p(x)$  by

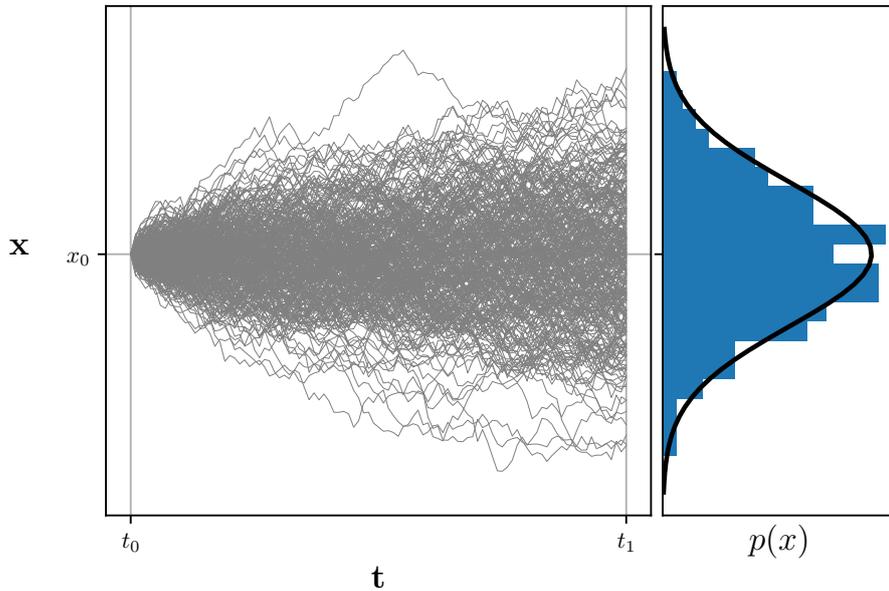
$$N([x_0, x_1]) = N_0 \int_{x_0}^{x_1} p(x)dx \quad (4.27)$$

where  $N_0$  is the total number of samples taken. If we assume that the chosen bins are sufficiently small, we can say that  $p(x)$  is approximately constant in the interval of the bin. Calling the center point of the bin  $x = (x_0 + x_1)/2$  we can write

$$\int_{x_0}^{x_1} p(x)dx = p(x) \int_{x_0}^{x_1} dx = p(x)(x_1 - x_0) = \frac{N([x_0, x_1])}{N_0} \quad (4.28)$$



a) A small number of realisations of the stochastic process that solves the given stochastic differential equation (4.31), one highlighted in red. All of those start at the value given by the initial condition and propagate until the final time  $t_1$ . The probability density estimated by counting the number of sample paths ending in the respective bins is shown in the right box.



b) The same procedure but showing 300 sample paths. For comparison the expected normal distribution is shown. It is clear that for an increasing number of sample paths considered the estimated distribution converges to this distribution.

**Figure 4.1.:** Schematic depiction of the process of sampling the solution of a stochastic differential equation which itself is a stochastic process and estimating the probability density.

This leads to

$$p(x) = \frac{N(x)}{N_0 \Delta x} =: H(x) \quad (4.29)$$

with the bin width  $\Delta x = x_1 - x_0$ . In other words if we normalize the histogram  $N(x)$  with the total sample count  $N_0$  and the bin width  $\Delta x$ , the histogram's values approximate the probability density  $p(x)$ . This of course also holds if the bin sizes are not constant which is the case for logarithmic axes scalings. The approximation is better when  $\Delta x$  is smaller since the actual variations of  $p(x)$  over the course of a bin have lesser effect. Smaller bins on the other side lead to fewer samples in each bin which in turn leads to larger (relative) statistical errors so a reasonable bin size has to be used.

This approach to solving a Fokker-Planck equation numerically using this scheme is depicted in Figure 4.1a for the very basic example equation

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} \quad (4.30)$$

which is simple Wiener diffusion with a constant diffusion coefficient  $b^2 = 1$ . The underlying stochastic differential equation is then simply given by  $dX_t = dW_t$ . To estimate the density we are integrating the stochastic integral

$$X_{t_1}(\omega) = X_{t_0} + \int_{t_0}^{t_1} dW_t(\omega), \quad X_{t_0} = x_0 \quad (4.31)$$

for numerous random sample paths  $\omega$  (the details of how exactly this integration is taking place are discussed in the following chapters) and then estimate the density  $p(t = t_1, x)$  at some later point in time  $t_1$  by binning the  $x$ -dimension and counting how many sample paths are in each bin. In Figure 4.1a a small number of sample paths are depicted with the resulting estimation for the probability density. Naturally using only few sample paths this estimate is not very useful but as shown in Figure 4.1b sampling some more realisations gives a good estimate of the expected density. Writing this procedure down we get

$$\tilde{p}(t_1, x) = \frac{1}{2\Delta x} \int_{x-\Delta x}^{x+\Delta x} p(t_1, x) = \frac{1}{2\Delta x} \frac{\text{Number of sample paths with } x - \Delta x \leq X_1(\omega) < x + \Delta x}{\text{Number of sample paths}} \quad (4.32)$$

and the result is in the form of a density histogram as discussed above. The fraction on the rightmost side of this equation directly corresponds to the fraction in equation (4.13) while replacing the probability density there with actual number of events observed, similar to the most basic interpretation of probability.

Since the stochastic process  $X_t$  is a function  $X(t)$  (for each realisation) and in physical terms its values  $x$  usually resemble a quantity like location or momentum it seems natural to call every sample path a “particle” with a trajectory, propagating as the stochastic process evolves. This notion is also used by Strauss and Effenberger (2017) who call it a *pseudo particle* since in spite of similar properties it is not directly related to a physical, existing particle. In the further course we shall adopt this terminology.

So far it has been implied that the time integration is done in forward direction. This means that pseudo particles are released by some source at some time and then propagate and are “detected” (their values are read and analysed) at a later time. Forward integration is useful if the state of the particles is relevant and will be evaluated in a large part (or the whole of) phase space. In contrast the integration of the stochastic differential equation can also be performed backwards in time instead of, as discussed, forwards in time. In this case we start at  $t_1$  and work our way back to  $t_0$ . The probability density is then no more described by the Kolmogorov forward equation (4.26) but by the Kolmogorov backward equation, which reads

$$\frac{\partial p(t, x)}{\partial t} = -A_i(t, x) \frac{\partial p(t, x)}{\partial x_i} - B_{ij}(t, x) B_{jk}(t, x) \frac{\partial^2 p(t, x)}{\partial x_i \partial x_k} \quad (4.33)$$

and is the adjoint equation of the forward equation (Kopp et al., 2012). This description is useful in situations where the detection of particles takes place in a small part of phase space so that after pseudo particles propagated (in the forward picture) most of them will not arrive at the detector and are therefore thrown away. To save computation time the roles of detection and particle source are reversed in the time backwards picture. Pseudo particles start at the detector and then propagate backwards in time to eventually reach the source. The computational effort decreases of course only if the source region (*i.e.* the effective detection region in the time-backwards case) is larger than the detection region. This is the case for example in the modelling of cosmic ray propagation of cosmic rays originating from the sun to Earth, where the Earth is smaller than the emitting region. In the context of this work we deal only with highly localized sources (*i.e.* particles originating at a single point in phase space) but are interested in their propagation regardless from their final position in phase space the time-forward integration is the only sensible variant.

## 5. Numerical solution of the transport equation

Starting from the transport equation (3.39) which incorporates spatial diffusion ( $\kappa$ ), convection with the bulk velocity  $u$ , momentum diffusion ( $a_1$ ,  $a_2$ ), losses ( $\dot{p}$ ) and particle sources or sinks ( $S$ )

$$\begin{aligned} \frac{\partial F}{\partial t} = & -u \frac{\partial F}{\partial z} + \frac{\partial}{\partial z} \left( \kappa \frac{\partial F}{\partial z} \right) - \frac{1}{p^2} \frac{\partial (p^2 a_1)}{\partial p} \frac{\partial F}{\partial z} \\ & + \left( \frac{p}{3} \frac{du}{dz} + \frac{\partial a_1}{\partial z} \right) \frac{\partial F}{\partial p} + \frac{1}{p^2} \frac{\partial}{\partial p} \left( p^2 a_2 \frac{\partial F}{\partial p} \right) + \frac{1}{p^2} \frac{\partial}{\partial p} (p^2 \dot{p} F) \\ & + S \end{aligned} \quad (5.1)$$

we are looking for a set of stochastic differential equations describing the same system. It is convenient to use the shock rest frame as inertial system so that the shock front itself is always at  $z = z_0 = 0$  and the bulk velocity is given relative to the shock front's movement in the jet. The source term is given in terms of a typical spatial number density  $n_0$  as

$$S = \frac{c\beta_0 n_0}{4\pi p_{\text{inj}}^2} \delta(z - z_0) \delta(p - p_{\text{inj}}) \quad (5.2)$$

representing continuous injection of particles at the shock front  $z_0$  with the injection momentum  $p_{\text{inj}}$ . First it is necessary to transform into a system of ‘‘code units’’ in which all quantities are dimensionless. Following Krülls and Achterberg (1994) we define the code units as

$$\tau = \frac{c^2 \beta_0^2}{\kappa_0} t, \quad x = \frac{c\beta_0}{\kappa_0} z, \quad y = \frac{p}{p_{\text{inj}}}, \quad (5.3)$$

write the bulk velocity in terms of the light speed  $\beta = u/c$  and normalize the coefficients with typical values  $\beta_0$  and  $\kappa_0$ :

$$\bar{\beta} = \frac{\beta}{\beta_0}, \quad \bar{\kappa} = \frac{\kappa}{\kappa_0}, \quad \bar{a}_1 = \frac{a_1}{c\beta_0 p_{\text{inj}} y}, \quad \bar{a}_2 = \frac{\kappa_0 a_2}{c^2 \beta_0^2 p_{\text{inj}}^2 y^2}. \quad (5.4)$$

The momentum losses written in terms of the dimensionless momentum  $y$  read

$$\dot{y} = \frac{\kappa_0}{c^2 \beta_0^2 p_{\text{inj}}} \dot{p} \quad (5.5)$$

and finally the distribution and the source term must be written in dimensionless form as

$$\bar{F} = \frac{4\pi p_{\text{inj}}^3}{n_0} F, \quad \bar{S} = \delta(x - x_0) \delta(y - 1). \quad (5.6)$$

Plugging those relationships into equation (5.1) the resulting equation is not substantially different and reads

$$\begin{aligned} \frac{\partial \bar{F}}{\partial t} = & -\bar{\beta} \frac{\partial \bar{F}}{\partial x} + \frac{\partial}{\partial x} \left( \bar{\kappa} \frac{\partial \bar{F}}{\partial x} \right) - \frac{1}{y^2} \frac{\partial (y^2 \bar{a}_1)}{\partial y} \frac{\partial \bar{F}}{\partial x} \\ & + \left( \frac{y}{3} \frac{du}{dx} + \frac{\partial \bar{a}_1}{\partial x} \right) \frac{\partial \bar{F}}{\partial y} + \frac{1}{y^2} \frac{\partial}{\partial y} \left( y^2 \bar{a}_2 \frac{\partial \bar{F}}{\partial y} \right) + \frac{1}{y^2} \frac{\partial}{\partial y} (y^2 \dot{y} \bar{F}) \\ & + \delta(x - x_0) \delta(y - 1). \end{aligned} \quad (5.7)$$

In order to write this equation in terms of a system of stochastic differential equations it needs to be expressed in conservation law form. With the differential number density  $N = y^2 \bar{F}$  this reads:

$$\begin{aligned} \frac{\partial N}{\partial \tau} = & - \frac{\partial}{\partial x} \left[ \left( \bar{\beta} + \frac{\partial \bar{\kappa}}{\partial x} + y \frac{\partial \bar{a}_1}{\partial y} + 3\bar{a}_1 \right) N \right] + \frac{\partial^2}{\partial x^2} (\bar{\kappa} N) \\ & - \frac{\partial}{\partial y} \left[ \left( -\frac{y}{3} \frac{d\bar{\beta}}{dx} - y \frac{\partial \bar{a}_1}{\partial x} + y^2 \frac{\partial \bar{a}_2}{\partial y} + 4\bar{a}_2 y - \dot{y} y^2 \right) N \right] + \frac{\partial^2}{\partial y^2} (\bar{a}_2 y^2 N) \\ & + \delta(x - x_0) \delta(y - 1) . \end{aligned} \quad (5.8)$$

From this equation it is straightforward to extract the coefficients of the corresponding stochastic differential equation. The multi-dimensional coupled Itô process shown in equation (4.23) reads in two dimensions  $x, y$

$$d \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} A_0(x, y) \\ A_1(x, y) \end{pmatrix} d\tau + \begin{pmatrix} B_{00}(x, y) & B_{01}(x, y) \\ B_{10}(x, y) & B_{11}(x, y) \end{pmatrix} d\vec{W} \quad (5.9)$$

which incorporates the drift vector  $\vec{A}$  and the diffusion matrix  $B$  and where  $d\vec{W}$  is a two-dimensional Wiener process. These coefficients can be taken from equation (5.8) and we can directly see that the diffusion matrix  $B$  is diagonal:

$$A_0(x, y) = \bar{\beta} + \frac{\partial \bar{\kappa}}{\partial x} + y \frac{\partial \bar{a}_1}{\partial y} + 3\bar{a}_1 , \quad (5.10)$$

$$A_1(x, y) = -\frac{y}{3} \frac{d\bar{\beta}}{dx} - y \frac{\partial \bar{a}_1}{\partial x} + y^2 \frac{\partial \bar{a}_2}{\partial y} + 4\bar{a}_2 y - \dot{y} y^2 , \quad (5.11)$$

$$B(x, y) = \begin{pmatrix} \sqrt{\bar{\kappa}} & 0 \\ 0 & y\sqrt{\bar{a}_2} \end{pmatrix} . \quad (5.12)$$

This system of two generally coupled stochastic differential equations is equivalent to equation (5.8) as shown in chapter 4 and one can attempt to find numerical solutions to it to approximate the transport equation. The details of this are outlined underneath. In general all coefficients ( $\bar{\beta}$ ,  $\bar{\kappa}$ ,  $\bar{a}_{1,2}$ ,  $\dot{y}$ ) can depend on both the spatial or momentum variable, which was omitted in the preceding equations for readability.

## 5.1. Explicit Euler scheme

The integrals of the stochastic differential equation have to be approximated numerically which can be realized with a variety of numerical schemes based on discretizing the integration variable. They can be categorized into explicit and implicit schemes where explicit schemes calculate the next step's value only using the state at the current step while implicit schemes are given in terms of systems of equations that contain both values for the current step and the next step. In the case of implicit methods the system of equations has to be (numerically) solved for the next step's variables which requires additional computation time. Every scheme also has an order since they are essentially truncations of a Taylor series. Therefore schemes with an order higher than 1 generally contain derivatives of the coefficients of the differential equation whose evaluation also increases the computational effort since they require additional evaluations of the coefficient functions.

The constant source term is realized in the simulation by giving each pseudo particles a different start time  $t_0$  and letting them all propagate until they reach the final time  $T$  at which their state is evaluated. They all start at the injection location and momentum  $x = 0$ ,  $y = 1$ . This setup

corresponds to the continuous injection of particles as it is assumed in our model because of the system's symmetry under time translations.

The simplest numerical scheme is the *Cauchy-Euler-Scheme* also known as *Euler-Mayurama-Scheme* which is a first-order explicit scheme. Discretizing the time by introducing some time step  $\Delta\tau$  leads to the solution of the integral equation as

$$\vec{X}_n = X_n + \vec{A}(t_n, \vec{X}_n)\Delta\tau + \mathbf{B}(t_n, \vec{X}_n)\Delta\vec{W}_n . \quad (5.13)$$

The increments of the Wiener process are normally distributed with a variance of  $\Delta\tau$  so they can be written as

$$\Delta W = \sqrt{\Delta\tau} \mathcal{G} \quad (5.14)$$

for a standard-normally distributed variable  $\mathcal{G}$  (which is with variance 1 and mean 0) (Strauss and Effenberger, 2017). This random number must be generated for each increment of  $\vec{X}$  and for each component of  $\Delta W$  independently and identically distributed which is achieved by using a separate random number generator seeded differently for each component. For this work the random number generators of the *PCG* family of generators is used (O'Neill, 2014). As described above a set of stochastically independent pseudo-particles are simulated and their positions in phase space after some time  $T$  can be analysed using histograms as discussed in the previous section. For this we will use a 1+1-dimensional model with  $x$  as spatial dimension and  $p$  as relativistic momentum.

The shock itself is in principle an infinitely thin gradient between the upstream and downstream velocity  $\beta_u$  and  $\beta_d$ . Since the pseudo-particles can only gain momentum if they experience the gradient at some point of their propagation, and since they propagate in discrete steps, an infinitely thin gradient would always be missed by the pseudo particles. Therefore a smoothed form of the shock profile which has a finite width and gradient has to be used to observe acceleration. The shock profile is approximated by a hyperbolic tangent function which extends the gradient to the shock width  $X_{\text{sh}}$  (Krülls and Achterberg, 1994):

$$\bar{\beta} = a - b \tanh \frac{x}{X_{\text{sh}}} . \quad (5.15)$$

The derivative of the profile is required in the drift coefficient and reads

$$\frac{d\bar{\beta}(x)}{dx} = \frac{-b}{X_{\text{sh}} \cosh^2(x/X_{\text{sh}})} . \quad (5.16)$$

The coefficients  $a$  and  $b$  are given in terms of the compression ratio  $r = \beta_u/\beta_d$  which gives the ratio of the bulk velocity upstream and downstream of the shock:

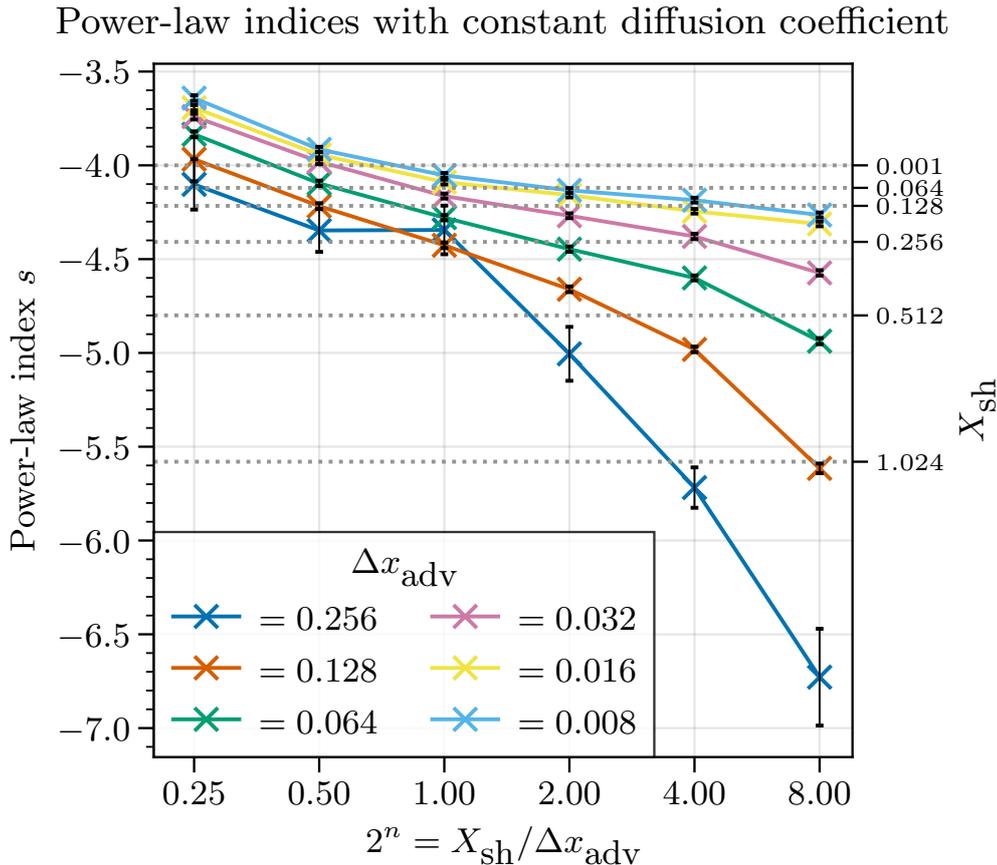
$$a = \frac{1}{2}\beta_u\left(1 + \frac{1}{r}\right), \quad b = \frac{1}{2}\beta_u \frac{r^2 - 1}{r^2 + r} = \frac{1}{2}\beta_u \frac{r - 1}{r} . \quad (5.17)$$

Those two parameters are the physical parameters governing the bulk velocity profile. Zhang (2000) showed that the need for an analytically smoothed shock profile can be mitigated by using skew brownian motion which essentially means that the particles "local" time is scaled after crossing the shock. Their algorithm basically counts the number of shock crossings a particle experiences during its trajectory and applies the expected momentum gains after the simulation finished.

The first model considered in this work is pure diffusive shock acceleration which means we neglect momentum diffusion and losses and the coefficients  $a_1 = a_2 = \dot{y} = 0$ . For these assumptions the stochastic differential equations simplify considerably and read

$$A_0(x, y) = \frac{\partial \bar{\kappa}}{\partial x} + \bar{\beta}, \quad A_1(x, y) = -\frac{1}{3} \frac{d\bar{\beta}}{dx} y , \quad (5.18)$$

$$B_{00}(x, y) = \sqrt{\bar{\kappa}}, \quad B_{10} = B_{01} = B_{11} = 0 . \quad (5.19)$$



**Figure 5.1:** Increasing the ratio of the shock width  $X_{\text{sh}}$  to the advective step size  $\Delta x_{\text{adv}}$  leads to steeper power laws since the velocity gradient is less steep and particles can only increase their energy by a smaller amount in crossing it. The compression ratio in these runs is  $r = 2$  and the expected comparison values shown as dashed lines are from Krülls and Achterberg (1994).

For the diffusion coefficient we follow Krülls and Achterberg (1994) and Toptygin (1980) and assume that it is proportional to the squared bulk velocity

$$\bar{\kappa}(x) = q\bar{\beta}(x)^2 \quad (5.20)$$

with the proportionality constant  $q$ . For comparison of the results of this work with previous results this assumption is applied in the following. In future applications, this coefficient is chosen (or prescribed by for example large-scale magnetohydrodynamic simulations) depending on the assumed physical turbulence situation and such a condition will generally not hold.

### 5.1.1. Numerical steps

The choice of the timestep  $\Delta\tau$ , which is the Monte-Carlo parameter, is crucial for efficient as well as accurate simulation of shock acceleration. If the timestep is chosen to be too large in comparison to the shock width  $X_{\text{sh}}$  most of the particles are not able to experience the shock gradient and

therefore will accelerate less than expected. On the other hand diffusive shock acceleration depends on the particle crossing the shock multiple times due to diffusion so either the diffusion coefficient needs to be high enough or the timestep large enough. This also justifies the approximation of the shock gradient as a continuous function since for large enough diffusive length scales the shock appears infinitely thin from the perspective of the particle. These effects reflect the balance between diffusion and advection necessary to observe efficient acceleration that is also a property of the physical (*i.e.* real particles instead of pseudo particles) situation. In terms of equations one needs the advective step size  $\Delta x_{\text{adv}}$  (*i.e.* the typical length a pseudo particle travels during one time step due to advection) to be smaller than the shock width  $X_{\text{sh}}$  and the latter again being smaller than the diffusive step size  $\Delta x_{\text{diff}}$  (*i.e.* the average length of propagation due to diffusion):

$$\Delta x_{\text{adv}} < X_{\text{sh}} < \Delta x_{\text{diff}} . \quad (5.21)$$

Setting the spatially variable diffusion coefficient to a constant value  $\bar{\kappa}$  we can observe the interplay of the advective step size and the shock width in Figure 5.1. Choosing a smaller shock width leads to flatter power laws as one would expect since it resembles a better approximation to the infinitely thin shock. The expected powerlaw index from equation (3.58) is  $s = -4$  for a compression ratio of  $r = 2$  used in this simulation.

In the case of spatial diffusion as given by equation (5.20) the step sizes are directly related to the spatial drift and diffusion coefficients which for pure diffusive shock acceleration means

$$\begin{aligned} \Delta x_{\text{adv}} &= A_0 \Delta \tau = \left( \frac{\partial \bar{\kappa}}{\partial x} + \bar{\beta} \right) \Delta \tau = \left( \frac{\bar{\kappa}}{X_{\text{diff}}} + \bar{\beta} \right) \Delta \tau \\ \Delta x_{\text{diff}} &= (B_{00} + B_{01}) \sqrt{\Delta \tau} = \sqrt{\bar{\kappa} \Delta \tau} \end{aligned} \quad (5.22)$$

where (following Krülls and Achterberg, 1994) the diffusive length scale  $X_{\text{diff}} = 4\Delta x_{\text{diff}}$  is introduced. Since in our model  $\bar{\kappa} = q\bar{\beta}^2$ , we can write

$$\begin{aligned} \Delta x_{\text{adv}} &= \left( \frac{q\bar{\beta}^2}{X_{\text{diff}}} + \bar{\beta} \right) \Delta \tau , \\ \Delta x_{\text{diff}} &= \sqrt{q\bar{\beta}^2 \Delta \tau} . \end{aligned} \quad (5.23)$$

To study the influence of varying numerical step sizes on the momentum power-law indices we solve this system of equations for  $\Delta \tau$  and  $q$  since  $\bar{\beta}$  is in practice a prescribed physical parameter. We can identify  $\Delta x_{\text{diff}}$  in the equation for  $\Delta x_{\text{adv}}$

$$\Delta x_{\text{adv}} = \frac{\Delta x_{\text{diff}}^2}{X_{\text{diff}}} + \bar{\beta} \Delta \tau = \frac{\Delta x_{\text{diff}}^2}{X_{\text{diff}}} + \Delta x_{\text{diff}} \sqrt{\frac{\Delta \tau}{q}} = \Delta x_{\text{diff}} \left( \frac{\Delta x_{\text{diff}}}{X_{\text{diff}}} + \sqrt{\frac{\Delta \tau}{q}} \right) \quad (5.24)$$

and denote the ratio of step sizes with  $\delta$ :

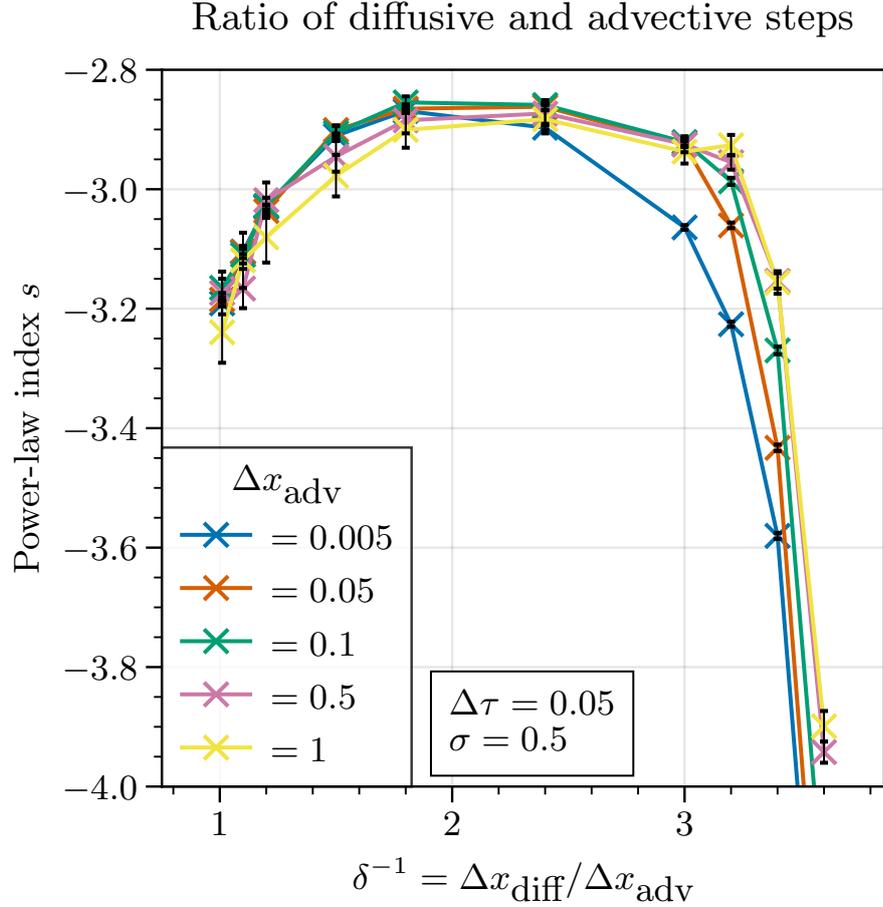
$$\delta := \frac{\Delta x'_{\text{adv}}}{\Delta x_{\text{diff}}} \quad (5.25)$$

which leads to

$$\sqrt{\frac{\Delta \tau}{q}} = \delta - \frac{\Delta x_{\text{diff}}}{X_{\text{diff}}} \equiv \tilde{\delta} . \quad (5.26)$$

Looking at the equation for the diffusive step size in equation (5.23) we find

$$\Delta \tau = \frac{\Delta x_{\text{diff}}^2}{q\bar{\beta}^2} . \quad (5.27)$$



**Figure 5.2.:** Powerlaw indices for different diffusion and advection steps.  $\delta^{-1} = \Delta x_{\text{diff}} / \Delta x_{\text{adv}}$  is the ratio of advective step size to diffusive step size as derived in the text. The remaining parameters are chosen according to the equations in the text. Both  $\bar{\beta}_s$  and  $q$  vary for different datapoints and  $r = 4$  which gives an expected power-law index of  $-2$ . The Cauchy-Euler scheme used in these runs is not able to reproduce this index.

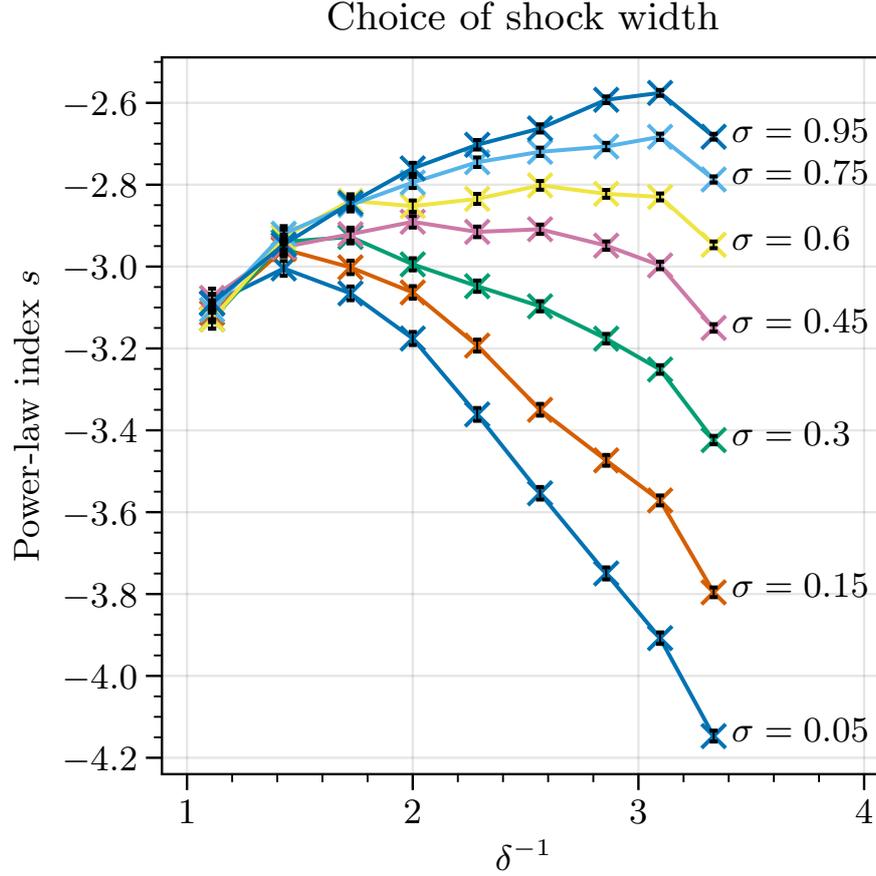
With this we can rewrite equation (5.26)

$$\tilde{\delta} = \frac{\Delta x_{\text{diff}}}{q\bar{\beta}} \quad \text{and} \quad \tilde{\delta} = \frac{\Delta\tau\bar{\beta}}{\Delta x_{\text{diff}}} . \quad (5.28)$$

From these two equations we can read off the desired parameters  $\Delta\tau$  and  $q$ :

$$\begin{aligned} q &= \frac{\Delta x_{\text{diff}}}{\bar{\beta}} \tilde{\delta}^{-1} = \frac{\Delta x_{\text{diff}}}{\bar{\beta}} \left( \delta - \frac{\Delta x_{\text{diff}}}{X_{\text{diff}}} \right)^{-1} , \\ \Delta\tau &= \frac{\Delta x_{\text{diff}}}{\bar{\beta}} \tilde{\delta} = \frac{\Delta x_{\text{diff}}}{\bar{\beta}} \left( \delta - \frac{\Delta x_{\text{diff}}}{X_{\text{diff}}} \right) . \end{aligned} \quad (5.29)$$

Since  $q$  and  $\bar{\beta}$  have to be positive in a physical environment and equation (5.21) must hold, the



**Figure 5.3.:** Influence of different shock widths  $X_{\text{sh}} = (1 - \sigma)\Delta x_{\text{adv}} + \sigma\Delta x_{\text{diff}}$ . The advective step size  $\Delta x_{\text{adv}} = 0.1$  for all runs since Figure 5.2 shows that the choice makes no significant difference. Other parameters are  $\Delta\tau = 0.05$ . Surprisingly, smoother shocks produce flatter power laws if other parameters are held constant.

following limits on  $\delta$  are imposed:

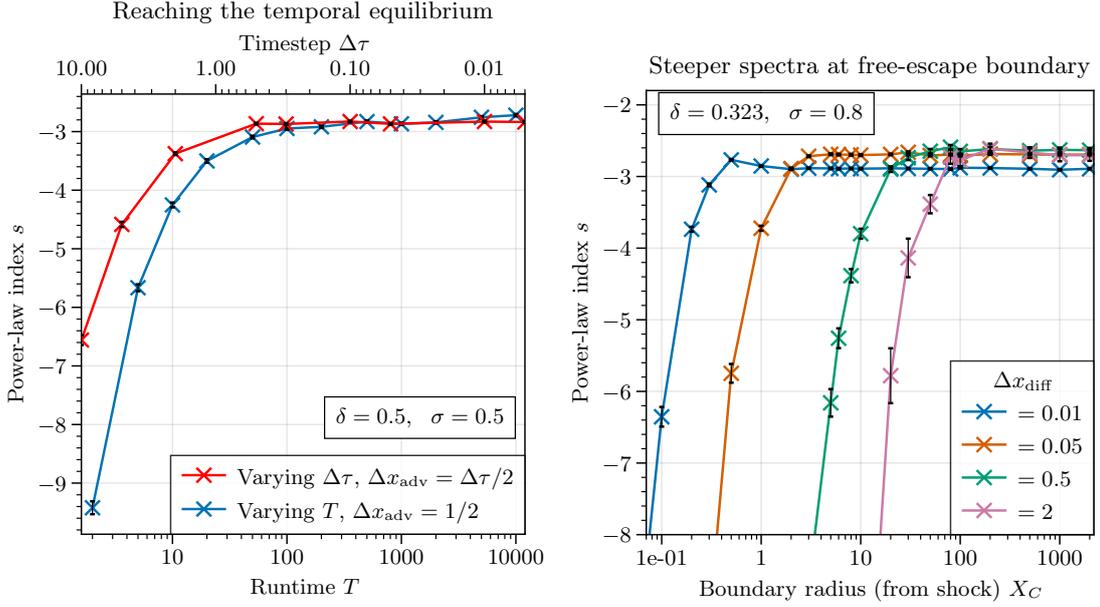
$$\frac{\Delta x_{\text{diff}}}{X_{\text{diff}}} < \delta = \frac{\Delta x_{\text{adv}}}{\Delta x_{\text{diff}}} < 1. \quad (5.30)$$

One can see that this implies the condition that  $\Delta x_{\text{adv}} < \Delta x_{\text{diff}}$ , and  $X_{\text{sh}}$  has to be chosen in between those two step sizes. In the special case of  $X_{\text{diff}} = 4\Delta x_{\text{diff}}$  the ratio  $\Delta x_{\text{diff}}/X_{\text{diff}} = 1/4$ , which imposes a lower limit of  $1/4$  on  $\delta$ . In terms of the inverse of  $\delta$  we have

$$1 < \delta^{-1} = \frac{\Delta x_{\text{diff}}}{\Delta x_{\text{adv}}} < \frac{X_{\text{diff}}}{\Delta x_{\text{diff}}} = 4 \quad (5.31)$$

which gives a range of sensible ratios between the advective and diffusive pseudo particle steps that can be studied. After choosing some value for one of those steps (which one is in principle not relevant) and a value for  $\delta$ , all other parameters can be calculated. Up until now all occurrences of  $\bar{\beta}$  have of course been dependent on  $x$ . Considering that

$$\bar{\beta}_d = \bar{\beta}_u/r < \bar{\beta}(x) < \bar{\beta}_u \quad (5.32)$$



- a) Varying the timestep and the runtime while keeping the other fixed ( $T = 300$  and  $\Delta\tau = 0.1$ ). The diffusion coefficient and bulk velocity are adjusted at each point when varying  $\Delta\tau$  to fulfill the discussed requirements on the step sizes. Below  $\Delta\tau \approx 0.5$  and above  $T \approx 100$ , that is  $\approx 600 - 1000$  single steps, no further flattening of the spectrum is observed.
- b) Particles crossing a boundary at  $x = \pm X_C$  are allowed to escape and are not taken into account for the calculation of the histograms. If the acceleration region is larger than  $X_C \approx 50\Delta x_{diff}$  the effects of a finite acceleration region do not show in the particle spectrum. This scales with the choice of diffusive step size.

**Figure 5.4.:** Influence of different timesteps and different spatial escape boundaries on the powerlaw index. Convergent behavior is observed if enough single time steps are simulated and the free-escape boundary is sufficiently far away from the shock. The finite boundary can be compared with a limited size of the physical acceleration region.

the parameters can be fixed using  $\bar{\beta}_u$  as upper limit. Then, the lower limits of  $q$  and  $\Delta\tau$  are used. Taking a look at the definitions of the step sizes in equation (5.23) at the other end of possible values of  $\bar{\beta}(x)$  it is clear that both  $\Delta x_{diff}$  and the second term of  $\Delta x_{adv}$  deviate at maximum by a factor of  $1/r$  whereas the first term of  $\Delta x_{adv}$  deviates by a factor of  $1/r^2$ . Since  $r > 1$ ,  $\Delta x_{adv}$  is even smaller in relation to  $\Delta x_{diff}$  for  $\bar{\beta}(x) \rightarrow \bar{\beta}_u/r$  which means the necessary constraints are still fulfilled.

This range for  $\delta$  is analysed in Figure 5.2 where the powerlaw index is determined for each set of parameters by using a weighted linear regression model on the histograms produced by the simulation of pseudo particles. The Poisson error of each bin is given by  $\sqrt{N}$  where  $N$  is the number of pseudo particles in this bin. For the logarithm of the bin values that is used for the linear regression the error is given by  $\sqrt{N}/N = 1/\sqrt{N}$  which gives the weights used. The results shown in Figure 5.2 show the expected behavior for different diffusive steps: if the diffusion step is too low, the efficiency of the acceleration decreases, which also happens if it is too high. In this and the following simulations the compression ratio is  $r = 4$  to evaluate a strong shock which gives  $s = -2$  and the physical parameters arise from the chosen numerical step sizes and relations for each datapoint. The results show that there is a broad region of nearly constant and flat powerlaws in the range  $1.8 < \delta^{-1} < 3.0$ . Furthermore the choice of the value of  $\Delta x_{adv}$  is

making little difference in the method's behavior despite being analysed over nearly three orders of magnitude. This is reasonable since all other parameters are adjusted to its value and the choice of  $\Delta x_{\text{adv}}$  is therefore not more than a choice of units.

### 5.1.2. Shock width

If the step sizes are fixed the shock width  $X_{\text{sh}}$  is constrained in the range  $[\Delta x_{\text{adv}}, \Delta x_{\text{diff}}]$  by equation (5.21). Therefore we can define another parameter  $0 \leq \sigma \leq 1$  which maps this range linearly and allows the discussion of the shock width independent from other parameters.

$$X_{\text{sh}} = (1 - \sigma)\Delta x_{\text{adv}} + \sigma\Delta x_{\text{diff}} . \quad (5.33)$$

For  $\sigma = 0$ ,  $X_{\text{sh}}$  is at the low end of the possible range and for  $\sigma = 1$  at the high end. In Figure 5.3 the effect of the shock width on the power-law indices is shown. The inverse value of  $\delta$  is shown to spread the interesting part of this plot further out. If  $\Delta x_{\text{adv}} \approx \Delta x_{\text{diff}}$  the choice of the shock width has few effects on the power-law index but if the step size ratio increases it gets clear that shock widths near the diffusive step size lead to more efficient acceleration. As in Figure 5.2 the spectrum steepens considerably for  $\delta \rightarrow 1/4$ . This holds for all choices of shock width.

### 5.1.3. Asymptotic behavior

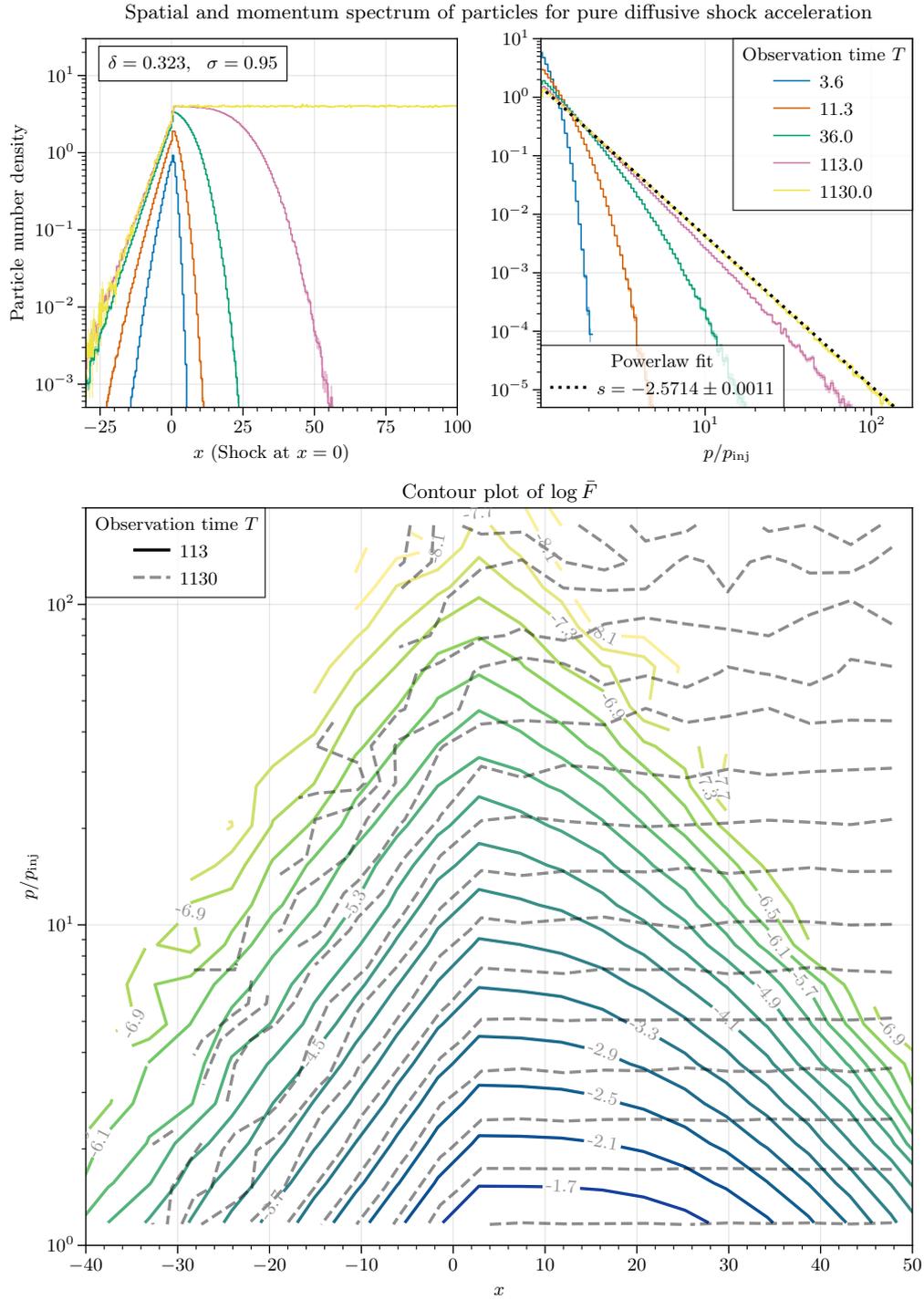
It is left to study how fast the numerical scheme converges depending on the finite timestep and when imposing a boundary condition. Considering the timestep it is clear from Figure 5.4a that the scheme reaches equilibrium for decreasing  $\Delta\tau$  in the same manner as it does for an increasing runtime  $T$  that the particles are allowed to propagate. There is therefore a minimum number of timesteps a pseudo particle has to propagate to allow the system to reach equilibrium which is around 1000 steps.

Up until now an infinite acceleration region has been assumed ( $X_C = \infty$ ). Introducing a spatial boundary at which particles escape the system does not affect the previous findings significantly. When particles reach a position  $|x| > X_C$  they are removed from the simulation which corresponds to the situation where the acceleration region is finite. For boundary radii below  $\approx 50\Delta x_{\text{diff}}$  a significant steepening in the momentum spectrum is observed in Figure 5.4b since particles escape before they can accelerate to higher momenta. As one would expect the point at which the boundary starts taking effect scales with the diffusive length  $\Delta x_{\text{diff}}$ .

The expected power-law index of  $-2$  could not be reproduced using this numerical scheme, contrary to the results of Krülls and Achterberg (1994). The closest reproduction is found for a value of  $\sigma$  close to 1, *i.e.* choosing the shock width  $X_{\text{sh}}$  close to the diffusive step size  $\Delta x_{\text{diff}}$ , and for  $\delta \approx 1/3$ . This can be inferred from Figure 5.3. The electron spectra in both momentum and spatial direction produced with these parameters are shown in Figure 5.5 where also the temporal evolution of the spectrum can be observed. The qualitative structure of the results for the spatial distribution and momentum spectrum closely resembles the results of Krülls and Achterberg (1994). Furthermore contour plots are also shown for the two longest observation times which visualize that particles reach a homogenous momentum distribution downstream after a time long enough, while the distribution upstream is not significantly changing and reaches equilibrium faster. These problems encountered with this method most likely lie in the choice of the numerical scheme used, which is why some alternative schemes are studied in the further course.

### 5.1.4. Momentum diffusion

The same model can also be used to study second-order Fermi acceleration by incorporating momentum diffusion in the transport equation (see section 3.2.3). In a simple model for pure



**Figure 5.5.:** Results from runs with the parameters that lead to the closest reproduction of the expected power-law index of  $-2$ . Parameters are given in the figure and are discussed in the text. The top two histograms are both integrated over the respective dimension not shown. Contours are shown for  $T = 113$  and  $T = 1130$ , the contours for the latter are more noisy because the same amount of pseudoparticles were used in each run of which more are outside of the field of view of the contour plot.

momentum diffusion  $\bar{\beta} = \bar{a}_1 = 0$  and the spatial diffusion coefficient  $\bar{\kappa} = 1$ . Incorporating free-escape boundaries and a constant momentum diffusion leads to the results shown in Figure 5.6.

## 5.2. Higher-order and implicit numerical schemes

The inability to predict the expected power-law index makes it necessary to use higher-order or implicit numerical schemes. Some possibilities are shown in the following and results on improvements are given.

### Predictor-corrector scheme

At first a second order predictor-corrector scheme is analysed as given by Kloeden and Platen (1992) as follows:

$$\vec{X}_{n+1} = \vec{X}_n + \left( \vec{A}(\vec{X}^s) + \vec{A}(\vec{X}_n) \right) \frac{\Delta\tau}{2} + \vec{D} \quad (5.34)$$

using a supporting value

$$\vec{X}^s = \vec{X}_n + \vec{A}(\vec{X}_n)\Delta\tau + \mathbf{B}(\vec{X}_n)\Delta\vec{W} \quad (5.35)$$

and the diffusive step  $D$  which for a diffusion matrix  $\mathbf{B}$  as in pure diffusive shock acceleration (only  $B_{xx} \neq 0$ ) reduces to

$$\begin{aligned} D_x &= \frac{1}{4} \left( B_{xx}(\vec{X}^+) + B_{xx}(\vec{X}^-) + 2B_{xx}(\vec{X}_n) \right) \Delta W_x \\ &\quad + \frac{1}{4} \left( B_{xx}(\vec{X}^+) - B_{xx}(\vec{X}^-) \right) \frac{(\Delta W_x^2 - \Delta\tau)}{\sqrt{\Delta\tau}} \\ D_y &= 0 \end{aligned} \quad (5.36)$$

The supporting value  $X^\pm$  are given by

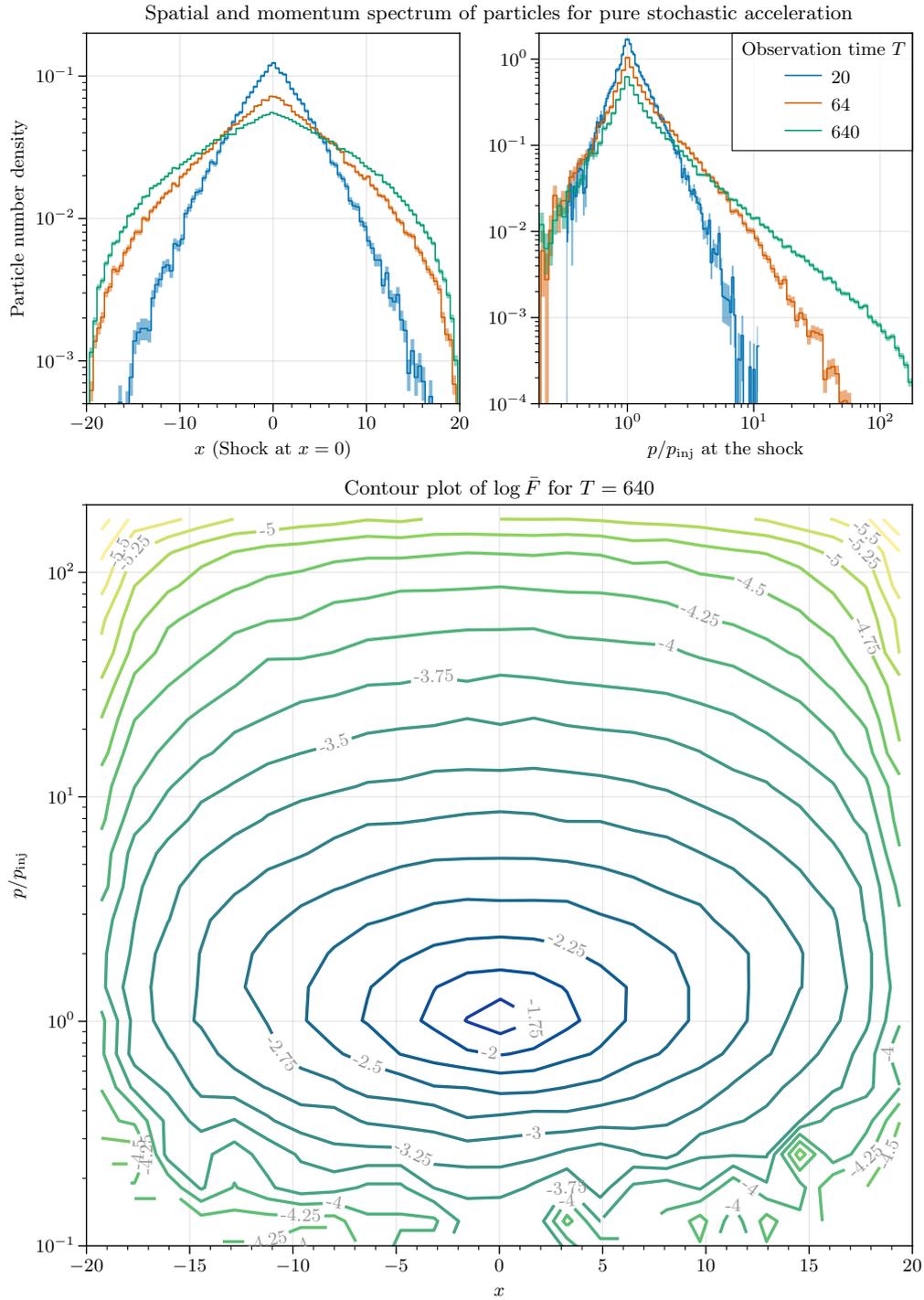
$$\vec{X}^\pm = \vec{X}_n + \vec{A}(\vec{X}_n)\Delta\tau \pm \sqrt{\Delta\tau} \begin{pmatrix} B_{xx}(\vec{X}_n) \\ B_{xy}(\vec{X}_n) \end{pmatrix} \quad (5.37)$$

with only the root mean square diffusion of the stochastic process. This scheme can be extended to a predictor-corrector scheme as used by Achterberg and Schure (2011) and also proposed by Kloeden and Platen (1992). Predictor-corrector schemes have the advantage of using an implicit scheme but without the need to solve a system of equations by predicting the next step's value using explicit scheme. The explicit scheme is therefore the ‘‘predictor’’ and the implicit scheme the ‘‘corrector’’. After generating one vector of random variables  $\Delta\vec{W}$  that is kept fixed throughout the calculation the predictor is calculated with the second-order explicit scheme mentioned above:

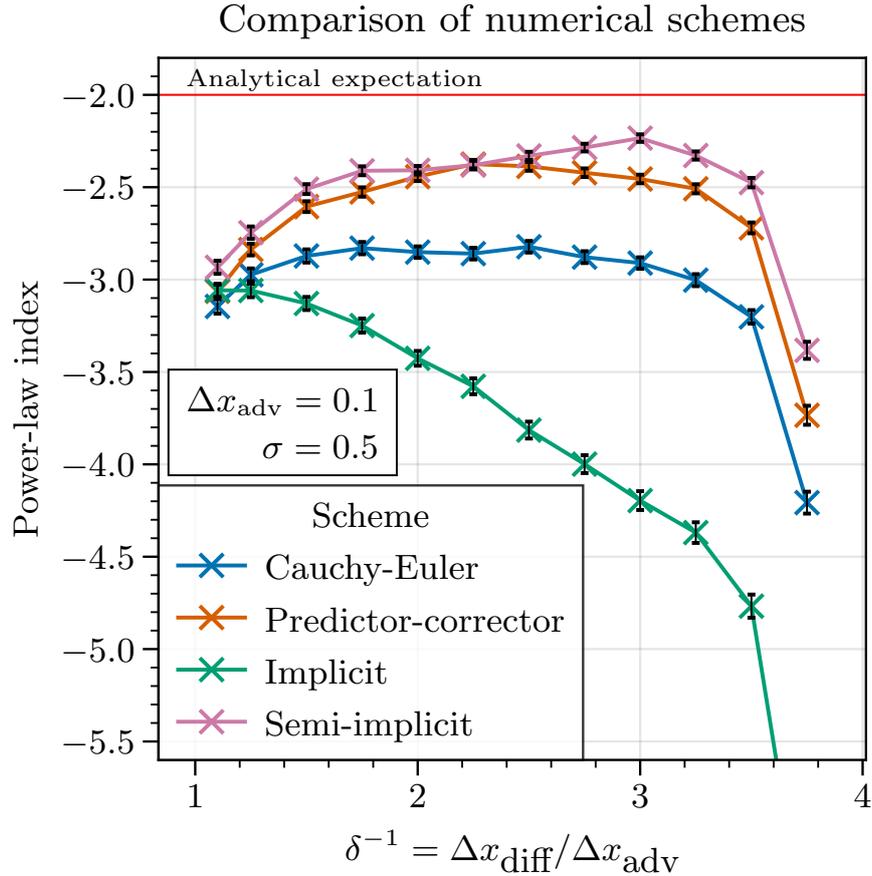
$$\vec{X}_{n+1}^p = \vec{X}_{n+1}^{2\text{nd order}} \quad (5.38)$$

The corrected final result of the scheme is then the average of the predicted value (replacing the implicit dependence) and the initial value:

$$\vec{X}_{n+1} = \vec{X}_n + \frac{\Delta\tau}{2} \left( \vec{A}(\vec{X}_{n+1}^p) + \vec{A}(\vec{X}_n) \right) + \vec{D} \quad (5.39)$$



**Figure 5.6.:** Pure second-order Fermi acceleration with constant momentum diffusion  $\bar{a}_2 = 0.01$  and free-escape boundaries at  $X_C = 20$ .



**Figure 5.7.:** Different numerical schemes are compared in their ability to reproduce flat power laws in the scenario of diffusive shock acceleration. While the implicit scheme performs worse than the standard Cauchy-Euler scheme, the other two are considerable improvements.

### Implicit Euler scheme

The implicit Euler scheme is the implicit counterpart of the explicit Euler scheme mentioned above with the evaluation of the advective step coefficient taking place at the new value of  $\vec{X}$  instead of the old one.

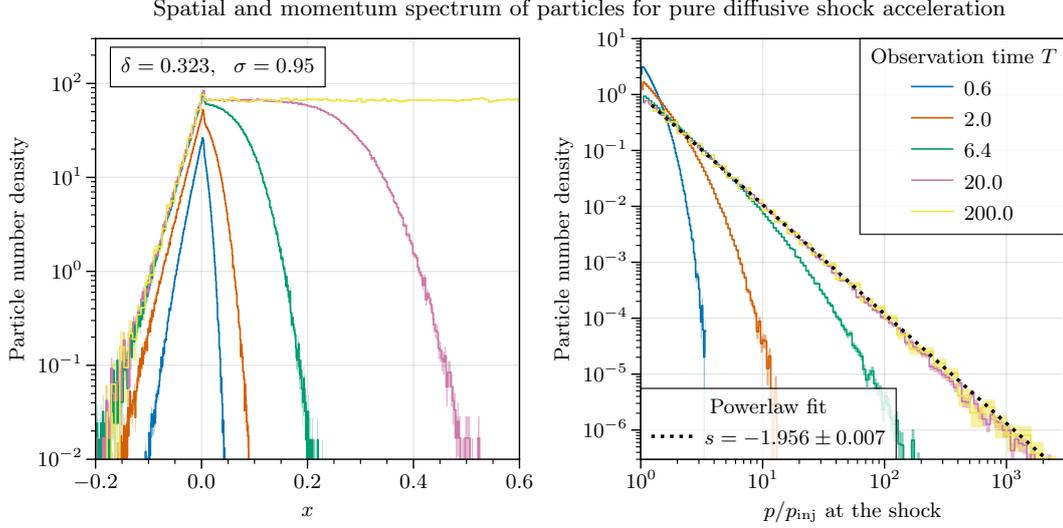
$$\vec{X}_{n+1} = X_n + \vec{A}(t_{n+1}, \vec{X}_{n+1})\Delta\tau + \mathbf{B}(t_n, \vec{X}_n)\Delta\vec{W}_n \quad (5.40)$$

This method needs an algorithm for solving the resulting equation for  $\vec{X}_{n+1}$  for which Broyden's method (Broyden, 1965) as root finding algorithm is used.

### Semi-implicit scheme

A semi-implicit scheme is a scheme that considers the advective step at the current value  $\vec{X}_n$  as well as the next value  $\vec{X}_{n+1}$ . The Euler scheme with a degree of implicitness  $\alpha$  reads

$$\vec{X}_{n+1} = X_n + \left( \alpha \vec{A}(t_{n+1}, \vec{X}_{n+1}) + (1 - \alpha) \vec{A}(t_n, \vec{X}_n) \right) \Delta\tau + \mathbf{B}(t_n, \vec{X}_n) \Delta\vec{W}_n \quad (5.41)$$



**Figure 5.8.:** Calculation of diffusive shock acceleration at five different times  $T$ . The histograms in the left panel are integrated over momentum, and the histograms right panel show the distribution near the shock. Using the semi-implicit second-order scheme the expected power law with index  $-2$  could be matched closely. Corresponding contour plots are shown in Figure 5.9

and the Euler-Mayurama scheme or the implicit Euler scheme are recovered for  $\alpha = 0$  or  $\alpha = 1$  respectively. Gardiner (2009) gives a version of this scheme for a vector-valued stochastic process which in general involves the calculation of a derivative of  $\mathbf{B}$  if no assumptions about  $\mathbf{B}$  can be made. The coefficients

$$C_{ij}(t, \vec{X}) = \sum_k B_{kj}(t, \vec{X}) \frac{\partial B_{ij}(t, \vec{X})}{\partial x_k} \quad (5.42)$$

are zero if the diffusion matrix  $\mathbf{B}$  is independent of all coordinates and can be used to form a vector

$$C_i = \sum_j C_{ij} \quad (5.43)$$

With this vector the new value  $\vec{X}_{n+1}$  is given in terms of two equations

$$\vec{X}_{n+1} = \vec{X}_n + \vec{A}(t^s, \vec{X}^s) \Delta\tau - \frac{1}{2} \vec{C}(t^s, \vec{X}^s) \Delta\tau + \mathbf{B}(t^s, \vec{X}^s) \Delta\vec{W} \quad (5.44)$$

$$\vec{X}_{n+1} = 2\vec{X}^s - \vec{X}_n \quad (5.45)$$

which can be combined into one equation that can be solved for  $\vec{X}^s$  again using Broyden's method as mentioned above

$$0 = 2\vec{X}_n - 2\vec{X}^s + \vec{A}(t^s, \vec{X}^s) \Delta\tau - \frac{1}{2} \vec{C}(t^s, \vec{X}^s) \Delta\tau + \mathbf{B}(t^s, \vec{X}^s) \Delta\vec{W} \quad (5.46)$$

giving the result using equation (5.45).

## Comparison

Results from the different numerical schemes discussed are shown in Figure 5.7 using the same shock acceleration model as before. The implicit scheme gives no improvements compared to the Cauchy-Euler scheme but instead misses the expected power-law index even more. Both the predictor-corrector scheme and the semi-implicit scheme are able to produce power laws that are considerably closer to the expectation, while for the chosen parameters  $-2$  is still not reached. The best scheme is the semi-implicit scheme, whose detailed histograms are presented in Figure 5.8 and corresponding contour plots are depicted in Figure 5.9. They match the analytical solutions of Toptygin (1980) closely.

All schemes of course come with a runtime penalty compared to the Cauchy-Euler scheme. Of the others the predictor-corrector scheme is by far the fastest since it avoids the need to iteratively find the roots of an equation in every step and it requires approximately twice the runtime on the CPU used. The remaining two schemes take  $\approx 11$  (implicit scheme) and  $\approx 21$  (semi-implicit scheme) times as long as the Cauchy-Euler scheme. Especially in the case of the semi-implicit scheme, investing time in the optimization of the code may improve the runtime, since assumptions about the diffusion coefficient (such as its diagonality) makes the possibly time-consuming calculation of parts of the required differentials obsolete.

## 5.3. Technologies used for numerical integration and code structure

The code created for integrating the stochastic differential equations is written in C++ (ISO, 2020) to achieve maximal computational speed with CYTHON bindings (Behnel et al., 2011) so that the integration process can be more conveniently steered from PYTHON (Python Software Foundation, 2022) where the evaluation and presentation of the results also take place. It is found under <https://github.com/pguenth/sde-crt/releases/tag/masterthesis>. The C++-library consists of a central class `PseudoParticleBatch` which is inherited to customize a particular simulation type. This class groups a set of `PseudoParticle` instances which run independently but have a set of common options that are added to the `PseudoParticleBatch`. When constructing a descendant of `PseudoParticleBatch` parameters defining the simulated situation can be set. These are

- breakpoints of type `TrajectoryBreakpoint` to impose conditions which stop the particles from propagating once reached. This is for example used to set the time of measurement and is checked at every iteration,
- boundaries (`TrajectoryBoundary`) that in general can be used to realize transformations which happen if the pseudo particle reaches some phase space point. An example is the implementation of cyclic boundary conditions while free-escape boundaries are implemented using breakpoints. No boundaries in this sense were used in this work,
- numerical schemes that can be implemented by inheriting `SDEScheme`. Schemes available in the library are the ones described in this work and can be found in the file `src/cpp/batch/scheme.h`,
- the stochastic process used. In this work only Wiener processes have been used,
- the start phase space points and times of the pseudo particles and their count,
- callbacks for the drift and diffusion coefficients denoted in vector and matrix form respectively. These are expected to receive a `SpaceTimePoint` which represents a point in time and phase space and return a vector or matrix. Additional (constant) parameters required to calculate



the coefficients can be added through binding of function arguments as supported by the C++ standard library's `std::bind`,

- a callback for retrieving the timestep which can be used for realizing variable timesteps. This also is not used in this work.

In principle the code is able to conduct simulations in arbitrary dimension, the limiting factor being the accuracy of the numerical schemes and their potential limit on dimensions supported. Tracking of the pseudo particle's trajectory through phase space is also possible but of course requires orders of magnitude more memory, so it was switched off for the simulations in this work.

For vector calculations in C++ the EIGEN library is used (Guennebaud, Jacob, et al., 2010). In the process of evaluating the results the popular PYTHON libraries NUMPY (Harris et al., 2020) and SCIPY (Virtanen et al., 2020) are used. The figures in this work are created using PROPLOT (Davis et al., 2021) which is a wrapper for the common MATPLOTLIB library (Hunter, 2007).

## 6. Radiative transfer

Astrophysicists face a central problem few physicists in other disciplines face: the subjects of their studies are located far away from the astrophysicist's scope of influence, in fact so far that humanity will most likely never be able to physically reach most of them. This limits the possibilities for observation dramatically, not only in that taking measurements gets technologically very intensive in relation to the achieved precision but also in that the experimental setup itself can not be designed to the researcher's need as it is the case in most research areas. To study a certain idea, theory or concept, one must first find objects in the universe that fit the chosen parameters and is observable given the technology available; maybe no such object might be available at all. There has been great technological advancement in the types of telescopes, detectors and techniques available, but the fundamental problem stands.

In astronomy and astrophysics observations are in most cases made by detecting electromagnetic radiation here on Earth. To draw conclusions about the mechanisms in astronomical objects that eventually lead to radiation it is crucial to be aware of the relation of the radiation received on Earth and the radiation emitted by the object under scrutiny. It is therefore a vital part of astrophysical modeling to also model the parameters which influence this relationship since they form the base of every statement about distant objects' radiative mechanisms. Many different techniques have evolved to achieve maximum spectral coverage and precision in angular resolution as well as in flux. Other than electromagnetic radiation leptons like myons and neutrinos, protons and other atomic nuclei arriving on Earth from a multitude of sources are also the subject of investigation but will not be further discussed here.

### 6.1. Elements of radiation physics

The fundamental quantity of electromagnetic radiation is the *intensity*  $I$  of radiation. It is defined as the radiated energy  $E$  of radiated photons of a certain photon energy  $\epsilon$  that passes through an area  $A$  and solid angle  $\Omega$  in a given amount of time  $t$ . In differential notation this reads

$$I_\epsilon = \frac{dE}{d\epsilon dA d\Omega dt} . \quad (6.1)$$

When a beam of parallel rays with intensity  $I$  is emitted through the area  $dA$  for some timespan  $dt$  photons are emitted inside a volume  $dV = c dt dA$ . Therefore one can define the spectral and angular energy density  $u$  as

$$u(\epsilon, \Omega) = \frac{I dA dt}{dV} = \frac{dE}{dV d\epsilon d\Omega} \quad (6.2)$$

Radiation propagating through space is generally described by the radiative transfer equation. This equation relates the intensity to absorption and emission regions and reads

$$\frac{dI_\epsilon}{dx} = j(\epsilon) - \kappa_\epsilon I_\epsilon \quad (6.3)$$

where the source of radiation is the emissivity  $j(\epsilon)$  and the absorption is written in terms of the spectral absorption coefficient  $\kappa_\epsilon$ . Introducing the optical depth  $d\tau_\epsilon = \kappa_\epsilon dx$  this equation can also be written as

$$\frac{dI_\epsilon}{d\tau_\epsilon} = S(\epsilon) - I_\epsilon . \quad (6.4)$$

The optical depth is proportional to the amount of radiation that is absorbed by a given region of space. A region with high optical depth is called optically thick and a region with low optical depth is called optically thin. To relate the intensity to observations the spectral energy flux density  $F(\epsilon)$  is defined as

$$F(\epsilon) = \int_{\Omega_s} \hat{n}_r \hat{n}_p I_\epsilon d\Omega \quad (6.5)$$

where we integrate over the emission solid angle and take into account that the direction of from source to observer  $\hat{n}_p$  lies at an angle to the normal of the receiving area  $\hat{n}_r$ . The angle between those two directions is given by  $\hat{n}_p \hat{n}_r = \cos \alpha$ . In astrophysics the quantity  $\epsilon F(\epsilon)$  or  $\nu F(\nu)$  is often used instead of  $F(\epsilon)$  or  $F(\nu)$  which has the advantage that this quantity has the same shape for different popular variable choices (Kulkarni, 2007):

$$\nu F(\nu) = \nu \left| \frac{d\lambda}{d\nu} \right| F(\lambda) = \lambda F(\lambda), \quad \nu F(\nu) = \nu \left| \frac{d\epsilon}{d\nu} \right| F(\epsilon) = \epsilon F(\epsilon) \quad (6.6)$$

### Relativistic effects

We assume a spherical emission region (blob) with a homogenous magnetic field  $B$  and a radius  $r'_b$  in the reference frame of the blob. From now on primed quantities describe quantities measured in the blob reference frame and unprimed quantities describe quantities measured in the observer's frame. The blob is located at a distance  $d_L$  from the observer which corresponds to a redshift  $z$ . It is moving relative to the observer with velocity  $\beta = v/c$  at an angle  $\theta$  measured between the line of sight and the direction of movement in the frame of the observer. This leads to a relativistic Doppler factor

$$\delta_D = \frac{1}{\Gamma(1 - \beta \cos \theta)} \quad (6.7)$$

as shown in equation (2.3), with the bulk Lorentz factor  $\Gamma = (1 - \beta^2)^{-1/2}$ . The relativistic Doppler effect leads to a change in frequency just as the classical Doppler effect for sound does, and to a change in radiation intensity which is often termed *boosting* or *de-boosting*. On cosmological distances and in the presence of strong gravitation the frequency of observed photons decreases additionally. The latter case is usually encountered near black holes, e.g. for radiation originating from the accretion disc, but since jets extend far from the black hole radiation from jets is not categorically influenced by gravity. In contrast, in the first case this is due to the space expanding during the time the photons travel to Earth which gets especially relevant for objects far away which is often the case in AGN. It is usually parametrized by the dimensionless parameter  $z$  given by the ratio of frequencies observed and emitted.

$$z = \frac{\nu'}{\nu} - 1 \quad (6.8)$$

where  $\nu'$  is the emitted and  $\nu$  the observed frequency. It is usually measured by observing characteristic emission or absorption lines of atoms. Accounting for relativistic Doppler effect and cosmological redshift the frequency of a photon in the observer frame is given by

$$\nu = \frac{\delta_D}{(1+z)} \nu' \quad (6.9)$$

The flux measured in the observer frame and therefore usually given in terms of spectral energy distributions (SED) from astronomical observations can be calculated from the total emissivity predicted by the model by Doppler-shifting and considering the observer's distance

$$f_\nu = \frac{\delta_D^4}{4\pi d_L^2} j'(\nu') \quad (6.10)$$

This formula holds for any homogenous source blob emitting isotropically (Dermer and Menon, 2009).

## 6.2. Synchrotron radiation

The results from the calculation of a distribution of charged particles in the jet cannot be compared directly to observations. Instead, only radiation eventually emitted from those particles is observable on Earth. Therefore it is necessary to discuss the processes that make particles lose their energy by emitting photons that we can detect as radiation. While the spectrum of stars is thermal, *i.e.* we observe a typical black-body radiation spectrum, this is not the case for many non-stellar objects in the universe like for example active galactic nuclei, where other emission mechanisms dominate the observations.

The presence of the magnetic field  $B$  causes moving electrons with a number distribution  $N'_e$  to emit synchrotron radiation. The emission from a single electron as seen in the blob reference frame has the form of a pulse (Rybicki and Lightman, 1986), its width being the synchrotron timescale. This timescale can be approximated with (Dermer and Menon, 2009)

$$\Delta t_{\text{syn}} \approx \frac{2}{\gamma'_s} \frac{2\pi}{\omega'_L} (1 - \beta'_e) \approx \frac{2\pi m_e c}{eB\gamma'^2_s} \quad (6.11)$$

using the electron's Lorentz factor  $\beta'_e$  and Larmor frequency  $\omega_L = eB/cm_e$  in Gaussian units. The factor  $1 - \beta'_e \approx 1/\gamma'^2$  accounts for the light travel time between the two positions of the electron at the beginning and end of the pulse, assuming large  $\gamma' \gg 1$ .

From this pulse width we can derive the characteristic (peak) synchrotron frequency  $\nu'^{\text{pk}}_{\text{syn}} = 1/\Delta t_{\text{syn}}$ . Introducing the critical magnetic field  $B_{\text{cr}} = m_e^2 c^3 / e\hbar$  we can write the peak frequency in terms of a dimensionless energy:

$$\epsilon'_{\text{syn}}{}^{\text{pk}} = \frac{\hbar}{m_e c^2} \nu'^{\text{pk}}_{\text{syn}} = \frac{B}{B_{\text{cr}}} \gamma'^2_s = \epsilon_B \gamma'^2_s \quad (6.12)$$

Up to this point only quantities in the blob reference frame have been considered. To compare models with observations we need to transform these quantities to the observers frame. The dimensionless photon energies transform to the observer's reference frame in the same way as the frequency in equation (6.9)

$$\epsilon = \frac{\delta_D}{(1+z)} \epsilon' \quad (6.13)$$

and therefore the electron Lorentz factor  $\gamma'_s$  given in terms of the observed peak photon energy reads

$$\gamma'_s = \sqrt{\frac{\epsilon'}{\epsilon_B}} = \sqrt{\frac{(1+z)\epsilon}{\delta_D \epsilon_B}} \quad (6.14)$$

The synchrotron power, averaged over all pitch angles (*i.e.* angles between the magnetic field and the electron velocity), emitted from a single relativistic electron is given by Rybicki and Lightman (1986) as

$$P' = \frac{4}{3} \sigma_T c \beta_e'^2 \gamma_s'^2 U_B = \frac{4}{3} \sigma_T c \gamma_s'^2 U_B \quad (6.15)$$

since for highly relativistic particles  $\beta_e' \approx 1$ . The magnetic field energy density is given by  $U_B = B^2/8\pi$  in Gaussian units and  $\sigma_T = 8\pi q^4/(3m^2 c^4)$  is the well-known Thomson scattering cross-section. This power can be incorporated into the transport equation as momentum loss  $\dot{p}$  by writing it as

$$\dot{p}' = m_e v \dot{\gamma}' \approx m_e c \dot{\gamma}' = \frac{P'}{c} = \frac{4}{3} \sigma_T \gamma_s'^2 U_B. \quad (6.16)$$

A simple approximation for the spectrum of a set of electrons with different momenta can be derived by assuming that the total synchrotron power emitted by one electron is radiated at the

characteristic synchrotron frequency (equation (6.12)) of this electron, *i.e.*

$$P' = m_e c^2 \dot{\gamma}' \stackrel{!}{=} m_e c^2 \int_0^\infty d\epsilon \epsilon \dot{N}'_{\text{ph}}(\epsilon) = m_e c^2 \int_0^\infty d\epsilon \epsilon \dot{N}'_{\text{ph},0} \underbrace{\delta(\gamma(\epsilon) - \gamma'_s)}_{N_e} \quad (6.17)$$

where  $\dot{N}'_{\text{ph}} = \dot{N}'_{\text{ph},0} N'_e(\gamma'(\epsilon'))$  represents change of the photon number which is proportional to the electron number  $N'_e = \delta(\gamma(\epsilon) - \gamma'_s) = \sqrt{\epsilon_B} \delta(\sqrt{\epsilon} - \sqrt{\epsilon'})$  which is split off. The latter is written in terms of a delta distribution since we are considering the case of only one electron at the characteristic Lorentz factor contributing to the radiation. Using a simple substitution  $u = \sqrt{\epsilon}$  the integral can be evaluated and the equation solved:

$$\int_0^\infty du \ 2u^3 \dot{N}'_{\text{ph},0} \sqrt{\epsilon_B} \delta(u - \sqrt{\epsilon'}) = 2\sqrt{\epsilon'^3 \epsilon_B} \dot{N}'_{\text{ph},0} \quad (6.18)$$

$$\Rightarrow \dot{n}'_{\text{ph}}(\epsilon') = \dot{N}'_{\text{ph},0} n_e(\gamma'_s) = \frac{2}{3} \frac{\sigma_T U_B}{m_e c} \frac{1}{\sqrt{\epsilon' \epsilon_B^3}} n_e(\gamma'_s) \quad (6.19)$$

The last expression is written in terms of photon and electron densities and is generalized for an arbitrary distribution of electrons. This approximation is referred to as delta approximation. From the photon number density we can derive other relevant quantities including the total emissivity, *i.e.* the emissivity integrated over volume:

$$j'^\delta(\epsilon') = \epsilon' J'^\delta(\epsilon') = m_e c^2 \epsilon'^2 V \dot{n}'_{\text{ph}}(\epsilon') = \frac{2}{3} c \sigma_T U_B \gamma'^3 N'_e(\gamma'_s). \quad (6.20)$$

The delta approximation can be used to quickly calculate the emissivity of an arbitrary electron distribution especially if computation time is of concern. An exact expression for the synchrotron emissivity in the blob frame of reference is given by Crusius and Schlickeiser (1986) and Finke, Dermer and Böttcher (2008):

$$j'(\epsilon') = \frac{\sqrt{3} \epsilon' e^3 B}{h} \int_1^\infty d\gamma' N'_e(\gamma') R(x) \quad (6.21)$$

with

$$x = \frac{4\pi \epsilon' m_e^2 c^3}{3e B h \gamma'^2} \quad (6.22)$$

$$R(x) = \frac{x}{2} \int_0^\pi d\theta \sin \theta \int_{x/\sin \theta}^\infty dt K_{5/3}(t) \quad (6.23)$$

and where the function  $K_{5/3}(t)$  is the modified Bessel function of the second kind. The total luminosity (in delta approximation) of the synchrotron emission can be calculated by integrating the emissivity equation (6.20) over all photon energies  $\epsilon'$ .

$$\begin{aligned} L'_{\text{tot}} &= \int_{-\infty}^\infty d\epsilon' J'_{\text{syn}}(\epsilon') = \int_{-\infty}^\infty d\epsilon' \frac{j'_{\text{syn}}(\epsilon')}{\epsilon'} \\ &= \frac{2}{3} c \sigma_T U_B \int_{-\infty}^\infty d\epsilon' \frac{1}{\epsilon_B \gamma_s'^2} \gamma_s'^3 N'_e(\gamma'_s) = \frac{2}{3} c \sigma_T U_B \int_{-\infty}^\infty d\gamma'_s (2\gamma'_s \epsilon_B) \frac{1}{\epsilon_B \gamma_s'^2} \gamma_s'^3 N'_e(\gamma'_s) \\ &= \frac{4}{3} c \sigma_T U_B \int_{-\infty}^\infty d\gamma'_s \gamma_s'^2 N'_e(\gamma'_s) \end{aligned} \quad (6.24)$$

The spectral energy density of the synchrotron photon field  $u'(\epsilon')$  can be written as (Finke, Dermer and Böttcher, 2008)

$$u'(\epsilon') = \epsilon' m_e c^2 n'_{\text{ph}}(\epsilon') = m_e c r'_b \epsilon' \dot{n}'_{\text{ph}}(\epsilon') = \frac{2}{3} \frac{\sigma_T r'_b U_B}{m_e c^2 \sqrt{\epsilon' \epsilon_B^3}} n'_e(\gamma'_s) \quad (6.25)$$

with the synchrotron photon density  $n'_{\text{syn}}$  and the electron density  $n'_e$ . ( $E'_{\text{ph}} = h\nu' = \epsilon' m_e c^2$ ) Integrating over  $\epsilon'$  gives the total energy density

$$u'_{\text{tot}} = m_e c r'_b \int d\epsilon' \epsilon' \dot{n}'_{\text{syn}}(\epsilon') = m_e c r'_b \frac{3}{4\pi r'^3_b m_e c^2} L'_{\text{tot}} = \frac{3}{4\pi r'^2_b c} L'_{\text{tot}} \quad (6.26)$$

where

$$L'_{\text{tot}} = \int d\epsilon' J'(\epsilon') = \int d\epsilon' \epsilon' m_e c^2 \dot{N}'(\epsilon') = \frac{4}{3} \pi r'^3_b m_e c^2 \int d\epsilon' \epsilon' \dot{n}'(\epsilon') \quad (6.27)$$

has been used. The timescale of synchrotron losses can be estimated by

$$\tau_{\text{syn}} = \frac{E'}{\dot{E}'} = \frac{3mc}{4\sigma_T \gamma' U_B} \quad (6.28)$$

using equation (6.15) (Gupta and Basumallick, 2017). Electrons with higher energies lose energy at a faster rate than lower-energy electrons.

### 6.3. Synchrotron self-Compton radiation

One observes radiation that lies above the extrapolated power laws starting in the radio band accounted to synchrotron radiation. These are commonly explained with inverse-Compton scattering which happens when electrons and photons scatter at each other inelastically in such a way that the electron loses part of its energy to the photon. This can of course happen with photons from a variety of sources as has been discussed in section 2.2.3 and in the further course the scattering of an electron population at the synchrotron photons it emits is discussed. The Compton scattering cross section can be given in the Thomson regime if the photon's energy is much smaller than the electron's rest mass:  $h\nu \ll m_e c^2$  which is the case if  $\nu \ll 1.24 \cdot 10^{20} \text{ s}^{-1}$ . Above this photon energy its wavelength is comparable to the de-Broglie wavelength of the electron and quantum effects set in. These can be incorporated by using the Klein-Nishina cross section. An expression for the inverse-Compton emissivity when an electron population  $N'_e(\gamma')$  upscatters a photon field  $u'(\epsilon')$  is given by Finke, Dermer and Böttcher (2008) in the Thomson and Klein-Nishina regime:

$$\epsilon'_s J'_{\text{ssc}}(\epsilon'_s) = \frac{3}{4} c \sigma_T \epsilon'^2 \int_0^\infty d\epsilon' \frac{u'(\epsilon')}{\epsilon'^2} \int_{\gamma'_{\text{min}}}^{\gamma'_{\text{max}}} d\gamma' \frac{N'_e(\gamma')}{\gamma'^2} F_C(\epsilon', \epsilon'_s, \gamma') \quad (6.29)$$

where  $\epsilon'_s$  is the dimensionless energy of the scattered photon,  $\epsilon'$  is the energy of the synchrotron photon and  $\gamma'$  is the Lorentz factor of the electron.  $F_C$  is the scattering kernel which is given by

$$F_C = 2q \ln q + (1 + 2q)(1 - q) + \frac{\Gamma_e^2 q^2 (1 - q)}{2(1 + \Gamma_e q)}, \quad q = \frac{\epsilon'_s / \gamma'}{\Gamma_e (1 - \epsilon'_s / \gamma')}, \quad \Gamma_e = 4\epsilon' \gamma' \quad (6.30)$$

in the range  $1/(4\gamma'^2) \leq q \leq 1$  and zero elsewhere. In the Thomson regime this emissivity can be approximated by a similar delta approximation like above. This leads to

$$j'_{\text{ssc}}(\epsilon'_s) = \frac{2}{3} c \sigma_T u'_{\text{syn}} \gamma'^3_T N'_e(\gamma'_T) \quad (6.31)$$

assuming a monochromatic synchrotron radiation field (Finke, Dermer and Böttcher, 2008). The electron Lorentz factor is given by  $\gamma'_T = \sqrt{\epsilon'_s / \epsilon'}$ . For the ratio of synchrotron luminosity and inverse-Compton luminosity a simple expression can be given (Condon and Ransom, 2015; Tavecchio, Maraschi and Ghisellini, 1998):

$$\frac{L_{\text{syn}}}{L_{\text{ssc}}} = \frac{U_B}{u'_{\text{tot}}} \quad (6.32)$$

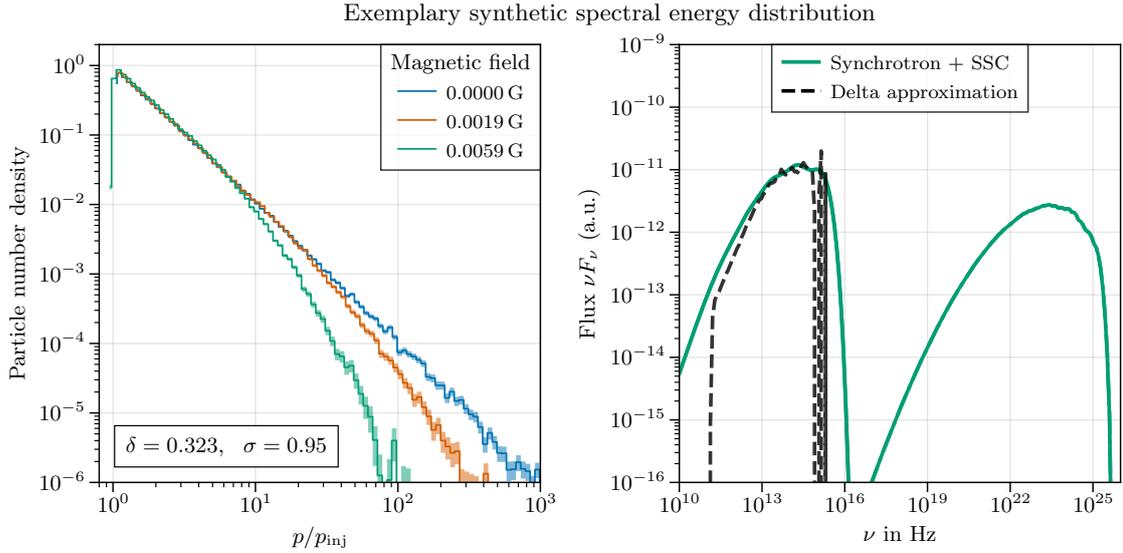
This statement rises by comparing the emissivities for both radiation processes in delta approximation.

## 6.4. Computation of synthetic SEDs

Using the numerical scheme discussed in the previous chapters and incorporating the radiation losses into the propagation of the electrons, synthetic spectral energy distributions can be calculated from the resulting electron spectra. For this we need an expression for  $\dot{y}$  in equation (5.8) which can be derived using  $p \approx \gamma mc$  and equations (5.5) and (6.15):

$$\dot{y} = \frac{\kappa_0}{c^2 \beta_0^2 p_{\text{inj}}} \dot{p} = \frac{\kappa_0}{c^2 \beta_0^2 p_{\text{inj}}} \dot{p} = \frac{4}{3} \frac{\kappa_0 \sigma_T U_B p_{\text{inj}}}{m^2 c^4 \beta_0^2} y^2 \quad (6.33)$$

For the calculation of the observed fluxes the library AGNPY (Nigro et al., 2022) is used which implements the equations shown in the preceding sections. With these results momentum distributions with synchrotron losses and synthetic spectral energy distributions that one would observe are calculated. These are shown in Figure 6.1. As expected, stronger magnetic fields lead to more losses in the electron distribution and losses affect electrons significantly from lower energies on. The spectral energy distribution shown in Figure 6.1 is for illustrative purposes only. To compare it with actual observations, the electron distribution would have to reach to higher momenta which requires much more calculation time. This is a major downside of the Monte-Carlo nature of the method. The particle density is expected to decrease exponentially which leads to strong Poisson noise if not enough pseudo particles are used. Additionally for every particle arriving at a high-energy bin orders of magnitude more particles will end up in lower-energy bins which renders the calculation inefficient.



**Figure 6.1.:** Momentum distribution at the shock with synchrotron losses and proof-of-concept synthetic spectral energy densities produced using the particle distribution for the strongest magnetic field. For better comparison and fitting with observational data, the momentum distribution would have to spread over more orders of magnitude.

## 7. Conclusion

Numerically solving transport equations is a key way of gaining insight of the physical conditions of AGN jets. While the method of solving the cosmic-ray transport equation with stochastic differential equation is in principle feasible, it shows several difficulties. When simulating diffusive shock acceleration the choice of numerical step sizes is crucial for arriving at a physical result. Since these directly depend on the physical parameters, difficulties arise when applying the method in prescribed circumstances. If one would apply the method to solve for a set of given plasma parameters, for example from large-scale magnetohydrodynamic simulations, care has to be taken that physical parameters do not exceed the tight constraints of the SDE method. In this case it may be a solution to use variable numerical parameters, depending on the physical parameters encountered in each time step, an idea also shown in Strauss and Effenberger (2017). Furthermore, the choice of integration scheme also crucial to the success of the method. Another downside is that for reproducing power-law spectra over many orders of magnitude comes with equal increases in runtime since Poisson noise must be suppressed and many pseudo particles have to run. This could eventually be mitigated in future work using particle splitting techniques. Nevertheless, it sensible results can be produced and the technical benefits of using stochastic differential equations can be exploited, for example in combining results from multiple runs with the same parameters or in distributed computing without additional effort.

In the future, it is planned to use the method shown here as a small-scale model for magnetohydrodynamic simulations. Exploiting the time dependence of the model, it may be possible to produce synthetic lightcurves to compare with observations of flares.

# A. Pitch angle- and gyrophase-averaged distribution function

Taking the average in  $\phi$  and  $\mu$  of equation (3.25) leads to

$$\begin{aligned} \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \left[ \frac{\partial F_i}{\partial t} + \frac{\partial F_a}{\partial t} + v\mu \left( \frac{\partial F_i}{\partial z'} + \frac{\partial F_a}{\partial z'} \right) - \epsilon\omega_p \left( \frac{\partial F_i}{\partial \phi} + \frac{\partial F_a}{\partial \phi} \right) - S \right] \\ = \sum_{\alpha,\beta} \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{1}{p^2} \frac{\partial}{\partial x_\alpha} \left[ p^2 D_{\alpha\beta} \left( \frac{\partial F_i}{\partial x_\beta} + \frac{\partial F_a}{\partial x_\beta} \right) \right] \end{aligned} \quad (\text{A.1})$$

On the left hand side of the equation the terms evaluate to:

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{\partial F_i}{\partial t} = \frac{\partial F_i}{\partial t} \quad (\text{Since } F_i \text{ is already the average}) \quad (\text{A.2})$$

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{\partial F_a}{\partial t} = \frac{\partial}{\partial t} \left( \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu F_a \right) = 0 \quad (\text{Because of equation (3.28)}) \quad (\text{A.3})$$

$$\frac{v}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \mu \frac{\partial F_i}{\partial z'} = \frac{\partial F_i}{\partial z'} \left( \frac{v}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \mu \right) = 0 \quad (\text{A.4})$$

$$\frac{\partial F_i}{\partial \phi} = 0 \quad (F_i \text{ is independent of } \phi) \quad (\text{A.5})$$

$$\frac{\epsilon\omega_p}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{\partial F_a}{\partial \phi} = \frac{\epsilon\omega_p}{4\pi} \int_{-1}^1 d\mu [F_a]_0^{2\pi} = 0 \quad (F_a \text{ is periodic: } F_a(\phi + 2\pi) = F_a(\phi)) \quad (\text{A.6})$$

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu S = S \quad (\text{A.7})$$

and the  $\partial F_a/\partial z'$  term remains as it is. On the right hand side there are 36 terms for each  $F_i$  and  $F_a$ . The diffusion coefficients with any  $\alpha = z'$  or  $\beta = z'$  vanish because  $g'_z = 0$ . For  $F_i$  terms with  $\beta = \mu$  and  $\beta = \phi$  vanish because  $F_i$  is independent from those variables. With a similar argument as in equation (A.6) all  $F_a$  terms with  $\beta = \phi$  vanish too.  $F_i$  terms with  $\alpha = \mu$  evaluate to

$$\frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{\partial}{\partial \mu} \left( D_{\mu\beta} \frac{\partial F_i}{\partial x_\beta} \right) = \frac{1}{4\pi} \int_0^{2\pi} d\phi \left[ D_{\mu\beta} \frac{\partial F_i}{\partial x_\beta} \right]_{\mu=-1}^{\mu=1} = 0 \quad (\text{A.8})$$

since  $D_{\mu\beta}(\mu = -1) = D_{\mu\beta}(\mu = 1) = 0$  because there is no pitch-angle diffusion if the particle is travelling parallel to the magnetic field. The same holds for  $F_a$ .

For the  $F_i$  terms with  $\alpha = \phi$  we have

$$\begin{aligned}
 \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \frac{\partial}{\partial \phi} \left( D_{\phi\beta} \frac{\partial F_i}{\partial x_\beta} \right) &= \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \left( \frac{\partial D_{\phi\beta}}{\partial \phi} \frac{\partial F_i}{\partial x_\beta} + D_{\phi\beta} \frac{\partial F_i}{\partial \phi \partial x_\beta} \right) \\
 &= \frac{1}{4\pi} \int_{-1}^1 d\mu \left[ \underbrace{\left( \frac{\partial F_i}{\partial x_\beta} \int_0^{2\pi} d\phi \frac{\partial D_{\phi\beta}}{\partial \phi} \right)}_{=0} - \int_0^{2\pi} d\phi \underbrace{\left( \frac{\partial F_i}{\partial \phi \partial x_\beta} \int_0^{2\pi} d\phi \frac{\partial D_{\phi\beta}}{\partial \phi} \right)}_{=0} \right] \\
 &\quad + \left( \frac{\partial D_{\phi\beta}}{\partial \phi} \underbrace{\int_0^{2\pi} d\phi \frac{\partial F_i}{\partial \phi \partial x_\beta}}_{=0} - \int_0^{2\pi} d\phi \underbrace{\left( \frac{\partial D_{\phi\beta}}{\partial \phi} \int_0^{2\pi} d\phi \frac{\partial F_i}{\partial \phi \partial x_\beta} \right)}_{=0} \right) \Bigg] \\
 &= 0
 \end{aligned} \tag{A.9}$$

since  $D_{\phi\beta}$  and  $\partial F_i / \partial x_\beta$  are periodic in  $\phi$  because it is a periodic coordinate. The same arguments as in equation (A.9) can be made for  $F_a$ . Summarizing these findings we have

$D$	$F_i$						$F_a$					
	$x'$	$y'$	$z'$	$p$	$\mu$	$\phi$	$x'$	$y'$	$z'$	$p$	$\mu$	$\phi$
$x'$	✓	✓	$g_{z'} = 0$	✓	$\partial_\mu F_i = 0$	$\partial_\phi F_i = 0$	✓	✓	$g_{z'} = 0$	✓	✓	(A.6)
$y'$	✓	✓	⋮	✓	⋮	⋮	✓	✓	⋮	✓	✓	⋮
$z'$	$g_{z'} = 0$	⋯	⋮	⋯	⋮	⋮	$g_{z'} = 0$	⋯	⋮	⋯	⋯	⋮
$p$	✓	✓	⋮	✓	⋮	⋮	✓	✓	⋮	✓	✓	⋮
$\mu$	(A.8)	⋯	⋮	⋯	⋮	⋮	(A.8)	⋯	⋮	⋯	⋯	⋮
$\phi$	(A.9)	⋯	⋮	⋯	⋮	⋮	(A.9)	⋯	⋮	⋯	⋯	⋮

**Table A.1.:** Eliminated terms. The terms left over are highlighted with ✓.

We are therefore left with

$$\begin{aligned}
 S &= \frac{\partial F_i}{\partial t} + \frac{v}{4\pi} \frac{\partial}{\partial z'} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu \mu F_a \\
 &- \sum_{\alpha=\beta=x',y',p} \frac{1}{4\pi p^2} \frac{\partial}{\partial x_\alpha} p^2 \frac{\partial F_i}{\partial x_\beta} \int_0^{2\pi} d\phi \int_{-1}^1 d\mu D_{\alpha\beta} \\
 &- \sum_{\substack{\alpha=x',y',p \\ \beta \neq z',\phi}} \frac{1}{4\pi p^2} \frac{\partial}{\partial x_\alpha} p^2 \int_0^{2\pi} d\phi \int_{-1}^1 d\mu D_{\alpha\beta} \frac{\partial F_a}{\partial x_\beta}
 \end{aligned} \tag{A.10}$$

# Bibliography

- Abdo, A. A. et al. (June 2010). “The Spectral Energy Distribution of Fermi Bright Blazars”. In: *The Astrophysical Journal* 716.1, pp. 30–70. DOI: [10.1088/0004-637X/716/1/30](https://doi.org/10.1088/0004-637X/716/1/30). arXiv: [0912.2040](https://arxiv.org/abs/0912.2040) [[astro-ph.CO](#)].
- Achatz, U., Steinacker, J. and Schlickeiser, R. (Oct. 1991). “Charged particle transport in a turbulent magnetized plasma - A reassessment and extension of quasilinear theory”. In: *Astronomy and Astrophysics* 250.1, pp. 266–279.
- Achterberg, A. and Schure, K. M. (Mar. 2011). “A more accurate numerical scheme for diffusive shock acceleration”. In: *Monthly Notices of the Royal Astronomical Society* 411.4, pp. 2628–2636. ISSN: 0035-8711. DOI: [10.1111/j.1365-2966.2010.17868.x](https://doi.org/10.1111/j.1365-2966.2010.17868.x). eprint: <https://academic.oup.com/mnras/article-pdf/411/4/2628/3053597/mnras0411-2628.pdf>. URL: <https://doi.org/10.1111/j.1365-2966.2010.17868.x>.
- Achterberg, A., Gallant, Y. A., et al. (Dec. 2001). “Particle acceleration by ultrarelativistic shocks: theory and simulations”. In: *Monthly Notices of the Royal Astronomical Society* 328.2, pp. 393–408. DOI: [10.1046/j.1365-8711.2001.04851.x](https://doi.org/10.1046/j.1365-8711.2001.04851.x). arXiv: [astro-ph/0107530](https://arxiv.org/abs/astro-ph/0107530) [[astro-ph](#)].
- Achterberg, A. and Krülls, W. M. (Nov. 1992). “A fast simulation method for particle acceleration.” In: *Astronomy and Astrophysics* 265.1, pp. L13–L16.
- Agudo, I. et al. (Jan. 2018). “POLAMI: Polarimetric Monitoring of Active Galactic Nuclei at Millimetre Wavelengths - III. Characterization of total flux density and polarization variability of relativistic jets”. In: *Monthly Notices of the Royal Astronomical Society* 473.2, pp. 1850–1867. DOI: [10.1093/mnras/stx2437](https://doi.org/10.1093/mnras/stx2437). arXiv: [1709.08744](https://arxiv.org/abs/1709.08744) [[astro-ph.GA](#)].
- Aharonian, F. A. (May 2002). “Proton-synchrotron radiation of large-scale jets in active galactic nuclei”. In: *Monthly Notices of the Royal Astronomical Society* 332.1, pp. 215–230. DOI: [10.1046/j.1365-8711.2002.05292.x](https://doi.org/10.1046/j.1365-8711.2002.05292.x). arXiv: [astro-ph/0106037](https://arxiv.org/abs/astro-ph/0106037) [[astro-ph](#)].
- Alfvén, H. and Herlofson, N. (June 1950). “Cosmic Radiation and Radio Stars”. In: *Physical Review* 78.5, pp. 616–616. DOI: [10.1103/PhysRev.78.616](https://doi.org/10.1103/PhysRev.78.616).
- Angel, J. R. P. and Stockman, H. S. (Jan. 1980). “Optical and infrared polarization of active extragalactic objects”. In: *Annual review of astronomy and astrophysics* 18, pp. 321–361. DOI: [10.1146/annurev.aa.18.090180.001541](https://doi.org/10.1146/annurev.aa.18.090180.001541).
- Baade, W. (May 1956). “Polarization in the Jet of Messier 87.” In: *The Astrophysical Journal* 123, pp. 550–551. DOI: [10.1086/146194](https://doi.org/10.1086/146194).
- Balogh, A. and Treumann, R. A. (2013). *Physics of Collisionless Shocks. Space Plasma Shock Waves*. Vol. 12. Springer New York. DOI: [10.1007/978-1-4614-6099-2](https://doi.org/10.1007/978-1-4614-6099-2).
- Bednarz, J. and Ostrowski, M. (Nov. 1996). “The acceleration time-scale for first-order Fermi acceleration in relativistic shock waves”. In: *Monthly Notices of the Royal Astronomical Society* 283.2, pp. 447–456. DOI: [10.1093/mnras/283.2.447](https://doi.org/10.1093/mnras/283.2.447). arXiv: [astro-ph/9608078](https://arxiv.org/abs/astro-ph/9608078) [[astro-ph](#)].
- Begelman, M. C. (Dec. 1995). “The Acceleration and Collimation of Jets”. In: *Proceedings of the National Academy of Science* 92.25, pp. 11442–11446. DOI: [10.1073/pnas.92.25.11442](https://doi.org/10.1073/pnas.92.25.11442).

- Begelman, M. C., Blandford, R. D. and Rees, M. J. (Apr. 1984). “Theory of extragalactic radio sources”. In: *Reviews of Modern Physics* 56.2, pp. 255–351. DOI: [10.1103/RevModPhys.56.255](https://doi.org/10.1103/RevModPhys.56.255).
- Begelman, M. C. and Kirk, J. G. (Apr. 1990). “Shock-Drift Particle Acceleration in Superluminal Shocks: A Model for Hot Spots in Extragalactic Radio Sources”. In: *The Astrophysical Journal* 353, p. 66. DOI: [10.1086/168590](https://doi.org/10.1086/168590).
- Behnel, S. et al. (2011). “Cython: The best of both worlds”. In: *Computing in Science & Engineering* 13.2, pp. 31–39.
- Bell, A. R. (Jan. 1978). “The acceleration of cosmic rays in shock fronts - I.” In: *Monthly Notices of the Royal Astronomical Society* 182, pp. 147–156. DOI: [10.1093/mnras/182.2.147](https://doi.org/10.1093/mnras/182.2.147).
- Bicknell, G. V., Dopita, M. A. and O’Dea, C. P. O. (Aug. 1997). “Unification of the Radio and Optical Properties of Gigahertz Peak Spectrum and Compact Steep-Spectrum Radio Sources”. In: *The Astrophysical Journal* 485.1, pp. 112–124. DOI: [10.1086/304400](https://doi.org/10.1086/304400).
- Biretta, J. A., Sparks, W. B. and Macchetto, F. (Aug. 1999). “Hubble Space Telescope Observations of Superluminal Motion in the M87 Jet”. In: *The Astrophysical Journal* 520.2, pp. 621–626. DOI: [10.1086/307499](https://doi.org/10.1086/307499).
- Biretta, J. A., Zhou, F. and Owen, F. N. (July 1995). “Detection of Proper Motions in the M87 Jet”. In: *The Astrophysical Journal* 447, p. 582. DOI: [10.1086/175901](https://doi.org/10.1086/175901).
- Blandford, R. D. and Königl, A. (Aug. 1979). “Relativistic jets as compact radio sources.” In: *The Astrophysical Journal* 232, pp. 34–48. DOI: [10.1086/157262](https://doi.org/10.1086/157262).
- Blandford, R. D. and Payne, D. G. (June 1982). “Hydromagnetic flows from accretion disks and the production of radio jets.” In: *Monthly Notices of the Royal Astronomical Society* 199, pp. 883–903. DOI: [10.1093/mnras/199.4.883](https://doi.org/10.1093/mnras/199.4.883).
- Blandford, R. D. and Znajek, R. L. (May 1977). “Electromagnetic extraction of energy from Kerr black holes.” In: *Monthly Notices of the Royal Astronomical Society* 179, pp. 433–456. DOI: [10.1093/mnras/179.3.433](https://doi.org/10.1093/mnras/179.3.433).
- Blandford, R. and Eichler, D. (Oct. 1987). “Particle acceleration at astrophysical shocks: A theory of cosmic ray origin”. In: *Physics Reports* 154.1, pp. 1–75. DOI: [10.1016/0370-1573\(87\)90134-7](https://doi.org/10.1016/0370-1573(87)90134-7).
- Błażejowski, M. et al. (Dec. 2000). “Comptonization of Infrared Radiation from Hot Dust by Relativistic Jets in Quasars”. In: *The Astrophysical Journal* 545.1, pp. 107–116. DOI: [10.1086/317791](https://doi.org/10.1086/317791). arXiv: [astro-ph/0008154](https://arxiv.org/abs/astro-ph/0008154) [astro-ph].
- Böttcher, M., Harris, D. and Krawczynski, H., eds. (2012). *Relativistic Jets from Active Galactic Nuclei*. Wiley. ISBN: 9783527410378.
- Broyden, C. G. (1965). “A Class of Methods for Solving Nonlinear Simultaneous Equations”. In: *Mathematics of Computation* 19.92, pp. 577–593. DOI: [10.1090/S0025-5718-1965-0198670-6](https://doi.org/10.1090/S0025-5718-1965-0198670-6). (Visited on 05/17/2022).
- Burbidge, G. R. (Sept. 1956). “On Synchrotron Radiation from Messier 87.” In: *The Astrophysical Journal* 124, p. 416. DOI: [10.1086/146237](https://doi.org/10.1086/146237).
- Bykov, A. M. et al. (Oct. 2013). “Microphysics of Cosmic Ray Driven Plasma Instabilities”. In: *Space Science Reviews* 178.2, pp. 201–232. ISSN: 1572-9672. DOI: [10.1007/s11214-013-9988-3](https://doi.org/10.1007/s11214-013-9988-3). URL: <https://doi.org/10.1007/s11214-013-9988-3>.
- Chatterjee, R. et al. (June 2011). “Connection Between the Accretion Disk and Jet in the Radio Galaxy 3C 111”. In: *The Astrophysical Journal* 734.1, 43, p. 43. DOI: [10.1088/0004-637X/734/1/43](https://doi.org/10.1088/0004-637X/734/1/43). arXiv: [1104.0663](https://arxiv.org/abs/1104.0663) [astro-ph.HE].

- Condon, J. J. and Ransom, S. M. (Oct. 2015). *Essential radio astronomy*. URL: <https://www.cv.nrao.edu/~sransom/web/xxx.html>.
- Courvoisier, T. J. -. (2013). *High Energy Astrophysics. An Introduction*. Springer-Verlag Berlin Heidelberg. ISBN: 978-3-642-30969-4. DOI: [10.1007/978-3-642-30970-0](https://doi.org/10.1007/978-3-642-30970-0).
- Crumley, P. et al. (June 2019). “Kinetic simulations of mildly relativistic shocks - I. Particle acceleration in high Mach number shocks”. In: *Monthly Notices of the Royal Astronomical Society* 485.4, pp. 5105–5119. DOI: [10.1093/mnras/stz232](https://doi.org/10.1093/mnras/stz232). arXiv: [1809.10809](https://arxiv.org/abs/1809.10809) [astro-ph.HE].
- Crusius, A. and Schlickeiser, R. (Aug. 1986). “Synchrotron radiation in random magnetic fields”. In: *Astronomy and Astrophysics* 164.2, pp. L16–L18.
- Curtis, H. D. (Jan. 1918). “Descriptions of 762 Nebulae and Clusters Photographed with the Crossley Reflector”. In: *Publications of Lick Observatory* 13, pp. 9–42.
- Davis, L. et al. (2021). *ProPlot*. Version 0.9.5. URL: <https://proplot.readthedocs.io/en/latest/index.html>.
- Dermer, C. D. and Atoyan, A. M. (Apr. 2002). “X-Ray Synchrotron Spectral Hardenings from Compton and Synchrotron Losses in Extended Chandra Jets”. In: *Astrophysical Journal Letters* 568.2, pp. L81–L84. DOI: [10.1086/340374](https://doi.org/10.1086/340374). arXiv: [astro-ph/0202406](https://arxiv.org/abs/astro-ph/0202406) [astro-ph].
- Dermer, C. D. and Menon, G. (2009). *High Energy Radiation from Black Holes: Gamma Rays, Cosmic Rays, and Neutrinos*. Princeton University Press. ISBN: 9780691144085. URL: <http://www.jstor.org/stable/j.ctt7rphp>.
- Dermer, C. D. and Schlickeiser, R. (Oct. 1993). “Model for the High-Energy Emission from Blazars”. In: *The Astrophysical Journal* 416, p. 458. DOI: [10.1086/173251](https://doi.org/10.1086/173251).
- Dihingia, I. K., Vaidya, B. and Fendt, C. (Aug. 2021). “Jets, disc-winds, and oscillations in general relativistic, magnetically driven flows around black hole”. In: *Monthly Notices of the Royal Astronomical Society* 505.3, pp. 3596–3615. DOI: [10.1093/mnras/stab1512](https://doi.org/10.1093/mnras/stab1512). arXiv: [2105.11468](https://arxiv.org/abs/2105.11468) [astro-ph.HE].
- Dröge, W. et al. (Feb. 2010). “Anisotropic Three-Dimensional Focused Transport of Solar Energetic Particles in the Inner Heliosphere”. In: *The Astrophysical Journal* 709.2, pp. 912–919. DOI: [10.1088/0004-637X/709/2/912](https://doi.org/10.1088/0004-637X/709/2/912).
- Drury, L. O., Markiewicz, W. J. and Voelk, H. J. (Nov. 1989). “Simplified models for the evolution of supernova remnants including particle acceleration”. In: *Astronomy and Astrophysics* 225.1, pp. 179–191.
- Event Horizon Telescope Collaboration et al. (Apr. 2019). “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole”. In: *Astrophysical Journal Letters* 875.1, L1, p. L1. DOI: [10.3847/2041-8213/ab0ec7](https://doi.org/10.3847/2041-8213/ab0ec7). arXiv: [1906.11238](https://arxiv.org/abs/1906.11238) [astro-ph.GA].
- Event Horizon Telescope Collaboration et al. (May 2022). “First Sagittarius A\* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way”. In: *Astrophysical Journal Letters* 930.2, L12, p. L12. DOI: [10.3847/2041-8213/ac6674](https://doi.org/10.3847/2041-8213/ac6674).
- Fanaroff, B. L. and Riley, J. M. (May 1974). “The morphology of extragalactic radio sources of high and low luminosity”. In: *Monthly Notices of the Royal Astronomical Society* 167, 31P–36P. DOI: [10.1093/mnras/167.1.31P](https://doi.org/10.1093/mnras/167.1.31P).
- Fermi, E. (Apr. 1949). “On the Origin of the Cosmic Radiation”. In: *Physical Review* 75.8, pp. 1169–1174. DOI: [10.1103/PhysRev.75.1169](https://doi.org/10.1103/PhysRev.75.1169).

- 
- Finke, J. D., Dermer, C. D. and Böttcher, M. (Oct. 2008). “Synchrotron Self-Compton Analysis of TeV X-Ray-Selected BL Lacertae Objects”. In: *The Astrophysical Journal* 686.1, pp. 181–194. DOI: [10.1086/590900](https://doi.org/10.1086/590900). URL: <https://doi.org/10.1086/590900>.
- Fitzpatrick, R. (2014). *Plasma Physics: An Introduction*. CRC Press, Boca Raton. ISBN: 9781466594272.
- Fokker, A. D. (1914). “Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld”. In: *Annalen der Physik* 348.5, pp. 810–820. DOI: <https://doi.org/10.1002/andp.19143480507>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19143480507>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19143480507>.
- Fraix-Burnet, D. (Jan. 1990). “Superluminal motions with oblique shocks”. In: *Astronomy and Astrophysics* 227.1, pp. 1–5.
- Gardiner, C. (2009). *Stochastic Methods. A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer Berlin Heidelberg.
- Ghisellini, G., Celotti, A., et al. (Mar. 2014). “Radio-loud active galactic nuclei at high redshifts and the cosmic microwave background”. In: *Monthly Notices of the Royal Astronomical Society* 438.3, pp. 2694–2700. DOI: [10.1093/mnras/stt2394](https://doi.org/10.1093/mnras/stt2394). arXiv: [1311.7147](https://arxiv.org/abs/1311.7147) [astro-ph.CO].
- Ghisellini, G., Tavecchio, F., et al. (Nov. 2014). “The power of relativistic jets is larger than the luminosity of their accretion disks”. In: *Nature* 515.7527, pp. 376–378. DOI: [10.1038/nature13856](https://doi.org/10.1038/nature13856). arXiv: [1411.5368](https://arxiv.org/abs/1411.5368) [astro-ph.HE].
- Gopal-Krishna et al. (Oct. 2001). “Are the hotspots of radio galaxies the sites of in situ acceleration of relativistic particles?” In: *Astronomy and Astrophysics* 377, pp. 827–834. DOI: [10.1051/0004-6361:20011149](https://doi.org/10.1051/0004-6361:20011149). arXiv: [astro-ph/0108144](https://arxiv.org/abs/astro-ph/0108144) [astro-ph].
- Guennebaud, G., Jacob, B., et al. (2010). *Eigen v3*. <http://eigen.tuxfamily.org>.
- Gupta, N. and Basumallick, P. P. (Oct. 2017). “Extended Jets of AGN: Proton Synchrotron Emission”. In: *Proceedings of the 7th International Fermi Symposium*, 24, p. 24.
- Harris, C. R. et al. (Sept. 2020). “Array programming with NumPy”. In: *Nature* 585.7825, pp. 357–362. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2). URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- Hasegawa, A. (1975). *Plasma Instabilities and Nonlinear Effects*. Physics and Chemistry in Space. Springer Berlin Heidelberg. ISBN: 9783642659829.
- Hazeltine, R. D. and Waelbroeck, F. L. (2004). *The Framework of Plasma Physics*. 1. CRC Press. DOI: [10.1201/9780429502804](https://doi.org/10.1201/9780429502804).
- Hunter, J. D. (2007). “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- ISO (Dec. 2020). *ISO/IEC 14882:2020 Programming languages — C++*. URL: <https://www.iso.org/standard/68564.html>.
- Itô, K. (1944). “Stochastic Integral”. In: *Proceedings of the Imperial Academy* 20.8, pp. 519–524. DOI: [10.3792/pia/1195572786](https://doi.org/10.3792/pia/1195572786).
- Jester, S. et al. (Apr. 2002). “X-rays from the jet in  $\text{jASTROBJ}_i\text{3C 273}_i/\text{ASTROBJ}_i$ : Clues from the radio-optical spectra”. In: *Astronomy and Astrophysics* 385, pp. L27–L30. DOI: [10.1051/0004-6361:20020278](https://doi.org/10.1051/0004-6361:20020278). arXiv: [astro-ph/0202428](https://arxiv.org/abs/astro-ph/0202428) [astro-ph].
- Jones, F. C. (Feb. 1994). “A Theoretical Review of Diffusive Shock Acceleration”. In: *The Astrophysical Journal Supplement Series* 90, p. 561. DOI: [10.1086/191875](https://doi.org/10.1086/191875).
-

- 
- Jorstad, S. G. et al. (Oct. 2005). “Polarimetric Observations of 15 Active Galactic Nuclei at High Frequencies: Jet Kinematics from Bimonthly Monitoring with the Very Long Baseline Array”. In: *The Astronomical Journal* 130.4, pp. 1418–1465. DOI: [10.1086/444593](https://doi.org/10.1086/444593). arXiv: [astro-ph/0502501](https://arxiv.org/abs/astro-ph/0502501) [astro-ph].
- Kadler, M., Kerp, J., et al. (June 2004). “Jet emission in jASTROBJ<sub>i</sub>NGC 1052<sub>i</sub>/ASTROBJ<sub>i</sub> at radio, optical, and X-ray frequencies”. In: *Astronomy and Astrophysics* 420, pp. 467–474. DOI: [10.1051/0004-6361:20034126](https://doi.org/10.1051/0004-6361:20034126). arXiv: [astro-ph/0403165](https://arxiv.org/abs/astro-ph/0403165) [astro-ph].
- Kadler, M., Ros, E., et al. (June 2008). “The Trails of Superluminal Jet Components in 3C 111”. In: *The Astrophysical Journal* 680.2, pp. 867–884. DOI: [10.1086/529539](https://doi.org/10.1086/529539). arXiv: [0801.0617](https://arxiv.org/abs/0801.0617) [astro-ph].
- Kagan, D. et al. (Oct. 2015). “Relativistic Magnetic Reconnection in Pair Plasmas and Its Astrophysical Applications”. In: *Space Science Reviews* 191.1-4, pp. 545–573. DOI: [10.1007/s11214-014-0132-9](https://doi.org/10.1007/s11214-014-0132-9). arXiv: [1412.2451](https://arxiv.org/abs/1412.2451) [astro-ph.HE].
- Kirk, J. G. (1994). “Particle Acceleration”. In: Kirk, Melrose and Priest, 1994, pp. 225–314.
- Kirk, J. G., Melrose, D. B. and Priest, E. R. (1994). *Plasma Astrophysics*. Ed. by A. O. Benz and T. J.-L. Courvoisier. Berlin Heidelberg: Springer-Verlag.
- Kirk, J. G., Schlickeiser, R. and Schneider, P. (May 1988). “Cosmic-Ray Transport in Accelerating Flows”. In: *The Astrophysical Journal* 328, p. 269. DOI: [10.1086/166290](https://doi.org/10.1086/166290).
- Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Berlin, Heidelberg: Springer-Verlag.
- Kolmogorov, A. (Dec. 1931). “Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung”. In: *Mathematische Annalen* 104.1, pp. 415–458. ISSN: 1432-1807. DOI: [10.1007/BF01457949](https://doi.org/10.1007/BF01457949). URL: <https://doi.org/10.1007/BF01457949>.
- Kopp, A. et al. (Mar. 2012). “A stochastic differential equation code for multidimensional Fokker-Planck type problems”. In: *Computer Physics Communications* 183.3, pp. 530–542. DOI: [10.1016/j.cpc.2011.11.014](https://doi.org/10.1016/j.cpc.2011.11.014).
- Krülls, W. M. and Achterberg, A. (June 1994). “Computation of cosmic-ray acceleration by Ito’s stochastic differential equations”. In: *Astronomy and Astrophysics* 286, pp. 314–327.
- Kulkarni, S. R. (2007). “When is  $\nu F_\nu$  useful?” In: URL: <https://sites.astro.caltech.edu/~srk/ay122/PowerLaw.pdf>.
- Kulsrud, R. and Pearce, W. P. (May 1969). “The Effect of Wave-Particle Interactions on the Propagation of Cosmic Rays”. In: *The Astrophysical Journal* 156, p. 445. DOI: [10.1086/149981](https://doi.org/10.1086/149981).
- Lanzuisi, G. et al. (Feb. 2014). “Active Galactic Nucleus X-Ray Variability in the XMM-COSMOS Survey”. In: *The Astrophysical Journal* 781.2, 105, p. 105. DOI: [10.1088/0004-637X/781/2/105](https://doi.org/10.1088/0004-637X/781/2/105). arXiv: [1312.2264](https://arxiv.org/abs/1312.2264) [astro-ph.CO].
- Liao, M. et al. (May 2022). “Exploring the connection between ultraviolet/optical variations and radio emission in radio-quiet quasars: clues about the origin of radio emission”. In: *Monthly Notices of the Royal Astronomical Society* 512.1, pp. 296–303. DOI: [10.1093/mnras/stac266](https://doi.org/10.1093/mnras/stac266). arXiv: [2201.11529](https://arxiv.org/abs/2201.11529) [astro-ph.GA].
- Lister, M. L. et al. (Dec. 2009). “MOJAVE: Monitoring of Jets in Active Galactic Nuclei with VLBA Experiments. VI. Kinematics Analysis of a Complete Sample of Blazar Jets”. In: *The Astronomical Journal* 138.6, pp. 1874–1892. DOI: [10.1088/0004-6256/138/6/1874](https://doi.org/10.1088/0004-6256/138/6/1874). arXiv: [0909.5100](https://arxiv.org/abs/0909.5100) [astro-ph.CO].
- Longair, M. S. (2011). *High energy astrophysics*. Cambridge: Cambridge University Press.
-

- Lynden-Bell, D. (June 2003). “On why discs generate magnetic towers and collimate jets”. In: *Monthly Notices of the Royal Astronomical Society* 341.4, pp. 1360–1372. DOI: [10.1046/j.1365-8711.2003.06506.x](https://doi.org/10.1046/j.1365-8711.2003.06506.x). arXiv: [astro-ph/0208388](https://arxiv.org/abs/astro-ph/0208388) [[astro-ph](#)].
- Mannheim, K. (Mar. 1993). “The proton blazar.” In: *Astronomy and Astrophysics* 269, pp. 67–76. arXiv: [astro-ph/9302006](https://arxiv.org/abs/astro-ph/9302006) [[astro-ph](#)].
- Mannheim, K., Krüß, W. M. and Biermann, P. L. (Nov. 1991). “A novel mechanism for nonthermal X-ray emission”. In: *Astronomy and Astrophysics* 251.2, pp. 723–731.
- Marcowith, A. et al. (Apr. 2016). “The microphysics of collisionless shock waves”. In: *Reports on Progress in Physics* 79.4, 046901, p. 046901. DOI: [10.1088/0034-4885/79/4/046901](https://doi.org/10.1088/0034-4885/79/4/046901). arXiv: [1604.00318](https://arxiv.org/abs/1604.00318) [[astro-ph.HE](#)].
- Matthews, J. H., Bell, A. R. and Blundell, K. M. (Sept. 2020). “Particle acceleration in astrophysical jets”. In: *New Astronomy Reviews* 89, 101543, p. 101543. DOI: [10.1016/j.newar.2020.101543](https://doi.org/10.1016/j.newar.2020.101543). arXiv: [2003.06587](https://arxiv.org/abs/2003.06587) [[astro-ph.HE](#)].
- Merloni, A., Heinz, S. and di Matteo, T. (Nov. 2003). “A Fundamental Plane of black hole activity”. In: *Monthly Notices of the Royal Astronomical Society* 345.4, pp. 1057–1076. DOI: [10.1046/j.1365-2966.2003.07017.x](https://doi.org/10.1046/j.1365-2966.2003.07017.x). arXiv: [astro-ph/0305261](https://arxiv.org/abs/astro-ph/0305261) [[astro-ph](#)].
- Mignone, A., Bodo, G., Massaglia, S., et al. (May 2007). “PLUTO: A Numerical Code for Computational Astrophysics”. In: *The Astrophysical Journal Supplement Series* 170.1, pp. 228–242. DOI: [10.1086/513316](https://doi.org/10.1086/513316). arXiv: [astro-ph/0701854](https://arxiv.org/abs/astro-ph/0701854) [[astro-ph](#)].
- Mignone, A., Bodo, G., Vaidya, B., et al. (May 2018). “A Particle Module for the PLUTO Code. I. An Implementation of the MHD–PIC Equations”. In: *The Astrophysical Journal* 859.1, p. 13. DOI: [10.3847/1538-4357/aabccd](https://doi.org/10.3847/1538-4357/aabccd). URL: <https://doi.org/10.3847/1538-4357/aabccd>.
- Nigro, C. et al. (Jan. 2022). *agnpy*. Version 0.1.8. DOI: [10.5281/zenodo.5932850](https://doi.org/10.5281/zenodo.5932850). URL: <https://doi.org/10.5281/zenodo.5932850>.
- O’Neill, M. E. (Sept. 2014). *PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation*. Tech. rep. HMC-CS-2014-0905. Claremont, CA: Harvey Mudd College.
- Perlman, E. (2012). “Optical, Infrared and UV Observations”. In: Böttcher, Harris and Krawczynski, 2012. Chap. 6, pp. 153–184.
- Petrosian, V. (Nov. 2012). “Stochastic Acceleration by Turbulence”. In: *Space Science Reviews* 173.1-4, pp. 535–556. DOI: [10.1007/s11214-012-9900-6](https://doi.org/10.1007/s11214-012-9900-6). arXiv: [1205.2136](https://arxiv.org/abs/1205.2136) [[astro-ph.HE](#)].
- Planck, M. (1917). “Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie”. In: *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin* 24, pp. 324–341. URL: <https://www.biodiversitylibrary.org/item/92720>.
- Porth, O. and Fendt, C. (Feb. 2010). “Acceleration and Collimation of Relativistic Magnetohydrodynamic Disk Winds”. In: *The Astrophysical Journal* 709.2, pp. 1100–1118. DOI: [10.1088/0004-637X/709/2/1100](https://doi.org/10.1088/0004-637X/709/2/1100). arXiv: [0911.3001](https://arxiv.org/abs/0911.3001) [[astro-ph.HE](#)].
- Prandini, E. and Ghisellini, G. (Feb. 2022). “The Blazar Sequence and Its Physical Understanding”. In: *Galaxies* 10.1, p. 35. DOI: [10.3390/galaxies10010035](https://doi.org/10.3390/galaxies10010035). arXiv: [2202.07490](https://arxiv.org/abs/2202.07490) [[astro-ph.HE](#)].
- Pudritz, R. E. et al. (Jan. 2007). “Disk Winds, Jets, and Outflows: Theoretical and Computational Foundations”. In: *Protostars and Planets V*. Ed. by B. Reipurth, D. Jewitt and K. Keil, p. 277. arXiv: [astro-ph/0603592](https://arxiv.org/abs/astro-ph/0603592) [[astro-ph](#)].

- Python Software Foundation (2022). *Python Language Reference*. Version 3.10.4. URL: <https://docs.python.org/3/reference/index.html>.
- Ramsey, J. P. and Clarke, D. A. (Apr. 2019). “MHD simulations of the formation and propagation of protostellar jets to observational length-scales”. In: *Monthly Notices of the Royal Astronomical Society* 484.2, pp. 2364–2387. DOI: [10.1093/mnras/stz116](https://doi.org/10.1093/mnras/stz116). arXiv: [1901.02845](https://arxiv.org/abs/1901.02845) [astro-ph.SR].
- Romero, G. E., Böttcher, M., et al. (July 2017). “Relativistic Jets in Active Galactic Nuclei and Microquasars”. In: *Space Science Reviews* 207.1-4, pp. 5–61. DOI: [10.1007/s11214-016-0328-2](https://doi.org/10.1007/s11214-016-0328-2). arXiv: [1611.09507](https://arxiv.org/abs/1611.09507) [astro-ph.HE].
- Romero, G. E. and Vila, G. S. (2014). *Introduction to Black Hole Astrophysics*. Vol. 876. DOI: [10.1007/978-3-642-39596-3](https://doi.org/10.1007/978-3-642-39596-3).
- Rybicki, G. B. and Lightman, A. P. (1986). *Radiative Processes in Astrophysics*. Wiley-VCH.
- Sambruna, R. and Harris, D. (2012). “Optical, Infrared and UV Observations”. In: Böttcher, Harris and Krawczynski, 2012. Chap. 7, pp. 185–213.
- Sambruna, R. M., Gambill, J. K., et al. (June 2004). “A Survey of Extended Radio Jets with Chandra and the Hubble Space Telescope”. In: *The Astrophysical Journal* 608.2, pp. 698–720. DOI: [10.1086/383124](https://doi.org/10.1086/383124). arXiv: [astro-ph/0401475](https://arxiv.org/abs/astro-ph/0401475) [astro-ph].
- Sambruna, R. M., Maraschi, L. and Urry, C. M. (June 1996). “On the Spectral Energy Distributions of Blazars”. In: *The Astrophysical Journal* 463, p. 444. DOI: [10.1086/177260](https://doi.org/10.1086/177260).
- Sbarrato, T. et al. (Apr. 2012). “The relation between broad lines and  $\gamma$ -ray luminosities in Fermi blazars”. In: *Monthly Notices of the Royal Astronomical Society* 421.2, pp. 1764–1778. DOI: [10.1111/j.1365-2966.2012.20442.x](https://doi.org/10.1111/j.1365-2966.2012.20442.x). arXiv: [1108.0927](https://arxiv.org/abs/1108.0927) [astro-ph.HE].
- Schlickeiser, R. (2002). *Cosmic Ray Astrophysics*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg.
- Schwartz, D. A. et al. (Sept. 2000). “Chandra Discovery of a 100 kiloparsec X-Ray Jet in PKS 0637-752”. In: *Astrophysical Journal Letters* 540.2, pp. 69–72. DOI: [10.1086/312875](https://doi.org/10.1086/312875). arXiv: [astro-ph/0005255](https://arxiv.org/abs/astro-ph/0005255) [astro-ph].
- Snios, B. et al. (July 2019). “Detection of Superluminal Motion in the X-Ray Jet of M87”. In: *The Astrophysical Journal* 879.1, 8, p. 8. DOI: [10.3847/1538-4357/ab2119](https://doi.org/10.3847/1538-4357/ab2119). arXiv: [1905.04330](https://arxiv.org/abs/1905.04330) [astro-ph.HE].
- Sol, H., Pelletier, G. and Asseo, E. (Mar. 1989). “Two-flow model for extragalactic radio jets.” In: *Monthly Notices of the Royal Astronomical Society* 237, pp. 411–429. DOI: [10.1093/mnras/237.2.411](https://doi.org/10.1093/mnras/237.2.411).
- Strauss, R. D. T. and Effenberger, F. (Oct. 2017). “A Hitch-hiker’s Guide to Stochastic Differential Equations”. In: *Space Science Reviews* 212.1, pp. 151–192. ISSN: 1572-9672. DOI: [10.1007/s11214-017-0351-y](https://doi.org/10.1007/s11214-017-0351-y). URL: <https://doi.org/10.1007/s11214-017-0351-y>.
- Tavecchio, F., Maraschi, L. and Ghisellini, G. (Dec. 1998). “Constraints on the Physical Parameters of TeV Blazars”. In: *The Astrophysical Journal* 509.2, pp. 608–619. DOI: [10.1086/306526](https://doi.org/10.1086/306526). arXiv: [astro-ph/9809051](https://arxiv.org/abs/astro-ph/9809051) [astro-ph].
- Thomas, T. and Pfrommer, C. (Jan. 2019). “Cosmic-ray hydrodynamics: Alfvén-wave regulated transport of cosmic rays”. In: *Monthly Notices of the Royal Astronomical Society* 485.3, pp. 2977–3008. ISSN: 0035-8711. DOI: [10.1093/mnras/stz263](https://doi.org/10.1093/mnras/stz263). eprint: <https://academic.oup.com/mnras/article-pdf/485/3/2977/28079488/stz263.pdf>. URL: <https://doi.org/10.1093/mnras/stz263>.

- Toptygin, I. N. (June 1980). “Acceleration of particles by shocks in a cosmic plasma.” In: *Space Science Reviews* 26.2, pp. 157–213. DOI: [10.1007/BF00167370](https://doi.org/10.1007/BF00167370).
- Urry, C. M. and Padovani, P. (Sept. 1995). “Unified Schemes for Radio-Loud Active Galactic Nuclei”. In: *Publications of the Astronomical Society of the Pacific* 107, p. 803. DOI: [10.1086/133630](https://doi.org/10.1086/133630). arXiv: [astro-ph/9506063](https://arxiv.org/abs/astro-ph/9506063) [[astro-ph](#)].
- Vaidya, B. et al. (Oct. 2018). “A Particle Module for the PLUTO Code. II. Hybrid Framework for Modeling Nonthermal Emission from Relativistic Magnetized Flows”. In: *The Astrophysical Journal* 865.2, 144, p. 144. DOI: [10.3847/1538-4357/aadd17](https://doi.org/10.3847/1538-4357/aadd17). arXiv: [1808.08960](https://arxiv.org/abs/1808.08960) [[astro-ph.HE](#)].
- Virtanen, P. et al. (2020). “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17, pp. 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- Vlasov, A. A. (June 1968). “The Vibrational Properties Of An Electron Gas”. In: *Soviet Physics Uspekhi* 10.6, pp. 721–733. DOI: [10.1070/pu1968v010n06abeh003709](https://doi.org/10.1070/pu1968v010n06abeh003709). URL: <https://doi.org/10.1070/pu1968v010n06abeh003709>.
- Wilson, A. S. and Colbert, E. J. M. (Jan. 1995). “The Difference between Radio-loud and Radio-quiet Active Galaxies”. In: *The Astrophysical Journal* 438, p. 62. DOI: [10.1086/175054](https://doi.org/10.1086/175054). arXiv: [astro-ph/9408005](https://arxiv.org/abs/astro-ph/9408005) [[astro-ph](#)].
- Winner, G. et al. (July 2019). “Evolution of cosmic ray electron spectra in magnetohydrodynamical simulations”. In: *Monthly Notices of the Royal Astronomical Society* 488.2, pp. 2235–2252. ISSN: 0035-8711. DOI: [10.1093/mnras/stz1792](https://doi.org/10.1093/mnras/stz1792). eprint: <https://academic.oup.com/mnras/article-pdf/488/2/2235/28974006/stz1792.pdf>. URL: <https://doi.org/10.1093/mnras/stz1792>.
- Woo, J.-H. and Urry, C. M. (Nov. 2002). “Active Galactic Nucleus Black Hole Masses and Bolometric Luminosities”. In: *The Astrophysical Journal* 579.2, pp. 530–544. DOI: [10.1086/342878](https://doi.org/10.1086/342878). arXiv: [astro-ph/0207249](https://arxiv.org/abs/astro-ph/0207249) [[astro-ph](#)].
- Zhang, M. (Sept. 2000). “Calculation of Diffusive Shock Acceleration of Charged Particles by Skew Brownian Motion”. In: *The Astrophysical Journal* 541.1, pp. 428–435. DOI: [10.1086/309429](https://doi.org/10.1086/309429).

# Acknowledgements

First I would like to thank my supervisors Karl Mannheim and Sarah Wagner. Many times, Karl gave me the right amount of motivation (and sometimes pressure) to stay on track, which I am grateful for, and always had precise explanations at hand when I was lost in confusion about physics. Thank you, Sarah, for the hour-long discussions in our office until late at night and for the orientation in the vast field of astronomy in general and jet physics in particular. It was a pleasure to work with you.

I'm also very thankful for the help of Sarah, Bernd Schleicher, Christoph Wendel and Maurice Kruse in proof-reading this thesis on very short notice and for their very helpful comments.

Thank you to everyone I met in the astronomy department over the last months for their warm welcome to the fascinating world of astronomy, and for many interesting and helpful discussions.

I express gratitude to my parents for their way of raising me, with open eyes for the workings of the world and for always supporting and embracing my curiosity. Finally I deeply have to thank my girlfriend Theresa for her unconditional support and for the freedom she never failed to give. Without those people, this thesis would not have been possible.

Last but not least, I have to acknowledge the incredible work the open-source/free software community accomplishes. I could not work the way I (want to) work without the countless software projects I use day to day, made by countless developers. On the same account a big thanks goes out to the authors of Wikipedia, which is always there providing a quick refreshment on long-forgotten knowledge and overview over big topics. Research would be a much slower process without such a project.

# Declaration

I, Patrick Günther, declare that this thesis and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

---

Date, Signature