Investigation of the solar atmosphere using machine learning techniques

Author:
Daniel ZAHND

Supervisor:
Prof. Dr. Lucia KLEINT

DRAFT

December 15, 2023

Abstract

As a source of not only all energy needed on earth to enable plants to grow or weather phenomena to form, but also of inspiration and fascination, the sun and its underlying physics has ever been of utmost interest. There are certain solar phenomena such as solar flares or coronal mass ejections, which pose potential threats to technical endeavors on earth and also in space, such as power grids, spacecraft and ultimately humans serving as astronauts in space. Therefore, in analogy to meteorological forecasts on earth, there is a desire to be able to forecast solar flares and coronal mass ejections; this field of research is commonly called space weather. In order to make predictions of solar flares or coronal mass ejections, certain precursors in the solar atmosphere causally related or correlated to these phenomena have to be identified and studied. Aiming for this goal, inversions of so-called atmospheric models of stars are needed to obtain and investigate those physical parameters and to assess, whether some could be identified as precursors or not. Many stellar atmospheres are not invertible in an analytical way, which is to say that inversions of more complex stellar models are only possible by lots of computational power. Therefore, an alternative, in terms of computational power less demanding way of inverting stellar atmospheres is explored in this thesis. In the thesis at hand, normalizing flows are applied to observations of the Swedish Solar Telescope (SST) in combination with the Milne-Eddington model as the sun's atmospheric model.

Contents

Chapter 1

Introduction

For many people, looking up at the night sky is a source of utmost excitation and inspiration. By naked eye, the most prominent objects in the night sky are of course the moon, the planets of our solar system and many stars. The most interesting star for us humans is the Sun, because it drives and powers our everyday life. Without it, there would be no growth of plants and there would ultimately be no life, as we know it, possible on Earth. Apart from mere satisfaction of human curiosity and fascination, understanding stellar and especially solar physics is also key to many viable applications for mankind. The field of space weather is concerned with understanding and predicting solar eruptions, so-called solar flares, which according to [Stix, 2002, p.436] can, but do not have to be correlated with coronal mass ejections.

1.1 The Sun

Among the many characteristics of the Sun, there are at least five phenomenological properties as visualized in fig. 1.1; these include its mass $M_\odot = 1.989 \times 10^{30}$ kg, its radius $R_\odot = 696\,340$ km, its luminosity

$$L_\odot = 4\pi R_\odot^2 \tilde{\sigma} T_\odot^4 = 3.846 \times 10^{26}\,\text{W}, \tag{1.1}$$

its effective temperature $T_\odot \approx 5800$ K and its distance to Earth $d_{\odot,E} \approx 150 \times 10^6$ km. Thereby, $\tilde{\sigma} = 5.670 \times 10^{-8}\,\text{Wm}^{-2}\text{K}^{-4}$ is the Stefan-Boltzmann constant. Comparing these characteristics of the Sun to those of Earth with mass $M_E = 5.972 \times 10^{24}$ kg, radius $R_E = 6341$ km and distance to the Sun $d_{\odot,E}$, one can see that the Sun is approximately 0.3 million times more massive and roughly 1.3 million times more voluminous than Earth.
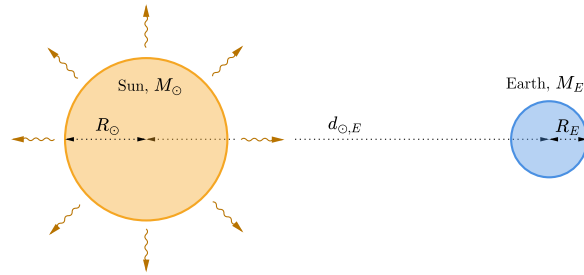
Figure 1.1: Visualization of the basic characteristics of the Sun and Earth.

The outer layers of the Sun can be roughly divided into three regions, the photosphere, the chromosphere and the corona. The photosphere is the innermost of these outer layers. Acoording to [Weigert et al., 2006, p.135], its depth is determined by the optical depth being such, that photons from below the photosphere cannot escape into space anymore; this is to say, that all observable photons from outside the Sun originate at or above the photosphere, which is its namegiving property. Above the photosphere, the chromosphere is located Since the density is very low in this layer, there is not much contribution to the overall emitted energy of the Sun coming from this layer despite its high temperature of up to several 10^5 K. The outermost layer of the solar atmosphere is the corona with even higher temperatures. The identification of the heating mechanism leading to these high temperatures in the solar corona is still one of the major open questions of stellar physics, known as the coronal heating problem. The boundary of the corona to interplanetary space is of a continuous nature.

Those characteristics alone however provide little insight to the formation of Sunspots, solar flares or coronal mass ejections. In order to gain insight to the physics relevant for those phenomena, a detailed theory of radiative transfer is necessary, which then provides a useful tool to extract physical parameters from solar observations of a certain region of the solar atmosphere. Analyzing those parameters for so-called active regions of the Sun - these are regions, where Sunspots, solar flares and coronal mass ejections form - provides insight to the physical processes going on during these phenomena.

1.1.1 Sunspots

Apart from the granulated surface of the Sun, there sometimes are certain dark spots. These dark spots are identified either as Sunspots or

pores. Both phenomena are visible in fig. 1.2. Sunspots consist of a dark
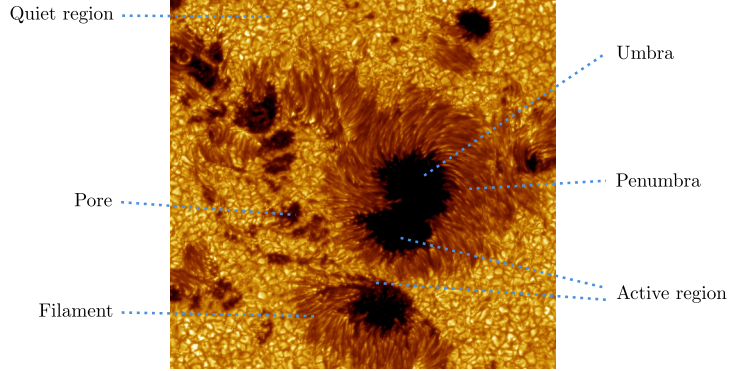
Figure 1.2: Sunspots in the active region 10030, observed on July 15, 2002. Image credit: NASA Goddard Space Flight Center, `https://www.flickr.com/photos/gsfc/5510488494`, accessed on November 07, 2023. Adapted by the author.

region, called the umbra, which is surrounded by a brighter region exhibiting a filament-like structure, known as the penumbra. Pores however just consist of an umbra without a penumbra. The filaments constituting the penumbra are usually approximately radially aligned around the Sunspot.

In fig. 1.3, a schematic representation of a Sunspot is shown. Sunspots form below the photosphere and then rise to the latter due to buoyancy. The relative darkness of the umbra within a Sunspot is due to lower tem-
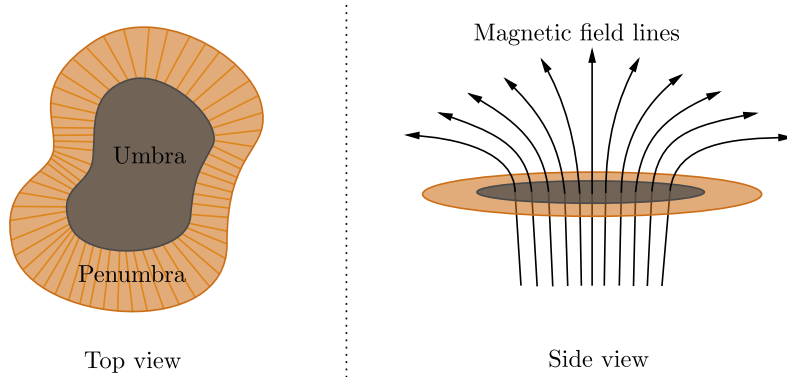
Figure 1.3: Schematic visualization of a Sunspot. Note, that areas on the Sun which show Sunspots and/or pores are called active regions, whereas other realms on the Sun's surface are called quiet regions.

peratures compared to the undisturbed photosphere. [Weigert et al., 2006,

3

p.142] suggest, that the explanation for the temperature difference is found in strong magnetic fields penetrating the photosphere in the region of the umbra and partially of the penumbra, as it is seen in the right panel of fig. 1.3. As known from classical electrodynamics, Maxwell's equations state that there are no magnetic monopoles, i.e. $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ for any magnetic field $\boldsymbol{B}$. Hence, [Stix, 2002, p.342] mentions, that if there is a magnetic field penetrating the photosphere from below to above, it must again penetrate from above to below, since magnetic field lines always have to be loops. This gives rise to the so-called bipolar structure/arrangement of Sunspots, which is indeed characteristic to many observed active regions containing Sunspots. However, [Stix, 2002, p.342] states, that a bipolar structure is not necessary to observed Sunspots; this is because sometimes the emerging magnetic flux is more concentrated where it leaves the photosphere than where it reenters; thus leading to only one visible Sunspot formed. A bipolar structure of Sunspots is schematically shown in fig. 1.4. Sunspots are particularly interesting for space
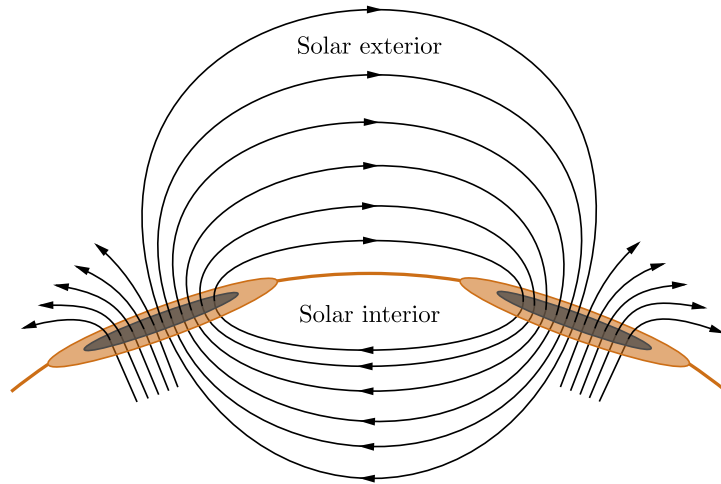
Figure 1.4: Schematic representation of a bipolar Sunspot arrangement.

weather, because they are correlated with eruptive events of the Sun, such as solar flares or coronal mass ejections. Both solar flares and coronal mass ejections belong to the category of solar eruptions, which are are tied to instabilities in the magnetic field configuration within the solar atmosphere.

1.1.2 Solar flares

As [Stix, 2002, p.432] remarks, solar flares originate around Sunspots and oftentimes occur right after a filament in the penumbra of a Sunspot suddenly disappears. A solar flare commonly lasts for about 30 minutes to 1 hour in total, where the first several minutes are dedicated to a fast rise in brightness of the solar flare. The brightness then slowly decreases for the remaining active time of the solar flare. A solar flare essentially is the ejection of a large packet of energy in the form of high-energy photons and some particle radiation in the form of electrons and protons. The electromagnetic radiation caused by a solar flare travels at the speed of light c, since photons are massless.

1.1.3 Coronal mass ejections

The term coronal mass ejection refers to a phenomenon, where the Sun ejects large amounts of mass into interplanetary space. As [Stix, 2002, p.436] points out, the ejected mass does not necessarily have to originate in the solar corona, as the term could misleadingly suggest; the term reflects the fact, that the mass is seen to be ejected by the corona. Often, coronal mass ejection follows a solar flare, but this does not need to be the case. A coronal mass ejection can be described as the sudden ejection of large amounts of energy by the Sun in the form of particle radiation (mass). Since these particles do have a mass, this particle radiation cannot travel at the speed of light.

1.2 Why bother about stellar atmosphere inversions?

The ultimate goal of space weather is to identify the triggers of solar flares and coronal mass ejections, since they pose potential threats to power grids on Earth, spacecraft and possibly astronauts themselves. In order to look for possible precursors causally related to flaring activity or coronal mass ejections of the Sun, the physical parameters describing the conditions in the Sun's atmosphere need to be known, e.g. the magnetic field strength, its direction, the temperature, pressure and so on. Since in-situ measurement is not possible on the Sun because of its extreme environment, those parameters need to be inferred by means of remote sensing techniques. Such techniques involve measuring Stokes profiles, which can be turned into atmospheric parameters by means of

an inversion of a stellar atmospheric model. Those models encompass the equations of radiative transfer; inversions of those models usually cannot be inverted analytically, depending on their degree of complexity. But there are some models, which can be inverted analytically, such as the Milne-Eddington approximation studied in this thesis.

Notation

At this point, a brief remark about the notation convention used in the following material is in order. Scalar quantities are denoted by an normal mathematics letter Q. For vectorial quantities, a bold mathematics letter $\boldsymbol{Q}$ is used. Finally, for tensorial quantities such as a matrix, a non-italic bold mathematics letter $\mathbf{Q}$ is made use of.

Chapter 2

Solar physics

All stars, and herewith also the Sun, are sources of gigantic amounts of energy. This energy is set free by means of atomic fusion processes happening within the star and is conveyed to an observer through electromagnetic or particle radiation. In order to infer properties of radiation source, e.g. a star, the properties of the emitted radiation need to be observed and analyzed using the theoretical framework of physics.

In the case of solar atmosphere research, there are three main tools for investigation available to the physicist; first, the theory of radiative transfer, which is concerned with the propagation of electromagnetic radiation in some medium. Second, there is spectroscopy, which is about how electromagnetic radiation interacts with matter as a function of wavelength. Third, polarimetry is necessary, which addresses how and why electromagnetic radiation is polarized as it is.

2.1 Radiative transfer

Because electromagnetic radiation is the dominant radiation source in solar atmosphere research, radiative transfer is integral to investigating the solar atmosphere.

2.1.1 Quantities in radiative transfer

In the theory of radiative transfer, the definition of several measurable quantities is necessary. If some quantity is a called a spectral quantity, this means that the quantity is measured at a certain wavelength λ or frequency ν; spectral quantity will be denoted by a subscript λ or ν.

The radiant energy E with units $[E] = $ J is defined as the energy

radiated away from some source. The radiant flux Φ with units $[\Phi] = $ Js$^{-1} = $ W however is defined as the radiant energy per unit time and is equal to the spectral flux integrated over all wavelengths. Furthermore, the spectral flux Φ_λ having units $[\Phi_\lambda] = $ Wm^{-1} is the radiant flux per unit wavelength. Additionally, the radiance I with units $[I] = $ Wsr^{-1}m^{-2} is the radiant flux per unit solid angle and per unit projected area; similarly the spectral radiance I_λ with units Wsr^{-1}m^{-3} is the radiance per unit wavelength. Finally, the irradiance F with units $[F] = $ Wm^{-2} is given by the radiant flux per unit projected area; it is also known as the flux density and is equal to the radiance integrated over all solid angles of a hemisphere. Associated to the radiance, there is also the spectral irradiance F_λ, with is nothing but the radiance per unit wavelength and therefore with units $[F_\lambda] = $ Wm^{-3}.

2.1.2 Thermal and thermodynamic equilibrium

The thermal equilibrium (TE) is defined as a state of a system, such that there are no macroscopic flows of heat. That is the case, whenever the temperature T is homogeneously constant throughout the system under consideration.

The thermodynamic equilibrium however is defined as a state of a system, in which there are no macroscopic flows of matter in addition to the thermal equilibrium condition.

Furthermore, there is the concept of local thermal equilibrium (LTE). A local thermal equilibrium is defined as a thermal equilibrium confined to local space, whereby local space refers to a parcel consisting of significantly more than one molecule, but much smaller in extent than the global system.

A system in local thermal equilibrium implies that a parcel within the considered system can locally be described as being in the state of a thermal equilibrium, which is to say, that Planck's law holds for this considered parcel.

2.1.3 Planck's law

The Planck law can be derived considering the physics of thermodynamics under the assumption of a thermal equilibrium for the considered system. It describes the spectral radiance emitted by a blackbody; that is some object, that perfectly absorbs and emits radiation of all wavelengths, and is at a thermal equilibrium with its environment. The Planck

law in both the frequency and wavelength domains is given by

$$B_\nu(T)\,\mathrm{d}\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_B T}}-1}\,\mathrm{d}\nu, \quad B_\lambda(T)\,\mathrm{d}\lambda = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_B T}}-1}\,\mathrm{d}\lambda, \quad (2.1)$$

where B_ν and B_λ are spectral radiances with units $[B_\nu] = \mathrm{Wm}^{-2}\mathrm{sr}^{-1}\mathrm{Hz}^{-1}$ and $[B_\lambda] = \mathrm{Wm}^{-3}\mathrm{sr}^{-1}$.

2.1.4 Blackbodies and Kirchhoff's law

A blackbody is a hypothetical object, that absorbs all radiation at all wavelengths completely. As a result, this body will at some point reach a thermal equilibrium with its environment, since absorption of radiation leads to a warming of the body which causes emission of thermal radiation and hence loss of energy and therefore heat. Eventually, absorbed and emitted energy will be balanced and thermal equilibrium is reached; in this state, the blackbody emits and absorbs radiation isotropically and exactly according to the Planck formula $B_\lambda(T)$[1].

Let the emissivity ε_λ and absorptivity A_λ be defined as

$$\varepsilon_\lambda(T) \doteq \frac{\mathcal{E}_\lambda(T)}{B_\lambda(T)}, \qquad A_\lambda \doteq \frac{\mathcal{A}_\lambda(T)}{B_\lambda(T)}, \qquad (2.2)$$

where $\mathcal{E}_\lambda(T)$ is the emittive spectral radiance and $\mathcal{A}_\lambda(T)$ is the absorptive spectral radiance of a body at a given temperature T. In thermal equilibrium, the emittive spectral radiance must be equal to the absorptive spectral radiance, for if this would not be the case, macroscopic flow of heat would result. From this consideration, Kirchhoff's law follows as

$$\mathcal{E}_\lambda(T) = \mathcal{A}_\lambda(T) \quad \Leftrightarrow \quad \varepsilon_\lambda(T) = A_\lambda(T), \qquad (2.3)$$

which in words states, that the emissivity is equal to the absorptivity at every particular wavelength λ for a system in thermal equilibrium at temperature T.

For a blackbody in thermal equilibrium, the absorptive spectral radiance is maximal, which is to say that $\mathcal{A}_\lambda(T) = B_\lambda(T)$ holds, leading to $A_\lambda(T) = 1$. Since the blackbody is in thermal equilibrium, the absorptive spectral radiance is equal to the emissive spectral radiance and hence $\mathcal{E}_\lambda(T) = B_\lambda(T)$ must hold, wherefrom $\varepsilon_\lambda(T) = 1$ follows. Therefore, a

---

[1]Using the transformation relation $\nu = \frac{c}{\lambda}$ and $\int_0^\infty B_\lambda\,\mathrm{d}\lambda \stackrel{!}{=} \int_0^\infty B_\nu\,\mathrm{d}\nu$, one can go back and forth between the formulation of Planck's law in the wavelength or frequency domain.

blackbody in thermal equilibrium can be described mathematically by means of the relation

$$\varepsilon_\lambda(T) = A_\lambda(T) = 1. \tag{2.4}$$

2.1.5 Rayleigh-Jeans and Wien approximations

Consider the Planck law eq. (2.1). The Rayleigh-Jeans law is the Planck law for the limit of low energies, whereas the Wien law is the Planck law for the limit of high energies. For low energies $h\nu/k_B T \to 0$, the Rayleigh-Jeans law

$$B_\nu(T)\, \mathrm{d}\nu = \frac{2\nu^2}{c^2} k_B T\, \mathrm{d}\nu, \quad \frac{h\nu}{k_B T} \to 0 \tag{2.5}$$

is obtained, whereas for high energies $\nu \to \infty$, the Wien law

$$B_\nu(T)\, \mathrm{d}\nu = \frac{2h\nu^3}{c^2} e^{-\frac{h\nu}{k_B T}}\, \mathrm{d}\nu, \quad \nu \to \infty \tag{2.6}$$

follows.

In order to derive the Rayleigh-Jeans law from the Planck law, one has to do the substitution $x = h\nu/k_B T$ in the Planck law and perform a Taylor expansion of the resulting expression around $x_0 = 0$. This corresponds to the case $h\nu/k_B T \to 0$.

Moreover, to arrive at the Wien law, the substitution $y = e^{-h\nu/k_B T}$ in the Planck law has to be made and a Taylor expansion of the resulting expression around $y_0 = 0$ has to be performed. This corresponds to the case $\nu \to \infty$.

2.1.6 Stefan-Boltzmann law

Consider some isotropic spectral radiance I_λ along a ray, flowing through a surface A as depicted in fig. 2.1. If one calculates the flux of energy per wavelength through the surface A, one gets a quantity known as the spectral irradiance F_λ, also known as spectral flux density. This can be done by integrating the normal component $I_{\lambda,n} = I_\lambda \cos(\theta)$ of I_λ over the whole half-sphere, that is over the parameter space $\theta \in [0, \pi/2]$ and $\varphi \in [0, 2\pi]$. Furthermore, the total flux F can be obtained from F_λ by integrating over all wavelengths. Therefore, the irradiance (flux density) F of isotropic electromagnetic radiation can be calculated by means of the integral

$$F = \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi/2} \sin(\theta)\cos(\theta)\, \mathrm{d}\theta \int_0^\infty I_\lambda\, \mathrm{d}\lambda = \pi \int_0^\infty I_\lambda\, \mathrm{d}\lambda, \tag{2.7}$$

10
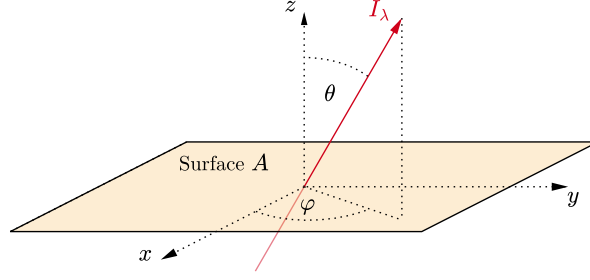
Figure 2.1: Visualization for a derivation of the Stefan-Boltzmann law.

where the factor $\sin(\theta)$ originates in the transformation from cartesian to polar coordinates. If one now assumes, that I_λ describes a system in thermal equilibrium, $I_\lambda = B_\lambda(T)$ follows and hence the Stefan-Boltzmann law

$$F = \tilde{\sigma} T^4 \tag{2.8}$$

for a system in thermal equilibrium.

The Stefan-Boltzmann law is a useful means to determine the effective temperature of stars, such as that of the Sun $T_\odot$. Measuring the solar constant $S_\odot \approx 1361\,\mathrm{Wm^{-2}}$, i.e. how much solar power per unit square the Earth receives on average, one can find out the effective temperature $T_\odot$ of the Sun by means of relating the two quantities as

$$S_\odot = \frac{L_\odot}{4\pi d_{\odot,E}^2} = \tilde{\sigma} T_\odot^4 \left(\frac{R_\odot}{d_{\odot,E}}\right)^2 \quad \Rightarrow \quad T_\odot = \left(\frac{S_\odot}{\tilde{\sigma}}\right)^{1/4} \left(\frac{d_{\odot,E}}{R_\odot}\right)^{1/2}, \tag{2.9}$$

which gives $T_\odot = 5777\,\mathrm{K} \approx 5800\,\mathrm{K}$ as stated in the introduction.

2.1.7 Radiative transfer equation

The radiative transfer equation is usually given in terms of spectral irradiance I_λ. The basic two forms of the radiative transfer equation shall now be briefly derived, following the material in [Rutten, 2015, pp.27-40].

Consider as a first step fig. 2.2. A beam of spectral irradiance I_λ passes through a slab of matter with thickness $\mathrm{d}z$, along a path s with length $\mathrm{d}s$. The question now is, how the spectral irradiance I_λ after the slab can be calculated. In order to quantify the spectral irradiance after passing through the slab of matter, two contributions due to electromagnetic interaction between radiation and matter within the slab have to be considered, namely extinction and emission. Extinction is loss of radiative energy along a path s due to scattering or absorption and it is accounted

for by the extinction coefficient α_λ with units $[\alpha_\lambda] = \mathrm{m}^{-1}$. Emission is gain of radiative energy along a path s due to spontaneous emission of photons caused by decay of excited states in molecules or atoms within the slab of matter; it is quantified by the emission coefficient j_λ having units $[j_\lambda] = \mathrm{W sr}^{-1}\mathrm{m}^{-4}$.
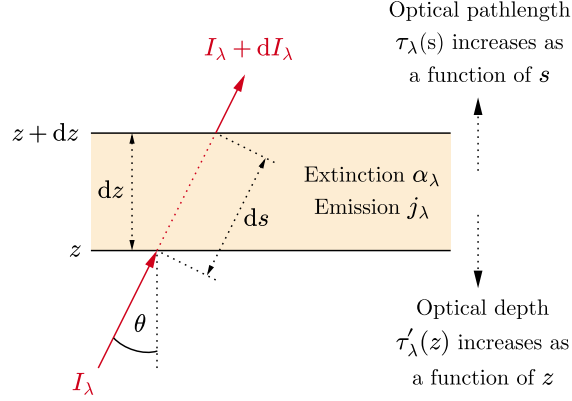
Figure 2.2: Illustration for the derivation of the radiative transfer equation. $[I_\lambda] = \mathrm{W\,sr}^{-1}\,\mathrm{m}^{-3}$ denotes the spectral radiance measured along a ray of inclination θ with respect to a horizontal plane in the coordinate system under consideration.

The difference $\mathrm{d}I_\lambda$ in spectral irradiance along the path of propagation must be the sum of the extinction and emission contributions. Considering the units of the extinction and emission coefficients, these contributions can be written as $-\alpha_\lambda I_\lambda\,\mathrm{d}s$ for the loss due to extinction and as $j_\lambda\,\mathrm{d}s$ for the gain due to emission. Herewith, one arrives at

$$\mathrm{d}I_\lambda = j_\lambda\,\mathrm{d}s - \alpha_\lambda I_\lambda\,\mathrm{d}s \quad \Leftrightarrow \quad \frac{\mathrm{d}I_\lambda}{\mathrm{d}s} = j_\lambda - \alpha_\lambda I_\lambda. \tag{2.10}$$

Defining the optical pathlength τ_λ by means of $\mathrm{d}\tau_\lambda \doteq \alpha_\lambda\,\mathrm{d}s$ and furthermore the so-called source function S_λ as $S_\lambda \doteq \frac{j_\lambda}{\alpha_\lambda}$, this equation can also be written as

$$\frac{\mathrm{d}I_\lambda(\tau_\lambda)}{\mathrm{d}\tau_\lambda} = S_\lambda(\tau_\lambda) - I_\lambda(\tau_\lambda), \tag{2.11}$$

which is called the transport equation in differential form. Defining the optical depth τ'_λ as $\mathrm{d}\tau'_\lambda \doteq \alpha_\lambda\,\mathrm{d}z = -\alpha_\lambda \cos(\theta)\,\mathrm{d}s = -\cos(\theta)\,\mathrm{d}\tau_\lambda$, the transport equation can also be written in terms of optical depth rather than optical pathlength, namely

$$\cos(\theta)\frac{\mathrm{d}I_\lambda(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} = I_\lambda(\tau'_\lambda) - S_\lambda(\tau'_\lambda). \tag{2.12}$$

The change in sign is due to the fact, that the optical depth τ'_λ is measured from the point of view of an observer looking vertically upon the matter, from which the radiation is emerging from, while the optical pathlength τ_λ measures the distance of propagation in the direction of the ray. Adopting the boundary conditions $\tau_\lambda(0) = \tau'_\lambda(0) = 0$, the optical pathlength $\tau_\lambda(S)$ at a position $s = S$ aswell as the optical depth $\tau_\lambda(Z)$ at the position $z = Z$ are e given by integration as

$$\tau_\lambda(S) = \int_{\tau_\lambda(0)}^{\tau_\lambda(S)} \mathrm{d}\tau_\lambda = \int_0^S \alpha_\lambda(s)\,\mathrm{d}s \qquad (2.13)$$

and

$$\tau'_\lambda(Z) = \int_{\tau'_\lambda(0)}^{\tau'_\lambda(Z)} \mathrm{d}\tau'_\lambda = \int_0^Z \alpha_\lambda(z)\,\mathrm{d}z. \qquad (2.14)$$

Formal solution of the radiative transfer equation

Both of the above formulations for the transfer equation in differential form can also be written in an integral form, as it is demonstrated as follows; consider for that matter fig. 2.3.
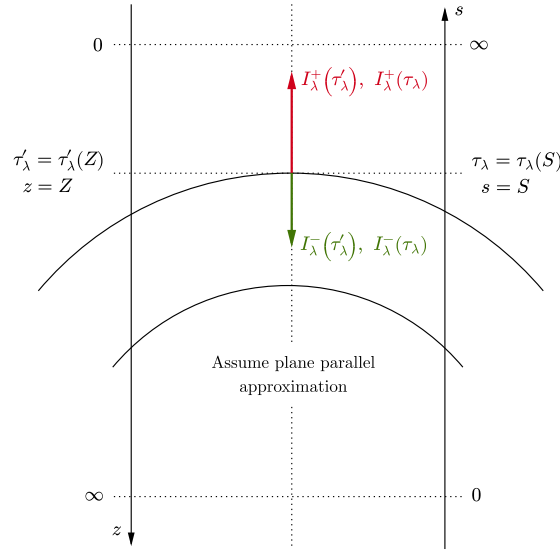
Figure 2.3: Illustration of coordinate systems for optical pathlength τ_λ parametrized by s and optical depth τ'_λ parametrized by z. Note, that the optical pathlength axis does not necessarily need to be parallel to the optical depth axis, the factor $\cos(\theta)$ corrects for the case where these axes are not parallel. The curved lines represent layers of a stellar atmosphere, which are locally assumed as plane parallel.

In order to derive integral forms of the transfer equation with respect to the optical pathlength τ_λ, the derivative

$$\frac{d}{d\tau_\lambda}\left[e^{\tau_\lambda}I_\lambda(\tau_\lambda)\right] = e^{\tau_\lambda}\left[I_\lambda(\tau_\lambda) + \frac{dI_\lambda(\tau_\lambda)}{d\tau_\lambda}\right] = e^{\tau_\lambda}S_\lambda(\tau_\lambda) \qquad (2.15)$$

is first calculated. Finding the outward spectral radiance $I_\lambda^+(\tau_\lambda)$ at a given optical pathlength τ_λ requires integration of eq. (2.15) from 0 to τ_λ, namely

$$\int_0^{\tau_\lambda}\frac{d}{dt}\left[e^t I_\lambda^+(t)\right]\,dt = e^{\tau_\lambda}I_\lambda^+(\tau_\lambda) - I_\lambda^+(0) = \int_0^{\tau_\lambda} e^t S_\lambda(t)\,dt. \qquad (2.16)$$

From this expression it follows by multiplication of the above equation with $e^{-\tau_\lambda}$ and rearrangement of terms, that the outward spectral radiance at a certain optical pathlength τ_λ from the origin of the coordinate system under consideration is given by

$$I_\lambda^+(\tau_\lambda) = I_\lambda^+(0)e^{-\tau_\lambda} + \int_0^{\tau_\lambda} e^{t-\tau_\lambda}S_\lambda(t)\,dt, \qquad (2.17)$$

which is what is commonly called the formal solution of the transfer equation with respect to optical pathlength. If one wants to calculate the inward spectral radiance $I_\lambda^-(\tau_\lambda)$ however, integration of eq. (2.15) is required to be carried from ∞ to τ_λ, such that

$$-\int_{\tau_\lambda}^\infty \frac{d}{dt}\left[e^t I_\lambda^-(t)\right]\,dt = e^{\tau_\lambda}I_\lambda^-(\tau_\lambda) = -\int_{\tau_\lambda}^\infty e^t S_\lambda(t)\,dt \qquad (2.18)$$

results. Multiplying this equation by $e^{-\tau_\lambda}$ results in the inward spectral radiance at a given optical pathlength τ_λ is calculable as

$$I_\lambda^-(\tau_\lambda) = -\int_{\tau_\lambda}^\infty e^{t-\tau_\lambda}S_\lambda(t)\,dt. \qquad (2.19)$$

For both formulations of $I_\lambda^+(\tau_\lambda)$ and $I_\lambda^-(\tau_\lambda)$ the boundary conditions

$$I_\lambda^-(\infty) = 0, \quad e^t S_\lambda(t) \xrightarrow{t\to\infty} 0 \qquad (2.20)$$

were used, which mean that at the position $\tau_\lambda = \infty$ there is no radiation directed towards the source moreover that the source function $S_\lambda(t)$ decreases faster with increasing optical pathlength than e^{-t}.

14

The integral representations of the transfer equation with respect to the optical depth τ'_λ can also be obtained in a similar way; for this purpose, the derivative

$$\frac{\mathrm{d}}{\mathrm{d}\tau'_\lambda}\left[\cos(\theta)e^{-\frac{\tau'_\lambda}{\cos(\theta)}}I_\lambda(\tau'_\lambda)\right] = e^{-\frac{\tau'_\lambda}{\cos(\theta)}}\left[\cos(\theta)\frac{\mathrm{d}I_\lambda(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} - I_\lambda(\tau'_\lambda)\right]$$
$$= -e^{-\frac{\tau'_\lambda}{\cos(\theta)}}S_\lambda(\tau'_\lambda) \tag{2.21}$$

is needed. Obtaining an expression for the outward spectral radiance $I_\lambda^+(\tau'_\lambda)$ at a given optical depth τ'_λ requires integration of eq. (2.21) from τ'_λ to ∞, namely

$$\int_{\tau'_\lambda}^{\infty}\frac{\mathrm{d}}{\mathrm{d}t}\left[\cos(\theta)e^{\frac{t}{\cos(\theta)}}I_\lambda^+(t)\right]\mathrm{d}t = -\cos(\theta)e^{-\frac{\tau'_\lambda}{\cos(\theta)}}I_\lambda^+(\tau'_\lambda)$$
$$= -\int_{\tau'_\lambda}^{\infty}e^{-\frac{t}{\cos(\theta)}}S_\lambda(t)\,\mathrm{d}t. \tag{2.22}$$

Dividing by $\cos(\theta)$ and multiplying by $e^{\frac{\tau'_\lambda}{\cos(\theta)}}$ yields

$$I_\lambda^+(\tau'_\lambda,\theta) = \frac{1}{\cos(\theta)}\int_{\tau'_\lambda}^{\infty}e^{\frac{\tau'_\lambda-t}{\cos(\theta)}}S_\lambda(t)\,\mathrm{d}t \tag{2.23}$$

as the expression for the outward spectral radiance at a given optical depth τ'_λ. Moreover, if the inward spectral radiance $I_\lambda^-(\tau'_\lambda)$ is to be calculated, integration of eq. (2.21) is required from τ'_λ to zero. Written down, this amounts to

$$-\int_0^{\tau'_\lambda}\frac{\mathrm{d}}{\mathrm{d}t}\left[e^{-\cos(\theta)\frac{t}{\cos(\theta)}}I_\lambda^-(t)\right] = -\cos(\theta)e^{-\frac{\tau'_\lambda}{\cos(\theta)}}I_\lambda^-(\tau'_\lambda)$$
$$= \int_0^{\tau'_\lambda}e^{-\frac{t}{\cos(\theta)}}S_\lambda(t)\,\mathrm{d}t. \tag{2.24}$$

Finally, dividing again by $-\cos(\theta)$ and multiplying by $e^{\frac{\tau'_\lambda}{\cos(\theta)}}$ leads to the inward spectral radiance at a given optical depth τ'_λ as

$$I_\lambda^-(\tau'_\lambda,\theta) = -\frac{1}{\cos(\theta)}\int_0^{\tau'_\lambda}e^{\frac{\tau'_\lambda-t}{\cos(\theta)}}S_\lambda(t)\,\mathrm{d}t. \tag{2.25}$$

Again, for both expressions $I_\lambda^+(\tau'_\lambda,\theta)$ and $I_\lambda^-(\tau'_\lambda,\theta)$ boundary conditions had to be applied, namely

$$I_\lambda^-(0,\theta) = 0, \quad e^{-t}S_\lambda(t) \xrightarrow{t\to\infty} 0. \tag{2.26}$$

These conditions mean, that at the location $\tau'_\lambda = 0$, there is no radiation directed towards the source; furthermore the source function $S_\lambda(t)$ is required not grow faster than e^t as optical depth τ'_λ increases.

2.1.8 Solutions for a homogeneous medium

In the case, where radiative transfer through a homogeneous material is considered, the extinction coefficient α_λ aswell as the emission coefficient j_λ are constant everywhere in the material. Therefore, the source function does not depend on any spatial parameter either.

Taking the optical pathlength τ_λ as the spatial parameter, the outward spectral radiance $I_\lambda^+(\tau_\lambda)$ can be calculated using eq. (2.17) and a constant source function $S_\lambda(\tau_\lambda) = \text{const}$. Since in that case the source function does not integrate with respect to τ_λ, performing the integration yields

$$
\begin{aligned}
I_\lambda^+(\tau_\lambda) &= I_\lambda^+(0)e^{-\tau_\lambda} + S_\lambda \int_0^{\tau_\lambda} e^{t-\tau_\lambda}\,\mathrm{d}t \\
&= I_\lambda^+(0)e^{-\tau_\lambda} + S_\lambda\left(1 - e^{-\tau_\lambda}\right),
\end{aligned}
\tag{2.27}
$$

which is the solution of the radiative transfer equation with respect to optical pathlength τ_λ in the case of a homogeneous medium. There are two approximations for this solution, namely for an optically thick or thin material.

The first approximation for an *optically thick* material determined by the requirement $\tau_\lambda \gg 1$, which is to say, that the optical pathlength through the medium is very long. In this case, the exponential factors $e^{-\tau_\lambda}$ approach zero, as $\lim_{\tau_\lambda \to \infty} e^{-\tau_\lambda} = 0$ holds. This is to say, that the solution to the radiative transfer equation in the optically thick case is given by

$$
I_\lambda^+(\tau_\lambda) \approx S_\lambda, \quad \tau_\lambda \gg 1, \quad S_\lambda = \text{const.}
\tag{2.28}
$$

Note, that the outward spectral radiance in this case does not vary with optical pathlength τ_λ.

The second approximation for an *optically thin* material is given by the condition $\tau_\lambda \ll 1$. Using a Taylor series expansion of $e^{-\tau_\lambda}$ up to first order in τ_λ for $\tau_\lambda \ll 1$ yields $e^{-\tau_\lambda} \approx 1 - \tau_\lambda$, in the case where the optical pathlength τ_λ is indeed very short. Inserting this expression into eq. (2.27), the solution to the radiative transfer equation in the optically thin case is accounted for by

$$
I_\lambda^+(\tau_\lambda) \approx I_\lambda^+(0) + \left[S_\lambda - I_\lambda^+(0)\right]\tau_\lambda \quad \tau_\lambda \ll 1, \quad S_\lambda = \text{const.}
\tag{2.29}
$$

Note, that the outward spectral radiance in this case varies linearly with optical pathlength τ_λ.

2.1.9 The Eddington-Barbier approximation

The Eddington-Barbier approximation is a more realistic simplification of the radiative tranfer equation than the assumption of a homogeneous medium, for it requires a source function S_λ that varies as a power series in τ'_λ rather than being constant throughout the medium. In order to derive the exact form of this approximation, optical depth τ'_λ is used rather than optical pathlength τ_λ, because in this way one arrives at an expression for the emergent spectral radiance, which is an actually measurable quantity, rather than obtaining an expression for the spectral radiance at an invisible layer within the medium. The optical depth τ'_λ at a coordinate of $z = Z$ is calculated by means of integrating the extinction coefficient $\alpha_\lambda(z)$ as $\tau'_\lambda(Z) = \int_0^Z \alpha_\lambda(z)\,\mathrm{d}z$, where the position $z = 0$ refers to the outermost layer of the considered medium - e.g. the position of an observer and $z = \infty$ denotes the innermost layer of said medium; consider fig. 2.3 for a visual clarification of that matter.

Now, the outward (emergent) spectral radiance at a position $z = Z$ and a local angle of θ with respect to the normal vector of the plane parallel medium layers is accounted for by eq. (2.23). Assuming that a considered source function $S_\lambda(\tau'_\lambda)$ satisfies the necessary requirements to be written as a power series, one can write down such an expansion in powers of τ'_λ as

$$S_\lambda(\tau'_\lambda) = \sum_{n=0}^{\infty} a_n \tau'^n_\lambda \tag{2.30}$$

and plug this expression into eq. (2.23). Carrying out this substitution yields

$$
\begin{aligned}
I^+_\lambda(\tau'_\lambda, \theta) &= \frac{1}{\cos(\theta)} \int_{\tau'_\lambda}^{\infty} e^{\frac{\tau'_\lambda - t}{\cos(\theta)}} S_\lambda(t)\,\mathrm{d}t \\
&= e^{\frac{\tau'_\lambda}{\cos(\theta)}} \sum_{n=0}^{\infty} a_n [\cos(\theta)]^n \int_{\frac{\tau'_\lambda}{\cos(\theta)}}^{\infty} e^{-x} x^n\,\mathrm{d}x.
\end{aligned}
\tag{2.31}
$$

Using the mathematical identity $\int_0^{\infty} e^{-x} x^n\,\mathrm{d}x = n!$ for $n \in \mathbb{N}$, an evaluation of the above expression at $\tau'_\lambda = 0$ brings forth

$$
\begin{aligned}
I^+_\lambda(\tau'_\lambda = 0, \theta) &= \sum_{n=0}^{\infty} a_n [\cos(\theta)]^n \int_0^{\infty} e^{-x} x^n\,\mathrm{d}x \\
&= \sum_{n=0}^{\infty} a_n [\cos(\theta)]^n\, n! \approx a_0 + a_1 \cos(\theta).
\end{aligned}
\tag{2.32}
$$

Recalling the above given power series form for the source function $S_\lambda(\tau'_\lambda)$, an evaluation at $\tau'_\lambda = \cos(\theta)$ yields

$$S_\lambda(\tau'_\lambda = \cos(\theta)) = \sum_{n=0}^{\infty} a_n \left[\cos(\theta)\right]^n = a_0 + a_1 \cos(\theta) + \mathcal{O}[\cos^2(\theta)]$$
$$\approx a_0 + a_1 \cos(\theta)$$
(2.33)

and therefore one obtains a relation for the outward spectral radiance $I_\lambda^+(\tau'_\lambda = 0, \theta)$ in the so-called Eddington-Barbier approximation, namely

$$I_\lambda^+(\tau'_\lambda = 0, \theta) \approx S_\lambda(\tau'_\lambda = \cos(\theta)).$$
(2.34)

If a ray propagates in a vertical direction with respect to the plane parallel layers of the medium against an observer, $\cos(\theta = 0) = 1$ and therefore

$$I_\lambda^+(\tau'_\lambda = 0, \theta = 0) \approx S_\lambda(\tau'_\lambda = 1)$$
(2.35)

hold. If the source function S_λ is given exactly as an affine function in τ'_λ, namely $S_\lambda(\tau'_\lambda) = a_0 + a_1\tau'_\lambda$, then the above two relations do not only hold approximately, but exactly, as it follows directly from eq. (2.32) and eq. (2.33).

2.2 Spectral line formation

In order to provide a qualitative understanding of spectral line formation, one can use the Eddington-Barbier relation with $\theta = 0$, that is

$$I_\lambda^+(\tau'_\lambda = 0) \approx S_\lambda(\tau'_\lambda = 1),$$
(2.36)

which is an approximative solution to the radiative transfer equation

$$\cos(\theta)\frac{\mathrm{d}I_\lambda(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} = I_\lambda(\tau'_\lambda) - S_\lambda(\tau'_\lambda).$$
(2.37)

Recall, that $S_\lambda = \frac{j_\lambda}{\alpha_\lambda}$ with j_λ the emission and α_λ the extinction coefficient is the source function and of optical depth τ'_λ is defined as $\tau'_\lambda(Z) = \int_0^Z \alpha_\lambda(z)\,\mathrm{d}z$.

Electrons in atoms and molecules as described by the Schrödinger equation have discrete energy states; these states encompass discrete rotational, vibrational and orbital energy states. Note, that electromagnetic radiation interacts with atoms atoms and molecules in such a way, that those energy states are either excited, deexcited or even ionized.

Furthermore, electromagnetic radiation can also scatter in matter. Excitation of discrete atomic or molecular energy states by photons of suitable wavelengths, combined with ionization and scattering effects, defines the extinction coefficient α_λ, while deexcitation, combined with recombination effects determines the emission coefficient j_λ. Hereby, excitation is to mean, that an electron of energy E within an atom or molecule transitions to a quantum state of higher energy $E + \Delta E$ by means of absorbing a photon of corresponding wavelength $\lambda_\gamma = \frac{hc}{\Delta E}$; hence deexcitation refers to a transition of an electron within an atom or molecule in energy state $E + \Delta E$ to the energy state E by means of emitting a photon of wavelength E_γ. Consider for an illustration of these processes fig. 2.4.
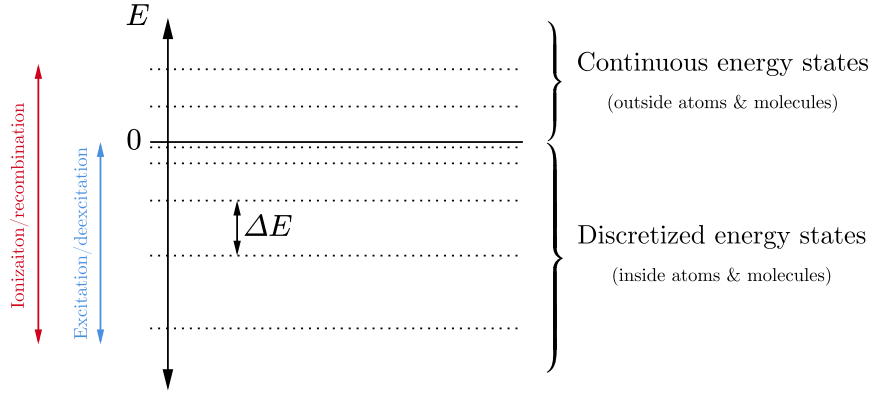
Figure 2.4: Illustration of discretized and continuous energy states in matter. Note, that the dotted lines for negative energies represent possible energy states within an atom or molecule (bound states), whereas there is actually a continuum of dotted lines for positive energies.

Consider now fig. 2.5. Consider a hypothetical atom, which is such that it contains only one electron and which has only two distinct energy states available to the electron. These energy states are separated by the energy ΔE corresponding to the wavelength $\lambda_\gamma = hc/\Delta E$. Assume further, that there is some hypothetical medium consisting of only those hypothetical atoms along with a electromagnetic radiation source; and that scattering is negligible. At precisely the wavelength λ_γ, this hypothetical medium will perfectly absorb all photons, as long as there are still some energy state changes available. This is to say, that there will also be some spike in the curve of α_λ, the extinction coefficient, as it is shown in the top left panel of fig. 2.5 - a spectral line is formed. Now, as such, the spectral line will be just a delta function; in reality however different
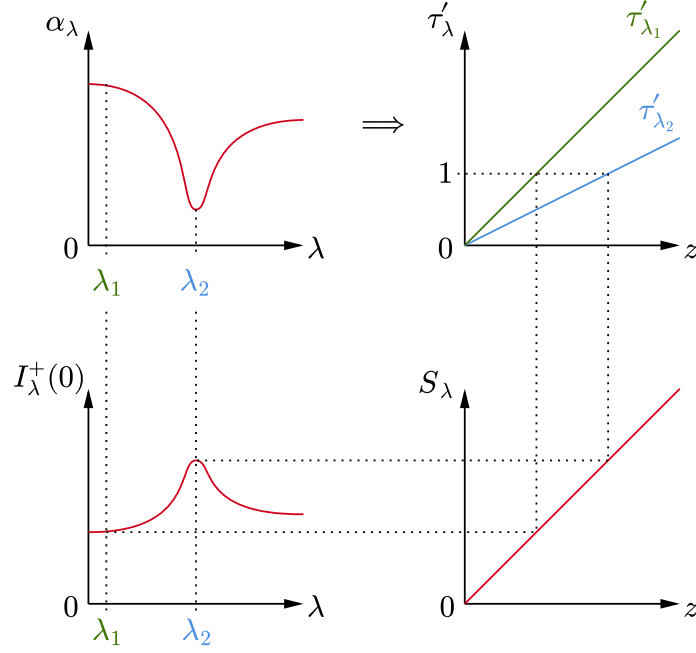
Figure 2.5: Illustration of a spectral line formation using the Eddington-Barbier relation as an explanatory tool. Note, that in this picture one assumes, that the extinction coefficient α_λ is independent of depth z, and furthermore, that the source function $S_\lambda = S$ is independent of wavelength and furthermore linear in z. This means, that the Eddington-Barbier relation does not only hold approximatively, but exactly. Furthermore, the optical depth is then just given as $\tau'_\lambda(Z) = \int_0^Z \alpha_\lambda \, \mathrm{d}z = \alpha_\lambda Z$. Figure inspired by [Rutten, 2015, p.38].

broadening mechanisms such as Doppler or Lorentz broadening of spectral lines apply, furthermore there are many more than just two energy levels available to electrons in a real medium. Therefore, in fig. 2.5 the extinction coefficient α_λ is plotted for some arbitrary, non-hypothetical medium. One can see, that the absorption line is centered as λ_2. Assuming, that the extinction coefficient is not dependent upon height z and taking into account the definition of the optical depth $\tau'_\lambda(Z) = \int_0^Z \alpha_\lambda \, \mathrm{d}z$, one deduces $\tau'_{\lambda_1}(z) = \alpha_{\lambda_1} z$ and $\tau'_{\lambda_2}(z) = \alpha_{\lambda_2} z$. Since $\alpha_{\lambda_1} \gg \alpha_{\lambda_2}$ clearly τ'_{λ_1} must be a steeper function of z than τ'_{λ_2}. This is to say, that the medium is more opaque at wavelength λ_2 than at wavelength λ_1. Taking into account now the Eddington-Barbier relation, which states, that the source function S_λ at optical depth $\tau'_\lambda = 1$ is equal to the outward spectral radiance $I_\lambda^+(0)$ one arrives at a spectral line in emission for an observer measuring $I_\lambda^+(0)$. If however, α_λ would have a spike upwards, the spectral line would appear in absorption for an observer at $\tau'_\lambda = 0$. A

study of spectral line properties and an analysis, which parameters and how they influence spectral line formation can lead to insights about the medium, in which they form. Using the theory of radiative transfer and especially simplified atmospheric models of stars for example, one can infer physical properties of an atmosphere, simply by investigating one or more spectral lines formed in the atmosphere under consideration.

2.3 Polarimetry

The subject of polarimetry deals with the polarization properties of electromagnetic waves and how they are affected by the properties of the media they are propagating in. In particular, the presence of a magnetic field in the propagating medium of electromagnetic waves leads to the so-called Zeeman splitting of spectral lines, from which the presence and magnitude of a magnetic field can be deduced. Polarimetry furthermore links observable quantities like the Stokes parameters to the parameters of spectral lines, thus providing insight to the medium, where those spectral lines formed. In the context of stellar atmospheres, polarimetry provides the link between the observable Stokes parameters and the parameters defining the stellar atmosphere, such as the magnetic field strength, inclination and azimuth.

2.3.1 Zeeman effect

The Zeeman effect for strong external magnetic fields allows inference about the external magnetic field strength. The presence of such an external magnetic field leads to the splitting of energy levels in atoms and molecules. This is to say, that also one spectral line splits into several spectral lines. According to [del Toro Iniesta, 2003, p.123], the Zeeman wavelength splitting λ_B can be quantified by means of

$$\lambda_B = \frac{\lambda^2 \nu_L}{c}, \quad \nu_L = \frac{e|\boldsymbol{B}|}{4\pi mc}, \tag{2.38}$$

where ν_L is the Larmor frequency, e is the elementary charge, m is the particle mass, λ is the wavelength of a spectral line in an inertial frame and $\boldsymbol{B}$ is the external magnetic field giving rise to the Zeeman splitting. These formulae can be derived by a rigorous calculation involving perturbation theory in quantum mechanics.

2.3.2 The Stokes parameters and Stokes vector

Consider a monochromatic electromagnetic wave $\boldsymbol{E}$. In general, $\boldsymbol{E}$ is a function of position $\boldsymbol{r} \in \mathbb{R}^3$ in euclidean space and time t, such that $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{r}, t)$. Without loss of generality, the propagation direction of the monochromatic electromagnetic wave can be defined in z-direction, hence [Stix, 2002, p.121] gives a monochromatic wave or arbitrary polarization state as

$$\boldsymbol{E}(\boldsymbol{r}, t) = \begin{pmatrix} E_x \cos(kz - \omega t) \\ E_y \cos(kz - \omega t + \xi) \\ 0 \end{pmatrix}, \qquad k = \frac{2\pi}{\lambda}. \tag{2.39}$$

The quantity ξ is the phase difference between the amplitude of the electromagnetic wave in x-direction and y-direction, which together with E_x and E_y therefore accounts for the polarization state of the wave. As a side note, notice that the dispersion relation $\omega = \omega(k)$ for a wave propagating in vacuum can be given as $\omega(k) = 2\pi\nu = ck$. For the above specified electromagnetic wave, the Stokes parameters are defined as

$$\boldsymbol{I} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} E_x^2 + E_y^2 \\ E_x^2 - E_y^2 \\ 2 E_x E_y \cos(\xi) \\ 2 E_x E_y \sin(\xi) \end{pmatrix}, \tag{2.40}$$

as given in [Stix, 2002, p.121]. Hereby, the identity $I^2 = Q^2 + U^2 + V^2$ holds, as it is shown by the calculation

$$\begin{aligned} Q^2 + U^2 + V^2 &= E_x^4 - 2E_x^2 E_y^2 + E_y^4 + 4E_x^2 E_y^2 \left[\cos^2(\xi) + \sin^2(\xi) \right] \\ &= E_x^4 + 2E_x^2 E_y^2 + E_y^4 = (E_x^2 + E_y^2)^2 = I^2. \end{aligned} \tag{2.41}$$

Since these parameters were defined for a monochromatic electromagnetic wave, the Stokes parameters are in general functions of wavelength; the wavelength dependency of $\boldsymbol{I} = \boldsymbol{I}(\lambda)$ shall henceforth be indicated by the subscript λ, such that the Stokes parameters are denoted by

$$\boldsymbol{I}(\lambda) = \boldsymbol{I}_\lambda = (I_\lambda, Q_\lambda, U_\lambda, V_\lambda)^\top, \tag{2.42}$$

which is called the Stokes vector, whereby also $I_\lambda^2 = Q_\lambda^2 + U_\lambda^2 + V_\lambda^2$ holds. In fig. 2.6, a schematic representation of the definition of the Stokes parameters (vector) is given.

2.3.3 The generalized radiative transfer equation

In the preceding material, the radiative transfer equation was already derived and solved for the spectral radiance $I_\lambda(\tau_\lambda', \theta)$ as a function of the

22

Stokes vector $\boldsymbol{I}_\lambda = (I_\lambda, Q_\lambda, U_\lambda, V_\lambda)^\top$
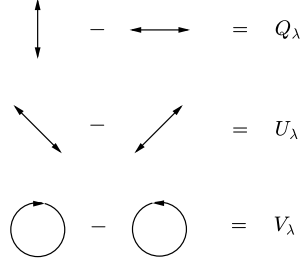
Figure 2.6: Schematic representation of the general Stokes vector $\boldsymbol{I}_\lambda$ containing the Stokes parameters I_λ, Q_λ, U_λ and V_λ. The observer is looking into negative z-direction with respect to the electromagnetic wave as defined by eq. (2.39). Figure inspired by [Degl'Innocenti, 2005, p.17].

continuum optical depth τ'_λ and local inclination θ of an electromagnetic radiation ray. This radiative transfer equation was given by

$$\cos(\theta)\frac{\mathrm{d}I_\lambda(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} = I_\lambda(\tau'_\lambda) - S_\lambda(\tau'_\lambda). \tag{2.43}$$

This equation can be generalized to account not only for changes of the spectral radiance (intensity) I_λ, but for changes of every Stokes parameter $\boldsymbol{I}_\lambda = (I_\lambda, Q_\lambda, U_\lambda, V_\lambda)^\top$. This can be done by means of introducing a propagation matrix $\mathbf{K}_\lambda(\tau'_\lambda) = \mathbb{1} + \mathbf{N}_\lambda(\tau'_\lambda)$, where $\mathbf{N}_\lambda(\tau'_\lambda)$ accounts for extinction and emission properties of the medium giving rise to spectral line formation; and where $\mathbb{1}$ is the unity matrix. Using the wavelength-dependent propagation matrix $\mathbf{K}_\lambda$, the generalized radiative transfer equation for the Stokes vector can be written as

$$\frac{\mathrm{d}\boldsymbol{I}_\lambda(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} = \mathbf{K}_\lambda(\tau'_\lambda)\left[\boldsymbol{I}_\lambda(\tau'_\lambda) - \boldsymbol{S}_\lambda(\tau'_\lambda)\right], \tag{2.44}$$

as given by [del Toro Iniesta, 2003, p.150]. Note that the factor of $\cos(\theta)$ is absorbed in propagation matrix $\mathbf{K}_\lambda$ in the above equation. In the above formulation of the generalized radiative transfer equation also the source function $S_\lambda(\tau'_\lambda)$ was generalized to the source function vector $\boldsymbol{S}_\lambda(\tau'_\lambda) = S_\lambda(\tau'_\lambda)(1,0,0,0)^\top$. It is evident from this form of the equation, that the unity matrix $\mathbb{1}$ of the propagation matrix takes care of the extinction and emission in the continuum, while the matrix $\mathbf{N}_\lambda$ takes into account extinction and emission pertaining to spectral lines.

Formal solution of the generalized radiative transfer equation

In order to derive a formal solution to the generalized radiative transfer equation following the treatment in [del Toro Iniesta, 2003, pp.152-157], the parameter $\kappa \doteq \tau_\lambda'$ shall be introduced to temporarily simplify the notation, hence yielding the generalized radiative transfer equation to be

$$\frac{\mathrm{d}\boldsymbol{I}_\lambda(\kappa)}{\mathrm{d}\kappa} = \mathbf{K}_\lambda(\kappa)\left[\boldsymbol{I}_\lambda(\kappa) - \boldsymbol{S}_\lambda(\kappa)\right]. \tag{2.45}$$

Consider now for the moment only the homogeneous differential equation

$$\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\kappa)}{\mathrm{d}\kappa} = \mathbf{K}_\lambda(\kappa)\boldsymbol{I}_{\lambda,h}(\kappa). \tag{2.46}$$

As it is done in [del Toro Iniesta, 2003, p.152], a so-called evolution operator $\mathbf{P}(\kappa, \kappa')$ can be defined, such that it gives the transformation $\boldsymbol{I}_{\lambda,h}(\kappa') \rightarrow \boldsymbol{I}_{\lambda,h}(\kappa)$ of the solution $\boldsymbol{I}_{\lambda,h}(\kappa')$ to the homogeneous equation from position κ' to κ, i.e.

$$\boldsymbol{I}_{\lambda,h}(\kappa) = \mathbf{P}(\kappa, \kappa')\boldsymbol{I}_{\lambda,h}(\kappa'). \tag{2.47}$$

This operator is linear, because the differential equation governing $\boldsymbol{I}_{\lambda,h}(\kappa)$ is also linear. Furthermore, because of the way it was defined, it has the properties

$$\mathbf{P}(\kappa, \kappa) = \mathbb{1}, \qquad \mathbf{P}(\kappa, \kappa')\mathbf{P}(\kappa', \kappa'') = \mathbf{P}(\kappa, \kappa''). \tag{2.48}$$

Differentiating eq. (2.47) with respect to κ, namely

$$\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\kappa)}{\mathrm{d}\kappa} = \frac{\mathrm{d}\mathbf{P}(\kappa, \kappa')}{\mathrm{d}\kappa}\boldsymbol{I}_{\lambda,h}(\kappa') = \mathbf{K}_\lambda(\kappa)\mathbf{P}(\kappa, \kappa')\boldsymbol{I}_{\lambda,h}(\kappa'), \tag{2.49}$$

the identity

$$\frac{\mathrm{d}\mathbf{P}(\kappa, \kappa')}{\mathrm{d}\kappa} = \mathbf{K}_\lambda(\kappa)\mathbf{P}(\kappa, \kappa') \tag{2.50}$$

results. Furthermore differentiating eq. (2.47) with respect to κ' yields

$$\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\kappa)}{\mathrm{d}\kappa'} = \mathbf{P}(\kappa, \kappa')\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\kappa')}{\mathrm{d}\kappa'} + \frac{\mathrm{d}\mathbf{P}(\kappa, \kappa')}{\mathrm{d}\kappa'}\boldsymbol{I}_{\lambda,h}(\kappa') = 0. \tag{2.51}$$

By means of rearranging this equation and inserting the definition of the homogeneous equation (2.46) the equation

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{P}(\kappa, \kappa')}{\mathrm{d}\kappa'}\boldsymbol{I}_{\lambda,h}(\kappa') &= -\mathbf{P}(\kappa, \kappa')\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\kappa')}{\mathrm{d}\kappa'} \\
&= -\mathbf{P}(\kappa, \kappa')\mathbf{K}_\lambda(\kappa')\boldsymbol{I}_{\lambda,h}(\kappa')
\end{aligned} \tag{2.52}$$

obtains, wherefrom the identity

$$\frac{\mathrm{d}\mathbf{P}(\kappa, \kappa')}{\mathrm{d}\kappa'} = -\mathbf{P}(\kappa, \kappa')\mathbf{K}_\lambda(\kappa') \tag{2.53}$$

is deduced.

Now, the generalized radiative transfer equation eq. (2.45) is multiplied with the operator $\mathbf{P}(\kappa', \kappa)$ from the left, thus obtaining

$$\begin{aligned}
\mathbf{P}(\kappa', \kappa)\frac{\mathrm{d}\boldsymbol{I}_\lambda(\kappa)}{\mathrm{d}\kappa} &= \mathbf{P}(\kappa', \kappa)\mathbf{K}_\lambda(\kappa)\left[\boldsymbol{I}_\lambda(\kappa) - \boldsymbol{S}_\lambda(\kappa)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}\kappa}\left[\mathbf{P}(\kappa', \kappa)\boldsymbol{I}_\lambda(\kappa)\right] - \frac{\mathrm{d}\mathbf{P}(\kappa', \kappa)}{\mathrm{d}\kappa}\boldsymbol{I}_\lambda(\kappa) \\
&= -\frac{\mathrm{d}\mathbf{P}(\kappa', \kappa)}{\mathrm{d}\kappa}\left[\boldsymbol{I}_\lambda(\kappa) - \boldsymbol{S}_\lambda(\kappa)\right],
\end{aligned} \tag{2.54}$$

where the identity eq. (2.53) was employed in the last step. Rearranging the above equation yields

$$\frac{\mathrm{d}}{\mathrm{d}\kappa}\left[\mathbf{P}(\kappa', \kappa)\boldsymbol{I}_\lambda(\kappa)\right] = \frac{\mathrm{d}\mathbf{P}(\kappa', \kappa)}{\mathrm{d}\kappa}\boldsymbol{S}_\lambda(\kappa) = -\mathbf{P}(\kappa', \kappa)\mathbf{K}_\lambda(\kappa)\boldsymbol{S}_\lambda(\kappa), \tag{2.55}$$

where again identity eq. (2.53) was made use of. One has now an integratable equation with respect to κ. Integrating the left hand side of the above equation from κ_0 to κ_1 gives

$$\begin{aligned}
\int_{\kappa_0}^{\kappa_1} \frac{\mathrm{d}}{\mathrm{d}\kappa}\left[\mathbf{P}(\kappa', \kappa)\boldsymbol{I}_\lambda(\kappa)\right]\mathrm{d}\kappa &= \mathbf{P}(\kappa', \kappa_1)\boldsymbol{I}_\lambda(\kappa_1) - \mathbf{P}(\kappa', \kappa_0)\boldsymbol{I}_\lambda(\kappa_0) \\
&= -\int_{\kappa_0}^{\kappa_1} \mathbf{P}(\kappa', \kappa)\mathbf{K}_\lambda(\kappa)\boldsymbol{S}_\lambda(\kappa)\,\mathrm{d}\kappa.
\end{aligned} \tag{2.56}$$

Through rearrangement of the equation and multiplication with $\mathbf{P}(\kappa_1, \kappa')$, the expressions

$$\begin{aligned}
\mathbf{P}(\kappa_1, \kappa')\mathbf{P}(\kappa', \kappa_1)\boldsymbol{I}_\lambda(\kappa_1) &= \mathbf{P}(\kappa_1, \kappa')\mathbf{P}(\kappa', \kappa_0)\boldsymbol{I}_\lambda(\kappa_0) \\
&\quad - \int_{\kappa_0}^{\kappa_1} \mathbf{P}(\kappa_1, \kappa')\mathbf{P}(\kappa', \kappa)\mathbf{K}_\lambda(\kappa)\boldsymbol{S}_\lambda(\kappa)\,\mathrm{d}\kappa,
\end{aligned} \tag{2.57}$$

results, which can due to $\mathbf{P}(\kappa, \kappa')\mathbf{P}(\kappa', \kappa) = \mathbf{P}(\kappa, \kappa) = \mathbb{1}$ and $\mathbf{P}(\kappa'', \kappa')\mathbf{P}(\kappa', \kappa) = \mathbf{P}(\kappa'', \kappa)$ further simplified to a first formal solution to the generalized radiative transfer equation, namely

$$\boldsymbol{I}_\lambda(\kappa_1) = \mathbf{P}(\kappa_1, \kappa_0)\boldsymbol{I}_\lambda(\kappa_0) - \int_{\kappa_0}^{\kappa_1} \mathbf{P}(\kappa_1, \kappa)\mathbf{K}_\lambda(\kappa)\boldsymbol{S}_\lambda(\kappa)\,\mathrm{d}\kappa. \tag{2.58}$$

Resubstituting $\tau'_\lambda = \kappa$ and hence $\mathrm{d}\kappa = \mathrm{d}\tau'_\lambda$ and identifying $\kappa_0 = \tau'_{\lambda,0} \to \infty$ and $\kappa_1 = \tau'_{\lambda,1} \to \tau'_\lambda$, one gets the general formal solution

$$\boldsymbol{I}_\lambda(\tau'_\lambda) = \mathbf{P}\left(\tau'_\lambda, \infty\right) \boldsymbol{I}_\lambda(\infty) - \int_\infty^{\tau'_\lambda} \mathbf{P}\left(\tau'_\lambda, t\right) \mathbf{K}_\lambda(t) \boldsymbol{S}_\lambda(t)\, \mathrm{d}t \qquad (2.59)$$

to the generalized radiative transfer equation. Considering fig. 2.3, the optical depth $\tau'_\lambda \to \infty$ can be understood to be located at the limit, where no photon can ever escape, e.g. at the center of a star, whose radiative transfer is described by eq. (2.59). Further taking a closer look at fig. 2.3, the optical depth $\tau'_\lambda = 0$ can be understood as the position of an observer located outside or right at boundary of the radiative transfer system, that is for example the position of an observer looking at a stellar atmosphere from outside the atmosphere. In this case, the measurable Stokes vector at the position of an observer outside the radiative transfer system is given as

$$\boldsymbol{I}_\lambda(\tau'_\lambda = 0) = \int_0^\infty \mathbf{P}\left(0, t\right) \mathbf{K}_\lambda(t) \boldsymbol{S}_\lambda(t)\, \mathrm{d}t. \qquad (2.60)$$

This result is obtained from eq. (2.59), where [del Toro Iniesta, 2003, p.153] gives the limit of the evolution operator $\mathbf{P}(0, t)$ for $t \to \infty$ to be

$$\lim_{t \to \infty} \mathbf{P}(0, t) = 0. \qquad (2.61)$$

This result makes sense insofar, as this limit can be represented as $\boldsymbol{I}_{\lambda,h}(0) = \mathbf{P}(0, \infty) \boldsymbol{I}_{\lambda,h}(\infty)$ by means of using the definition eq. (2.47) of the evolution operator; it should account for the change in $\boldsymbol{I}_{\lambda,h}$ from $\tau'_\lambda \to \infty$ to $\tau'_\lambda = 0$. Since the solution $\boldsymbol{I}_{\lambda,h}$ to the homogeneous radiative transfer equation eq. (2.46) does not contain any source term, the homogeneous solution must change to zero as one goes from optical depth infinity to optical depth zero. This is because in this case, a ray travels a theoretically infinite distance through the radiative transfer system without any sources along the way, therefore extincting all photons on the way to optical depth zero.

2.3.4 Absorption, emission and dispersion profile functions

Absorption can be defined as the removal of energy from an electromagnetic field by interaction with matter [del Toro Iniesta, 2003, p.87]. Emission however, is the opposite of absorption, namely contribution of

energy to an electromagnetic field by means of photons emerging from electronic transitions in atoms and molecules. There is a third effect due to the interaction of matter with electromagnetic fields, namely dispersion; [del Toro Iniesta, 2003, p.87] defines it as the dephasing of the electric field components as the radiation propagates through matter.

Spectral lines due to absorption and emission properties of matter can be accounted for by so-called profile functions. These functions are not just delta functions, as the discrete energy levels in atoms and molecules would suggest, but are always broadened by different processes. Among these processes are the so-called pressure broadening (Lorentz broadening) and the thermal broadening (Doppler broadening). Doppler broadening is caused by the thermal velocities of atoms and molecules in matter, which move in arbitrary directions in general. These effects are accounted for by distinct profile functions; these two profile functions can however be combined to account for both pressure and thermal broadening by means of the mathematical operation of a convolution; thus yielding the Voigt profile function $H(u, a)$ as

$$H(u,a) = \frac{a}{\pi} \int_{-\infty}^{\infty} e^{-y^2} \frac{1}{(u-y)^2 + a^2} \, \mathrm{d}y, \qquad (2.62)$$

as given by [del Toro Iniesta, 2003, p.100]. Furthermore, the dispersion effects are accounted for by the Faraday-Voigt function $F(u, a)$ given by

$$F(u,a) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{-y^2} \frac{u-y}{(u-y)^2 + a^2} \, \mathrm{d}y. \qquad (2.63)$$

The Voigt and the Faraday-Voigt profile functions are visualized in fig. 2.7.

2.3.5 The propagation matrix

The propagation matrix $\mathbf{K}_\lambda(\tau'_\lambda)$ accounts for absorption and dispersion effects of electromagnetic waves in matter within the framework of the generalized radiative transfer equation. In particular, it accounts for spectral lines in a radiation continuum. Explicitly, [Stix, 2002, p.154] gives the propagation matrix to be defined as

$$\mathbf{K}_\lambda = \begin{pmatrix} \eta_I & \eta_Q & \eta_U & \eta_V \\ \eta_Q & \eta_I & \rho_V & -\rho_U \\ \eta_U & -\rho_V & \eta_I & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & \eta_I \end{pmatrix}. \qquad (2.64)$$
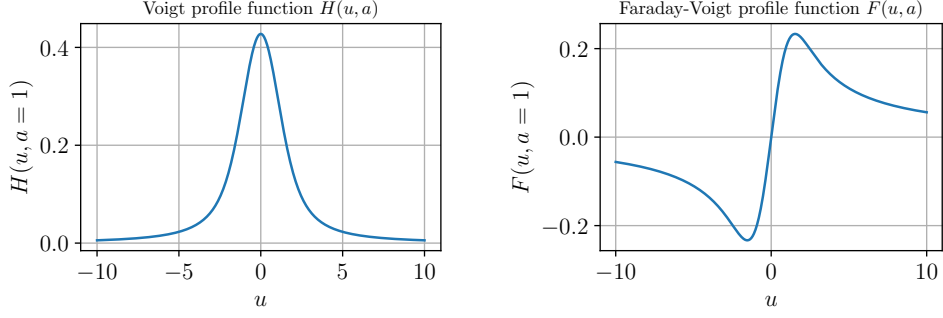
(a) Voigt profile function for $a = 1$.   **(b)** Faraday-Voigt profile function for $a = 1$.

Figure 2.7: Voigt and Faraday-Voigt profile functions for $a = 1$.

The individual entries of this matrix [del Toro Iniesta, 2003, p.116] accounts for by

$$
\begin{aligned}
\eta_I &= 1 + \frac{\eta_0}{2}\left[\phi_0 \sin^2(\theta) + \frac{1}{2}(\phi_{+1} + \phi_{-1})(1 + \cos^2(\theta))\right], \\
\eta_Q &= \frac{\eta_0}{2}\left[\phi_0 - \frac{1}{2}(\phi_{+1} + \phi_{-1})\right]\sin^2(\theta)\cos(2\varphi), \\
\eta_U &= \frac{\eta_0}{2}\left[\phi_0 - \frac{1}{2}(\phi_{+1} + \phi_{-1})\right]\sin^2(\theta)\sin(2\varphi), \\
\eta_V &= \frac{\eta_0}{2}(\phi_{-1} - \phi_{+1})\cos(\theta), \\
\rho_Q &= \frac{\eta_0}{2}\left[\psi_0 - \frac{1}{2}(\psi_{+1} + \psi_{-1})\right]\sin^2(\theta)\cos(2\varphi), \\
\rho_U &= \frac{\eta_0}{2}\left[\psi_0 - \frac{1}{2}(\psi_{+1} + \psi_{-1})\right]\sin^2(\theta)\sin(2\varphi), \\
\rho_V &= \frac{\eta_0}{2}(\psi_{-1} - \psi_{+1})\cos(\theta).
\end{aligned}
\tag{2.65}
$$

Hereby, the angles θ and φ are called inclination and azimuth and are defined as seen in fig. 2.8. The ranges for those angles are $\theta \in [0, \pi]$ and $\varphi \in [0, \pi]$. In the case of solar observations, $\theta = 0$ means that a vector field $\boldsymbol{V}$ under consideration is oriented towards an observer on-looking the Sun, whereas $\varphi =$ does either mean, that the vectorial quantity $\boldsymbol{V}$ faces either perfectly west or east on the solar surface. $\theta = \pi$ then does mean, that $\boldsymbol{V}$ points exactly away from the observer.

The quantities ϕ_α and ψ_α for $\alpha \in \{0, -1, 1\}$ are left to be defined;
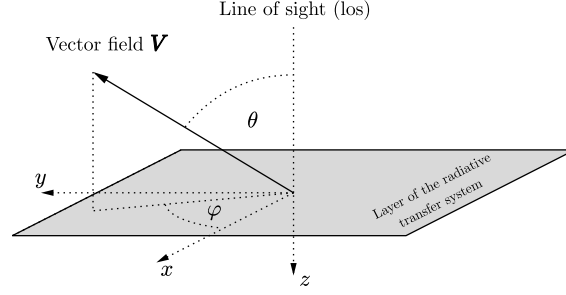
Figure 2.8: Visualization of the definitions for the inclination and azimuth angles θ and φ.

[del Toro Iniesta, 2003, p.123] defines these parameters as

$$\phi_\alpha(u, a) = \frac{1}{\sqrt{\pi}} H(u + \alpha u_B - u_{los}, a),$$
$$\psi_\alpha(u, a) = \frac{1}{\sqrt{\pi}} F(u + \alpha u_B - u_{los}, a). \tag{2.66}$$

Herewith, u, u_B and u_{los} aswell as a are dimensionless quantities determining the the Voigt and Faraday-Voigt functions. In these functions u is related to the wavelength λ and is to be regarded as a fixed value for a certain spectral line under consideration. [Uitenbroek, 2020, p.21] gives the parameter u_{los} as

$$u_{los} = \lambda \frac{\boldsymbol{u} \cdot \boldsymbol{n}}{c\Delta\lambda_D}, \tag{2.67}$$

where $\boldsymbol{n}$ defines the line of sight vector and $\boldsymbol{u}$ the velocity of the medium constituents relative to an inertial observer outside the radiative transfer system. This is a dimensionless number related to the actual line of sight velocity v_{los} as $v_{los} = c u_{los}$. Furthermore, $\Delta\lambda_D$ refers to the Doppler broadening of a spectral line due to the thermal velocity

$$v_{th} = \sqrt{\frac{2k_B T}{m}} \tag{2.68}$$

of the particles of mass m in the medium; it is given by the expression

$$\Delta\lambda_D = \frac{v_{th}\lambda}{c} = \frac{\lambda}{c}\sqrt{\frac{2k_B T}{m}}. \tag{2.69}$$

Since the thermal velocity of the particles is responsible for the Doppler broadening, one can also call the thermal velocity v_{th} the Doppler velocity $v_{dop} \doteq v_{th}$. Last but not least, u_B is a factor originating in the

29

Zeeman effect; therefore it connects the radiative transfer equation with an external magnetic field $\boldsymbol{B}$ via the relation

$$u_B = g_L \frac{e\lambda^2 |\boldsymbol{B}|}{4\pi mc\Delta\lambda_D}, \tag{2.70}$$

where g_L is the Landé factor. In total therefore, the propagation matrix $\mathbf{K}_\lambda$ for a certain wavelength λ can be fully determined by means of the seven independent parameters $\{|\boldsymbol{B}|, \theta, \varphi, v_{los}, v_{dop}, a, \eta_0\}$.

2.3.6 The Milne-Eddington approximation

The Milne-Eddington approximation of radiative transfer allows for the construction of a relatively simple stellar atmosphere model $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$, where $\boldsymbol{x}$ would be the atmospheric parameters, such as magnetic field strength, inclination and so forth, and where $\boldsymbol{y}$ would be the predicted observational data, namely the Stokes vectors $\boldsymbol{I}_\lambda$ for a certain wavelength interval $\lambda \in [\lambda_-, \lambda_+]$ giving so-called Stokes profiles. A Milne-Eddington model for a stellar atmosphere can even be analytically inverted, which means that there is a mathematically exact way to calculate atmospheric parameters $\boldsymbol{x}$ based on observational data $\boldsymbol{y}$ by means of mathematically inverting the model, i.e. $\boldsymbol{x} = \boldsymbol{M}^{-1}(\boldsymbol{y})$. In the case of real and not synthetic observational data $\boldsymbol{y}$, the best fit to the data given the model $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$ is to be found by iteratively changing the parameters; that is by means of synthesizing observations, comparing them to the real observations and adjusting the parameters, such that a better fit to the real observations is expected in the next iteration.

The simplest possible solution for the generalized radiative transfer equation eq. (2.44) is the so-called Milne-Eddington solution. Its derivation relies on the Milne-Eddington approximation, which basically consists of three fundamental assumptions given in [Degl'Innocenti, 2005, p.411] profoundly simplifying the radiative transfer problem:

(1) The radiative transfer system is assumed to be of a plane parallel nature and in a local thermal equilibrium.

(2) All parameters relevant to spectral line formation are required to be independent of all measures of height within the system.

(3) The source function is required to be affinely dependent upon the continuum optical depth.

The second assumption is to say, that the propagation matrix $\mathbf{K}_\lambda$ is isotropically constant, i.e.

$$\mathbf{K}_\lambda = \mathbf{K}_{\lambda,0}. \tag{2.71}$$

The third ingredient to the Milne-Eddington model means, that the source function vector can be written as

$$\boldsymbol{S}_\lambda(\tau'_\lambda) = \boldsymbol{S}_{\lambda,0} + \tau'_\lambda \boldsymbol{S}_{\lambda,1}, \qquad (2.72)$$

where $\boldsymbol{S}_{\lambda,0} = S_{\lambda,0}(1,0,0,0)^\top$ and $\boldsymbol{S}_{\lambda,1} = S_{\lambda,1}(1,0,0,0)^\top$ are constant vectors. Using the first assumption of the Milne-Eddington model, one can furthermore make the identifications $S_{\lambda,0} = B_\lambda(T)$ and $S_{\lambda,1} = B_\lambda(T)\beta$ with $\beta \in \mathbb{R}$ a constant. This is because local thermal equilibrium implies, that the source function is locally equal to the Planck function.

In the Milne-Eddington case, the homogeneous generalized radiative transfer equation eq. (2.46) reads as

$$\frac{\mathrm{d}\boldsymbol{I}_{\lambda,h}(\tau'_\lambda)}{\mathrm{d}\tau'_\lambda} = \mathbf{K}_{\lambda,0}\boldsymbol{I}_{\lambda,h}(\tau'_\lambda). \qquad (2.73)$$

According to standard mathematical literature, the matrix exponential $e^\mathbf{A}$ for a matrix $\mathbf{A}$ is defined as

$$e^\mathbf{A} = \mathbb{1} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \cdots = \sum_{k=1}^\infty \frac{\mathbf{A}^k}{k!}, \qquad (2.74)$$

Analogously to the scalar case, a solution to the generalized homogeneous radiative transfer equation can be given as

$$\boldsymbol{I}_\lambda(\tau'_\lambda = 0) = e^{-\mathbf{K}_{\lambda,0}\tau'_\lambda}\boldsymbol{I}_\lambda(\tau'_\lambda). \qquad (2.75)$$

Comparing this to the definition eq. (2.47) of the evolution operator, one finds

$$\mathbf{P}(0,\tau'_\lambda) = e^{-\mathbf{K}_{\lambda,0}\tau'_\lambda} \qquad (2.76)$$

in the Milne-Eddington case. One can now calculate the solution $\boldsymbol{I}_\lambda(\tau'_\lambda = 0, \theta)$ using the special case eq. (2.60) of the formal solution to the generalized radiative transfer equation, namely

$$\boldsymbol{I}_\lambda(\tau'_\lambda = 0, \theta) = \int_0^\infty \mathbf{P}(0,t)\,\mathbf{K}_\lambda(t)\boldsymbol{S}_\lambda(t)\,\mathrm{d}t. \qquad (2.77)$$

First, note that the derivative of $e^{-\mathbf{K}_{\lambda,0}\tau'_\lambda}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}\tau'_\lambda}e^{-\mathbf{K}_{\lambda,0}\tau'_\lambda} = -e^{-\mathbf{K}_{\lambda,0}\tau'_\lambda}\mathbf{K}_{\lambda,0}, \qquad (2.78)$$

because

$$\frac{\mathrm{d}}{\mathrm{d}x}e^{x\mathbf{A}} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathbb{1} + x\mathbf{A} + \frac{1}{2!}x^2\mathbf{A}^2 + \frac{1}{3!}x^3\mathbf{A}^3 + \dots\right)$$

$$= \mathbf{A} + x\mathbf{A}^2 + \frac{1}{2!}x^2\mathbf{A}^3 + \frac{1}{3!}x^3\mathbf{A}^4 + \dots \tag{2.79}$$

$$= \left(\mathbb{1} + x\mathbf{A} + \frac{1}{2!}x^2\mathbf{A}^2 + \frac{1}{3!}x^3\mathbf{A}^3 + \dots\right)\mathbf{A} = e^{x\mathbf{A}}\mathbf{A},$$

hence consequently also $\frac{\mathrm{d}}{\mathrm{d}x}e^{-x\mathbf{A}} = -e^{x\mathbf{A}}\mathbf{A}$ holds for any matrix $\mathbf{A}$ and real parameter x. Therefore, $-e^{-\mathbf{K}_{\lambda,0}\tau_\lambda'}\mathbf{K}_{\lambda,0}^{-1}$ is an antiderivative of $-e^{-\mathbf{K}_{\lambda,0}\tau_\lambda'}$. Now, one proceeds by calculating

$$\boldsymbol{I}_\lambda(\tau_\lambda' = 0) = \int_0^\infty \mathbf{P}\left(0,t\right)\mathbf{K}_\lambda(t)\boldsymbol{S}_\lambda(t)\,\mathrm{d}t$$

$$= \int_0^\infty \underbrace{e^{-\mathbf{K}_{\lambda,0}t}\mathbf{K}_{\lambda,0}}_{\doteq\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{F}(t)}\underbrace{\left(\boldsymbol{S}_{\lambda,0} + t\boldsymbol{S}_{\lambda,1}\right)}_{\doteq\boldsymbol{G}(t)}\,\mathrm{d}t$$

$$= [\mathbf{F}(t)\boldsymbol{G}(t)]_0^\infty - \int_0^\infty \mathbf{F}(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{G}(t)\,\mathrm{d}t$$

$$= \left[-e^{-\mathbf{K}_{\lambda,0}t}\mathbf{K}_{\lambda,0}^{-1}\mathbf{K}_{\lambda,0}(\boldsymbol{S}_{\lambda,0} + t\boldsymbol{S}_{\lambda,1})\right]_0^\infty \tag{2.80}$$

$$+ \int_0^\infty e^{-\mathbf{K}_{\lambda,0}t}\mathbf{K}_{\lambda,0}^{-1}\mathbf{K}_{\lambda,0}\boldsymbol{S}_{\lambda,1}\,\mathrm{d}t$$

$$= \boldsymbol{S}_{\lambda,0} + \int_0^\infty e^{-\mathbf{K}_{\lambda,0}t}\boldsymbol{S}_{\lambda,1}\,\mathrm{d}t$$

$$= \boldsymbol{S}_{\lambda,0} + \left[e^{-\mathbf{K}_{\lambda,0}t}\mathbf{K}_{\lambda,0}^{-1}\boldsymbol{S}_{\lambda,1}\right]_0^\infty$$

$$= \boldsymbol{S}_{\lambda,0} - \mathbf{K}_{\lambda,0}^{-1}\boldsymbol{S}_{\lambda,1}.$$

This is the fundamental result of the Milne-Eddington approximation and is sometimes also called the Unno-Rachkovsky solution, as stated by [del Toro Iniesta, 2003, p.160]. It is important to note here, that λ refers to the central wavelength of the considered spectral line(s), which means, that the parameters defining the Milne-Eddington atmosphere are considered to be constant for the wavelength interval(s) around the spectral line(s) under examination. Identifying therefore $\boldsymbol{S}_{\lambda,0} = S_{\lambda,0}(1,0,0,0)^\top = S_0(1,0,0,0)^\top$ and $\boldsymbol{S}_{\lambda,1} = S_{\lambda,1}(1,0,0,0)^\top = S_1(1,0,0,0)^\top$ and taking into account, that the propagation matrix $\mathbf{K}_{\lambda,0}$ depends upon the seven parameters as specified in section 2.3.5, the Stokes values $\boldsymbol{I}_\lambda(\tau_\lambda' = 0)$ around the central wavelength λ of a spectral line are determined fully by the set of nine parameters $|\boldsymbol{B}|$, θ, φ, v_{los}, v_{dop}, a, η_0, S_0, S_1. Note, that S_0

is related to the Planck function as $B_\lambda(T) = S_0$; therefore, one can also assign a temperature T to the system under investigation once one has a value for S_0 and knows which wavelength λ to consider.

The parameter vector $\boldsymbol{x}$ of the Milne-Eddington model is hence given as

$$\boldsymbol{x} = (|\boldsymbol{B}|, \theta, \varphi, v_{los}, v_{dop}, a, \eta_0, S_0, S_1)^\top \in \mathbb{R}^9. \qquad (2.81)$$

Let a number $D \in \mathbb{N}$ of wavelengths be sampled from a given wavelength interval $[\lambda_{min}, \ldots, \lambda, \ldots, \lambda_{max}]$ centered around λ, the central wavelength of a spectral line. Therefore, the observations vector $\boldsymbol{y}$ of the Milne-Eddington model is accounted for by

$$\boldsymbol{y} = (\boldsymbol{I}_{\lambda_{min}}, \ldots, \boldsymbol{I}_\lambda, \ldots, \boldsymbol{I}_{\lambda_{max}})^\top \in \mathbb{R}^D. \qquad (2.82)$$

Thus, the full Milne-Eddington model can be written as

$$\boldsymbol{M} : \mathbb{R}^9 \to \mathbb{R}^D, \quad \boldsymbol{x} = \begin{pmatrix} |\boldsymbol{B}| \\ \theta \\ \varphi \\ v_{los} \\ v_{dop} \\ a \\ \eta_0 \\ S_0 \\ S_1 \end{pmatrix} \mapsto \boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{I}_{\lambda_{min}} \\ \vdots \\ \boldsymbol{I}_\lambda \\ \vdots \\ \boldsymbol{I}_{\lambda_{max}} \end{pmatrix}, \qquad (2.83)$$

where eq. (2.80) is identified to define the model $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$.

2.4 Stellar atmosphere models

2.4.1 Stellar atmosphere inversions

Consider some observation $\boldsymbol{y}$ of a stellar atmosphere. Assume, that there is a model $\boldsymbol{M}$ with associated parameters $\boldsymbol{x}$, such that the model is given by the relationship $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{x}) = \boldsymbol{y}$.

So if one knows the parameter values $\boldsymbol{x}_0$ of the stellar atmosphere model $\boldsymbol{M}$, then observations $\boldsymbol{y}_0$ can be generated based on the parameter values $\boldsymbol{x}_0$ by means of the relation $\boldsymbol{y}_0 = \boldsymbol{M}(\boldsymbol{x}_0)$; this is called the forward model. However, if one only has observations $\boldsymbol{y}_0$ and the stellar model $\boldsymbol{M}$, one would like to find out the model parameters $\boldsymbol{x}_0$ associated to these observations; this is called the backward model and can be achieved by means of applying the relation $\boldsymbol{x}_0 = \boldsymbol{M}^{-1}(\boldsymbol{y}_0)$. As one

can see, this requires a so-called inversion of the atmosphere model $\boldsymbol{M}$, which can lead to degeneracy in the model parameters $\boldsymbol{x}_0$, meaning that more than one configuration of $\boldsymbol{x}_0$ can lead to the same observations $\boldsymbol{y}_0$. A normalizing flow helps to disentangle this degeneracy by means of introducing a probability density function for each model parameter, such that the more likely parameter configurations leading to some observations $\boldsymbol{y}_0$ can be distinguished from less likely parameter configurations leading to the same observations.

2.4.2 Solving a stellar atmosphere inversion problem

Consider a stellar model $\boldsymbol{M}(\boldsymbol{x})$. Given observations $\boldsymbol{y}_0$, one would like to infer the stellar parameters $\boldsymbol{x}_0$ associated to these observations by applying the relation

$$\boldsymbol{x}_0 = \boldsymbol{M}^{-1}(\boldsymbol{y}_0). \qquad (2.84)$$

Commonly, such inversion problems are carried out using a gradient search minimization algorithm which basically functions as follows. First, a function which should be minimized is defined; in our case

$$\boldsymbol{F}(\boldsymbol{x}) \doteq \boldsymbol{M}^{-1}(\boldsymbol{y}_0) - \boldsymbol{x}. \qquad (2.85)$$

By computing the gradient of $\boldsymbol{F}(\boldsymbol{x}) = (F_1(\boldsymbol{x}), \dots, F_n(\boldsymbol{x}))^\top$, $n \in \mathbb{N}$ with respect to the parameters $\boldsymbol{x}$, that is for every $F_i \in \boldsymbol{F} = (F_1, \dots, F_n)^\top$ the vector $\boldsymbol{\nabla}_{\boldsymbol{x}} F_i(\boldsymbol{x})$ is calculated, one approaches a minimum of $\|\boldsymbol{F}\|$, if $\boldsymbol{x}$ is iteratively changed in the direction of negative gradient of $\boldsymbol{F}(\boldsymbol{x})$.

But these types of algorithms are computationally costly, if one wants to infer some sort of distribution of model parameters based on more than just one observation; this is because in the case of multiple observations, as it is almost always the case in real-life applications, a possible solution to the minimization problem needs to be valid for all observations. Furthermore, this procedure just outputs one single vector as a solution. Here the normalizing flows technique comes into play, which provides a computationally more efficient way to do inversions and furthermore allows for the possibility to construct a probability density of possible solutions to the inversion problem, because the a trained normalizing flow outputs not just one single vector as gradient descent does, but a whole bunch of solution vectors. This means, that for each component of these solution vectors one can construct a probability density just by interpolating between the points of a normalized histogram for each component.

2.4.3 Calculation of synthesized datasets using a stellar atmosphere model

In the case, where data $\boldsymbol{x} \in \mathbb{R}^d$ is conditioned by context $\boldsymbol{y} \in \mathbb{R}^D$, it can be of interest to find a mathematical model implementing a transformation $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$. In order to be able to train a normalizing flow implementing a transformation

$$\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}), \tag{2.86}$$

if can be a commodity to know the true posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$, thus enabling a comparison to the posterior distribution $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ learned by the normalizing flow; this can be done by application of the Bayes theorem. Note, that $\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})$ are weights depending upon training data $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ associated to the implementation of the function $\boldsymbol{f}_{\boldsymbol{y},\phi}$ by a neural network. As to calculate the true posterior distribution by means of the model $\boldsymbol{M}(\boldsymbol{x}) = \boldsymbol{y}$ linking together the so-called context $\boldsymbol{y}$ with the input $\boldsymbol{x}$ to the normalizing flow, one can proceed as follows.

First, $\boldsymbol{x}$ is assumed to follow a specific probability density $p(\boldsymbol{x})$, for example a uniform distribution. Next, one can sample $L \in \mathbb{N}$ datapoints $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ from the prior $p(\boldsymbol{x})$. Using the model $\boldsymbol{M}$, for every $\boldsymbol{x}_k$ the corresponding value $\boldsymbol{y}_k$ can be calculated as $\boldsymbol{y}_k = \boldsymbol{M}(\boldsymbol{x}_k)$, leading to the likelihood distribution $p(\boldsymbol{y}|\boldsymbol{x})$. Using the Bayes theorem and the law of total probability, one can calculate $p(\boldsymbol{x}_k|\boldsymbol{y}_k)$ as

$$p(\boldsymbol{x}_k|\boldsymbol{y}_k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k)}{\sum_{j=1}^L p(\boldsymbol{y}_k|\boldsymbol{x}_j)p(\boldsymbol{x}_j)}, \tag{2.87}$$

leading to the probability distribution $p(\boldsymbol{x}|\boldsymbol{y})$ for $k \in \{1, \ldots, L\}$ and $L \to \infty$.

Chapter 3

Artificial intelligence

The field of artificial intelligence (AI) is relatively recent field of research dedicated to mimic the process of learning information. Hereby, inspiration for first principles of a sub-field of artificial, the so-called deep learning, comes from the brain structure of mammals.

As artificial intelligence tries to mimic the process of learning, it remains to be answered what learning actually is. In order to define how a learning process is constituted, another quantity has to be defined first, namely information. [Mitchell, 1997, p.2] defines learning such that a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. If the so-called information entropy, which will be defined at a later stage, serves as a performance measure P, the process of learning can also be described by the statement, that any learning process decreases the information entropy of some system under consideration.

3.1 Information theory

Following [Goodfellow et al., 2016, p.73], the fundamental intuition of information theory is the consideration, that the occurrence of an unlikely event should convey more information than the occurrence of a likely event. The quantification of information in the field of information theory therefore relies on three key requirements, as given by [Goodfellow et al., 2016, p.73]:

(1) Likely events should have low information content. Guaranteed events are required to have no information content.

(2) Unlikely events should have high information content.

(3) Independent events should have additive information; i.e. learning about two independent events of the same likelihood should convey twice as much information as learning about the occurrence of only one of these events.

3.1.1 Self-information

The above mentioned requirements are satisfied by the so-called self-information $\mathcal{I}[p(x)]$ as defined by [Goodfellow et al., 2016, p.73], namely

$$\mathcal{I}[p(x)] = -\log\left[p(x)\right], \tag{3.1}$$

with X being a random variable and $p(x)$ the probability density function thereof.

By this definition, information $\mathcal{I}[p(x = U)]$ of an unlikely event $x = U$ will be higher than information $\mathcal{I}[p(x = L)]$ of a likely event $x = L$; this can be demonstrated by assessing the self-information contents of both events. Because U is said to be an unlikely event, $p(U) \approx 0$ will hold and therefore $\mathcal{I}[p(x = U)] = -\log[p(U)] = \log[1/p(U)] \gg 1$. Since L is a likely event, the probability $p(L)$ of L occurring is estimated by $p(L) \approx 1$ and the self-information $\mathcal{I}[p(x = L)]$ will therefore be very low, because $\mathcal{I}[p(x = L)] = -\log[p(L)] = \log[1/p(L)] \approx 0$ holds. Moreover, information about two independent events $x = A$ and $x = B$ is additive, as the calculation

$$\begin{aligned} \mathcal{I}[p(A \cap B)] &= -\log[p(A \cap B)] = -\log[p(A)p(B)] \\ &= -(\log[p(A)] + \log[p(B)]) = \mathcal{I}[p(A)] + \mathcal{I}[p(B)] \end{aligned} \tag{3.2}$$

shows it to be the case.

3.1.2 Information entropy

One can describe the information content of a probability density function by the so-called information entropy given by [Goodfellow et al., 2016, p.74] as

$$H_p[p(x)] = E_{x \sim p}(\mathcal{I}[p(x)]) = -E_{x \sim p}\left(\log[p(x)]\right), \tag{3.3}$$

where $p(x)$ is a probability density function of a random variable X. If the random variable X is of a discrete nature, the information entropy is called Shannon entropy, whereas if X is continuous, the information entropy goes by the name of differential entropy.

Information entropy defined in this way is a measure of the amount of uncertainty in a probability distribution. The information entropy will

be very low, if $p(x)$ describes mostly very likely events, whereas it will be very high, if $p(x)$ describes mainly very unlikely events. This can also be stated in the following way: Probability distributions that are nearly deterministic (delta-function shaped) have a low entropy, whereas distributions that are closer to a uniform distribution have high entropy.

The field of artificial intelligence and especially its sub-field machine learning are concerned with learning the available self-information from the data. That is to say, the information entropy initially present within the machine learning entity is iteratively reduced by means of training it on a dataset, thus resulting in learned information on the part of the trained artificial intelligence.

3.2 Overview of artificial intelligence

According to [Amini, 2023, p.7], the field of artificial intelligence can be broadly described as any computational technique, that enables a computer to engage with information in a reasonable way, such as to mimic how a human would process the available data or information. Machine learning is more specialized sub-field of artificial intelligence, which is concerned with any routine, that gives a computer the ability to learn a task without being explicitly programmed for that particular task. Deep learning, which is ultimately about reflecting and rebuilding a simple representation the human brain working principle, is an even more specialized sub-field of artificial intelligence, which is itself a sub-field of machine learning.

The foundational tool of deep learning is the so-called neural network, also known as deep neural network. The basic building block of a neural network is the single layer perceptron (SLP). By combining several single layer perceptrons to a multi layer perceptron (MLP), a deep neural network can be constructed. Given an input vector $\boldsymbol{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ and an output vector $\bar{\boldsymbol{x}} = (\bar{x}_1, \ldots, \bar{x}_w)^\top \in \mathbb{R}^w$ with $d, w \in \mathbb{N}$, the working principle of single layer perceptron is shown in fig. 3.1, similarly for a multi layer preceptron in fig. 3.2.

3.3 Neural networks

3.3.1 Single layer perceptron

In fig. 3.1, a single layer perceptron is visualized. The quantity $\boldsymbol{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ with $d \in \mathbb{N}$ is an arbitrary input vector. The single

layer perceptron generates a single output $\bar{x}$. First, all input vector components x_i are multiplied with initially arbitrary weights $\phi_i \in \mathbb{R}$ for all $i \in \{1, \ldots, d\}$ and subsequently summed up to some sum $\tilde{s} = \sum_{i=1}^{d} x_i \phi_i$; then, an arbitrary bias value $b \in \mathbb{R}$ is added to the sum $\tilde{s}$ such as to arrive at a new sum $s = b + \tilde{s}$. Finally, an output scalar $\bar{x} = a(s)$ is generated by means of pushing the beforehand attained sum s through a so-called activation function $a : \mathbb{R} \to \mathbb{R}, s \mapsto a(x)$, which is nothing more than some arbitrary nonlinear function, which introduces nonlinearities to the otherwise linear operations; this enables the perceptron to learn nonlinear correlations. All quantities involving the letters b and ϕ

Figure 3.1: Working principle of a single layer perceptron; hereby, a represents a nonlinear activation function $a : \mathbb{R} \to \mathbb{R}, \ s \mapsto a(s)$.

compose the total vector $\boldsymbol{\phi}$ of weights for the perceptron.

3.3.2 Multi layer perceptron

A multi layer perceptron is the simplest form of a neural network. In fig. 3.2, such a multi layer perceptron is visualized. Again, $\boldsymbol{x} = (x_1, \ldots, x_d)^\top$ as an element of $\mathbb{R}^d$ with $d \in \mathbb{N}$ is an arbitrary input vector. Compared to the single layer perceptron, the multi layer perceptron however generates a vectorial output $\bar{\boldsymbol{x}} = (\bar{x}_1, \ldots, \bar{x}_w)^\top \in \mathbb{R}^w$ with $w \in \mathbb{N}$. Now, all input components x_i are connected to more than one perceptron; thus all perceptrons directly connected to the input components compose a so-called network layer, the first layer. There can be one or more network layers; generally there are $n \in \mathbb{N}$ layers in a neural network. At each perceptron of a certain fixed network layer, a single output is generated as with the single layer perceptron. All of these single outputs together

compose a new vector, which serves as the input vector for the next network layer composed of perceptrons. This general working principle is better explained visually, consider therefore fig. 3.2. Again, all quantities
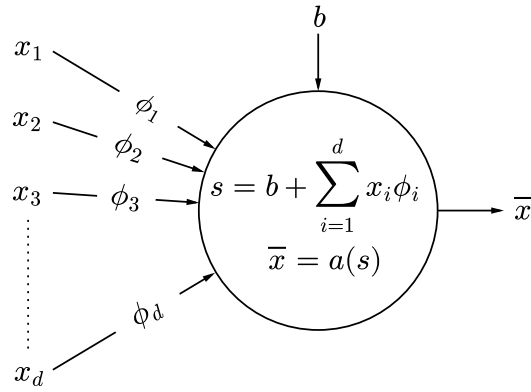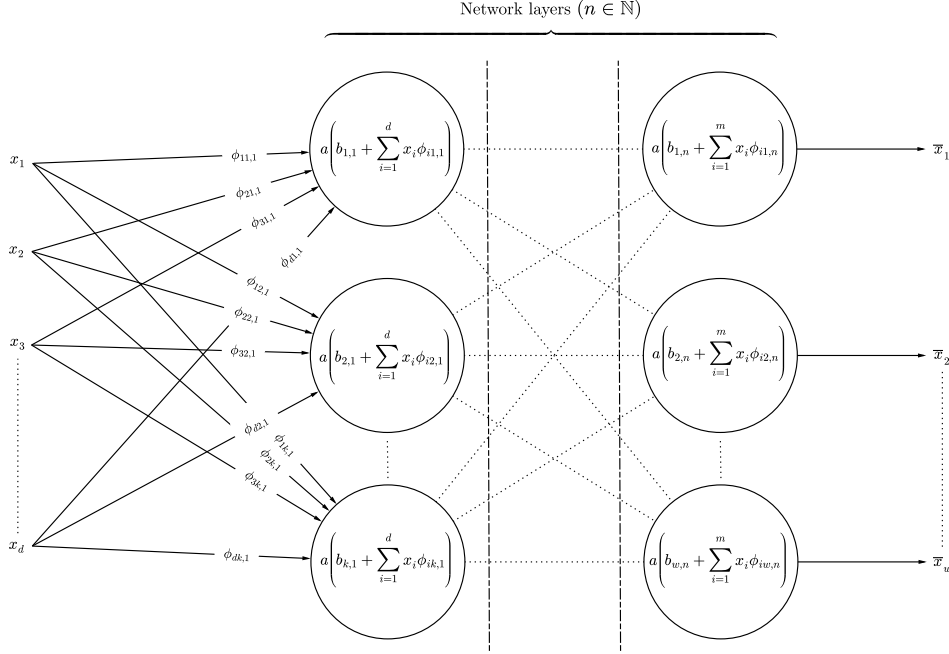

Figure 3.2: Working principle of a multi layer perceptron; hereby, a represents a nonlinear activation function $a : \mathbb{R} \to \mathbb{R}$, $s \mapsto a(s)$.

with letters b or ϕ compose the total weights vector $\boldsymbol{\phi}$ of the multi layer perceptron, which is the simplest possible neural network involving more than one perceptron.

3.3.3   Training process for neural networks

In order for a neural network to learn anything, it has to be conditioned by data, whose interpretation is known. This process is generally referred to as training of a neural network and requires a mathematical framework that allows for optimization. During the training process of a neural network, one would like to achieve the best possible performance of the neural network in interpreting the training data correctly, that is to say, that the input data $\boldsymbol{x}$ maps to the known output data $\bar{\boldsymbol{x}}$ with minimal discrepancies.

The mathematical framework for such an optimization relies on the definition of a loss function $\mathcal{L}_{\boldsymbol{\phi}}$. Such a loss function can be implemented

in many different ways; for example on can just take the norm of the difference between input data $\boldsymbol{x}$ and output data $\bar{\boldsymbol{x}}$ as a loss function, i.e. $\mathcal{L}_{\phi(\hat{\mathbf{x}})} = \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|$, where $\hat{\mathbf{x}}$ is a matrix of training data and $\phi(\hat{\mathbf{x}})$ the associated weights of the neural network described by this loss function. However, the loss function depends on the task at hand and needs to be chosen accordingly, such that it represents a suitable measure on how well a neural network learns from the training data. Once one has defined a loss function $\mathcal{L}_\phi$, the key to training a neural network resides within something called gradient descent. The idea behind gradient descent is quite simple; one just calculates the gradient of the loss function $\mathcal{L}_\phi$ with respect to the weights ϕ, i.e. $\boldsymbol{\nabla}_\phi \mathcal{L}_\phi$ and updates all weights such that one moves into the direction of negative gradient. Iteratively repeating this process, eventually a local or global minimum of the loss function is found. In terms of mathematics, this procedure translates as follows. Consider some starting point ϕ_i and the loss function at this point, i.e. $\mathcal{L}_{\phi_i}$. The gradient $\boldsymbol{\nabla}\mathcal{L}_\phi$ of the loss function with respect to the weights ϕ indicates, in which direction of ϕ the loss function increases or decreases. The negative gradient points into the direction of a local/global minimum of $\mathcal{L}_\phi$, whereas the positive gradient points towards a local/global maximum. Hence, $\phi_{i+1} = \phi_i - \gamma \boldsymbol{\nabla}_\phi \mathcal{L}_{\phi_i}$ with $i \in \mathbb{N}$ and $\gamma \in \mathbb{R}$ a suitable step size provides a rule for eventually finding a global or local minimum of the loss function. One can therefore define an iterative procedure as follows:

(1) Take a start value ϕ_0 and calculate the value of the loss function $\mathcal{L}_{\phi_0}$ at this point.

(2) Update the weights ϕ_i according to the rule

$$\phi_{i+1} = \phi_i - \gamma \boldsymbol{\nabla}_\phi \mathcal{L}_{\phi_i} \tag{3.4}$$

for $i \in \mathbb{N}$ individual steps and $\gamma \in \mathbb{R}$ a suitable step size as long as $\epsilon \leq |\mathcal{L}_{\phi_{i+1}} - \mathcal{L}_{\phi_i}|$ holds for $\epsilon \in \mathbb{R}$ an accuracy threshold.

(3) Terminate, if $\epsilon > |\mathcal{L}_{\phi_{i+1}} - \mathcal{L}_{\phi_i}|$.

This iterative process is found visualized in fig. 3.3; note, that this is a simplified picture, as the horizontal axis would in reality be multidimensional.

3.4 Deep generative models

Generative models - as their name already suggests - generate data. These models learn how to do so by extracting defining features of the
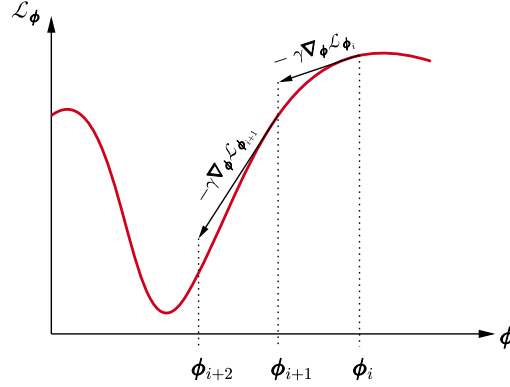
Figure 3.3: Visualization of a gradient descent routine. Note, that the horizontal axis is in reality multidimensional, whereas it is pictured here as just being one-dimensional. The vertical axis represents the value of the loss function $\mathcal{L}_{\phi}$ in dependence of the weights ϕ.

data they are trained on, such as the data's probability density. Consider a matrix of training data $\hat{\mathbf{x}}$, where the individual columns follow a probability distribution $\boldsymbol{x} \sim p_{\boldsymbol{x}}(\boldsymbol{x})$. According to [Raschka, 2022, p.593], a generative model in deep learning is then defined such, that it learns parameters $\boldsymbol{\phi}(\hat{\mathbf{x}})$ to construct a probability density function $p_{\boldsymbol{\phi}(\hat{\mathbf{x}})}(\boldsymbol{x}) \approx p_{\boldsymbol{x}}(\boldsymbol{x})$. This is to say, that one can then draw samples from the learned probability density $p_{\boldsymbol{\phi}(\hat{\mathbf{x}})}(\boldsymbol{x})$, which essentially are generated data imitating the training data.

One can distinguish at least three basic architectures of generative models, namely variational autoencoder networks (VAE's), generative adversarial networks (GAN's) and flows. These types of generative models are visualized in fig. 3.4, fig. 3.5 and fig. 3.6. The parallelograms used to visualize the working principles of these models represent dimensionality reduction/increase of an input variable compared to an output variable; if the parallelogram narrows towards the output, it means dimensionality reduction, whereas if it broadens towards the output, increase in dimensionality is visualized.

3.4.1 VAE's

Variational autoencoders work as sketched in fig. 3.4. A variational autoencoder consists of an encoder and a decoder. The encoder implements a conditional probability distribution $q_{\phi}(\boldsymbol{z}|\boldsymbol{x})$ for a latent variable $\boldsymbol{z}$, based on an input variable $\boldsymbol{x}$, where $\boldsymbol{\phi}$ are the weights of the neural network constituting the encoder. The decoder however models the

probability $p_{\boldsymbol{\theta}}(\boldsymbol{x}|\boldsymbol{z})$ of the output variable $\boldsymbol{x}$ based on the latent variable $\boldsymbol{z}$, where again $\boldsymbol{\theta}$ are weights associated to the neural network of the decoder. The variable $\boldsymbol{x}$ represents the input variable with probability density $\boldsymbol{x} \sim p_{\boldsymbol{x}}(\boldsymbol{x})$, whereas the variable $\boldsymbol{z}$ is a latent variable and is chosen to fit the probability density of a standard normal distribution, i.e. $\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z}) = \mathcal{N}(0,1)^d$, where $d \in \mathbb{N}$ is the dimension of the latent variable.
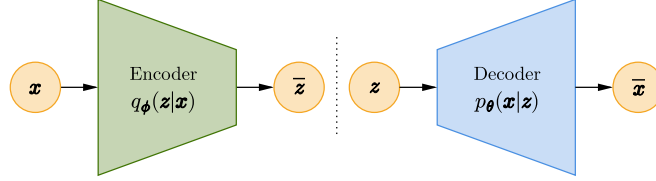
Figure 3.4: Working principle of a variational autoencoder (VAE). Figure inspired by [Weng, 2018].

A variational autoencoder is trained in such a way, that the encoder learns how to transform an input $\boldsymbol{x}$ such that the output $\bar{\boldsymbol{z}}$ follows the chosen latent probability density. Furthermore, the decoder has to learn, how to transform a variable $\boldsymbol{z}$ following a standard normal distribution to an output $\bar{\boldsymbol{x}}$, such that the output variable follows the probability density of the original input variable $\boldsymbol{x}$.

A VAE in principle can be said to learn the most important information from the training data $\hat{\mathbf{x}}$, as it has to pass this information from the input variable $\boldsymbol{x}$ through a low-dimensional bottleneck to $\bar{\boldsymbol{z}}$, from which again the original input $\bar{\boldsymbol{x}}$ has to be reconstructed as best as possible, such that $\boldsymbol{x} \approx \bar{\boldsymbol{x}}$. Once the VAE is trained sufficiently, one can just use the decoder to generate outputs $\bar{\boldsymbol{x}}$ based on $\boldsymbol{z}$ sampled from a multidimensional standard normal distribution.

A VAE could be used for stellar atmosphere inversion purposes, but it has several disadvantages in this regard compared to normalizing flows. First of all, VAE's have to be trained using an evidence lower bound (ELBO), because the exact likelihood of $\boldsymbol{x}$ given $\boldsymbol{z}$ is not tractable. A normalizing flow however allows for exact likelihood estimation of $\boldsymbol{x}$ given $\boldsymbol{z}$ because of its architecture. Furthermore, again because of its inherent architecture, a normalizing flow allows to perform exact inversions of data, since it is basically built of invertible functions. This is not the case for VAE's; yes, encoder and decoder are sort of inverse to one another, but it cannot be an exact inversion in the mathematical sense, since the VAE relies on a smaller dimension of the latent variable and the

input/output variable.

3.4.2 GAN's

Generative adversarial networks function as sketched in fig. 3.5. A generative adversarial network consists of a generator and a discriminator. The generator is basically implements a function $\bar{x} = G_\phi(z)$, where $\bar{x}$ is the output of the generator, ϕ are the weights associated to the neural network implementing the generator function and z is a latent variable, which is chosen to follow a standard normal distribution, i.e. $z \sim p_z(z) = \mathcal{N}(0,1)^d$ with $d \in \mathbb{N}$ again the dimensionality of the latent space. The purpose of the generator is to generate fake data $\bar{x}$ based on samples z of the latent probability density $p_z(z)$. The discriminator however takes an input x, such as the output of the generator, and outputs a value 0 or 1, where 0 stands for «input is fake» and 1 represents the situation «input is real». The task of the discriminator is therefore to decide, if an input x is fake or real.
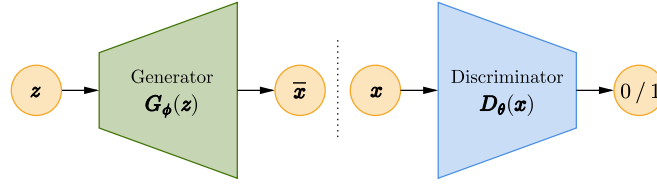
Figure 3.5: Working principle of a generative adversarial network (GAN). Figure inspired by [Weng, 2018].

A generative adversarial network is trained in such a way, that the generator learns how to generate the best possible real-looking data $\bar{x}$, whereas the discriminator learns as good as possible to distinguish between real data contained in the training set $\hat{x}$ and fake data $\bar{x}$ generated by the generator. Herein lies the adversarity of the model; the generator and the discriminator are actually competing against each other; the generator learns how to create the best possible fake data, whereas the discriminator learns, how to best distinguish between real and fake data - in this way, both the generator and the discriminator are optimized for their respective task. The loss function for training therefore has to be defined accordingly in a game-theoretical manner, where both the discriminator and the generator try to optimize their outcomes. Once trained, one can just use the generator to create fake data $\bar{x} = G_\phi(z)$ from samples z of $p_z(z)$ at will.

Despite the many interesting applications of GAN's, they cannot really be used for stellar atmosphere inversion purposes, since they do not suit the problem at hand, since stellar inversion problems do not involve classification between real and fake data.

3.4.3 Flows

Flow-based generative models work as sketched in fig. 3.6. A flow just consists of a bijective function $\bar{z} = \boldsymbol{f}_\phi(\boldsymbol{x})$, where ϕ refers to the weights of the neural network(s) used to implement this function. One of the main differences between flows and GAN's on the one hand and VAE's on the other hand is the fact, that flows are bijective in architectural nature, while GAN's and VAE's are not. This is to say that they are perfectly adapted to an inversion problem, such as a stellar atmosphere inversion task. Because of the bijective nature of flows, there is no dimensionality reduction or increase from input/output to latent space.
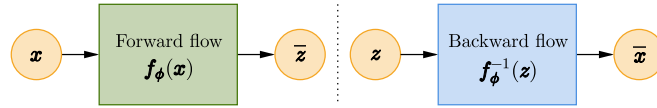
Figure 3.6: Working principle of a flow-based generative model. Figure inspired by [Weng, 2018].

Flows can be trained by exact likelihood calculation; that is to say, the loss function can be mathematically derived and implemented using the change of variable theorem from probability calculus in combination with the specific bijective transformations constituting the flow. A flow is called a normalizing flow, if the latent probability density function is chosen as a standard normal distribution - hence an arbitrary probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ of input data is mapped into the density of a standard normal distribution. In that way, the input data can be said to be normalized to the latent space. As flows and especially normalizing flows are best suited to stellar atmosphere inversions compared to the other available generative models, this technique will be the tool of choice in the material to follow.

3.5 Introduction to normalizing flows

Normalizing flows are useful to learn complex data distributions from available samples, be they experimentally obtained or of a synthesized

nature.

The goal of a normalizing flow is to learn a certain probability density function $p_{\boldsymbol{x}}(\boldsymbol{x})$ of some quantity represented by a potentially multivariate random variable $\boldsymbol{x}$. Given samples from $p_{\boldsymbol{x}}(\boldsymbol{x})$ obtained by either experimental or synthetical procedure, the normalizing flow can learn the probability density function by means of finding a diffeomorphism, which maps samples $\boldsymbol{z}$ from a standard normal distribution to $p_{\boldsymbol{x}}(\boldsymbol{x})$. Since a diffeomorphism is invertible by definition, a trained normalizing flow can then be used to draw samples from $p_{\boldsymbol{x}}(\boldsymbol{x})$ at will; this can be done by sampling from a standard normal distribution and subsequent application of the inverse learned diffeomorphism to the generated samples. This will result in a set of samples from $p_{\boldsymbol{x}}(\boldsymbol{x})$.

This technique can be used for various purposes; for example, for inversion problems in (stellar) atmospheric physics, as it is being explored in the subsequent material.

3.5.1   Why use normalizing flows for inversions?

If one has an analytical model $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$ of a stellar atmosphere, it can usually be inverted too, such that $\boldsymbol{x} = \boldsymbol{M}^{-1}(\boldsymbol{y})$. This raises the question, why one should even bother using normalizing flows to learn such inversions.

First of all, only simple models of radiative transfer, such as the Milne-Eddington atmosphere can be inverted analytically. More complex models require resource-demanding calculations by means of Monte Carlo methods to do inversions. The advantage of a normalizing flow resides in the fact, that inversions done by it would be comparatively much faster, once the normalizing flow is trained.

A second incentive to use normalizing flow for atmospheric inversions is, that a normalizing flow not only provides one with one solution to the inversion problem for each atmospheric parameter, but rather with a probability distribution for each parameter. This allows for error estimation and the determination of the most likely solution to an inversion problem, if there is more than one solution possible.

3.6   Change of variable formula in probability theory

The change of variable formula can be considered as the heart of the normalizing flow technique, since it guides the implementation of a loss

function to train a flow. Because of its importance for normalizing flows, a full proof of this theorem will therefore be given for the univariate case, which will be generalized to the multivariate case subsequently.

Theorem 3.6.0.1. *Let x and z be continuous random variables with associated probability density functions $p_x(x)$ and $p_z(z)$. Furthermore, let z be the transformation of x by an invertible and strictly increasing continuously differentiable function $f(x)$, such that $z = f(x)$ and $x = f^{-1}(z)$ hold; in this case the change of variable formula from probability theory applies as*

$$p_x(x) = p_z(f(x))\frac{\mathrm{d}f(x)}{\mathrm{d}x}. \tag{3.5}$$

Proof. First, let with a and b be defined real numbers with $a < b$. By definition of the transformation $f(x)$ let with c and d be defined the real numbers $c = f(a)$ and $d = f(b)$, hence

$$P(a \leq x \leq b) = P(c \leq z \leq d) \tag{3.6}$$

must hold, where this quantity is defined as $P(a \leq x \leq b) \doteq \int_a^b p_x(x)\,\mathrm{d}x$ and therefore represents the probability mass for an experiment outcome of x to be between a and b. Since x and z are continuous random variables with associated valid probability density functions, one can write

$$P(c \leq z \leq d) = \int_c^d p_z(z)\,\mathrm{d}z = \int_{f(a)}^{f(b)} f_z(z)\,\mathrm{d}z. \tag{3.7}$$

By the substitution rule for integrals one can substitute in the above integral $z(x) = f(x)$ and

$$\frac{\mathrm{d}z(x)}{\mathrm{d}x} = \frac{\mathrm{d}f(x)}{\mathrm{d}x} \quad \Leftrightarrow \quad \mathrm{d}z = \mathrm{d}z(y) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}\,\mathrm{d}x. \tag{3.8}$$

Therefore, the relation

$$\begin{aligned} P(c \leq z \leq d) &= \int_{f(a)}^{f(b)} p_z(z)\,\mathrm{d}z \\ &= \int_a^b p_z(f(x))\frac{\mathrm{d}f(x)}{\mathrm{d}x}\,\mathrm{d}x = P(a \leq x \leq b) \end{aligned} \tag{3.9}$$

obtains, thus

$$p_x(x) = p_z(f(x))\frac{\mathrm{d}f(x)}{\mathrm{d}x} \tag{3.10}$$

must hold, where $\mathrm{d}f(x)/\mathrm{d}x$ acts as a normalizing factor for the new probability density $p_x(x)$. $\qquad\square$

Let now $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^d$, $d \in \mathbb{N}$ be continuous multivariate random variables with associated probability density functions $p_{\boldsymbol{x}}(\boldsymbol{x})$ and $p_{\boldsymbol{z}}(\boldsymbol{z})$. Furthermore, let

$$\boldsymbol{f} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}) \tag{3.11}$$

be a diffeomorphism transforming $\boldsymbol{x}$ to $\boldsymbol{z}$. In this case, the formula generalizes to the multivariate case as

$$p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\boldsymbol{z}}[\boldsymbol{f}(\boldsymbol{x})] \left| \det \left(\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} \right) \right|$$

$$= p_{\boldsymbol{z}}[\boldsymbol{f}(\boldsymbol{x})] \left| \det \begin{pmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\boldsymbol{x})}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_d(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_d(\boldsymbol{x})}{\partial x_d} \end{pmatrix} \right|. \tag{3.12}$$

3.7 General principle of a normalizing flow

3.7.1 Conceptual formulation of a normalizing flow

Consider some vector $\boldsymbol{x} \in \mathbb{R}^d$ with $d, q \in \mathbb{N}$ and a latent vector $\boldsymbol{z} \in \mathbb{R}^d$. The probability densities for these vectors shall be given by

$$\boldsymbol{x} \sim p_{\boldsymbol{x}}(\boldsymbol{x}), \quad \boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z}) = \mathcal{N}(0,1)^d. \tag{3.13}$$

The aim of a normalizing flow is to learn a diffeomorphism

$$\boldsymbol{f}_{\phi(\hat{\mathbf{x}})} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x}) \tag{3.14}$$

transforming the probability density function $p_{\boldsymbol{x}}$ to $p_{\boldsymbol{z}}$. The conceptual principle of a normalizing flow is visualized in fig. 3.7. Hereby, $\boldsymbol{\phi}(\hat{\mathbf{x}})$ are weights depending on training data $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$, which are associated to a deep neural network implementing the diffeomorphic function $\boldsymbol{f}_{\phi(\hat{\mathbf{x}})}$. Given such a function, using the change of variable formula the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ can be calculated as

$$p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\boldsymbol{z}}[\boldsymbol{f}_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x})] \left| \det \left(\frac{\partial \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x})}{\partial \boldsymbol{x}} \right) \right|. \tag{3.15}$$

Knowing the transformation $\boldsymbol{z} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x})$ and given that $\boldsymbol{z} \sim \mathcal{N}(0,1)^d$, one can easily calculate samples from the distribution $p_{\boldsymbol{x}}(\boldsymbol{x})$ by means of sampling $\boldsymbol{z}$ from a d-dimensional standard normal distribution and obtaining the corresponding $\boldsymbol{x}$ plugging $\boldsymbol{z}$ into the inverse function $\boldsymbol{f}_{\phi(\hat{\mathbf{x}})}^{-1}$, such that $\boldsymbol{x} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}^{-1}(\boldsymbol{z})$.
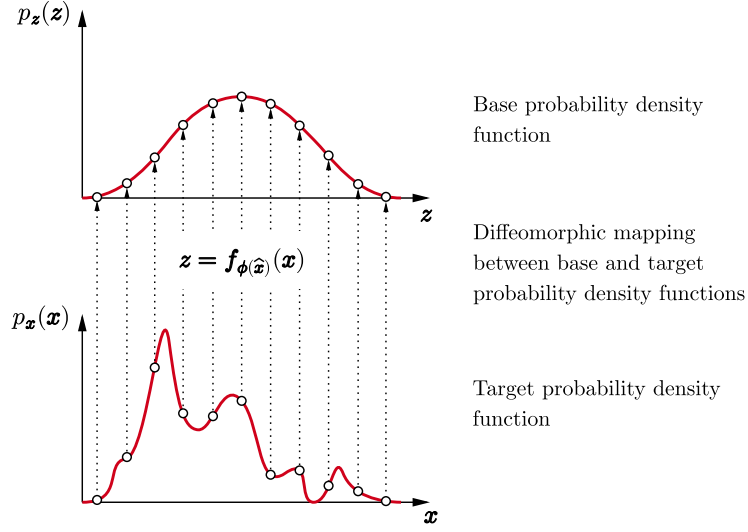
Figure 3.7: Conceptual principle of a normalizing flow $\boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{\phi}(\hat{\mathbf{x}})}(\boldsymbol{x})$ transforming a target probability density function $p_{\boldsymbol{x}}(\boldsymbol{x})$ to a base probability density function $p_{\boldsymbol{z}}(\boldsymbol{z})$.

3.7.2 Requirements for a normalizing flow

Consider a function

$$\boldsymbol{f} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}) \tag{3.16}$$

composed of several functions $\boldsymbol{f}_{(k)}$, $k \in \{1, \ldots, n\}$, $n \in \mathbb{N}$, such that

$$\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}_{(n)} \circ \cdots \circ \boldsymbol{f}_{(1)}(\boldsymbol{x}). \tag{3.17}$$

In this case, the transformation functions $\boldsymbol{f}_{(k)}$ aswell as the function $\boldsymbol{f}$ as a whole should satisfy several conditions to serve purposefully as a tool to sample from a probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ given samples $\boldsymbol{z}$ obtained from a standard normal distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$. According to [Kobyzev et al., 2021], all functions $\boldsymbol{f}_{(k)}$ should satisfy the following conditions:

(1) All transformation functions $\boldsymbol{f}_{(k)}$ need to be invertible and differentiable. In order to transform one probability density into another density, it is necessary to calculate the Jacobian of the transformation function, as stated by the change of variable theorem. Furthermore, the transformation functions should be invertible, because one would like to sample from a standard normal distribution, plug the obtained values $\boldsymbol{z}$ into the inverse transformation function and thereby obtaining samples $\boldsymbol{x}$. Both of these conditions - invertibility and differentiabiliby - are satisfied by diffeomorphic functions.

49

(2) The transformation functions $\boldsymbol{f}_{(k)}$ need to be expressive enough to model real data distributions. The normalizing flow composed of these transformation functions needs to learn a real data distribution; hence enough flexibility of the transformation functions is required in order to allow the neural networks composing the flow to learn actual and potentially complicated probability distributions.

(3) The transformation functions $\boldsymbol{f}_{(k)}$ should be computationally efficient. That is, the Jacobian determinant should be calculable in an as efficient as possible way; furthermore also the application of the trained forward and inverse normalizing flow should consume as little as possible time. In order to train and apply the normalizing flow, i.e. the transformation functions composing the normalizing flow, many Jacobian determinants need to be calculated, as required by the change of variable formula. Therefore, transformation functions with a simple Jacobians allowing for quick and easy calculations of the corresponding determinants are preferable.

3.7.3 Normalizing flows with conditioning data

Consider a model $\boldsymbol{M}(\boldsymbol{x}) = \boldsymbol{y}$, that produces observations $\boldsymbol{y} \in \mathbb{R}^D$, $D \in \mathbb{N}$ based on model parameters $\boldsymbol{x} \in \mathbb{R}^d$, $d, q \in \mathbb{N}$. Furthermore, consider a latent vector $\boldsymbol{z} \in \mathbb{R}^d$. The probability densities for these vectors are given by

$$\boldsymbol{x} \sim p_{\boldsymbol{x}}(\boldsymbol{x}), \quad \boldsymbol{y} \sim p_{\boldsymbol{y}}(\boldsymbol{y}), \quad \boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z}) = \mathcal{N}(0, 1)^d. \tag{3.18}$$

The aim of the normalizing flow is to model the probability density $p_{\boldsymbol{\phi}(\hat{\mathbf{x}}, \hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$, namely the probability density function of $\boldsymbol{x}$, given a certain observation $\boldsymbol{y}$, thereby defining an inversion of the model $\boldsymbol{M}(\boldsymbol{x}) = \boldsymbol{y}$, see fig. 3.8 for a visualization of a normalizing flow with conditioning data; which from now on will be called a conditional normalizing flow. The multidimensional quantity $\boldsymbol{\phi}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ is a set of parameters learned by a neural network, which is used to implement the normalizing flow.

If there is no conditioning data $\boldsymbol{y}$ available, the normalizing flow can directly be applied to learn the probability distribution $p_{\boldsymbol{\phi}(\hat{\mathbf{x}})}(\boldsymbol{x})$ based on standard normally distributed data $z \sim p_{\boldsymbol{z}}(\boldsymbol{z}) = \mathcal{N}(0, 1)^d$. In this case, the conditioning data $\boldsymbol{y}$, i.e. the symbols $\boldsymbol{y}$ and $\hat{\mathbf{y}}$, can be omitted in all formulae pertaining to the current chapter of this document.
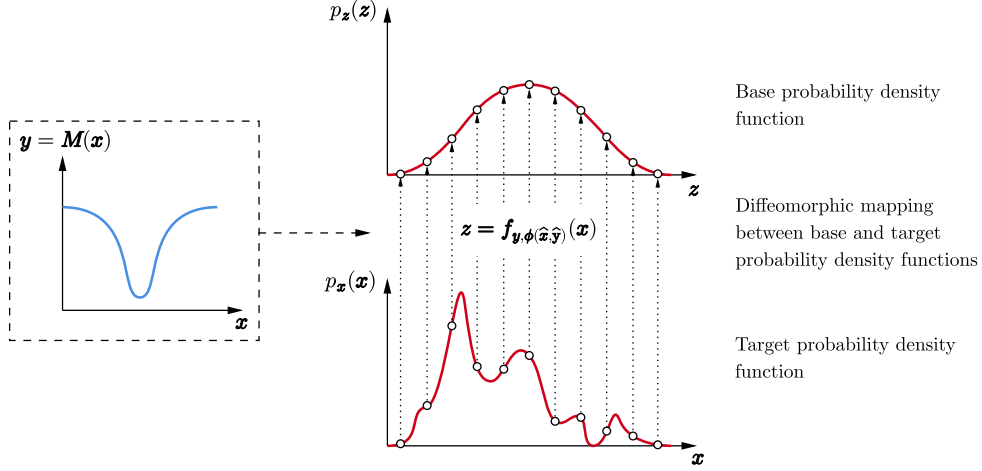
Figure 3.8: Conceptual principle of a conditional normalizing flow $\boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})$, transforming a target probability density function $p_{\boldsymbol{x}}(\boldsymbol{x})$ to a base probability density function $p_{\boldsymbol{z}}(\boldsymbol{z})$ conditional on $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$, where $\boldsymbol{M}$ is a model that relates the data $\boldsymbol{x}$ to the conditioning data $\boldsymbol{y}$.

3.7.4 Coupling layer normalizing flows

Consider a diffeomorphism

$$\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}), \tag{3.19}$$

where $\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})$ are weights depending on training data $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$ and $\hat{\mathbf{y}} \in \mathbb{R}^{D \times q}$, which are associated to a deep neural network implementing the diffeomorphic function $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}$. The probability density for $\boldsymbol{x}$ as calculated via $\boldsymbol{x} = \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}^{-1}(\boldsymbol{z})$ therefore becomes conditional on the context $\boldsymbol{y}$, because the network parameters $\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})$ themselves depend on the context $\hat{\mathbf{y}}$ used to train the network. Hence for $p_{\boldsymbol{x}}(\boldsymbol{x})$, the notation $p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ is introduced.

With the change of variable theorem, one obtains

$$\begin{aligned} 1 = \int_{\mathbb{R}^d} p_{\boldsymbol{z}}(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{z} &= \int_{\mathbb{R}^d} p_{\boldsymbol{z}}[\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})] \left| \det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}} \right) \right| \, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{R}^d} p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y}) \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{3.20}$$

from which it follows, that the probability density $p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ can be expressed by means of the diffeomorphism $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}$, the Jacobian $\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}}$ and the probability density $p_{\boldsymbol{z}}(\boldsymbol{z})$, namely

$$p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y}) = p_{\boldsymbol{z}}[\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})] \left| \det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}} \right) \right|. \tag{3.21}$$

This identity is to be considered as the key to implementing a loss function for the normalizing flow technique.

Since a concatenation of diffeomorphisms is again a diffeomorphism, the function $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}$ can be composed of many diffeomorphisms. Let $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(1)}, \ldots, \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},n}$ with $n \in \mathbb{N}$ be a number of diffeomorphisms, such that

$$\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})} = \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(n)} \circ \cdots \circ \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(1)}. \tag{3.22}$$

Introducing the notation $\boldsymbol{x} \doteq \boldsymbol{z}_{(0)}$, $\boldsymbol{z} \doteq \boldsymbol{z}_{(n)}$ and $\boldsymbol{z}_{(k)} \doteq \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})$ with $k \in \{1, \ldots, n\}$, the function $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}$ can be written as

$$\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{z}_{(0)}) = \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(n)} \circ \cdots \circ \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(1)}(\boldsymbol{z}_{(0)}). \tag{3.23}$$

In fig. 3.9, an overview of the working principle of a coupling layer normalizing flow is visualized. Because $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})$ is a concatenation of
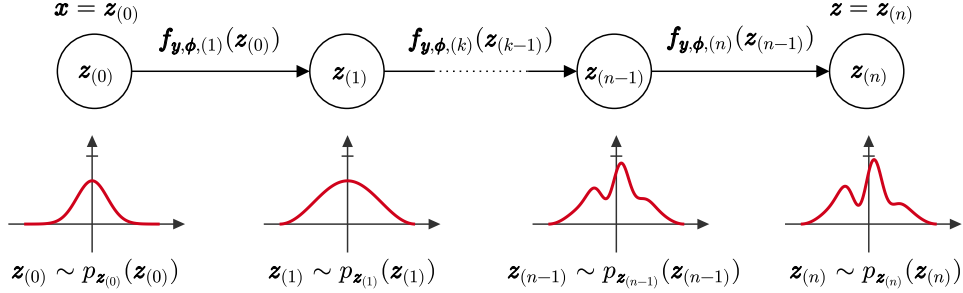
Figure 3.9: Visualization of the working principle of coupling layer normalizing flows.

functions $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})$ with $k \in \{1, \ldots, n\}$, the determinant of the Jacobian for $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})$ with respect to $\boldsymbol{x}$ is given by

$$\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right) = \prod_{k=1}^{n} \det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})}{\partial \boldsymbol{z}_{(k-1)}}\right), \tag{3.24}$$

where the Jacobians on the right-hand side take the form

$$\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})}{\partial \boldsymbol{z}_{(k-1)}} = \begin{pmatrix} \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),1}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),1}} & \cdots & \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),1}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),d}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),d}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),1}} & \cdots & \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),d}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),d}} \end{pmatrix} \tag{3.25}$$

with $k \in \{1, \ldots, n\}$. Every function $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})$ for $k \in \{1, \ldots, n\}$ represents one of the n so-called coupling layers constituting the total

flow function $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$. Every composite function $\boldsymbol{f}_{\boldsymbol{y},\phi,(k)}$ is implemented by application of neural networks depending on $\boldsymbol{x}$ and $\boldsymbol{y}$, such that each composite function is fully invertible and differentiable. The total set of parameters learned by the neural networks is then identified with $\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})$, hence defining the function $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$.

3.8 Affine coupling layer normalizing flow

Affine coupling layer normalizing flow are one of the most simple types of normalizing flows. They basically rely on a concatenation of many affine transformations, where the scale and shift parameters are implemented as neural networks. Although quite easy to implement, these type of flows lack flexibility, as noted by [Durkan et al., 2019, p.3]. However, they provide good insight to the working principle of a normalizing flow and are sufficiently flexible for some less complex applications as shown in chapter 4.

3.8.1 Construction of an affine coupling layer normalizing flow

Consider a diffeomorphism

$$\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})} = \boldsymbol{f}_{\boldsymbol{y},\phi,(n)} \circ \cdots \circ \boldsymbol{f}_{\boldsymbol{y},\phi,(1)} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}), \quad (3.26)$$

where all quantities are given as defined in section 3.7.4.

In the case of $\boldsymbol{f}_{\boldsymbol{y},\phi,(k)}$ being so-called affine coupling layers, the functions can explicitly written by means of two arbitrary functions

$$\boldsymbol{\mu}_{(k)} : \mathbb{R}^{m+D} \to \mathbb{R}^{d-m}, \quad \boldsymbol{\sigma}_{(k)} : \mathbb{R}^{m+D} \to \mathbb{R}^{d-m} \quad (3.27)$$

with $m \in \{1,\ldots,d\}$ and $k \in \{1,\ldots,n\}$ implemented as deep neural networks. That is to say, that the neural networks $\boldsymbol{\mu}_{(k)}$ and $\boldsymbol{\sigma}_{(k)}$ take a vector of dimension $m+D$ as an input and return a vector of dimension $d-m$ as an output; there can be an arbitrary sequence of network layers and activation functions in between input and output. Every mapping $\boldsymbol{z}_{(k)} = \boldsymbol{f}_{\boldsymbol{y},\phi,(k)}(\boldsymbol{z}_{(k-1)})$ for $k \in \{1,\ldots,n\}$ is now defined as an affine transformation, such that

$$f_{\boldsymbol{y},\phi,(k),l}(\boldsymbol{z}_{(k-1)}) = z_{(k-1),l}, \quad (3.28)$$
$$f_{\boldsymbol{y},\phi,(k-1),l}^{-1}(\boldsymbol{z}_{(k)}) = z_{(k),l} \quad (3.29)$$

for $l \in \{1, \ldots, m\}$ and

$$f_{\boldsymbol{y},\boldsymbol{\phi},(k),l}\big(\boldsymbol{z}_{(k-1)}\big) = \mu_{(k),l}\big(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y}\big) + e^{\sigma_{(k),l}(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y})} \cdot z_{(k-1),l}, \quad (3.30)$$

$$f^{-1}_{\boldsymbol{y},\boldsymbol{\phi},(k-1),l}\big(\boldsymbol{z}_{(k)}\big) = \big[z_{(k),l} - \mu_{(k),l}\big(\boldsymbol{z}_{(k),1:m},\boldsymbol{y}\big)\big] \cdot e^{-\sigma_{(k),l}(\boldsymbol{z}_{(k),1:m},\boldsymbol{y})} \quad (3.31)$$

for $l \in \{m+1, \ldots, d\}$, where $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)} = (f_{\boldsymbol{y},\boldsymbol{\phi},(k),1}, \ldots, f_{\boldsymbol{y},\boldsymbol{\phi},(k),d})^\top$ and $\boldsymbol{z}_{(k),1:m} = (z_{(k),1}, \ldots, z_{(k),m})^\top$.

It is a convenient property of such affine functions $\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}\big(\boldsymbol{z}_{(k-1)}\big)$ for $k \in \{1, \ldots, n\}$, that the Jacobians with respect to $\boldsymbol{z}_{(k-1)}$ are triangular, that is

$$\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}\big(\boldsymbol{z}_{(k-1)}\big)}{\partial \boldsymbol{z}_{(k-1)}} = \begin{pmatrix} \mathbb{1}_m & \mathbb{O}_m \\ \frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k),m+1:d}(\boldsymbol{z}_{(k-1)})}{\partial \boldsymbol{z}_{(k-1),m+1:d}} & \operatorname{diag}\left(e^{\sigma_{(k),m+1:d}(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y})}\right) \end{pmatrix}, \quad (3.32)$$

where $\mathbb{1}_m$ is a unity matrix and $\mathbb{O}_m$ is a zero matrix of dimension $m \times m$. The lower left quantity in the above matrix is again a Jacobian and takes the form

$$\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k),m+1:d}\big(\boldsymbol{z}_{(k-1)}\big)}{\partial \boldsymbol{z}_{(k-1),m+1:d}} = \begin{pmatrix} \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),m+1}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),m+1}} & \cdots & \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),m+1}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),d}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),d}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),m+1}} & \cdots & \frac{\partial f_{\boldsymbol{y},\boldsymbol{\phi},(k),d}(\boldsymbol{z}_{(k-1)})}{\partial z_{(k-1),d}} \end{pmatrix}; \quad (3.33)$$

the lower right quantity however is a diagonal matrix with the elements $e^{\sigma_{(k),l}(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y})}$ for $l \in \{m+1, d\}$ on the diagonal. Therefore, the Jacobian $\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}(\boldsymbol{z}_{(k-1)})/\partial \boldsymbol{z}_{(k-1)}$ is triangular and hence its determinant is given by the product of the diagonal elements, that is

$$\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}\big(\boldsymbol{z}_{(k-1)}\big)}{\partial \boldsymbol{z}_{(k-1)}}\right) = \prod_{j=m+1}^{d} e^{\sigma_{(k),j}(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y})}. \quad (3.34)$$

Taking the logarithm of this expression results in a transformation of the product to a sum, such that

$$\log\left[\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}\big(\boldsymbol{z}_{(k-1)}\big)}{\partial \boldsymbol{z}_{(k-1)}}\right)\right] = \sum_{j=m+1}^{d} \sigma_{(k),j}\big(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y}\big) \quad (3.35)$$

holds. Recall, that the complete normalizing flow is given as a concatenation of affine coupling functions (layers), such that

$$\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right) = \prod_{k=1}^{n} \det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi},(k)}\big(\boldsymbol{z}_{(k-1)}\big)}{\partial \boldsymbol{z}_{(k-1)}}\right). \quad (3.36)$$

Taking the logarithm of this expression and inserting the previous result eq. (3.35), this leads to

$$\log\left[\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right)\right] = \sum_{k=1}^{n} \log\left[\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\phi,(k)}(\boldsymbol{z}_{(k-1)})}{\partial \boldsymbol{z}_{(k-1)}}\right)\right]$$
$$= \sum_{k=1}^{n} \sum_{j=m+1}^{d} \sigma_{(k),j}(\boldsymbol{z}_{(k-1),1:m}, \boldsymbol{y}). \tag{3.37}$$

3.8.2 Training process for affine coupling layer normalizing flows

Recalling the diffeomorphism $\boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})$ and the change of variable formula for the multivariate case, the probability density $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ can be written as found in eq. (3.21). Taking the natural logarithm of expression eq. (3.21) and inserting the result eq. (3.37) into it, one obtains

$$\log\left[p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})\right] = \log\left(p_{\boldsymbol{z}}[\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})]\right)$$
$$+ \left|\sum_{k=1}^{n} \sum_{j=m+1}^{d} \sigma_{(k),j}(\boldsymbol{z}_{(k-1),1:m}, \boldsymbol{y})\right|. \tag{3.38}$$

Note, that the natural logarithm $\log(a)$ is strictly monotonic increasing for $a > 0$; furthermore, $-\infty < \log(a) \leq 0$ for $0 < a \leq 1$. Therefore, maximizing the so-called likelihood $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ is equivalent to maximizing $\log\left[p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})\right]$, which is named the log-likelihood. Maximizing $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ corresponds to finding optimal network parameters $\phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ for the deep neural networks $\boldsymbol{\mu}_{(k)}$ and $\boldsymbol{\sigma}_{(k)}$ for $k \in \{1, \ldots, n\}$ constituting the flow function $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$, such that the most suitable model parameters $\boldsymbol{x}$ given some context $\boldsymbol{y}$ are found. In the case of a stellar atmosphere inversion, $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ represents a probability distribution for the solutions $\boldsymbol{x}$ of an inversion $\boldsymbol{x} = \boldsymbol{M}^{-1}(\boldsymbol{y})$, where $\boldsymbol{M}$ is the atmosphere model and $\boldsymbol{y}$ are observations.

That is to say, that the normalizing flow, i.e. the neural networks constituting it, are trained by means of minimizing the negative log-likelihood $-\log\left[p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})\right]$, which is equivalent to maximizing positive log-likelihood and hence also to maximizing the positive likelihood. Therefore, the loss function to minimize by any gradient-descent and

backpropagation algorithm employed can be defined as

$$\mathcal{L}_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x},\boldsymbol{y}) = -\log\left(p_{\boldsymbol{z}}[\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})]\right) - \log\left(\left|\det\left(\frac{\partial \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right)\right|\right)$$

$$= -\log\left(p_{\boldsymbol{z}}[\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x})]\right) - \left|\sum_{k=1}^{n}\sum_{j=m+1}^{d}\sigma_{(k),j}(\boldsymbol{z}_{(k-1),1:m},\boldsymbol{y})\right|. \tag{3.39}$$

This loss function is minimized with respect to the neural network parameters $\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})$.

3.9 Piecewise rational quadratic spline normalizing flow

Affine coupling layer normalizing flows as describe in detail above are simple and relatively easy to implement, but they lack flexibility. As [Durkan et al., 2019, p.3] note, affine coupling layer flows may struggle to model especially multimodal density functions using just affine transformations. Recall, that as a base density, usually a standard normal distribution is taken. If there are now multiple modes present in the target distribution of the data to model, the flow needs to learn how to transform just one mode present in the base distribution to the multiple modes in the multimodal target distribution, thus presenting the flow with the need to learn potentially highly nonlinear transformation. This is where affine coupling layer normalizing flows are at their limit.

Hence, [Durkan et al., 2019, p.4] introduce piecewise rational quadratic spline normalizing flows, which are more flexible in nature, but mathematically still relatively simple and furthermore fully invertible and differentiable, thus diffeomorphic.

This type of normalizing flow is therefore the flow of choice for the more complex applications as presented and discussed in chapter 5 and chapter 6. These kind of flows were however not implemented from scratch by the author, but the already existing software [Durkan et al., 2020] was made use of; and the theoretical considerations below follow the material as presented by [Durkan et al., 2019].

3.9.1 Construction of a piecewise rational quadratic spline normalizing flow

Consider a diffeomorphism

$$f_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})} = f_{\boldsymbol{y},\phi,(n)} \circ \cdots \circ f_{\boldsymbol{y},\phi,(1)} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{x} \mapsto \boldsymbol{z} = f_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}). \tag{3.40}$$

In the case of this function being implemented by piecewise rational quadratic spline coupling layers, these individual layers can be defined as follows.

Again, $\boldsymbol{x} \sim p_{\boldsymbol{x}}(\boldsymbol{x})$ and $\boldsymbol{y} \sim p_{\boldsymbol{y}}(\boldsymbol{y})$ hold, whereas the modelled probability density of $\boldsymbol{x}$ conditional on $\boldsymbol{y}$ as learned by a normalizing flow is denoted as $\boldsymbol{x} \sim p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$, where $\boldsymbol{\phi}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ are weights associated to the neural network(s) used to implement the flow. The quantities $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$ and $\hat{\mathbf{y}} \in \mathbb{R}^{D \times q}$ are just training data.

Consider an input vector $\boldsymbol{x} \in \mathbb{R}^d$ and a conditioning vector $\boldsymbol{y} \in \mathbb{R}^D$. Let now be a vector $\boldsymbol{s}$ defined, such that it contains the first m entries of the input vector $\boldsymbol{x}$ and the whole conditioning vector $\boldsymbol{y}$, that is to say, $\boldsymbol{s} \doteq (x_1, \ldots, x_m, y_1, \ldots, y_D) \in \mathbb{R}^{m+D}$. In order to define a single layer $f_{\boldsymbol{y},\boldsymbol{\phi},(k)}$ with $k \in \{1, \ldots, n\}$ of the flow, one requires $(d-m)(3K-1)$ parameters per layer k These parameters are given for each layer k by means of a neural network

$$\boldsymbol{\mu}_{(k)} : \mathbb{R}^{m+D} \to \mathbb{R}^{(d-m)(3K-1)}, \quad \boldsymbol{s}_{(k-1)} \mapsto \boldsymbol{\phi}_{(k)} = \boldsymbol{\mu}_{(k)}\big(\boldsymbol{s}_{(k-1)}\big) \qquad (3.41)$$

whereby $m \in \{1, \ldots, d\}$, $k \in \{1, \ldots, n\}$ and $K \in \mathbb{N}$. The integer K is the number of spline knots of choice. The precise nature of the mapping $\boldsymbol{\mu}_{(k)}$ is that of an arbitrary neural network, which therefore cannot be explicitly written, but can be accounted for by

$$\boldsymbol{\phi}_{(k)} = \boldsymbol{\mu}_{(k)}\big(\boldsymbol{s}_{(k-1)}\big) = \begin{pmatrix} \phi^w_{(k),m+1,1} \\ \vdots \\ \phi^w_{(k),i,j} \\ \vdots \\ \phi^w_{(k),d,K} \\ \phi^h_{(k),m+1,1} \\ \vdots \\ \phi^h_{(k),i,j} \\ \vdots \\ \phi^h_{(k),d,K} \\ \phi^\delta_{(k),m+1,1} \\ \vdots \\ \phi^\delta_{(k),i,j} \\ \vdots \\ \phi^\delta_{(k),d,K-1} \end{pmatrix} \in \mathbb{R}^{(d-m)(3K-1)} \qquad (3.42)$$

in the general case. Let furthermore $\boldsymbol{\phi}_{(k),(i)}$ with $i \in \{m+1,\ldots,d\}$ be defined as

$$\boldsymbol{\phi}_{(k),(i)} \doteq (\phi^w_{(k),i,1},\ldots,\phi^w_{(k),i,K},\phi^h_{(k),i,1},\ldots,\phi^h_{(k),i,K},\phi^\delta_{(k),i,1},\ldots,\phi^\delta_{(k),i,K-1})^\top. \tag{3.43}$$

The letter w denotes the with of bins, h the height of bins and δ the derivatives at the internal points.

The last ingredients to be finally able to define the individual transformations $\boldsymbol{f}_{\boldsymbol{y},\phi,(k)}$ are $d-m$ actual quadratic functions

$$g_{\boldsymbol{\phi}_{(k),(i)}} : \mathbb{R} \to \mathbb{R} \tag{3.44}$$

for each coupling layer k. Now, each coupling layer k is defined as

$$\boldsymbol{f}_{\boldsymbol{y},\phi,(k)} : \mathbb{R}^d \to \mathbb{R}^d, \quad \boldsymbol{z}_{(k-1)} \mapsto \boldsymbol{z}_{(k)} = \begin{pmatrix} z_{(k-1),1} \\ \vdots \\ z_{(k-1),m} \\ g_{\boldsymbol{\phi}_{(k),(m+1)}}\big(z_{(k-1),m+1}\big) \\ \vdots \\ g_{\boldsymbol{\phi}_{(k),(d)}}\big(z_{(k-1),d}\big) \end{pmatrix}, \tag{3.45}$$

where $\boldsymbol{x} = \boldsymbol{z}_{(0)}$ and $\boldsymbol{z} = \boldsymbol{z}_{(n)}$. Thus, the total diffeomorphism $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$ is just given by the concatenation of all $\boldsymbol{f}_{\boldsymbol{y},\phi,(k)}$ for $k \in \{1,\ldots,n\}$, namely

$$\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})} = \boldsymbol{f}_{\boldsymbol{y},\phi,(n)} \circ \cdots \circ \boldsymbol{f}_{\boldsymbol{y},\phi,(1)}, \tag{3.46}$$

whereby the piecewise rational quadratic spline normalizing flow is defined.

3.9.2 Training process for piecewise rational quadratic spline normalizing flows

As the precice tranformation rules for the piecewise rational quadratic spline flow are quite complex, the exact formula for the loss function allowing for a direct implementation is not given here. However, a few remarks shall be made with reference to training such a normalizing flow.

Since all involved functions used to define the piecewise rational quadratic spline flow are fully differentiable and invertible, one can just use eq. (3.21) as a means to implement the loss function $\mathcal{L}_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x},\boldsymbol{y})$. What one has to do to this end is to take negative logarithm of the probability density $p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})(\boldsymbol{x}|\boldsymbol{y})}$ as the loss function, i.e.

$$\mathcal{L}_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x},\boldsymbol{y}) = -\log\left[p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})\right]. \tag{3.47}$$

Minimizing the loss function now means to minimize $\log\left[1/p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\boldsymbol{y}})}(\boldsymbol{x}|\boldsymbol{y})\right]$ and therefore to maximize the likelihood $p_{\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\boldsymbol{y}})}(\boldsymbol{x}|\boldsymbol{y})$. This is to say, that one finds the most likely input $\boldsymbol{x}$ pertaining to a certain context $\boldsymbol{y}$, if the loss function is truly globally minimized. The logarithm comes into play, because multiplications then turn into additions, which are much easier to implement than multiplications. However, the exact calculation the determinant $\det\left(\partial\boldsymbol{f}_{\boldsymbol{y},\boldsymbol{\phi}(\hat{\mathbf{x}},\hat{\boldsymbol{y}})}(\boldsymbol{x})/\partial\boldsymbol{x}\right)$ is tedious and not insightful, therefore it is not given here.

Chapter 4

Proof of concept: Affine coupling layer normalizing flows

This section provides several experiments and applications of an affine coupling layer normalizing flow as described in detail in section 3.8. Hereby, the implementation was carried out from scratch in `Python` following section 3.8, using the `PyTorch` framework.

4.1 Moons

4.1.1 Conceptual formulation

In this experiment, a normalizing flow using affine coupling layers was tested on data $\boldsymbol{x} = (x_1, x_2)^\top \in \mathbb{R}^2$ given by a probability density function $p_{\boldsymbol{x}}(\boldsymbol{x})$ representing two moons. The goal was to map the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ representing the moons distribution to a standard normal distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$, i.e. to find a diffeomorphism

$$\boldsymbol{f}_{\phi(\hat{\mathbf{x}})} : \mathbb{R}^2 \to \mathbb{R}^2, \quad p_{\boldsymbol{x}}(\boldsymbol{x}) \sim \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x}) \sim \mathcal{N}(0,1)^2. \tag{4.1}$$

To this end, a normalizing flow to implement the function $\boldsymbol{f}_{\phi(\hat{\mathbf{x}})}$ was created as described in section 3.8. Herewith, $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$ denotes the training data, where $d = 2$ and $q \in \mathbb{N}$. It is to be remembered, that the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ can be written as being conditional on the network parameters $\boldsymbol{\phi}(\hat{\mathbf{x}})$. Hence, the notation $p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\phi(\hat{\mathbf{x}})}(\boldsymbol{x})$ is adopted to indicate, that for a trained normalizing flow, $\boldsymbol{x}$ can be written as $\boldsymbol{x} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}})}^{-1}(\boldsymbol{z})$ and that the density $p_{\boldsymbol{x}}(\boldsymbol{x})$ as sampled by application of

the inverse normalizing flow to normally distributed samples $\boldsymbol{z}$ is therefore dependent upon the network parameters $\boldsymbol{\phi}(\hat{\mathbf{x}})$.

4.1.2 Results

The subsequent results and figures were obtained[1] using a training sample size of 100000, a batch size of 512, a hidden layer size of 512, a coupling layer amount of 10, a learning rate of 0.001, a scheduling rate of 0.999 and 25 epochs to train the flow. The training process is visualized by
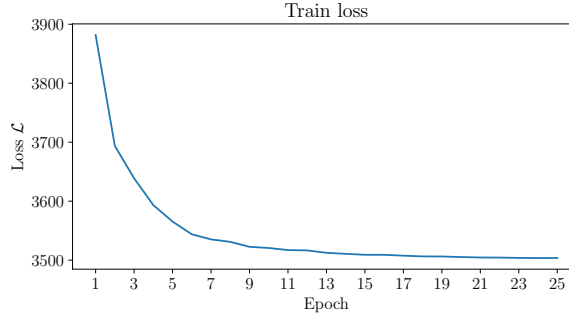

Figure 4.1: Visualization of the training process in terms of losses vs. epochs.

fig. 4.1.

The performance after training can be seen as shown in fig. 4.2. It is evident, that the normalizing flow clearly has learned most of the defining properties of the probability distribution for $\boldsymbol{z}$. While the presented results are not perfect, it can safely be assumed that nearly perfect results would obtain if the epoch number for the training process would be chosen as going to infinity. Perfect hereby means, that the flow $\boldsymbol{f}_{\boldsymbol{\phi}(\hat{\mathbf{x}})}$ maps samples $\boldsymbol{x}$ from the distribution $p_{\boldsymbol{x}}(\boldsymbol{x})$ to samples $\boldsymbol{z}$ from the standard normal distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$ without any mismaps.

---

[1]The code which produced the presented results for the moons example is available at `https://github.com/danielzahnd/nf-moons-example`.
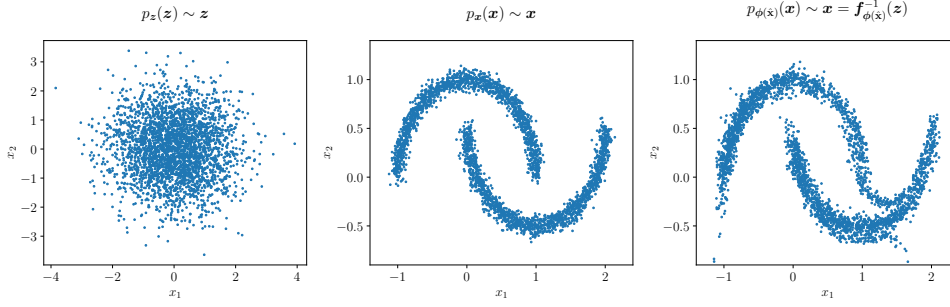
Figure 4.2: Results for a normalizing flow implementing the moons example.

4.2 Linear regression

4.2.1 Conceptual formulation

In this experiment, a normalizing flow using affine coupling layers was tested on data $\boldsymbol{x} = (x_1, x_2)^\top = (a, b)^\top \in \mathbb{R}^2$ and associated context $\boldsymbol{y} = (y_1, \ldots, y_{10})^\top \in \mathbb{R}^{10}$ given by probability density functions $p_{\boldsymbol{x}}(\boldsymbol{x})$ and $p_{\boldsymbol{y}}(\boldsymbol{y})$. Hereby, the context $\boldsymbol{y}$ is generated based on $\boldsymbol{x}$ by means of a model

$$\boldsymbol{M} : \mathbb{R}^2 \to \mathbb{R}^{10}, \quad \boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x}), \tag{4.2}$$

where the model $\boldsymbol{M}$ is defined as

$$y_j = M_j(\boldsymbol{x}) = x_1 u_j + x_2 = a u_j + b_j, \quad j \in \{1, \ldots, 10\} \tag{4.3}$$

with $\boldsymbol{u} = (u_1, \ldots, u_{10})^\top \in \mathbb{R}^{10}$ being a vector containing 10 equally spaced numbers between -5 and 5.

The goal was to map the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ representing the distributions of the slopes and intercepts of an affine function to a standard normal distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$, i.e. to find a diffeomorphism

$$\boldsymbol{f}_{\boldsymbol{y}, \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})} : \mathbb{R}^2 \to \mathbb{R}^2, \quad p_{\boldsymbol{x}}(\boldsymbol{x}) \sim \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y}, \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})}(\boldsymbol{x}) \sim \mathcal{N}(0, 1)^2. \tag{4.4}$$

To this end, a normalizing flow with context was created as described above and in section 3.8, in order to implement the function $\boldsymbol{f}_{\boldsymbol{y}, \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})}$. Herewith, $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$ and $\hat{\mathbf{y}} \in \mathbb{R}^{D \times q}$ denote the training data, where $d = 2$, $D = 10$ and $q \in \mathbb{N}$. It is to be remembered, that the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ can be written as being conditional on $\boldsymbol{y}$, since the two are intertwined by $\boldsymbol{x} = \boldsymbol{M}^{-1}(\boldsymbol{y})$. Furthermore and because this conditional probability is to be learned by the normalizing flow $\boldsymbol{f}_{\boldsymbol{y}, \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})}$, the notation $p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ is adopted to indicate, that for a trained

normalizing flow, $\boldsymbol{x}$ can be written as $\boldsymbol{x} = \boldsymbol{f}^{-1}_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{z})$ and that the density $p_{\boldsymbol{x}}(\boldsymbol{x})$ as sampled by application of the inverse normalizing flow to normally distributed samples $\boldsymbol{z}$ is therefore dependent upon the network parameters $\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})$.

4.2.2 Results

The following results and figures were obtained[2] using a training sample size of 90000, a batch size of 512, a hidden layer size of 128, a coupling layer amount of 8, a learning rate of 0.001, a scheduling rate of 0.999 and 25 epochs to train the flow. In fig. 4.3, the training process of the
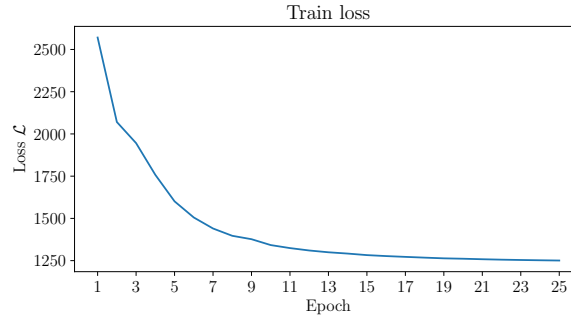

Figure 4.3: Visualization of the training process in terms of losses vs. epochs.

normalizing flows is visualized in terms of the training loss $\mathcal{L}$.

In fig. 4.4, the latent densities for the slope and intercept parameters can be seen; these density plots show how the training samples look like, if they are inserted into the normalizing flow function; that is to say, if $\hat{\boldsymbol{x}}_j$ and $\hat{\boldsymbol{y}}_j$ with $j \in \{1,\ldots,L\}$, $L \in \mathbb{N}$ denote training samples, then one obtains samples $\bar{\boldsymbol{z}}_j$ by means of calculating

$$\bar{\boldsymbol{z}}_j = \boldsymbol{f}_{\hat{\boldsymbol{y}},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\hat{\boldsymbol{x}}_j), \tag{4.5}$$

which should be standard normally distributed. As is readily seen, the mean and standard deviations for both the slope and intercept parameters are close to 0 and 1, as they should be. The standard deviation for the latent density of the intercept parameter b is however somewhat off of 1, as it should be; this could likely resolved by iteratively optimizing the training settings.

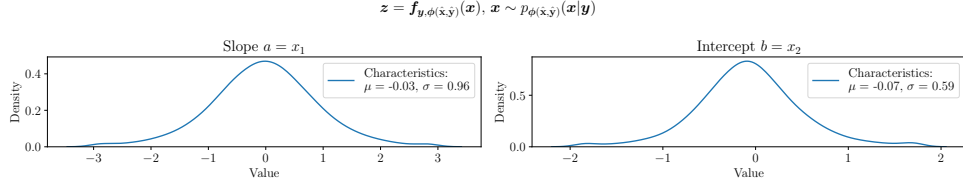[2]The code which produced the presented results for the linear regression example is available at `https://github.com/danielzahnd/nf-linear-regression-example`.

Figure 4.4: Calculated target densities.

$$p_{\phi(\hat{\mathbf{x}}, \hat{\boldsymbol{y}})}(\boldsymbol{x}|\boldsymbol{y}) \sim \boldsymbol{x} = \boldsymbol{f}_{\boldsymbol{y}, \phi(\hat{\mathbf{x}}, \hat{\boldsymbol{y}})}^{-1}(\boldsymbol{z})$$
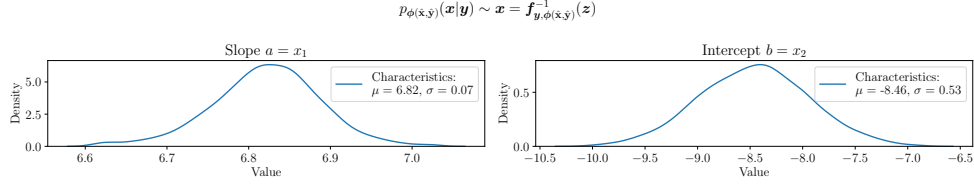

Figure 4.5: Calculated base densities.

Moreover, in fig. 4.5, the calculated densities for the slope and intercept parameters are depicted for a particular test observation $\boldsymbol{y}_{test}$ given to the normalizing flow as context. In particular, $\boldsymbol{y}_{test}$ consists of 10 values calculated by means of the above defined linear model eq. (4.2) based on test parameters $\boldsymbol{x}_{test} = (x_1, x_2)^\top = (a, b)^\top = (6.82, -8.56)$.

Furthermore, fig. 4.6 provides a view of the distributions aswell as of correlations for both the slope and intercept parameters. The corner and
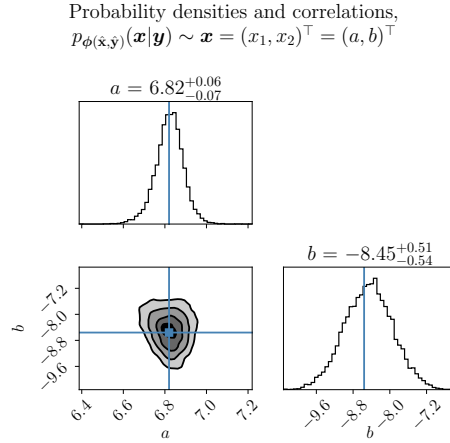

Figure 4.6: Corner plot for the calculated densities of slope a and intercept parameter b using the `corner` package.

pairs plots show nicely, that the normalizing flow indeed was successfully trained to perform linear regressions. First of all, the created Gaussian

noise on observations used to train the normalizing flow is reflected in the Gaussian-shaped densities for the slope and intercept parameters. Secondly, the mean values and the associated standard deviations show, that the true values for a and b indeed lie within the range covered by one standard deviation around the mean of the considered parameter a or b.

4.3 Polynomial regression

4.3.1 Conceptual formulation

Text.

4.3.2 Results

Text.

4.4 Feature extraction

4.4.1 Conceptual formulation

In these experiments, a normalizing flow using affine coupling layers was tested on data $\boldsymbol{x} = (x_1, \ldots, x_1 0)^\top \in \mathbb{R}^{10}$ and associated context $\boldsymbol{y} = (y_1, \ldots, y_{240})^\top \in \mathbb{R}^{240}$ given by probability density functions $p_{\boldsymbol{x}}(\boldsymbol{x})$ and $p_{\boldsymbol{y}}(\boldsymbol{y})$. Hereby, the context $\boldsymbol{y}$ are spectra of the Sun taken by IRIS satellite. From these spectra, altogether 10 features for each spectrum can be extracted by means of a program implementing a function

$$\boldsymbol{M} : \mathbb{R}^{240} \to \mathbb{R}^{10}, \quad \boldsymbol{y} \mapsto \boldsymbol{x} = \boldsymbol{M}(\boldsymbol{y}). \tag{4.6}$$

The goal was to map the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ representing the distributions of the extracted features to a standard normal distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$, i.e. to find a diffeomorphism

$$\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})} : \mathbb{R}^{10} \to \mathbb{R}^{10}, \quad p_{\boldsymbol{x}}(\boldsymbol{x}) \sim \boldsymbol{x} \mapsto \boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}) \sim \mathcal{N}(0,1)^{10}. \tag{4.7}$$

To this end, a normalizing flow with context was created as described above and in section 3.8, in order to implement the function $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$. Here, $\hat{\mathbf{x}} \in \mathbb{R}^{d \times q}$ and $\hat{\mathbf{y}} \in \mathbb{R}^{D \times q}$ denote the training data, where $d = 10$, $D = 240$ and $q \in \mathbb{N}$. It is to be remembered, that the probability density $p_{\boldsymbol{x}}(\boldsymbol{x})$ can be written as being conditional on $\boldsymbol{y}$, since the two are

intertwined by $\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{y})$. Furthermore and because this conditional probability is to be learned by the normalizing flow $\boldsymbol{f}_{\boldsymbol{y},\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}$, the notation $p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}(\boldsymbol{x}|\boldsymbol{y})$ is adopted to indicate, that for a trained normalizing flow, $\boldsymbol{x}$ can be written as $\boldsymbol{x} = \boldsymbol{f}_{\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})}^{-1}(\boldsymbol{z})$ and that the density $p_{\boldsymbol{x}}(\boldsymbol{x})$ as sampled by application of the inverse normalizing flow to normally distributed samples $\boldsymbol{z}$ is therefore dependent upon the network parameters $\phi(\hat{\mathbf{x}},\hat{\mathbf{y}})$.

Given the similarity of the feature extraction task to performing inversions on a stellar atmosphere, a successful outcome of corresponding experiments is to be judged as a powerful proof of concept for the application of normalizing flows to inversions of the Sun's atmosphere.

4.4.2 Results: Test 1

As a first test for the normalizing flow implemented as described above, the flow was trained on all available data; that is to say, that a train and test split of the data at hand was created, whereby the flow was subsequently trained on the train split of the data. The trained flow was then tested by means of propagating a spectrum from the test split through the trained normalizing flow. It was then tested, if the predicted feature distribution for the given test spectrum matched the exact feature values $\boldsymbol{x}$ as calculated by $\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{y})$.

The following results and figures were obtained[3] using a training sample size of 39952, a batch size of 512, a hidden layer size of 64, a coupling layer amount of 6, a learning rate of 0.001, a scheduling rate of 0.999 and 60 epochs to train the flow.

In fig. 4.7, the training process of the normalizing flows is visualized in terms of the train and test losses.

Moreover, in fig. 4.8, the calculated densities and correlations of feature parameters are depicted for a particular test observation $\boldsymbol{y}_{test}$ given to the normalizing flow as context. In particular, $\boldsymbol{y}_{test}$ belongs to the test set of the data only used for testing; which is to say, that the normalizing flow has no knowledge of the test observation. The blue lines in the mentioned plot indicate the true values. As it is evident from said figure, the normalizing flow predicts feature distributions agreeing with the true values for the features very well for most of the features. There only seem to be some inaccuracies with the line width and triplet emission features,

---

[3]The code which produced the presented results for the first test of the feature extraction example is available at `https://github.com/danielzahnd/nf-feature-extraction-example-1`.
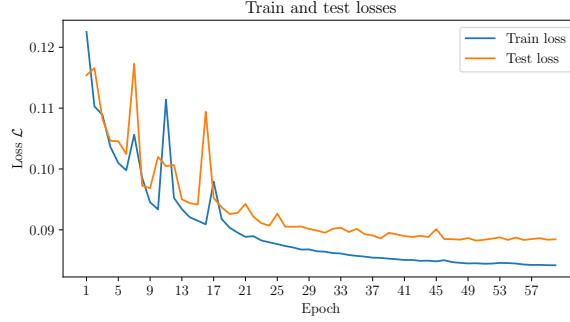
Figure 4.7: Visualization of the training process in terms of train and test losses vs. epochs.

but iteratively optimizing the training settings and the structure of the normalizing flow is likely to improve results even for these two mentioned features.

4.4.3   Results: Test 2

As a second test for the normalizing flow implemented as described above, the flow was trained on an artificially created dataset of spectra and corresponding features. This dataset was created by means of first calculating an average spectrum based on the dataset used in the first test mentioned in section 4.4.2. To this average spectrum, random Gaussian noise was added, thus creating a dataset of Gaussian distributed spectra centered around the average spectrum. This was the dataset used to train the normalizing flow on. After training, the normalizing flow was tested on the average spectrum; this is to say, that the average spectrum was propagated through the normalizing flow, which allows for checking, whether the predicted feature distribution for the average spectrum is indeed Gaussian or not. Based on the Gaussian distributed spectra used to train the flow, a Gaussian distribution of the features densities as predicted by the normalizing flow should be expected.

The following results and figures were obtained[4] using a training sample size of 44671, a batch size of 512, a hidden layer size of 64, a coupling layer amount of 6, a learning rate of 0.001, a scheduling rate of 0.999 and 60 epochs to train the flow.

---

[4]The code which produced the presented results for the first test of the feature extraction example is available at `https://github.com/danielzahnd/nf-feature-extraction-example-2`.
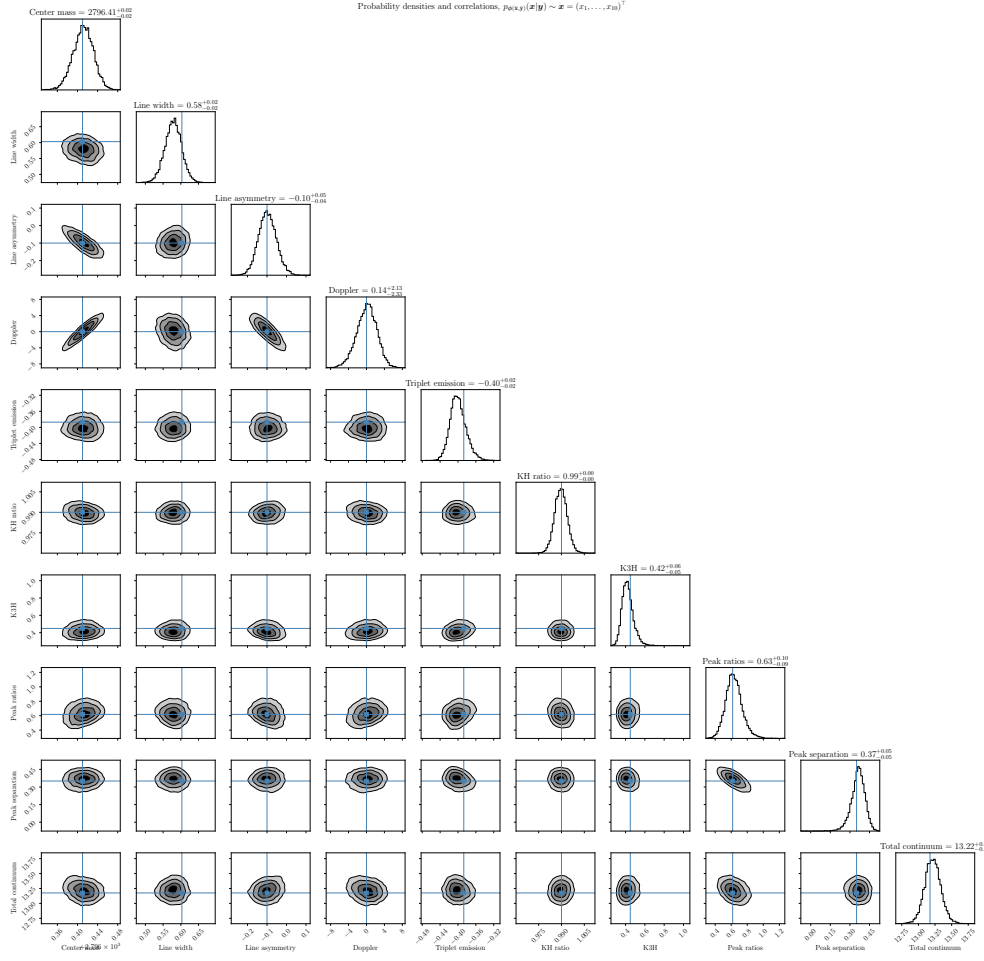
67

Figure 4.8: Corner plot for the calculated densities of feature parameters using the `corner` package.

In fig. 4.9, the training process of the normalizing flows is visualized in terms of the train and test losses.

Moreover, in fig. 4.10, the calculated densities and correlations of feature parameters are depicted for a particular test observation $\boldsymbol{y}_{test}$ given to the normalizing flow as context; in this case, $\boldsymbol{y}_{test}$ is identical to the average spectrum. The blue lines in the mentioned plot indicate the true values. As it is evident from said figure, the normalizing flow predicts feature distributions agreeing with the true values for the features very well for most of the features. There only seem to be some inaccuracies with the line width and triplet emission features, but iteratively optimizing the training settings and the structure of the normalizing flow is
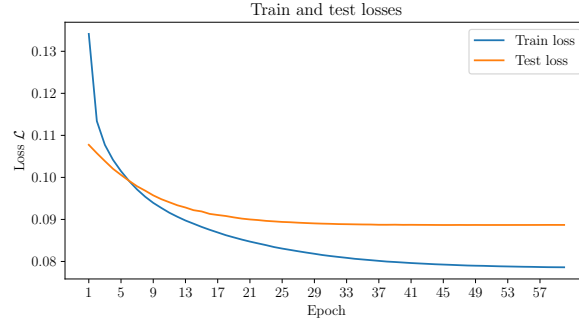
Figure 4.9: Visualization of the training process in terms of train and test losses vs. epochs.

likely to improve results even for these two mentioned features. As for the total continuum feature, there seem to be two peaks in the density distribution, the true values for the features corresponding to the average spectrum situated nicely between those peaks, as it would be expected.

4.4.4 Results: Test 3

As a third test for the normalizing flow implemented as described above, the flow was trained on an a dataset containing spectra and corresponding features of two very different categories. This dataset was created by means of first clustering the available data into several clustering algorithm[5]. After that, the two most different clusters of spectra and corresponding features were combined to one dataset, which was subsequently split into a train and test dataset. The normalizing flow was subsequently trained on the train split of the data. The trained flow was then tested by means of propagating the centroid spectra for both clusters of spectra contained in the train dataset through the trained normalizing flow. It was then checked, whether the predicted feature distributions for these two centroid spectra matched the exact feature values $\boldsymbol{x}$ as calculated by $\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{y})$.

The following results and figures were obtained[6] using a training sample size of 4018, a batch size of 512, a hidden layer size of 256, a coupling layer amount of 6, a learning rate of 0.001, a scheduling rate of 0.999 and 35 epochs to train the flow.

---

[5]In this case, the `KMeans` function of the `sklearn.cluster` package was used.

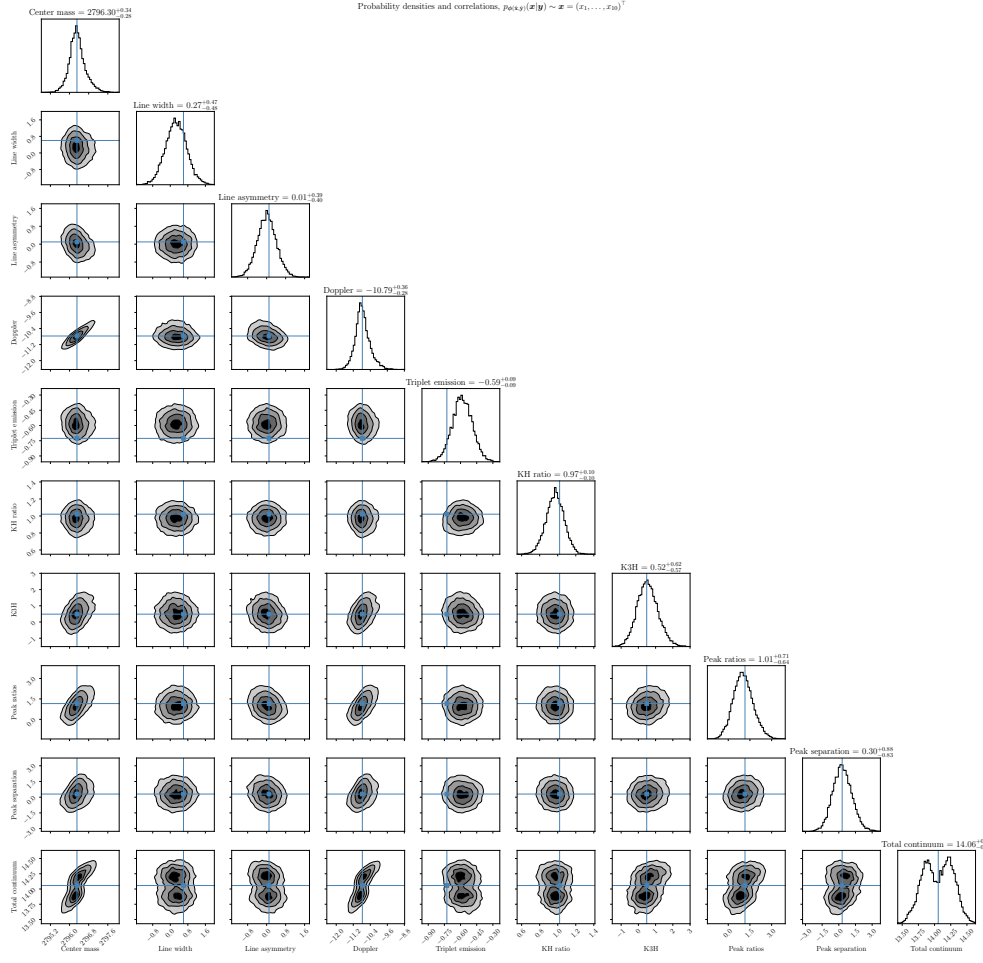[6]The code which produced the presented results for the first test of the feature extraction example is available at `https://github.com/danielzahnd/nf-feature-extraction-example-3`.

69

Figure 4.10: Corner plot for the calculated densities of feature parameters using the `corner` package.

In fig. 4.11, the training process of the normalizing flows is again visualized in terms the train and test losses.

Furthermore, in fig. 4.12, the calculated densities and correlations of feature parameters are depicted for two test observations given to the normalizing flow as context; in this case, the centroid spectra of both categories of spectra used to train the normalizing flow on were used as test observations. The blue lines in the mentioned plot indicate the true values for the feature distributions, whereas orange/grey differentiate, whether a feature distribution corresponds to either one or the other centroid spectrum. As it is evident from said figure, the normalizing flow predicts the feature distributions for both centroid spectra to agree with
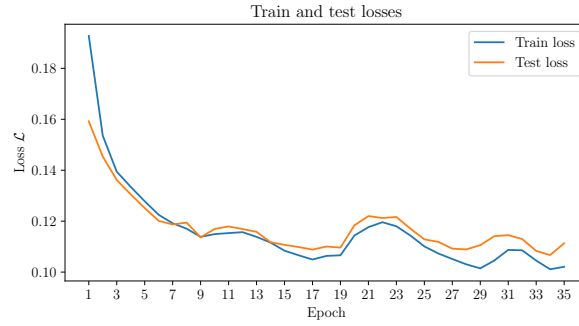
Figure 4.11: Visualization of the training process in terms of train and test losses vs. epochs.

the true values well for only about half of the features. This could however be resolved by means of using a larger dataset to train the normalizing flow on or by iteratively optimizing the structure of the flow and the training settings of it. In general however, the normalizing flow seems to be able to clearly differentiate, if an input spectrum belongs to one or the other category of spectra.
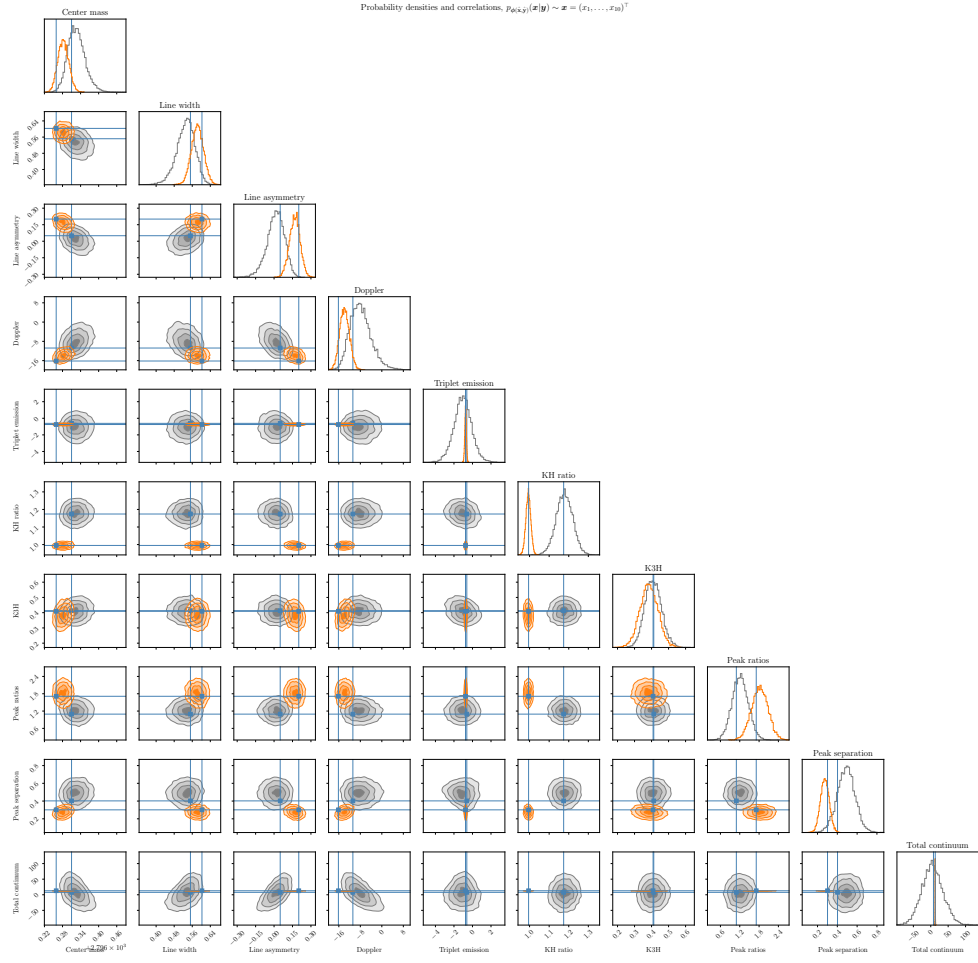
Figure 4.12: Corner plot for the calculated densities of feature parameters using the `corner` package.

Chapter 5

Results (tbd)

5.1 Experiment 1

Text.

5.2 Experiment 2

Text.

Chapter 6

Discussion and interpretation (tbd)

6.1 Experiment 1

Text.

6.2 Experiment 2

Text.

Chapter 7

Appendix

7.1 Calculating a posterior probability density function

Suppose, one has some data $\boldsymbol{x} \in \mathbb{R}^d$ and corresponding data $\boldsymbol{y} \in \mathbb{R}^D$ connected to one another by a model $\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})$, where $d, D \in \mathbb{N}$. Furthermore, assume that $\boldsymbol{x}$ and $\boldsymbol{y}$ follow the probability density functions $\boldsymbol{x} \sim p(\boldsymbol{x})$ and $\boldsymbol{y} \sim p(\boldsymbol{y})$. Given the Bayes theorem

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{\int_{\mathbb{R}^d} p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}}, \tag{7.1}$$

one can in principle sample from the posterior density function of $\boldsymbol{x}$ conditioned on $\boldsymbol{y}$, if both the likelihood density function $p(\boldsymbol{y}|\boldsymbol{x})$ and the prior density function $p(\boldsymbol{x})$ are known.

If one knows $p(\boldsymbol{x})$ or makes a credible assumption for it, one can sample $L \in \mathbb{N}$ datapoints $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ from it. For the likelihood density function, one oftentimes takes a Gaussian distribution with mean $\boldsymbol{M}(\boldsymbol{x})$ and standard deviation σ as an Ansatz, namely

$$p(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi\sigma^2)^D}} \exp\left(-\frac{\|\boldsymbol{y} - \boldsymbol{M}(\boldsymbol{x})\|^2}{2\sigma^2}\right). \tag{7.2}$$

Hereby, the components of $\boldsymbol{x}$ are assumed not being correlated with one another, which is the case with the components of $\boldsymbol{x}$ being independent parameters of the model $\boldsymbol{M}$. If there are $L \in \mathbb{N}$ samples $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ taken from the prior $p(\boldsymbol{x})$, one can write the posterior probability density

$p(\boldsymbol{x}_k|\boldsymbol{y}_k)$ for any $k \in \{1,\ldots,L\}$ as

$$p(\boldsymbol{x}_k|\boldsymbol{y}_k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k)}{\sum_{j=1}^{L} p(\boldsymbol{y}_k|\boldsymbol{x}_j)p(\boldsymbol{x}_j)} = \frac{\exp\left(-\frac{\|\boldsymbol{y}_k - \boldsymbol{M}(\boldsymbol{x}_k)\|^2}{2\sigma^2}\right)p(\boldsymbol{x}_k)}{\sum_{j=1}^{L} \exp\left(-\frac{\|\boldsymbol{y}_k - \boldsymbol{M}(\boldsymbol{x}_j)\|^2}{2\sigma^2}\right)p(\boldsymbol{x}_j)}, \quad (7.3)$$

where the Bayes theorem and the law of total probability were used. For $L \to \infty$, this leads to the exact posterior probability density $p(\boldsymbol{x}|\boldsymbol{y})$. Since the denominator in the above formula is just a normalizing constant, one can write

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}) = \exp\left(-\frac{\|\boldsymbol{y} - \boldsymbol{M}(\boldsymbol{x})\|^2}{2\sigma^2}\right)p(\boldsymbol{x}). \quad (7.4)$$

Taking the natural logarithm yields the log-likelihood as

$$\log\left[p(\boldsymbol{x}|\boldsymbol{y})\right] \propto \log\left[p(\boldsymbol{y}|\boldsymbol{x})\right] + \log\left[p(\boldsymbol{x})\right] = -\frac{\|\boldsymbol{y} - \boldsymbol{M}(\boldsymbol{x})\|^2}{2\sigma^2} + \log\left[p(\boldsymbol{x})\right]. \quad (7.5)$$

Assuming a prior $p(\boldsymbol{x})$ of the form

$$p(\boldsymbol{x}) = \begin{cases} W^{-d}, & \boldsymbol{x}_{min} \le \boldsymbol{x} \le \boldsymbol{x}_{max} \\ 0, & \text{otherwise} \end{cases} \propto \begin{cases} 1, & \boldsymbol{x}_{min} \le \boldsymbol{x} \le \boldsymbol{x}_{max} \\ 0, & \text{otherwise} \end{cases}, \quad (7.6)$$

with $\|\boldsymbol{x}_{min} - \boldsymbol{x}_{max}\| = W^d$ for a constant $1 < W \in \mathbb{R}$ satisfying the normalizing condition $1 = \int_{\mathbb{R}^d} p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$, one can write

$$\log\left[p(\boldsymbol{x})\right] = \begin{cases} -d\log(W), & \boldsymbol{x}_{min} \le \boldsymbol{x} \le \boldsymbol{x}_{max} \\ -\infty, & \text{otherwise} \end{cases}. \quad (7.7)$$

Using $\log(1) = 0$ and $\log(0) = -\infty$ one can establish the proportionality

$$\log\left[p(\boldsymbol{x})\right] \propto \begin{cases} 0, & \boldsymbol{x}_{min} \le \boldsymbol{x} \le \boldsymbol{x}_{max} \\ -\infty, & \text{otherwise} \end{cases} \quad (7.8)$$

and thus

$$\log\left[p(\boldsymbol{x}|\boldsymbol{y})\right] \propto -\frac{\|\boldsymbol{y} - \boldsymbol{M}(\boldsymbol{x})\|^2}{2\sigma^2} + \begin{cases} 0, & \boldsymbol{x}_{min} \le \boldsymbol{x} \le \boldsymbol{x}_{max} \\ -\infty, & \text{otherwise} \end{cases}. \quad (7.9)$$

The sampling process needed to obtain a posterior density close to the true one, Markov Chain Monte Carlo (MCMC) methods are usually applied.

Bibliography

[Amini, 2023] Amini, A. (2023). *Introduction to Deep Learning - Lecture notes.* Massachusetts Institute of Technology.

[Degl'Innocenti, 2005] Degl'Innocenti, E. L. (2005). *Polarization in Spectral Lines*, volume 307 of *Springer eBook Collection Physics and Astronomy*. Springer Netherlands, Dordrecht.

[del Toro Iniesta, 2003] del Toro Iniesta, J. C. (2003). *Introduction to Spectropolarimetry*. Cambridge University Press, Cambridge.

[Durkan et al., 2019] Durkan, C., Bekasov, A., Murray, I., and Papamakarios, G. (2019). Neural spline flows.

[Durkan et al., 2020] Durkan, C., Bekasov, A., Murray, I., and Papamakarios, G. (2020). nflows: normalizing flows in PyTorch.

[Goodfellow et al., 2016] Goodfellow, I., Courville, A., and Bengio, Y. (2016). *Deep learning.* Adaptive computation and machine learning. The MIT Press, Cambridge, Massachusetts.

[Kobyzev et al., 2021] Kobyzev, I., Prince, S. J. D., and Brubaker, M. A. (2021). Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11):3964–3979.

[Mitchell, 1997] Mitchell, T. M. (1997). *Machine learning.* McGraw-Hill international editions. McGraw-Hill, New York, NY, international ed. edition.

[Raschka, 2022] Raschka, S. (2022). *Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python.* Packt Publishing Limited, Birmingham, 1 edition.

[Rutten, 2015] Rutten, R. J. (2015). *Introduction to astrophysical radiative transfer: Lecture notes.*

[Stix, 2002] Stix, M. (2002). *The Sun: An Introduction.* Astronomy and Astrophysics Library. Springer, Berlin and Heidelberg, second edition edition.

[Uitenbroek, 2020] Uitenbroek, H. (2020). *Polarized Radiative Transfer: Milne-Eddington Inversions Workshop - Lecture notes.* National Solar Observatory.

[Weigert et al., 2006] Weigert, A., Wendker, H. J., and Wisotzki, L. (2006). *Astronomie und Astrophysik: Ein Grundkurs.* Lehrbuch Physik. Wiley-VCH Verlag, Weinheim, 1. nachdr. der 4., völlig überarb. und erw. aufl. edition.

[Weng, 2018] Weng, L. (2018). Flow-based deep generative models. *lilianweng.github.io.*