

Astrophysics III: Galactic Astronomy

Lecture, D-PHYS, ETH Zurich, Spring Semester 2016

Tuesday: 12.45–13.30, HIT F13, and Wednesday: 8.45–10.30, HIT J51, Höggerberg

Exercises: Wednesday: 10.45–12.30, HIT J51

Dates: Feb. 23 to June 1, 2016 (except for Easter break, March 27 – April 3)

Website: www.astro.ethz.ch/education/courses/Astrophysics_3

Lecturer: Prof. Dr. H.M. Schmid,

Office, HIT J22.2, Tel: 044-63 27386; e-mail: schmid@astro.phys.ethz.ch

Teaching Assistants and Co-Lecturers:

Natalia Engler, HIT J 41.2, natalia.engler@phys.ethz.ch

Bruderer Claudio, HIT J 41.2, claudio.bruederer@phys.ethz.ch

ETH Zurich, Institut für Astronomie,

Wolfgang Pauli Str. 27

ETH-Höggerberg, 8093 Zurich

Contents

1	Introduction	1
1.1	The Milky Way and the Universe	1
1.2	Short history of the research in galactic astronomy	3
1.3	Lecture contents and literature	4
2	Components of the Milky Way Galaxy	7
2.1	Geometric components	7
2.2	Stars	9
2.2.1	Properties of main-sequence stars	10
2.2.2	Observational Hertzsprung-Russell diagrams	12
2.2.3	Stellar clusters and associations	15
2.2.4	Globular clusters	17
2.2.5	Age and metallicity of stars	19
2.2.6	Cepheids and RR Lyr variables as distance indicators	21
2.2.7	Star count statistics	23
2.3	Stellar Dynamics	29
2.3.1	Velocity parameters relative to the sun	29
2.3.2	Solar motion relative to the local standard of rest	30
2.3.3	Velocity dispersion in the solar neighborhood	32
2.3.4	Moving groups	33
2.3.5	High velocity stars	33
2.3.6	Radial velocity dispersion in clusters	34
2.3.7	Kinematics of the galactic rotation	35
2.3.8	The GAIA revolution	41
2.4	Interstellar matter (ISM) in the Milky Way	43
2.4.1	The ISM in the solar neighborhood	43
2.4.2	Global distribution of the ISM in the Galaxy	45
2.4.3	Galactic rotation curve from line observations	45
2.4.4	HI and CO observations in other galaxies	47
3	Galactic dynamics	49
3.1	Potential theory	49
3.1.1	Basic equations for the potential theory	49
3.1.2	Newton's theorems	50
3.1.3	Equations for spherical systems	52
3.1.4	Simple spherical cases and characteristic parameters	53
3.1.5	Spherical power law density models	55

3.1.6	Potentials for flattened systems	56
3.1.7	The potential of the Milky Way	57
3.2	The motion of stars in spherical potentials	59
3.2.1	Orbits in a static spherical potential	59
3.2.2	Radial and azimuthal velocity component.	62
3.2.3	Motion in a Kepler potential	63
3.2.4	Motion in the potential of a homogeneous sphere	64
3.3	Motion in axisymmetric potentials	65
3.3.1	Motion in the meridional plane	65
3.3.2	Nearly circular orbits: epicycle approximation	66
3.3.3	Density waves and resonances in disks	69
3.4	Two-body interactions and system relaxation	71
3.4.1	Two-body interaction	71
3.4.2	Relaxation time	74
3.4.3	The dynamical evolution of stellar clusters	74
4	Physics of the interstellar medium	81
4.1	Gas	81
4.1.1	Description of a gas in thermodynamic equilibrium	81
4.1.2	Description of the diffuse gas	83
4.1.3	Ionization	85
4.1.4	H II-regions	87
4.2	Dust	89
4.2.1	Extinction, reddening and interstellar polarization	89
4.2.2	Particle properties	91
4.2.3	Temperature and emission of the dust particles	92
4.2.4	Evolution of the interstellar dust	93
4.3	Magnetic fields	95
4.4	Radiation field	96
4.5	Cosmic rays	96
4.5.1	Properties of the cosmic rays	96
4.5.2	Motion in the magnetic field	97
4.5.3	The origin of the cosmic rays	98
4.6	Radiation processes	99
4.6.1	Radiation transport	99
4.7	Spectral lines: bound-bound radiation processes	101
4.7.1	Rate equations for the level population	102
4.7.2	Collisionally excited lines	103
4.7.3	Collisionally excited molecular lines	108
4.7.4	Recombination lines: excitation through recombination	109
4.7.5	Absorption lines	110
4.8	Free-bound and free-free radiation processes	115
4.8.1	Recombination continuum	115
4.8.2	Photoionization or photo-electric absorption	116
4.9	Free-free radiation processes or bremsstrahlung	117
4.9.1	Radiation from accelerated charges	117
4.9.2	Thermal bremsstrahlung	118

4.10	Compton and Thomson scattering	120
4.11	Temperature equilibrium	122
4.11.1	Heating function H for neutral and photo-ionized gas	122
4.11.2	Cooling of the gas	123
4.11.3	The cooling function $\Lambda(T)$	123
4.11.4	Cooling time scale	126
4.11.5	Equilibrium temperatures.	126
4.12	Dynamics of the interstellar gas	128
4.12.1	Basic equations for the gas dynamics	128
4.12.2	Shocks	130
4.12.3	Example: supernova shells	133
5	Star formation	135
5.1	Molecular clouds.	135
5.2	Elements of star formation	136
5.2.1	Time scale for contraction	140
5.3	Initial mass function	141
5.4	Proto-stars	142
6	Milky Way formation and evolution	145
6.1	Virial theorem and galaxy formation	145
6.2	Timing the Milky Way evolution with high redshift observation	147
6.3	Gas infall and minor mergers today	148
6.3.1	Gas inflow	148
6.3.2	Mergers with dwarf galaxies	148
6.4	The chemical evolution of the Milky Way	149
6.4.1	Nucleosynthesis and stellar yields	149
6.4.2	The role of SN Ia.	151
6.4.3	Modelling the chemical evolution of the Milky Way	152

Chapter 1

Introduction

1.1 The Milky Way and the Universe

This lecture concentrates on the physical properties of the Milky Way galaxy and the processes which are important to understand its current structure and properties. Another strong focus is set on observational data which provide the basic empirical information for our models and theories of the Milky Way.

The place of our Galaxy in the Universe is roughly illustrated in the block diagram in Fig. 1.1.

- The Milky Way is a quite normal spiral galaxy among billions of galaxies in the observable Universe.
- The galaxies were born by the assembly of baryonic matter in the growing potential wells of dark matter concentrations in an expanding Universe. This process started about 14 billion years ago with the big bang. The galaxies evolved with time by assembling initially gas rich matter fragments, going through phases of strong star formation, having phases of high activity of the central black hole, and many episodes of minor and perhaps also major interactions with other galaxies. Although the Milky Way belongs to one of the frequent galaxy types, it represents just one possible outcome of the very diverse galaxy evolution processes.
- Initially, the big bang produced matter only in the form of hydrogen, helium and dark matter. The heavy elements which we see today were mainly produced in galaxies from H and He by nuclear processes in previous generations of intermediate and high mass stars (see Fig. 1.1). Stars form through the collapse of dense, cool interstellar clouds. Then they evolve due to nuclear reactions until they expel a lot of their mass at the end of their evolution in stellar winds or supernova (SN) explosions. This matter, enriched in heavy elements, goes back to the interstellar gas in the Milky Way and may form again a new generation of stars. The remnants of the stellar evolution, mostly white dwarfs (WD) and neutron stars (NS), contain also a lot of heavy elements which are no more available for the galactic nucleo-synthesis cycle.
- Many galaxies, including the Milky Way, have a super-massive black hole (SM-BH) in their center. The black hole grows by episodic gas accretion which may be triggered by galaxy interaction. Supernovae explosions, active phases of the central black hole,

or galaxy interactions are responsible for the loss of interstellar matter of a galaxy to the intergalactic medium. On the other side cold intergalactic matter (IGM), from either primordial origin or gas which was already in a galaxy, can fall onto the Milky Way and enhance the gas content.

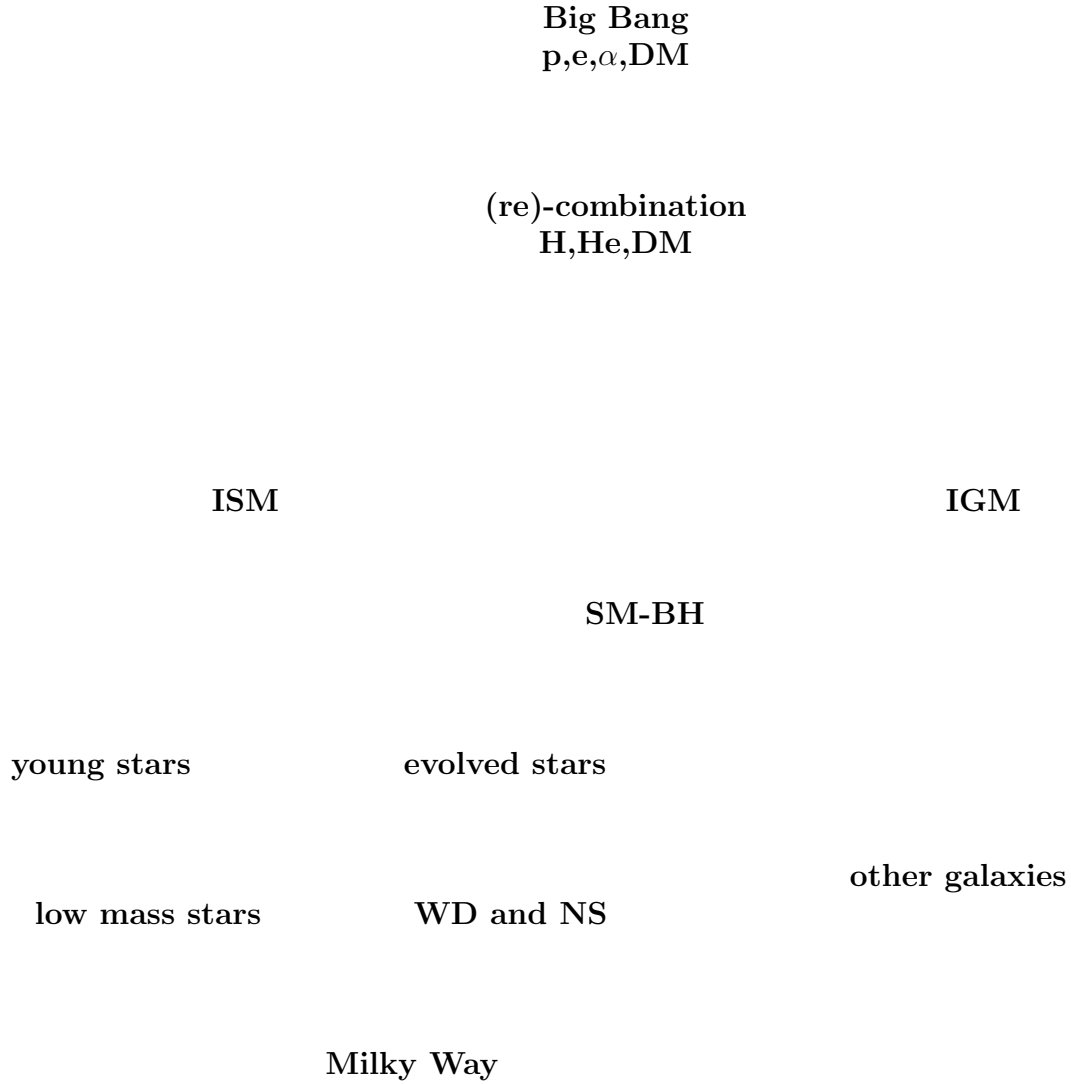


Figure 1.1: The Milky Way in relation to the big bang, the intergalactic matter (IGM), the internal interstellar matter (ISM), different types of stars (WD: white dwarfs; NS: neutron stars), the central, super-massive black hole (SM-BH), and other interacting galaxies.

1.2 Short history of the research in galactic astronomy

Our knowledge on the Milky Way is constantly improving. The Milky Way research profits also a lot from new results gained in other fields in astronomy, like stellar evolution theory, interstellar matter studies, extra-galactic astronomy, or dark matter research. Most important for the progress is the steady advance in observational techniques. The following Table 1.1 lists a few milestones in the evolution of our knowledge in Galactic astronomy.

Table 1.1: Chronology of important studies in Galactic astronomy.

year	important concept, theory, event, or observation
1610	Galileo resolves with his telescope the diffuse light of the Milky Way into countless faint stars.
around 1750	Thomas Wright and Emmanuel Kant describe the Milky Way as a disk of stars with the sun in its center. Kant also speculates that there might exist other Milky Ways similar to our own and that some of the known nebulae could be such galaxies, or “island universes”.
1785	Herschel counts stars in many hundred directions and concludes that the sun is close to the center of a flattened elliptical system which is 5 times larger in the Milky Way plane when compared to polar directions.
1838	Bessel measured for the first time the distance to a star, 61 Cyg at 3.5 pc, based on the yearly parallax measurements.
1845	Lord Rosse sees for the first time a spiral structure in a nebula (M51) which could be an external galaxy.
around 1890	Photography is introduced in astronomy and this allowed to record thousands or millions of stars on a single plate. Herschels Milky Way concept was quantified more accurately by the photographic studies of J. Kapteyn. In the Kapteyn model (1920) the sun is about 650 pc away from the galactic center. The star density drops steadily from the center to about 10 % of the central density at 2.8 kpc in the galactic plane and at 550 pc in polar direction (5:1 ratio).
1919	Shapley studies the distribution of the globular clusters and finds that they are equally frequent above and below the galactic plane but strongly concentrated towards the constellation Sagittarius. Shapley concludes that the sun is far away from the galactic center (he estimated 15 kpc instead of 8 kpc because the interstellar extinction was not known yet).
1923	Hubble detects Cepheid variables in M31 (Andromeda galaxy) and this provides very strong evidence that nebulae with spiral structure, but also other nebula, are galaxies like our Milky Way.
around 1928	Lindblad and Oort develop and prove the basic dynamical model for the Milky Way, in which most stars and the gas in the galactic disk rotate around the galactic center with a speed of about 200 km/s.

1930	Robert Trumpler describes the interstellar absorption due to interstellar dust. The extinction is in the disk plane about 1.8 mag / kpc in the V-band (reduces radiation flux by about a factor of 5/kpc). This effect explains many discrepancies of earlier studies.
1944	W. Baade notices that there exist different populations of stars in galaxies and in the Milky Way. Population I stars are young stars located in the spiral arms and population II stars are old stars predominant in elliptical galaxies, in the bulges of disk galaxies, and in globular clusters.
1951	Even and Purcell detect with Radio observations the H I 21 cm line emission which was predicted by van de Hulst in 1944. This line allows the observation of the diffuse interstellar gas in the Milky Way.
around 1970	Vera Rubin and others describe the galaxy rotation problem based on spectroscopic observations of disk galaxies. Since then more and more evidence was collected that this initially unexpected effect is due to the presence of dark matter as postulated first by Fritz Zwicky in 1933 for galaxy clusters.
around 1995	sensitive near-IR observations provide firm proof for the existence the central super-massive black hole in our Galaxy with measurements of the Keplerian motion of surrounding stars.
2014	the GAIA satellite start with the measurements of accurate distances, positions and proper motions of millions of stars in the Milky Way. Around 2020 there should exist for “most” stars on “our side” of the Milky Way a very accurate position map with stellar motion parameters.

1.3 Lecture contents and literature

Plan for this lecture: Important topics to be covered by this lecture are:

- components of the Milky Way,
- galactic dynamics,
- physics of the interstellar medium,
- star formation,
- origin and evolution of the Milky Way.

Textbooks:

- *Galactic Astronomy*. J. Binney & M. Merrifield, M. 1998, Princeton Series in Astrophysics
An introduction in galactic astronomy.
- *Galactic Dynamics*. J. Binney & S. Tremaine 2008 (2nd edition), Princeton Series in Astrophysics
The standard textbook for galactic dynamics.
- *Physical Processes in the Interstellar Medium*. L. Spitzer, Wiley & Sons, 1978
The classic collection of basic concepts, but the relation to observations are all outdated.
- *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. D. Osterbrock, University Science / Oxford Univ. Press, 1989 (2nd ed.)
Easily understandable textbook.

Review articles or collection of review articles on galactic astronomy: The review articles provide usually more detailed and more actual information on specific topics with the drawback that they are often more rapidly outdated than textbooks.

- *The Milky Way as a Galaxy.* G. Gilmore, I. King, P. van der Kruit, Saas-Fee Advanced Course 19, 1989, Geneva Observatory.
- *The Galactic Interstellar Medium.* W.B. Burton, B.G. Elmegreen & R. Genzel, Saas-Fee Advanced Course 21, Springer, 1992

On-line sources:

- **http://adsabs.harvard.edu/abstract_service.html**
NASA astrophysics database system contains essentially all scientific articles in astronomy and astrophysics. Many articles can be downloaded from this site and essentially all articles are available from an ETH account.

Chapter 2

Components of the Milky Way Galaxy

This chapter gives an overview of the two major baryonic constituents in our Galaxy; the stars, and the interstellar matter. This discussion describes mainly observational data which characterize well the Galaxy, its appearance, structure and dynamics. The first section gives an overview of modern all-sky observations of our Galaxy, and how these data illustrate the distribution of the stars and the interstellar matter.

The second section reminds basic properties of star and star clusters from the Astrophysics I lecture. Then it is discussed how stars can be used as test particles for tracing the galactic structure and the local dynamics in Section 2.3 including a description of the GAIA mission is given which will change this research field in the coming years with high precision measurements of hundreds of millions of galactic stars.

In Section 2.4 the main components of the interstellar matter are briefly described. Emission lines observations of the interstellar gas are very important in providing the large scale structure and the overall rotation of the galactic disk. Later, in Chapter 4, follows a much more detailed treatment of the physics of the interstellar matter.

2.1 Geometric components

The Milky Way is visible as a straight band extending along a great circle on the celestial sphere from a declination of $+63^\circ$ in the northern constellation Cas (Cassiopeia) to -63° in the southern constellation Crux (Cru). The Galactic center is in the direction of Sgr (Sagittarius) at the position $\alpha = 17^h46^m$, $\delta = -28^\circ56'$ in equatorial coordinates.

The galactic center is the zero point for the galactic coordinate system with longitude angle ℓ ($0^\circ \leq \ell \leq 360^\circ$) and latitude angle b ($-90^\circ \leq b \leq 90^\circ$). The galactic system is shown in Slide 2-1 within the equatorial coordinate system. Longitude increases from the center towards NE and the galactic anti-center is in Auriga (Aur). The galactic North pole is in Com (Coma Berenices) and the South pole in Scl (Sculptor).

The galactic structure is best illustrated in maps in galactic coordinates. Slide 2-2 to 2-5 shows modern all-sky maps (Mollweide or equal-area projections) of the Galaxy in different wavelength bands. They provide views of the different geometric structures and the distribution of different matter components.

The distribution of stars is best visible in the near-IR map in Slide 2-2 because the

absorption by the interstellar dust in this wavelength band is small. The stars trace nicely the galactic disk and the elongated central bulge region.

The distribution of cold gas can be seen in the radio map for the H I line emission in Slide 2-3. H I is a very good tracer of the diffuse, neutral interstellar gas.

The dust, absorbs the UV and visual light. Therefore, there are “dark” lanes and holes in the visual map along the Milky Way disk (Slide 2-4), where the dust in the solar neighborhood hides the background stars. The large scale distribution of the dust is better visible in the far-IR wavelength range (Slide 2-5), where the dust re-emits the absorbed radiation.

Schematically, the Milky Way can be divided into the components disk, bulge, and halo (see Fig. 2.1).

Figure 2.1: Schematic side view of the Milky Way.

Disk. The disk consists of stars, open star clusters and associations, H II regions, molecular clouds, and diffuse gas and dust. There is an overall galactic rotation with a velocity of about $v \approx 220$ km/s. The disk extends from about 3 – 17 kpc from the galactic center and the sun is located at about $r = 8$ kpc. The width of the disk, as measured from the star density, is of the order 100 pc at the location of the sun.

Bulge. The galactic bulge is a bar extending to about 3 kpc from the center. It consists mainly of old, metal-rich stars with randomly oriented orbits around the galactic center. There is essentially no cold gas in the bulge except for the very center of the galaxy where there exists a small gas disk with a radius of about 100 pc. In the very center of the Galaxy is a super-massive black hole.

Halo. The extended galactic halo has a much lower density of baryonic matter than the disk and the bulge. An important baryonic component of the halo are the globular clusters. They reside in a spherical distribution around the galactic center. About half of the globular clusters lie within 2 kpc from the galactic center but some are also further away than 10 kpc. There exists also a (low density) population of halo stars with a distribution similar to the globular clusters.

The nearest dwarf galaxies are also located in the galactic halo. The Canis Major and Sagittarius dwarf galaxies are currently colliding with the Milky Way at a distance of about ≈ 10 kpc from the galactic center. The galactic halo contains further clouds of neutral H I gas within a hot, low density gas.

The main mass component of the halo and the Milky Way is dark matter. It extends to a radius of about 100 kpc from the galactic center and dominates the galactic gravitational potential on large scales.

2.2 Stars

The stars are a major component of the Milky Way. Stars are ideal test particles which provide accurate positions, density distributions and motion information for the characterization of the Galactic potential and dynamical processes. In addition one can estimate for certain stars their age and/or their metallicity which provide further dynamical but also evolutionary information about the Milky Way system. On the other side the large scale Milky Way structure has a strong impact on the star formation which takes place in dense molecular clouds.

In this section the properties of stars are described with the particular focus on parameters which provide diagnostic information about the Milky Way system. Stellar astrophysics is a main topic of the ETH lecture Astrophysics I. Slide 2-6 provides as a reminder a short description of the evolution of a low and a high mass star together with the corresponding (schematic) evolutionary tracks in the theoretical Hertzsprung-Russell diagram. In the following we summarize basic formulae and a few important points on stellar parameters and evolution.

Stars can be characterized quite well by a few key parameters. The most basic quantities are L luminosity, R radius, T_{eff} effective surface temperature, M mass, and τ age. Another important parameter for galactic studies is the metallicity (e.g. Z). Further parameters are binarity and the corresponding binary parameters, stellar rotation, and magnetic fields. There exist several important relations between stellar parameters.

Black-body laws: For a sphere radiating like a black body there is according to the *Stefan-Boltzmann law*:

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4. \quad (2.1)$$

The Planck curve describes the spectral energy distribution of a black body

$$B_{T_{\text{eff}}}(\lambda) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda k T_{\text{eff}}} - 1}. \quad (2.2)$$

The wavelength spectrum has its maximum flux $B_{\text{max}} = B_{T_{\text{eff}}}(\lambda_{\text{max}})$ according to *Wien's law* at

$$\lambda_{\text{max}} = \frac{2.9\text{mm}}{T_{\text{eff}}[\text{K}]} . \quad (2.3)$$

For $\lambda \gg \lambda_{\text{max}}$ the spectral energy distribution can be described by the *Rayleigh-Jeans approximation*:

$$B_{T_{\text{eff}}}(\lambda) \approx \frac{2c}{\lambda^4} k T_{\text{eff}}, \quad (2.4)$$

2.2.1 Properties of main-sequence stars

Main sequence stars burn hydrogen to helium. This phase lasts about 90 % of the nuclear burning life time of a star. Therefore about 90 % of all stars are main sequence stars and their properties are therefore particularly relevant.

Mass-luminosity relations on the main sequence. The luminosity of main sequence stars is a strong function of mass which is described by a power law function with different exponent α for different mass ranges:

$$\frac{L}{L_{\odot}} \approx a \left(\frac{M}{M_{\odot}} \right)^{\alpha}, \quad (2.5)$$

$$\begin{aligned} \text{where } \alpha = 2.3, \quad a = 0.23 & \quad \text{for } M < 0.43 M_{\odot}, \\ \alpha = 4.0, \quad a = 1.0 & \quad \text{for } 0.43 M_{\odot} < M < 2 M_{\odot}, \\ \alpha = 3.5, \quad a = 1.5 & \quad \text{for } 2 M_{\odot} < M < 20 M_{\odot}, \\ \alpha = 1.0, \quad a = 3200 & \quad \text{for } M > 20 M_{\odot}. \end{aligned}$$

Main sequence lifetime. The main sequence lifetime of star ends when about 10 % of all H is burnt to He. High mass stars have a much larger luminosity and therefore they burn their fuel much faster than low mass stars. To first order one can write for example for higher mass stars

$$\tau_{\text{ms}} \propto \frac{M}{L} \propto \frac{1}{M^{2.5}} \quad \text{for } 20 M_{\odot} > M > 2 M_{\odot} \quad (2.6)$$

or for low mass stars

$$\tau_{\text{ms}} \propto \frac{1}{M^{1.3}} \quad \text{for } M < 0.43 M_{\odot}. \quad (2.7)$$

Stellar parameters for main-sequence stars. The following table lists main sequence parameters for different spectral types of stars.

Table 2.1: Parameters for main sequence stars: mass, luminosity, radius, effective surface temperature and main sequence life time.

sp.type	O5 V	B0 V	A0 V	G0 V	M0 V	M8 V
$M/M_{\odot}$	60	18	3.0	1.1	0.50	0.06
$L/L_{\odot}$	$8 \cdot 10^6$	$7 \cdot 10^4$	54	1.5	0.080	$1.2 \cdot 10^{-3}$
$R/R_{\odot}$	12	7.5	2.5	1.1	0.50	0.10
T_{eff} [K]	41'000	30'000	9500	6000	3800	2600
τ_{ms} [yr]	$8 \cdot 10^5$	$4 \cdot 10^6$	$6 \cdot 10^8$	$7 \cdot 10^9$	$6 \cdot 10^{10}$	$5 \cdot 10^{11}$

The parameters given in Table 2.1 are only approximative. The given value allow to construct a $\log T_{\text{eff}} - \log L/L_{\odot}$ plot or a “theoretical Hertzsprung-Russel diagram”. Detailed studies show that there are many subtle dependencies of the basic stellar parameters on e.g. age, metallicity, or rotation rate, but this is beyond the scope of this lecture.

Initial mass function (IMF). The initial mass function describes the mass distribution $N_S(M)$ of newly formed stars per mass bin ΔM . This distribution is quite universal and it will be an important topic in the Chapter 5 on star formation. However, it is useful for the understanding of galactic stellar populations to introduce the IMF in this introductory chapter. The standard IMF (Salpeter 1955) can be described by a power law distribution

$$\frac{dN_S}{dM} \propto M^{-2.35} \quad \text{for } M > 0.5 M_\odot. \quad (2.8)$$

This relation is often given as a logarithmic power law of the form

$$\frac{dN_S}{d \log M} \propto M^{-1.35} \quad \text{because} \quad \frac{dN_S}{dM} = \frac{dN_S}{d \log M} \frac{d \log M}{dM} = \frac{1}{M} \frac{dN_S}{d \log M}.$$

This is equivalent to a linear fit with slope -1.35 in $\log M$ - $\log N_S$ diagram (Figure 5.3). This law indicates, that the number of newly formed stars with a mass between 1 and 2 $M_\odot$ is about 20 times larger than the stars with masses between 10 and 20 $M_\odot$. One may also say that twice as much gas from a star-forming cloud ends up in stars between 1 and 2 $M_\odot$ when compared to stars with masses between 10 and 20 $M_\odot$. For low mass star the IMF power law has a steep cut-off for $M < 0.5 M_\odot$ where the general law does not apply.

Figure 2.2: Schematic illustration of the initial mass function (IMF) for stars.

Discussion on main sequence stars. Luminosity, effective surface temperature, and the life time of main-sequence stars are very important for the interpretation of stellar populations. The following points can be made:

- high mass stars are born much less frequently than low mass stars,
- high mass stars, although rare, dominate the luminosity of a new-born population of stars (a young association or star cluster),
- high mass stars are blue stars and therefore a young population has a blue color,
- after some time (e.g. > 1 Gyr) the yellow-red low mass stars dominate the main-sequence population because all short-lived high-mass stars are gone,
- the total luminosity of a stellar population decreases steadily with age.

2.2.2 Observational Hertzsprung-Russell diagrams

A stellar population can be characterized well if the stars can be placed into the Hertzsprung-Russell diagram (HR-diagram) or color-magnitude diagram. This requires the measurement of the absolute brightness which can be related to the absolute luminosity and the color index which can be related to the surface temperature. This information follows from accurate photometry and distance determinations. In astronomy many different photometric systems are used and each requires accurate calibration procedures. This subsection provides only a simplified description of the basic principles.

Measurements of magnitudes and colors. Photometric measurements are carried out typically in wavelength bands which are specific for each instrument used. As general photometric reference the Vega magnitude system is used.

All photometric measurements are related to the star Vega (α Lyr) by the definition that Vega has an apparent magnitude of

$$m_{\lambda}(\text{Vega}) = 0.0^m \quad (2.9)$$

in all photometric bands in the wavelength region from about 150 nm to 15 μm (UV - visual - IR range).

Photometric magnitude is a logarithmic quantity which relates the relative flux ratio of two measurements ℓ_1 and ℓ_2 by the relation

$$m_1 - m_2 = -2.5 \log \frac{\ell_1}{\ell_2}. \quad (2.10)$$

This means that a star with $m_2 = 2.5^m$ is 10 times fainter than a star with $m_1 = 0^m$.

Apparent colors or color indices CI between two wavelength filters λ_1 and λ_2 are also quantified as magnitude difference

$$\text{CI} = m_{\lambda_1} - m_{\lambda_2}, \quad (2.11)$$

e.g. the color B-V is the difference between the standard Johnson blue filter and visual filter $m_B - m_V$. B-V is positive for a star which is more “red” than Vega and negative for a star which is more “blue”. Colors for other filter pairs are defined according to the same principle.

Distances and interstellar extinction. The apparent magnitude m measured for stars must be converted in the next step into absolute stellar magnitudes M and intrinsic stellar colors. For this one needs to take into account the distance of the star and the possible interstellar extinction.

The relation between the apparent flux f_{λ} and absolute flux F_{λ} of a star depends on the distance d and the interstellar extinction τ_{λ}

$$f_{\lambda}(d) = \frac{F_{\lambda}}{4\pi d^2} e^{-\tau_{\lambda}}. \quad (2.12)$$

This relation can be expressed in magnitudes. For this the absolute magnitude M_{λ} is introduced, which is the apparent magnitude of an object at a distance of 10 pc without interstellar extinction:

$$M_{\lambda} = m_{\lambda}(f_{\lambda}(10 \text{ pc})). \quad (2.13)$$

For example, our sun has an absolute magnitude of $M_V = +4.5^m$ in the visual band. Vega is at a distance of about 10 pc and therefore also the absolute magnitude of Vega is approximately

$$M(\text{Vega}) \approx 0^m.$$

The general formula for the conversion of the apparent magnitude m of a star into absolute magnitudes M is given by the following formula:

$$m_\lambda = M_\lambda + 5 \log d [\text{pc}] - 5 + A_\lambda. \quad (2.14)$$

In this equation there are two terms:

- the *distance modulus*: $5 \log d [\text{pc}] - 5$ which follows from

$$m_\lambda - M_\lambda = -2.5 \log \frac{f_\lambda(d)}{f_\lambda(10 \text{ pc})} = -2.5 \log \frac{(10 \text{ pc})^2}{(d [\text{pc}])^2} = -(5 - 5 \log d [\text{pc}]), \quad (2.15)$$

- and the interstellar extinction: $A_\lambda \geq 0^m$.

The interstellar extinction is due to small $< 1 \mu\text{m}$ interstellar dust particles. Their absorption is stronger in the blue than in the visual $A_B > A_V$ and therefore the light is reddened. On average the following relation approximates quite well the extinction effect:

$$E_{B-V} = A_B - A_V \approx 3.1 A_V. \quad (2.16)$$

The color effect $E_{B-V} = A_B - A_V$ is according to this relation roughly proportional to the absolute extinction A_V and therefore one can use the reddening of a star as a measure for the extinction. The reddening follows from the measurements of the apparent color $m_B - m_V$ for a star for which the intrinsic color $M_B - M_V$ is known, for example from its spectral type. This method can also be applied to photometric measurements in other filters. Typically, the extinction is about $A_V \approx 1.8^m/\text{kpc}$ in the galactic disk and $A_V < 0.2^m$ for extragalactic observations in the direction of the galactic poles.

HR-diagram for the stars in the solar neighborhood. HR-diagrams for nearby stars have two advantages:

- the distances d are well known from parallax measurements (to a precision of 10 %), and
- the interstellar extinction is small $A_V < 0.2^m$ and can be neglected.

Slide 2-7 shows the Hertzsprung-Russell diagram as determined from data of the Hipparcos satellite. Hipparcos obtained between 1990 and 1993 accurate distances and photometry for about 100'000 stars up to a distance of about 120 pc and covered all stars brighter than $m_V = 7.2^m$ and selected additional stars of interest. Slide 2-7 shows the location of about 17'000 single stars in the HR-diagram which could be measured with the highest precision.

The Hipparcos HR-diagram has the following characteristics:

- the nearby stars are a good average sample for the stars in the Milky Way,
- for nearby stars it is possible to measure accurately the location of the main-sequence for low mass stars down to an absolute magnitude of $M_V = 12^m$,

- Hipparcos obtained for each star about 50 – 150 photometric measurements and could therefore measure the photometric variability and use well defined averaged values,
- the Hipparcos sample is not contaminated by foreground or background stars because the distances are well known for all objects,
- the nearby stars are well known and the sample can be cleaned from binaries which can spoil the stellar photometry of supposed single stars,
- a significant disadvantage of the Hipparcos sample is the lack of rare high mass objects.

The *Hipparcos HR-Diagram* shows that the local population of bright stars is mainly composed of low to intermediate mass main sequence stars ($M \approx 0.5 - 3 M_{\odot}$) and a significant population of evolved stars on the red giant branch. There are also some very low mass main sequence stars $M_V > 12^m$ and white dwarfs in the Hipparcos sample. There are many more such faint stars in the solar neighborhood, but they were too faint for the Hipparcos satellite.

HR-diagram for stellar clusters The stars in a stellar cluster have all essentially the same distance (same distance modulus $m - M$) and a similar amount of interstellar extinction A_{λ} . For this reason it is possible to determine observationally all features of the HR-diagram or color-magnitude diagram for the stellar population in the studied cluster without knowledge of the exact distance and interstellar extinction.

Slide 2-8 shows as example the color magnitude diagrams of the nearby Hyades and Pleiades clusters. We see for both clusters the stars on the main sequence, but simply shifted relative to each other because of the different distance modulus. The distance moduli are $m - M = 3.3^m$ (46 pc) for the Hyades and 5.65^m (135 pc) for the Pleiades.

A key parameter of the color-magnitude diagram is the upper end of the main sequence which provides the age of the cluster. One can assume that all stars in a cluster have essentially the same age. In young clusters the main-sequence extends to very bright stars while in older clusters all high mass stars have already evolved away from the main-sequence. In the case of the Hyades the turn-off point is around $M_V = +0.5^m$, while it is around $M_V = -2.5$ for the Pleiades.

The distribution of cluster stars in the color-magnitude diagrams provide very important information about stellar evolution because all stars have the same age. This allows to trace and establish the exact evolution of stars within the HR-diagram.

One difficulty to be considered for the analysis of observational color-magnitude diagrams is the contamination of the cluster sample by foreground or background stars. For this reason the data of rich clusters in low density fields (location at high galactic latitude or fields with high background absorption) provide good results with less contamination.

2.2.3 Stellar clusters and associations

Galactic clusters. There are more than 1000 galactic clusters (or open clusters) known and the total number is estimated to be about 100'000. Galactic clusters have a radius of the order of 10 pc and a wide range of star densities ranging from 0.3 stars/pc³ for the Hyades to about 1000 stars/pc³ at the center of the richest clusters. For comparison, the star density in the solar neighborhood is about 0.1 stars/pc³. Dense clusters are dynamically bound by the mutual gravitational attraction of the cluster stars, while lower density systems are in the process of dissolving themselves. The total masses of galactic clusters lie in the range of about 100 to 3000 M_⊙. The integrated brightness is typically $M_V \approx -5$, but can also be as high as $M_V \approx -10$ for the most extreme cases. Table 2.2 lists parameters and Slide 2-9 shows pictures of some well-known galactic clusters.

Table 2.2: Parameters for galactic clusters

name	dist. [pc]	age [Myr]	N_{stars}	turn-off stars
M67	900	4000	≈ 1000	F5
Hyades	46	625	≈ 200	A7
Pleiades	135	125	≈ 1000	B6
Orion (NGC 1976)	410	< 0.5	≈ 2500	O6

A few comments on the open clusters shown in Slide 2-9 (see also Table 2.2):

- M67 is one of the oldest open clusters known. It is the nearest of the old open clusters and therefore well studied. The main sequence turn-off is around spectral type F. Because of its age it contains more than 100 white dwarf stars.
- The Hyades is the nearest open cluster. The bright red star, α Tau, is a foreground object and does not belong to the cluster. The Hyades cluster shows a strong mass segregation. The central 2 pc of the cluster contains only systems with masses $> 1 M_{\odot}$ or white dwarfs. The cluster contains about 20 A, 60 F, 50 G, 50 K dwarfs, and about 10 white dwarfs but only about 15 M stars. It seems that lower mass stars have been lost.
- The Pleiades is the nearest cluster which is dominated by blue stars. It is a rich cluster with more than 1000 members. Because it is so close and young the full main sequence from B-stars down to brown dwarfs could be mapped.
- The Orion-(Trapezium) cluster, or NGC 1976, is part of the nearest high mass star forming region including the famous Trapezium stars and the Orion nebula. The brightest star, θ^1 Ori C is an O6 V star, which is responsible for the ionization of the Orion Nebula. The stars are younger than < 1 Myr and many stars are still forming or they are in their pre-main sequence phase. For such young clusters one cannot indicate a well defined age, because the duration of the star-formation process is of the same order as the cluster age. The presence of thick interstellar clouds make the derivation of the cluster parameters quite difficult because many stars are due to the dust not visible in the V-band. In any case the stellar density of the Orion-Nebula cluster is with ≈ 10000 stars/pc³ very high.

Clusters ages. Slide 2-10 shows schematically the distribution of star from different open clusters in the HR-diagram. Clearly visible is the difference in the main-sequence turn-off which is a good indicator of stellar age. The HR-diagram of young clusters has a main-sequence extending to O or early B stars, some A-F supergiants in the Hertzsprung gap (the low star density region in the HR-diagram between main sequence and red giant branch), and a concentration of M supergiants. Intermediate age clusters show still some late B or A stars on the main sequence and then a well developed red giant branch separated by a smaller Hertzsprung gap. Older galactic clusters (≈ 1 Gyr) show a main-sequence turn-off in the F-star region and a continuous sub-giant branch (without Hertzsprung gap) extending into a lower brightness red giant branch.

There exist almost no galactic open clusters with ages larger than 1 Gyr (thus, M67 is an exception). If the cumulative age distribution of galactic clusters is plotted then the 50 % mark is around 300 Myr (see Fig. 2.3). The large number of galactic cluster and their age distribution indicates the following *evolutionary scenario for galactic clusters*:

- new clusters are continuously formed in the galactic disk,
- after formation they loose stars and dissolve with time mainly due to dynamical interaction with interstellar clouds (see next Chapter),
- older clusters ($\tau > 1$ Gyr) are very rare because they were all disrupted,
- it is assumed that a large fraction of the stars in the Milky Way disk were initially formed in clusters.

Figure 2.3: Cumulative distribution of cluster ages (according to Binney and Tremaine based on data from Piskunov et al. (2007)).

Stellar associations and groups. A stellar association or group is a very loose assembly of about 100 or less stars which are not dynamically bound. The space density is lower than the typical density in the galactic disk, with perhaps 100 stars within a volume of 10^6 pc^3 . Associations and groups can often be identified because of a small concentration of young, rare stars. Two types of associations are well known:

- O- or OB-associations with an enhanced density of massive main sequence stars,
- T-associations, which contain an over-density of variable T Tauri type pre-main sequence stars.

The nearest examples are the Sco-Cen OB association and the Taurus-Auriga T association.

Associations are just transient groups of newly formed stars in the galactic disk (spiral arm) population. They are in the process of dispersing from a star forming region into the “galactic field”. OB associations may cover a very large sky region and individual O or B stars of an association may be members of a new formed cluster. In the Orion OB associations the Trapezium stars in the Orion cluster (NGC 1976) are such an example.

It is difficult to identify associations and quantify their frequency and lifetime in the galactic disk. For this reason it is not clear whether more stars in the Galaxy are formed in dense star clusters or in loose associations.

2.2.4 Globular clusters.

Globular clusters are spherical systems which contain typically 10^5 to 10^6 stars and a mass of $10^5 - 10^6 M_\odot$ in a volume with a radius of $r \approx 20 - 50$ pc. They have a high central star density of 100 to $> 10^4$ stars/pc³ and are dynamically very stable and long lived. The absolute brightness of globular clusters is on average $M_V \approx -8.5^m$. There are about 150 globular clusters known in our galaxy, and they are distributed in the galactic halo. Two examples for the globular clusters are shown in Slide 2-11. ω Cen is one of the brightest and best studied globular clusters. NGC 6522 is an example of an object very close to the galactic center, located in the low extinction region called “Baade’s window”, where the contamination by foreground and background stars is a severe complication for the investigation of this globular cluster.

Figure 2.4: Schematic HR-Diagram for globular clusters.

The Hertzsprung-Russell diagrams of globular clusters are special because they contain only old low mass stars. Figure 2.4 shows a schematic HR-diagram for globular clusters which has the following characteristics:

- the main sequence (MS) turn-off point is in the region of F and G stars, or at stellar masses $0.9 - 1.3 M_\odot$ indicating an age of the order 10 Gyr,
- there is a subgiant branch which joins the main sequence with the giant branch (RGB),

- near $M_V \approx +0.5$ there is a horizontal branch (HB), which contains pulsating RR Lyr stars and some blue hot stars. The HB extends toward the red until it rises in the so-called asymptotic giant branch (AGB) lying just above the giant branch.
- above the main-sequence turn-off point there are a few so-called blue stragglers stars, which are too bright for main-sequence stars with the age of the globular cluster,
- the horizontal branch may extend into the white dwarf cooling track for clusters which were observed with very high sensitivity.

Stellar evolution of globular cluster stars. According to the stellar evolution theory the stars with an initial mass just above the main-sequence turnoff stars have evolved to the red giant branch. Stars with even higher initial mass are now in the core helium burning phase on the horizontal branch. Even higher initial mass stars are either evolving up along the asymptotic red giant branch or they have already lost their envelope due to stellar winds so that their hot core becomes visible. They then evolve to the blue part of the horizontal branch where they stop nuclear burning and enter the white dwarf cooling track.

The “blue stragglers” stars are special cases. They were probably low mass binaries which merged after some time ($\approx$ Gyr) to a higher mass, rapidly rotating star. These stars are therefore still on the main sequence because of the late merging event. All single stars with the same mass have already evolved to an advanced evolutionary stage.

Metallicity of globular clusters. Globular clusters are very old, and have a very special space distribution. Another very important property is the very low metallicity of the stars in a large fraction of globular clusters. A low metallicity means that the abundance of heavy elements is 10 to 100 times lower than in the sun. This indicates that all the star in a globular cluster were born in a well mixed gas clouds and that no additional stars were formed in a later generation from gas with different metallicity.

An indicator for low metallicity is the color of the main sequence. High metallicity stars have atmospheres with more heavy elements (e.g. Fe) producing many absorption lines in the UV and blue spectral region (see Slide 2–12). For this reason they emit for a given luminosity less blue light because the UV and blue radiation cannot escape from the deep, hot layers of the stellar atmosphere. The radiation escapes only from higher cooler layers and the resulting spectral energy distribution is redder than for low metallicity stars.

For this reason, the main sequence of globular clusters is shifted in the HR-diagram towards the blue. The stars appear for a given color less luminous (in fact they are for a given luminosity just more blue) and are therefore called subdwarfs (main-sequence stars are dwarfs). A subdwarf branch indicates therefore a low metallicity.

A similar line opacity effect occurs for the red giant branch. For metal poor clusters the red giant branch is shifted significantly to the blue. Slide 2–13 illustrates the location of the main-sequence and the giant branch for clusters with different metallicities.

With modern large telescopes like the VLT it is possible to take accurate spectroscopic measurements of individual stars in globular clusters so that the elemental abundances can be derived from a detailed spectral analysis.

Origin of globular clusters. The metal-poor globular clusters are probably relics of the Milky Way formation process, because they are old and have preserved the gas abundance

pattern which dominated in the early Universe.

The globular clusters with higher elemental abundances ($\approx$ solar) may have formed during phases of extreme star formation, e.g. induced by a galaxy merging event. The bright globular cluster ω Cen may be the dense center of a tidally disrupted galaxy.

Similar evolutionary histories are put forward for globular clusters seen in other galaxies. It should be noted that these are only tentative evolutionary scenarios because our understanding of globular clusters is still incomplete.

2.2.5 Age and metallicity of stars

Stars serve as test bodies for deriving the galactic dynamics and the galactic gravitational potential. In addition we can also derive or at least constrain the age and metallicity of the stars. This provides information about the evolution of the distribution and dynamics of stars from their formation in an interstellar cloud to the present day. Similarly we can use the metallicity of stars as a second parameter for constraining the time and region where they were formed. Thus, selecting stars with a certain age or metallicity can provide important information about earlier epochs and long term evolutionary processes of our Galaxy.

Stellar ages. The age of a star or a stellar group can be estimated from the following age indicators:

- the determination of the main-sequence turn-off age for stellar clusters or groups is a very reliable age indicator for ages from 10 Myr to 13 Gyr,
- high mass stars, such as O stars and early B stars, as well as classical Cepheids, bright giants, or Wolf-Rayet stars are always young $\tau < 100$ Myr objects,
- the stellar rotation speed and coronal activity are useful age indicator for low mass stars of spectral type G, K, and M; fast rotating, active stars are relatively young $\tau \lesssim 1$ Gyr, while slowly rotating, quiet stars are old $\tau \gtrsim 1$ Gyr,
- low luminosity red giants, planetary nebulae, and white dwarfs are typically evolved intermediate or low mass stars which are older than $\tau > 500$ Myr,
- RR Lyr variables are very old $\tau \gtrsim 10$ Gyr objects and they are very reliable indicators for an old population.

Stellar metallicity. The metallicity of a star is often indicated with one of the following two parameters:

- Z is the mass fraction of all elements heavier than H ($= X$) and He ($= Y$). The sun has $X = 0.70$, $Y = 0.28$ and $Z = 0.02$, a metal rich galactic disk star has $Z = 0.05$, while stars in metal-poor globular cluster have $Z \approx 0.002 - 0.0002$.
- $[\text{Fe}/\text{H}]$ is the logarithmic iron abundance relative to hydrogen and in relation to the solar value

$$[\text{Fe}/\text{H}] = (\log \text{Fe}/\text{H})_{\text{star}} - (\log \text{Fe}/\text{H})_{\odot}.$$

A globular cluster star (as example) with an iron underabundance of 100 with respect to the sun has the value $[\text{Fe}/\text{H}] = -2.0$. Often the $[\text{Fe}/\text{H}]$ value is a good indicator of the overall metallicity of a star. This definition can also be used to quantify specific elemental abundance ratios for stars such as e.g. $[\text{Ca}/\text{Fe}]$ or $[\text{O}/\text{Si}]$ and others.

The best method for a metallicity determination are spectroscopic abundance determinations from high resolution spectra of well understood stars. These are stars where the elemental abundances of the surface layers are representative for the initial composition of the star. Many main-sequence stars, but also certain giant stars fulfill this criterion. Slide 2–14 illustrates the dependence of the line strengths with metallicity or the presence of specific abundance patterns (e.g. for HE 0107-5240). High resolution spectroscopy requires time consuming observations and stars $m_v > 17^m$ might be too faint even for a large telescope.

For stars with well known distances, clusters, or for stellar groups the metallicity can also be derived from photometry as described in the subsection for globular clusters (or Slide 2-13). The method is based on the strength of line opacities in the blue-UV spectral region, which is high for high metallicity objects and weak for low metallicity objects. The corresponding color effect provides then a measure for the metallicity. This technique is very powerful if a cluster is investigated for the presence of two populations of stars with different metallicities.

Metallicity gradients in the Milky Way. Metallicity determinations from cluster photometry and spectroscopic studies provide a quite detailed picture of the different metallicity gradients in the Galaxy:

- for young disk stars there is a metallicity gradient where the metallicity is higher $[\text{Fe}/\text{H}] \gtrsim 0.0$ for regions closer to the galactic center than the sun and lower $[\text{Fe}/\text{H}] \lesssim 0.0$ further out; the metallicity gradient is of the order

$$\frac{\Delta[\text{Fe}/\text{H}]}{\Delta d} \approx \frac{-0.05}{\text{kpc}}.$$

- old galactic open clusters have a lower metallicity than young clusters and the temporal gradient is of the order

$$\frac{\Delta[\text{Fe}/\text{H}]}{\Delta \tau} \approx \frac{-0.05}{\text{Gyr}}.$$

- globular cluster have typically a much lower metallicity, if they are located at large galacto-centric distances; a rough statement for the metallicity is:

$$[\text{Fe}/\text{H}] > -1.0 \text{ for clusters at } d \lesssim 3 \text{ kpc},$$

$$[\text{Fe}/\text{H}] < -1.0 \text{ for clusters at } d \gtrsim 3 \text{ kpc}.$$

- the metallicity of the galactic bulge is not well known, but it is approximately solar ($[\text{Fe}/\text{H}] \approx 0.0$).

2.2.6 Cepheids and RR Lyr variables as distance indicators

Distance determinations are required for the 3-dimensional mapping of the distribution of objects. A very basic method for the determination of distance modulus $m - M$ is the main-sequence fitting for stellar clusters. This method works well for good observations of clusters, where the main sequence can be observed over a significant color range. This requires photometry of F-G stars in open clusters because all O, B, and A stars have similar colors. For globular cluster one needs to reach even K dwarfs for the main sequence fitting.

Pulsating Cepheid variables provide a very powerful alternative for the distance determination because their pulsation period is an indicator of the stellar type and its absolute luminosity. The calibration of the period-luminosity or P-L relation has a very interesting history since the first detection of such a relation for Cepheids in the Small Magellanic Cloud by Henriette Leavitt in 1912. Initially the size of our Galaxy, or the distance to the M31 based on Cepheids were estimated wrongly by about a factor of two by Shapley, Hubble and others until Baade recognized in 1952 that there are two different types of Cepheids:

- the population II metal poor, old, low mass RR Lyr variables with periods $P \lesssim 1^d$ and a pulsation brightness amplitude of $\Delta m \approx 1^m$. They are low mass $\approx 0.7 M_\odot$ horizontal branch stars which are in their helium burning phase. RR Lyr variables are further divided into subgroups which are defined according to subtle differences in evolutionary phase and metallicity.
- the population I, metal rich, young, high mass classical Cepheids with periods in the range $3^d \leq P \leq 40^d$. They are evolved high mass stars crossing the HR-diagram.
- in addition there are several other groups of Cepheid-type pulsating variables like W Vir stars, δ Scuti stars (main sequence pulsators), or RV Tau variables, which are not discussed here.

Cepheids are A to K giants or supergiants located in the (vertical) pulsation instability strip in the middle of the Hertzsprung-Russell diagram. These stars pulsate because of an opacity effect or κ -*mechanism* due to He-ionization. The process works as follows:

- the slightly enhanced temperature in the stellar envelope leads to the additional ionization of He^+ to He^{+2} ,
- the He^{+2} ionization enhances the opacity and the outward radiation transport (= energy transport) is reduced, the star heats further up and starts to expand,
- with the expansion the gas density and temperature drops, He^{+2} recombines to He^+ , the opacity drops and the radiation can escape,
- the stellar envelope cools rapidly, contracts, heats up and the He ionization increases again,
- the opacity and temperature rises again, and a new cycle begins.

For pulsating variables there exists a simple relationship between the mean density of a star and pulsation period:

$$P \left(\frac{\bar{\rho}}{\rho_\odot} \right)^{1/2} \approx Q, \quad (2.17)$$

where P is the pulsation period, $\bar{\rho} \propto M/R^3$ the mean density, and Q the pulsation constant. This relation indicates basically that the pulsation period P is roughly the time required for a sound wave to move through the star.

Cepheid variables are ideal objects for distance determinations. They are bright objects and it is easy to identify pulsating variables in a crowded field of stars with repeated observations. The properties of Cepheids variables have been studied in much detail. For this lecture we consider only a rough relation for their absolute magnitude:

- The classical Cepheids (pop. I) are F to K supergiants and the most luminous Cepheids have an absolute magnitude $M_V \approx -6^m$. They belong to the brightest stars (in m_v) in a galaxy. A simple empirical period-luminosity relation valid for $3^d \leq P \leq 40^d$ is

$$M_V(\text{max}) = -2.0^m - 2.8 \log P [\text{d}].$$

- The RR Lyr variables are old (pop. II) A to G horizontal branch stars which are in the He-burning phase. Their absolute magnitude is

$$M_V \approx +0.5^m.$$

In globular cluster or in a similarly old stellar population they are about 5 mag brighter than the main-sequence turn-off.

Classical Cepheids are and will remain in near future important distance indicator for young clusters in the Milky Way and the distances to other galaxies. They are an important part for the distance ladder in extra-galactic astronomy and cosmology.

The RR Lyr variables are important tracers of the old galactic population, and therefore ideal for globular cluster studies and for determinations of distances to objects in the galactic halo.

2.2.7 Star count statistics

Star counts provide information about the distribution and frequency of stars in our galaxy. This is a very basic technique in Astronomy, which was introduced initially for studies of the Milky Way. The technique is now also applied to other objects like galaxies for studies in cosmology and extra-galactic astronomy, or asteroids for solar system studies.

Different types of star counts are used for studies in Galactic astronomy.

- Determination of the number of all stars brighter than a given limit in a brightness limited sample; the comparison for different sky regions provides the overall geometric distribution of objects.
- Determination of the number of stars in a volume limited sample (e.g. out to a distance of d_{lim} , or a particular cluster) for determining the volume density of stars which can then be compared with the volume density of other regions in the Milky Way.

These types of studies can be refined by the determination of the space distributions for different stellar types. The distribution of stars can be described by:

- $A(m, S)$: the differential star counts, which is the number of stars of type S , at apparent magnitude m , per unit magnitude interval (e.g. from $[m - 0.5, m + 0.5]$ and per solid angle $d\omega$, e.g. square degree.
- $N(m_{\text{lim}}, S)$: the integrated star counts for stars of type S down to the magnitude limit m_{lim} , e.g. $m_{\text{lim}} + 0.5$, and $d\omega$:

$$N(m_{\text{lim}}, S) = \int_{-\infty}^{m_{\text{lim}}} A(m, S) dm. \quad (2.18)$$

Homogeneous distribution. We calculate first the volume limited number of stars for a homogeneous distribution for a given star density D [stars/pc³] as function of distance r :

$$N(r_{\text{lim}}) = \omega D \int_0^{r_{\text{lim}}} r^2 dr = \frac{\omega D}{3} r_{\text{lim}}^3$$

The corresponding magnitude limited number follows then from the relation between radius limit in [pc] and magnitude limit:

$$m_{\text{lim}} = M + 5 \log r_{\text{lim}} - 5 \quad \text{or} \quad r_{\text{lim}} = 10^{0.2(m_{\text{lim}} - M + 1)}$$

Combining these two equation yields $N(m_{\text{lim}}) = 10^{0.6 m_{\text{lim}} + C}$ or

$$\log N(m_{\text{lim}}) = 0.6 m_{\text{lim}} + C, \quad (2.19)$$

where C is a constant that depends on D , ω , and M .

This equation states that:

- a homogeneous distribution of stars produces a line in a $m_{\text{lim}} - \log N$ star count diagram,
- for a homogeneous distribution the number of stars increase by $\Delta \log N = 0.6$ or a factor 4.0 if the count limits m_{lim} are one magnitude deeper.

Realistic distributions. In reality, a detailed count statistics needs to consider that the stellar luminosity function and the star density is a function of distance and that there is interstellar absorption.

- $\Phi(M, r, A(r), S)$: the luminosity function of a selected stellar type S Φ is in general also a function of the distance and should consider the interstellar absorption along the line of sight,
- $D(r, S)$: a star density which depends on the distance within the selected field.

The general formula for the differential star density is:

$$A(m, S) = \omega \int_0^\infty D(r, S) \Phi(M, r, A(r), S) r^2 dr. \quad (2.20)$$

This is essentially a convolution integral of the density function and the luminosity function. The effective width of the luminosity functions determines the range of distances that can contribute to the observed number of stars with magnitude m . Obviously, it is not easy to “deconvolve” the problem and derive the line of sight distribution D from the absorption affected luminosity function Φ .

Selecting carefully the stellar type and understanding the selection bias is the real challenge for the interpretation of star count data. The analysis can be strongly simplified if the selected star type S fulfills certain conditions:

- $D(r, S)$ can be well determined, if the luminosity function does not depend on the distance and if also the interstellar extinction can be neglected $\Phi(M, r, A(r), S) = \Phi(M, 0, A(r), S)$. This is essentially the case for:
 - stellar types with narrow luminosity functions like e.g. F, G and K-type main sequence stars or RR Lyr variables, and
 - sight-lines perpendicular to the galactic plane which are barely affected by interstellar extinction.
- The properties of the luminosity function $\Phi(M, r, A(r), S)$ can be quite well determined if the stellar density does not depend on distance $D(r, S) = D(0, S)$. This is essentially the case for
 - the determination of $\Phi(M, 0, 0, S)$ in the solar neighborhood where changes in the luminosity function and effects due to interstellar extinction can be neglected,
 - the determination of the luminosity function $\Phi(M, r_c, A_c, S)$ of a cluster where all stars have essentially the same distance and extinction.

The following paragraphs summarize a few basic results of stellar count statistics.

Integrated star counts. The star counts show that the number of stars is higher in the galactic plane when compared to the galactic poles (see Table 2.3 and Fig. 2.5). The difference is about a factor of 5 for stars brighter than $\lesssim 10^m$ in agreement with the historical results from Herschel and Kapteyn. For fainter magnitudes the stellar density is much higher in the galactic plane when compared to the poles.

Table 2.3: Integrated star counts in the solar neighborhood per deg^2 and mag in the Galactic plane $N(m, 0^\circ)$ and towards the North Galactic pole $N(m, 90^\circ)$, the ratio of these two values, and the total number of stars $N_{\text{tot}}(m)$ over the entire sky.

m_V	$\log N(m, 0^\circ)$	$\log N(m, 90^\circ)$	$N(m, 0^\circ)/N(m, 90^\circ)$	$\log N_{\text{tot}}(m)$
5	-1.08	-1.69	4.1	3.20
10	1.25	0.55	5.6	5.52
15	3.42	2.27	14	7.56
20	5.0	3.4	40	9.0

Another important result is that the number counts increase less than expected for a homogeneous star distribution (factor 4 per mag or $4^5 = 1024$ per 5 mag). In the case of the polar direction this is due to strong decrease in stellar density with distance. In the galactic plane it is mainly due to the interstellar absorption.

Figure 2.5: Total star number counts for stars in the galactic plane and towards the galactic poles and comparison with the slope of a homogeneous star distribution.

Table 2.3 list also the total number of stars over the entire sky. The celestial sphere has $41'253 \text{ deg}^2$ or $\log 41253 = 4.6$, and therefore the total star counts lie in the range

$$\log N(m, 90^\circ) + 4.6 < \log N_{\text{tot}}(m) < \log N(m, 0^\circ) + 4.6.$$

The luminosity function and the integrated luminosity and mass for the stars in the solar neighborhood are presented in Fig. 2.6 or Table 2.4. This statistics show that the most frequent stars have an absolute magnitude of about $M_V \approx 15$ which is about 10 mag fainter than the sun. These are M-type low mass stars and white dwarfs.

The integrated luminosity or the integrated mass of the stars with the same absolute magnitude M_V are not well represented by the luminosity function. The luminosity of the stars in the solar neighborhood is mainly produced by stars with $M_V \approx 0^m$ which are A and F main sequence stars and G and K giants (see Hipparcos HR-Diagram in Slide 2-7). Contrary to this, the mass is in the stars with $M_V \approx +5$ to $+15$ which are the low mass main sequence stars (K and early M) and the white dwarfs.

Table 2.4: General luminosity function $\Phi(M_V)$, integrated luminosity $L/L_\odot(M_V)$, and integrated mass $M/M_\odot(M_V)$ per 10^3 pc^3 and magnitude for the stars in the galactic disk near the sun.

M_V	$\Phi(M_V)$	$L/L_\odot(M_V)$	$M/M_\odot(M_V)$
-5	$6 \cdot 10^{-5}$	0.6	0.002
0	0.1	11.2	0.4
+5	3.4	3.4	3.7
+10	7.8	0.1	3.4
+15	12.5		2.0
+20	3.0		0.2
total	131	54	44

Figure 2.6: Star counts luminosity function, integrated luminosity and mass for the stars in the solar neighborhood.

Mass to light ratio. The total values for the luminosity and mass of stars can be used to determine a mass to light ratio for the stellar population in the solar neighborhood:

$$R_{M/L} = \frac{(M/M_{\odot})}{(L/L_{\odot})} = 0.8.$$

For a young cluster this value is much smaller and for a globular cluster much larger.

The volume density of different stellar types in the solar neighborhood are listed in Table 2.5. The table gives number counts for main sequence stars, red giants and white dwarfs. The used volume of 10^6 pc^3 corresponds to a sphere with a radius of 62 pc and it contains more than 10^5 stars. However, essentially all M type main sequence stars and white dwarfs are faint stars $M > 10^m$. Observations which pick only objects with $m < 10^m$, e.g. the HD-star catalog or the Hipparcos catalog miss all these faint stars, or more than 80 % of all stars. Therefore one needs to go very deep to produce a complete star catalog.

Table 2.5: Mean number densities $N(S)$ in stars/ 10^6 pc^3 for the stars of the different spectral types.

Spec.Type	main seq.	giants	white dwarfs
O stars	0.02		
B stars	100		6300
A stars	500		10000
F stars	2500	50	5000
G stars	6300	160	5000
K stars	10000	400	2500
M stars	63000	30	
total	82500	640	28800

The star numbers in Table 2.5 for the solar neighborhood indicate:

- the distribution of main-sequence stars has a a very larger fraction of low mass stars which can be expected from the IMF and main-sequence lifetime,
- evolved giants are of the order 10 times less frequent than main-sequence stars of spectral type B, A, F and early G, and this represents roughly the 10 times shorter red giant phase when compared to the main-sequence life-time.
- there is a large number of white dwarfs, the remnants of previous B to early G main-sequence stars. The high number of white dwarfs proves that there were already several previous generations of stars in the galactic disk.

These points illustrate that the number counts in the solar neighborhood are very important for quantifying the density of the faint low mass stars and white dwarfs in the Milky Way disk.

The star distributions vertical to the disk shows a very strong dependence on stellar type. This property is not surprising because the cold gas, star forming regions, and young stars are strongly concentrated towards the disk mid-plane, much more than the overall star distribution. Therefore, the average star and particularly older stars must have on average a wider distribution than the young stars.

With number counts for different stellar types perpendicular to the disk one can derive in detail their vertical or “z”-distribution. The distribution can be approximated with an exponential law

$$D(z, S) = D(0, S) e^{-z/\beta}, \quad (2.21)$$

where β is the disk scale height. Table 2.6 gives disk scale heights β_S and disk surface densities Σ_S for various stellar types.

Table 2.6: Vertical scale heights β_S perpendicular to the Galactic disk for various stellar types and other tracers. For frequent stars also the disk surface density Σ_S is given

stellar type	β_S [pc]	Σ_S [stars/pc ²]	stellar type	β_S [pc]
O main seq.	50	$1.5 \cdot 10^{-6}$	clas. Cepheids	50
B main seq.	60	$6 \cdot 10^{-3}$	open cluster	80
A main seq.	120	$6 \cdot 10^{-2}$	interstellar gas	120
F main seq.	190	0.6	planetary nebulae	260
G main seq.	340	2.0	RR Lyr variables	2000
K main seq.	350	3.5	subdwarfs	2000
M main seq.	350	20	globular clusters	3000
G giants	400	0.06		
K giants	270	0.0012		
white dwarfs	500	12.5		

Table 2.6 shows for normal stars a clear correlation between average age and disc scale height. This indicates that older objects have a larger vertical dispersion. Exceptions are the RR Lyr variables, the subdwarfs, and the globular clusters which belong to the halo, and they have therefore a much larger disk scale height.

Another interesting fact is that the surface density of white dwarfs is more than 50 % of the M dwarfs. The average white dwarf has a mass of $\approx 0.5 M_\odot$, while the mean M-dwarf mass is more like $\approx 0.3 M_\odot$ and therefore both groups of stars contribute a similar amount to the stellar mass of the galactic disk. Roughly the mass share of the stars in the Galactic disk is:

- ≈ 30 % M dwarfs,
- ≈ 30 % white dwarfs,
- ≈ 30 % G, and K main sequence stars.

2.3 Stellar Dynamics

All stars in the Milky Way move in the galactic potential. Most disk stars move with the same circular direction in more or less co-planar orbits around the galactic center. Halo stars and stars in galactic bulge have orbits with a wide distribution of orbital plane orientations and eccentricities. The orbits of stars can also be stochastically deflected by small scale mass concentrations due to stellar clusters or massive interstellar clouds or by the dynamical interactions between individual stars.

2.3.1 Velocity parameters relative to the sun

The space motion v_s of a star relative to the sun consists of the radial velocity component v_r and the tangential or transverse velocity component v_t according to

$$v_s = \sqrt{v_r^2 + v_t^2}. \quad (2.22)$$

Typical relative space velocities for stars are:

- about 220 km/s for the orbital motion around the galactic center,
- about 5 – 50 km/s for the velocity dispersion of corotating stars in the disk,
- about 0.2 – 20 km/s for the velocity dispersion in groups and clusters,
- up to 500 km/s for stars on counter-rotating orbits,
- up to and beyond 1000 km/s for stars in close orbits around the super-massive black hole in the galactic center.

Thus one needs to reach a measuring precision of about ± 1 km/s for the investigation of the velocity dispersion in the galactic disk and in clusters. A lower precision (± 10 km/s) is sufficient for the investigations of the galactic rotation. The observations provide the radial velocity and the angular proper motion:

The radial velocity v_r is measured in km/s via the Doppler shift of spectral lines of the star. The measured values must be corrected by up to ± 30 km/s for the Earth motion around the sun in order to get the stellar motion relative to the sun. A positive v_r means that the object is moving away from the sun (it's spectrum is red-shifted).

The radial velocity can be measured for all stars which are bright enough for a spectrometric measurements and which have well defined spectral features. In particular, it is possible to measure very accurate radial velocities for very distant stars (e.g. at 50 kpc in the Magellanic Clouds) if they are bright and have suitable spectra.

The angular proper motion is measured in units of arcsec/yr in the direction of the right ascension μ_α (E–W) and declination μ_δ (N–S). Positive values are given for objects which move towards E and N, respectively.

At least two measurement taken at two epochs separated by a few year, better many years, are required to determine the proper motion. In addition one needs also to correct for the yearly parallax π because of the Earth's motion around the sun (see Slide 2–15). This parallax correction depends on the distance of the star. Therefore a proper motion measurement should include a parallax measurement π (trigonometric distance measurement) or at least an lower limit for the distance of the object.

The angular proper motion is large and easy to measure for nearby stars. In fact a large proper motion is often used as criterion for the search or selection of nearby stars. The 5 stars with the largest proper motion, Barnard's star, Kapteyn's star, Groombridge 1830, Lacaille 9352, and Gliese 1, have a proper motion of about 5 – 10 arcsec/yr and are all nearby ($d < 10$ pc) stars.

Transverse velocity. From the measured angular motion $\mu = (\mu_\alpha^2 + \mu_\delta^2)^{1/2}$ and parallax π the tangential velocity can be derived

$$v_t = 4.74 \text{ km/s} \frac{\mu [\text{arcsec/yr}]}{\pi [\text{arcsec}]} \quad (2.23)$$

The determination of v_t is much less accurate for distant stars. For a given transverse velocity, say 10 km/s, the measurable angular proper motion μ and the parallax π decrease both proportional with the distance and therefore the measuring uncertainty rises rapidly.

2.3.2 Solar motion relative to the local standard of rest

The sun has like all other stars peculiar motion components. For this reason the solar motion is not an ideal reference.

For studying the dynamics of the galactic disk it is therefore useful to define a reference system which is more convenient. For this it is assumed that the galactic disk geometry and rotation is rotationally symmetric. In this system one can define a motion vector composed of a radial component Π , an azimuthal component Θ ($= R \cdot \Omega$) and a vertical component Z (Fig. 2.7) which are defined by

$$\Pi = \frac{dR}{dt}, \quad \Theta = R \frac{d\theta}{dt}, \quad Z = \frac{dz}{dt}.$$

Figure 2.7: Illustration of the motion of the local standard of rest and the solar motion.

With this definition we can define for the solar neighborhood a velocity vector

$$(\Pi, \Theta, Z)_{\text{l sr}} = (0, \Theta(R), 0).$$

which represents the mean velocity of the stars near the sun on their circular orbit around the galactic center. This velocity vector is called local standard of rest (l sr). It represents a very useful reference for the investigation of the stellar dynamics in the galactic disk.

The vector $(u, v, w)_{\text{lsr}}$ is used for the peculiar velocity components of a star relative to the local standard of rest. For the peculiar velocity of the sun, one needs to measure the motion components (u, v, w) of a large group of stars relative to the sun for deriving the average velocity of the sun with respect to this sample:

$$\langle u_{\odot} \rangle = -\frac{1}{N} \sum_{i=1}^N u_i,$$

and similarly for $\langle v_{\odot} \rangle$, $\langle w_{\odot} \rangle$. The resulting mean values are about $(u_{\odot}, v_{\odot}, w_{\odot})_{\text{lsr}} \approx (-10, 10, 5)$ (in km/s) using a sample of disk stars as reference. However, the derived results vary by about ± 5 km/s depending on the selected type of star. The obtained values for the radial and vertical components $u_{\odot}$ and $w_{\odot}$ represent well the solar motion with respect to the local standard of rest. However, the azimuthal velocity Θ must be corrected for a bias effect because the orbits of the stars in the galactic disk have an eccentricity (see Fig. 2.8).

Figure 2.8: Orbits around the galactic center for stars near the sun with different eccentricities.

Stars with an orbits of type *a* in Fig. 2.8 have near the sun their maximum distance from the galactic center and will therefore have an azimuthal velocity component which is smaller than the local standard of rest $\Theta_a < \Theta_{\text{lsr}}$. Contrary to this the stars with orbits of type *c* have their innermost point near the sun and will therefore move faster $\Theta_c > \Theta_{\text{lsr}}$. Stars with circular orbits (b) move with the same speed as the standard of rest $\Theta_b = \Theta_{\text{lsr}}$. Because there are more stars at small galacto-centric radius the average azimuthal speed of the stars will be slower than the local standard of rest $\langle \Theta \rangle < \Theta_{\text{lsr}}$. This bias effect, although difficult to quantify, needs to be considered for the definition of the Θ -component of the solar motion with respect to the local standard of rest.

The result, which is finally found, for the peculiar motion of the sun relative to the local standard of rest is:

$$(u_{\odot}, v_{\odot}, w_{\odot})_{\text{lsr}} = (-9, 12, 7) \quad [\text{km/s}].$$

This implies that the sun is moving with 16.5 km/s in the direction

$$\ell = -\arctan(v_{\odot}/u_{\odot}) = 53^{\circ} \quad \text{and} \quad b = \arcsin(w_{\odot}/(u_{\odot}^2 + v_{\odot}^2)^{1/2}) = 28^{\circ}$$

with respect to the local standard of rest. This is a slow motion toward smaller galactic radii and upwards towards the north galactic pole, in the direction of the star Vega. One should expect much improved values for the solar motion after the GAIA mission.

2.3.3 Velocity dispersion in the solar neighborhood

The velocity dispersion of galactic disk stars can be determined from the measured space motion and applying a correction for the solar motion with respect to the local standard of rest. If only stars in the solar neighborhood are considered then one can neglect the effects of the differential rotation in the Galaxy.

The data show that the peculiar velocities of the stars (with respect to lsr) show essentially a random or Gaussian distribution. It is useful to fit the measurements with a 3-dimensional ‘‘Gaussian ellipsoid’’ function

$$n(u, v, w) = \frac{\nu}{(8\pi^3 \langle u^2 \rangle \langle v^2 \rangle \langle w^2 \rangle)^{1/2}} \exp \left[- \left(\frac{u^2}{2\langle u^2 \rangle} + \frac{v^2}{2\langle v^2 \rangle} + \frac{w^2}{2\langle w^2 \rangle} \right) \right], \quad (2.24)$$

$n(u, v, w)$: is the number of stars per unit volume with velocities in an interval (du, dv, dw) ,

$\langle u^2 \rangle^{1/2}, \langle v^2 \rangle^{1/2}, \langle w^2 \rangle^{1/2}$: are the velocity dispersions along the three axes.

For one component of the velocity dispersion there is

$$\langle u^2 \rangle = \frac{1}{\nu} \int_{-\infty}^{+\infty} dw \int_{-\infty}^{+\infty} dv \int_{-\infty}^{+\infty} n(u, v, w) u^2 du,$$

and similarly for $\langle v^2 \rangle$ and $\langle w^2 \rangle$. This assumes that the principle axis of the velocity ellipsoid are along the coordinate axes. This is a simplification which is reasonable for basic results, but may be an over-simplification for more subtle studies.

Table 2.7: Velocity dispersion in km/s for different types of stars in the solar neighborhood.

stellar type	$\langle u^2 \rangle^{1/2}$	$\langle v^2 \rangle^{1/2}$	$\langle w^2 \rangle^{1/2}$
B0 main seq.	10	9	6
A0 main seq.	15	9	9
F0 main seq.	24	13	10
G0 main seq.	26	18	20
K0 main seq.	28	16	11
M0 main seq.	32	21	19
Class. Cepheids	13	9	5
G giants	26	18	15
M giants	31	23	16
planetary nebulae	45	35	20
white dwarfs	50	30	25
RR Lyr Var. (halo)	160	100	120

Velocity dispersion for different stellar types. The components of the velocity dispersion have been determined for many different types of stars and a few results are given in Table 2.7. Interesting properties of the stellar velocity dispersion are:

- for all stellar types the three dispersion components behave like

$$\langle u^2 \rangle^{1/2} > \langle v^2 \rangle^{1/2} > \langle w^2 \rangle^{1/2},$$

with roughly $\langle w^2 \rangle^{1/2} \approx 0.5 \langle u^2 \rangle^{1/2}$. Thus the dispersion in the radial velocity component is twice as large as for the vertical velocity component.

- B and A type main-sequence stars and classical Cepheids have the smallest velocity dispersion while evolved low mass stars, like M giants, planetary nebulae and white dwarfs show a much larger dispersion. Thus, there exists for different stellar types a clear correlation between the average age of the sample and the velocity dispersion. This is similar to the typical disk scale height for the different stellar types (Table 2.6).
- halo stars, in Table 2.7 represented by the RR Lyr stars, show a completely different velocity distribution than the disk stars.

2.3.4 Moving groups

Star clusters are gravitationally bound systems. For this reason all cluster members have essentially the same space velocity. The analysis of space velocities is therefore a very powerful tool to separate cluster members from non-cluster members.

The same method can also be applied to moving groups and associations of young stars. These systems are not gravitationally bound, but they were probably bound some time ago when they were formed in an interstellar cloud. The stars in a moving group are therefore moving still in the same space direction. Based on this property it is possible to identify young stars with the same age. Famous examples are the β Pic (age ≈ 20 Myr) and TW Hya (age ≈ 5 Myr) moving groups. Members of these two groups belong to the nearest young stars in the solar neighborhood.

Slide 2–16 shows the proper motion of the Hyades cluster. Because the stars have parallel space motions, they seem to converge in a common vertex point on the sky.

2.3.5 High velocity stars

A small fraction of galactic stars have a very large space motion. They are obviously not moving with the general flow of stars around the galactic center. The orbits of these high velocity stars can be characterized from their motion in the solar neighborhood. According to their velocity vector (u, v, w) (radial, azimuthal, vertical) one can say qualitatively:

- a star with an azimuthal component $v < -250$ km/s is on a retrograde orbit,
- stars on a prograde orbit but with a large space velocity $v_s > 100$ km/s are on elliptical orbits with $e > 0.3$,
- the highest eccentricity have stars with $v \approx -250$ km/s because then they have only a very small azimuthal velocity component. They are on orbits which go close to the galactic center.

A classical analysis of high-velocity stars is shown in Slide 2–17 with a so-called Bottlinger diagram. The diagram distinguishes between old disk stars which are on prograde orbits with enhanced eccentricities $e \approx 0.4$, and stars which belong to the halo population which can have very high orbital eccentricities $e \approx 0.5 - 0.9$ or even retrograde orbits. Again, the space velocity is a good tool to distinguish and characterize different stellar populations.

2.3.6 Radial velocity dispersion in clusters

The measurable angular proper motion decreases rapidly with distance and therefore it becomes impossible to determine the tangential velocity component v_t for distant stars and clusters. However, it is still possible to measure spectroscopically the radial velocity of stars and the velocity dispersion of a system. We consider this technique here for the measurement of the velocity dispersion in stellar clusters.

From the observational data we can distinguish three different cases:

- In nearby clusters it is possible to measure for many individual stars (essentially all bright objects) their radial velocity v_r and also their transverse velocity v_t from the proper motion. In this case one can carry out a detailed kinematic study of the cluster and investigate radial dependencies and anisotropies in the velocity ellipsoid for the cluster stars. Nearby stellar clusters such as the Hyades or Pleiades can be studied in this way.
- In more distant clusters only radial velocities v_r of individual stars can be measured because the proper motion components are too small. For such cases one should consider the distribution of the measured stars within the cluster. For example, in globular clusters, it is often difficult to get spectra of stars in the crowded central region and only the velocities of stars further out are measured. In this case one needs to consider the dependence of the stellar velocity with distance to the cluster center.
- For very distant clusters, e.g. in other galaxies, the individual stars cannot be resolved. In this case one can just measure the line width of the integrated cluster spectrum and determine a rough line broadening due to the velocity dispersion in the cluster. This analysis needs to consider which stars contribute most to the integrated spectrum, and preferentially they should have stable (non-pulsating atmospheres), narrow spectral lines.

From the measured radial velocities one can derive the systemic radial velocity $\langle v_r \rangle$ and velocity dispersion $\langle v_r^2 \rangle^{1/2}$ by fitting the data with a Gaussian distribution:

$$f(v) = \frac{1}{(2\pi\langle v_r^2 \rangle)^{1/2}} \exp\left[-\frac{(v - \langle v_r \rangle)^2}{2\langle v_r^2 \rangle}\right]. \quad (2.25)$$

Table 2.3.6 gives some values for the measured stellar velocity dispersion in clusters. It is visible that the dispersion is smaller for open clusters and larger for globular clusters. According to the virial theorem the velocity dispersion v_r is a measure of the ratio between cluster mass M_{cl} and the cluster radius, e.g. the core radius r_c :

$$\langle v_r^2 \rangle \propto \frac{M_{cl}}{r_c}.$$

Measurements of the radial velocity dispersion $\langle v_r^2 \rangle^{1/2}$ has the following diagnostic potential.

- estimates of the cluster mass including the invisible mass,
- investigation of the radial mass segregation in the cluster,
- search for signatures indicating transient processes.

Table 2.8: Key parameters for open and globular clusters: stellar velocity dispersion $\langle v_r^2 \rangle^{1/2}$ in km/s core radius and total mass.

cluster	type	$\langle v_r^2 \rangle^{1/2}$	r_c [pc]	M [$M_\odot$]
Pleiades	open	0.5	1.4	800
Praesepe	open	0.5	3.5	550
ω Cen	globular	9.8	3.8	$5 \cdot 10^6$
NGC 6388	globular	18.9	0.5	$1.3 \cdot 10^6$

2.3.7 Kinematics of the galactic rotation

Qualitative expectations for the solar neighborhood. For the solar neighborhood we consider the effect of the galactic rotation on the systematic motion of stars for the radial velocity $v_r(\ell)$ and transverse velocity $v_t(\ell)$ direction as function of galactic longitude. For this discussion we assume that the Milky Way is rotating differentially, in the sense that the orbital period is shorter for an object closer to the galactic center. Further, we assume that the stars move on circular orbits.

Figure 2.9: Sketch of the systematic radial velocity of stars relative to the local standard of rest because of the differential galactic rotation.

Radial velocity. Because of the differential rotation (shorter orbital period for smaller r) the stars with $R < R_0$ will overtake the sun, while stars with $R > R_0$ will be overtaken by the sun. According to Fig. 2.9 the systematic radial velocity is:

- v_r is positive for $0^\circ < \ell < 90^\circ$ and $180^\circ < \ell < 270^\circ$, and
- v_r is negative for $90^\circ < \ell < 180^\circ$ and $270^\circ < \ell < 360^\circ$,
- v_r is zero for $\ell = 0^\circ, 90^\circ, 180^\circ$, and 270° ,
- $v_r(\ell)$ is roughly a sine-curve $v_r(\ell) \approx c_1 \sin 2\ell$.

Transverse velocity. Because all stars inside the solar orbit $R < R_0$ overtake the sun they move towards larger longitude. The stars outside the sun $R > R_0$ move backwards, but

this is again in the direction of the galactic longitude angle and they have therefore again a positive angular motion. According to Fig. 2.10 the systematic transverse velocity is:

- v_t is zero or positive for all longitudes $0^\circ \leq \ell \leq 360^\circ$
- v_t has a maximum for the direction $\ell \approx 0^\circ$ and 180° ,
- v_t is zero for $\ell \approx 90^\circ$ and $\ell \approx 270^\circ$,
- $v_t(\ell)$ behaves like a shifted double cosine curve $v_t(\ell) \approx c_2 \cos 2\ell + c_3$

Figure 2.10: Sketch of the systematic transverse velocity of stars relative to the local standard of rest because of the differential galactic rotation.

General rotation formula. We will now derive the general formula for the radial velocity v_r , the transverse velocity v_t and the angular proper motion in galactic longitude μ_ℓ for a differentially rotating disk. We consider again only circular orbits for the derivation. The definition of used parameters are given in the schematic geometric sketch shown in Fig. 2.11.

Radial velocity. The measured radial velocity for a star at point P is:

$$v_r = \Theta \cos \alpha - \Theta_0 \sin \ell. \quad (2.26)$$

According to the law of sines, and $\sin(\alpha + 90^\circ) = \cos \alpha$ there is:

$$\frac{\sin \ell}{R} = \frac{\sin(\alpha + 90^\circ)}{R_0} = \frac{\cos \alpha}{R_0}.$$

We can replace in Eq. 2.26 $\cos \alpha$ and use the angular orbital velocity $\Omega = \Theta/R$ or $\Omega R = \Theta$

$$v_r = \Theta \frac{R_0}{R} \sin \ell - \Theta_0 \sin \ell = \Omega R_0 \sin \ell - \Omega_0 R_0 \sin \ell = (\Omega - \Omega_0) R_0 \sin \ell. \quad (2.27)$$

This is a general result which assumes only that the galactic rotation is circular.

Figure 2.11: Geometry for the derivation of the relative stellar motions in a differentially rotating galactic disk.

From Eq. 2.27 we can derive the change in the angular rotation rate for different galactic radii $\Omega(R)$ for stars with known distances (if we know Ω_0 and R_0). We can also use this formula for the derivation of the galactic rotation curve from emission line observations of the interstellar gas (see next section).

Figure 2.12: Illustration for the radial velocity of stars as function of distance due to their orbital rotation in the galactic disk.

Qualitatively, we can say for the dependence of the radial velocity $v_r(d)$ as function of distance (see Fig. 2.12):

- for the quadrant $0^\circ < \ell < 90^\circ$ the nearby stars are closer to the galactic center and they have therefore all a positive radial velocity.
- The maximum radial velocity $v_r(\text{max})$ is reached for the point along the sight-line where the distance to the galactic center is minimal,
- the radial velocity is zero $v_r = 0$ where the sight-line intersects the solar circle “on the other side”,
- further out the radial velocity v_r is negative and decreases further with distance.

- the quadrant $360^\circ > \ell > 270^\circ$ behaves similar, only the sign of the radial velocity is reversed.
- for the quadrant $90^\circ < \ell < 180^\circ$ we see only regions further out in the galaxy and therefore the difference $\Omega(R) - \Omega_0$ and the radial velocity v_r is always negative and decreases with distance.
- the quadrant $270^\circ > \ell > 180^\circ$ behaves similar, only the sign of the radial velocity is reversed.

Tangential velocity. For the tangential or transverse velocity of a star at point P there is:

$$v_t = \Theta \sin \alpha - \Theta_0 \cos \ell, \quad (2.28)$$

where v_t is positive for direction towards larger longitude ℓ . From Fig. 2.11 it is visible that

$$R \sin \alpha = R_0 \cos \ell - d.$$

One can now replace in Eq. 2.28 the term $\sin \alpha$, rearrange similar to the case of v_r , and it results

$$v_t = \frac{\Theta}{R} (R_0 \cos \ell - d) - \Theta_0 \cos \ell = (\Omega - \Omega_0) R_0 \cos \ell - \Omega d. \quad (2.29)$$

A similar discussion as for the radial velocity could be made for the transverse velocity of the stars in the galactic disk as function of distance. Essentially all stars on the other side of the galactic center would have a negative transverse velocity. However, the proper motions are very difficult to measure for distant stars and therefore this topic is not discussed here.

Oort's constants. In order to get accurate numerical values of the differential galactic rotation we can evaluate the general formulae 2.27 and 2.29 for the solar neighborhood and use the available, more accurate, velocity measurements of nearby stars.

Radial velocity. We consider first the radial velocity v_r given in Eq. 2.27 for a fixed longitude ℓ at the position of the sun R_0 . The only term which depends on distance is $\Omega - \Omega_0$. This difference can be approximated by a first-order Taylor expansion

$$(\Omega - \Omega_0) \approx \left(\frac{d\Omega}{dR} \right)_{R_0} (R - R_0). \quad (2.30)$$

The derivative of the angular rotation is:

$$\frac{d\Omega}{dR} = \frac{d}{dR} \left(\frac{\Theta}{R} \right) = \frac{1}{R} \frac{d\Theta}{dR} - \frac{\Theta}{R^2}$$

so that

$$\left(\frac{d\Omega}{dR} \right)_{R_0} = \frac{1}{R_0} \left(\frac{d\Theta}{dR} \right)_{R_0} - \frac{\Theta_0}{R_0^2}.$$

We can write Eq. 2.27 to first order

$$v_r \approx \left[\left(\frac{d\Theta}{dR} \right)_{R_0} - \frac{\Theta_0}{R_0} \right] (R - R_0) \sin \ell. \quad (2.31)$$

For $d \ll R_0$ the difference between R_0 and R can be approximated by

$$R_0 - R \approx d \cos \ell$$

and it results using $\sin \ell \cos \ell = (1/2) \sin 2\ell$

$$v_r \approx \frac{1}{2} \left[\frac{\Theta_0}{R_0} - \left(\frac{d\Theta}{dR} \right)_{R_0} \right] d \sin 2\ell. \quad (2.32)$$

We then obtained the double sine-wave variation of the radial velocity with galactic longitude as derived before from a qualitative discussion

$$v_r \approx A d \sin 2\ell \quad \text{with} \quad A = \frac{1}{2} \left[\frac{\Theta_0}{R_0} - \left(\frac{d\Theta}{dR} \right)_{R_0} \right], \quad (2.33)$$

where A is called the Oort's constant A .

Transverse velocity: Similarly we can evaluate the equation for the transverse velocity component v_t (Eq. 2.27) and get

$$v_t \approx \left[\left(\frac{d\Theta}{dR} \right)_{R_0} - \frac{\Theta_0}{R_0} \right] (R - R_0) \cos \ell - \Omega_0 d \approx \left[\frac{\Theta_0}{R_0} - \left(\frac{d\Theta}{dR} \right)_{R_0} \right] d \cos^2 \ell - \left(\frac{\Theta_0}{R_0} \right) d. \quad (2.34)$$

Using the trigonometric identity $\cos^2 \ell = (1 + \cos 2\ell)/2$ yields

$$v_t \approx \frac{1}{2} \left[\frac{\Theta_0}{R_0} - \left(\frac{d\Theta}{dR} \right)_{R_0} \right] d \cos 2\ell - \frac{1}{2} \left[\frac{\Theta_0}{R_0} + \left(\frac{d\Theta}{dR} \right)_{R_0} \right] d.$$

It results the shifted double wave cosine curve with the Oort's A defined above and Oort's B constant

$$v_t \approx d (A \cos 2\ell + B) \quad \text{with} \quad B = -\frac{1}{2} \left[\frac{\Theta_0}{R_0} + \left(\frac{d\Theta}{dR} \right)_{R_0} \right]. \quad (2.35)$$

This derivation was obtained in 1927 by Oort, who could prove that the Galaxy has a differential rotation.

Local rotation constants. Important is the result that the local value of the angular rotation rate and the radial derivative of the azimuthal velocity can be expressed with the Oort's constants:

$$\Omega_0 = \frac{\Theta_0}{R_0} = A - B \quad \text{and} \quad \left(\frac{d\Theta}{dR} \right)_{R_0} = -(A + B) \quad (2.36)$$

This describes well the local galactic rotation parameters.

Oort's rotation constant A can be derived from measurements of the radial velocity v_r of stars with known distance d for different longitudes ℓ using the formula

$$A = \frac{v_r}{d \sin 2\ell}.$$

Measured values for the Oort's constant A are

$$A \approx +15 \frac{\text{km}}{\text{s kpc}}.$$

The determination requires the measurement of radial velocities of a large, unbiased sample of stars. The measurement of radial velocities of many stars with an absolute precision of a few km/s is relatively easy. More problematic are the stellar distance determinations for this sample. Uncertainties for the A -value are of the order $\pm 1 \text{ km}/(\text{s kpc})$.

Oort's constant B needs measurements of the transverse velocity (proper motion) of stars with known distance and for different galactic longitude ℓ using the formula

$$B = \frac{v_t}{d} - A \cos 2\ell$$

Because there is no transverse velocity component $v_t \approx 0$ for the direction of the galactic rotation and $\cos 2\ell \approx -1$, there is roughly $B \approx -A$. The transverse velocity of many stars has been measured with the Hipparcos satellite and the resulting B -value obtained is

$$B \approx -12 \frac{\text{km}}{\text{s kpc}}$$

with a similar uncertainty as for the A -constant.

Angular rotation rate: Estimates of the angular rotation rate for the sun follow from

$$\Omega_0 = A - B \approx 27 \frac{\text{km}}{\text{s kpc}} = 27 \frac{10^{-9}}{\text{yr}}$$

The orbital period of the sun (or better the local standard of rest) around the galactic center is then roughly

$$P_0 = \frac{2\pi}{\Omega_0} = 230 \text{ Myr}.$$

Local velocity gradient. Also from the Oort's constants follows that the galactic rotation curve is essentially flat at the location of the sun. Explicitly:

$$\left(\frac{d\Theta}{dR}\right)_{R_0} \approx -(A + B) \approx -3(\pm 2) \frac{\text{km}}{\text{s kpc}}.$$

Galactic rotation velocity. The distance of the sun from the galactic center R_0 has been determined with various methods, like globular clusters, the motion of the star around the central black hole, and others which are not discussed here. The typical result of such studies yields $R_0 \approx 8 \text{ kpc}$. Because we know the angular velocity Ω_0 from the Oort's constant we get also the velocity of the galactic rotation Θ_0 for the local standard of rest:

$$\Theta_0 = R_0 \Omega = R_0(A - B) \approx 8.2 \text{ kpc} \cdot 27 \frac{\text{km}}{\text{s kpc}} = 220 \text{ km/s}$$

2.3.8 The GAIA revolution

Galactic astronomy will be revolutionized in a few years by the results from the GAIA space mission. The GAIA instrument is measuring now for billion of stars very accurate positions, parallax distances, proper motions, radial velocities, photometric brightness, color, and photometric variability as well as spectral types. More accurately GAIA will measure the following:

- photometry, colors, stellar positions, proper motion, parallax distances for all (≈ 1 billion) stars down to magnitude $m_V \approx 20$,
- radial velocities with a precision of a few km/s for all stars (about 100 million) down to $m_V \approx 17^m$,
- spectroscopy for millions of stars will be obtained for metallicity determinations.

The data quality of the GAIA mission will be extremely good. We pick as an illustration only one example, the measurements of distances by the annual parallax:

- about 340'000 stars down to $m_V = 10^m$ will have a parallax uncertainties of about 5–10 μas (micro-arcsec). The parallax of a star at 1 kpc is 1000 μas allowing thus distance measurements with a precision better than 1 %.
- a precision of 25 μas will be achieved for 30 million stars with $m_V < 15^m$. This limit includes very bright stars in the galactic bulge, many halo stars, and countless stars in the Magellanic Clouds.
- for 1 billion stars with $m_V < 20^m$ the achievable precision is 300 μas . This will provide accurate distances for all faint stars to distances up to 1 kpc.

This should be compared to the currently available Hipparcos distance parallaxes which reached a precision of a few mas (milli-arcsec) for about 100'000 stars with $m_V < 7.5^m$. This provided a mapping of all bright stars to about 100 pc. GAIA will go about 300 times further in distance.

Beside this, GAIA will also detect about $> 100\,000$ quasars, $> 100\,000$ asteroids, detect the reflex motion of > 1000 stars because of the presence of an extrasolar planet, measure the astrometric light bending due to General Relativity by the sun and planets. GAIA has also the potential to uncover new phenomena we are not aware of yet.

Expected results for galactic astronomy. GAIA will use many of the described methods discussed in this chapter for the study of the Milky Way. The much improved quality of the data will clarify or at least provide important progress for the following topics:

- we will get a synoptic picture of the evolution of our Galaxy from its detailed geometric and dynamic structure and the distribution of stellar metallicities as function of age,
- trace accurately the distribution of the invisible dark matter from the motion of stars out to distances $\gtrsim 10$ kpc from the sun,
- map the spiral structure and define its dynamics in much detail from the distribution and velocities of young stars,
- measure the bar and inner bulge dynamics,

- provide a comprehensive inventory of galactic star clusters and associations, measure the velocity dispersion of their star and identify the processes which leads to the cluster disruption,
- measure the distribution of dust with photometric measurements of the interstellar absorption (reddening) for millions of stars,
- describe the interaction of halo stars and globular clusters with the galactic disk based on the positions and motions of halo stars,
- identify the star streams which are the debris of dwarf galaxies which were tidally disrupted by an interaction with the Milky Way.

Beside this, there are countless other questions and problems in galactic astronomy which will be solved by the data of this mission.

Current status of the mission. GAIA was launched in December 2013 and has started its science operation in mid 2014. The mission duration is 5 year and up to now (early 2016) things are working fine. GAIA needs to take measurements for 5 years before the final science results can be produced. Before the end of the mission there are some data releases where some first results (source list, positions, preliminary brightness) are already distributed to the community.

Measuring principles. GAIA has two telescopes which observe simultaneously the sky in two observing directions with a fixed angle of 106.5° between. The spacecraft rotates continuously around an axis perpendicular to the line of sight of the two telescopes and together with a slight spacecraft precession the whole sky is scanned many time during the 5 year life-time.

The objects from the two observing directions are registered by the same detector so that their relative positions can be determined accurately.

The two telescopes have primary mirrors of 1.45×0.5 m each. The detector system consists of 106 CCD detectors with 4500×1966 pixel each what gives in total a system with 10^9 pixels

Because of the spacecraft rotation all stars move in the same direction over the focal plane. First the stars “hit” the wave front sensor and telescope angle monitor. The sky mapper detects the targets and this defines then the data to be stored by the following systems in order to reduce the data downlink rate. The large array is used for accurate position measurements, then low resolution prism spectro-photometry is taken, before the high resolution spectrograph takes $\lambda/\Delta\lambda \approx 10000$ spectra for radial velocity measurements and the spectral characterization of the brighter targets.

2.4 Interstellar matter (ISM) in the Milky Way

In this section we consider the distribution of the interstellar matter in different regions of the Milky Way disc. For this we distinguish five gas-components:

- three diffuse components; atomic gas, photo-ionized gas, and collisionally ionized gas,
- two higher density components, molecular clouds and H II regions which are usually associated with star forming regions.

Table 2.9: Components of the ISM in the Milky Way disc.

	T [K]	$N(\text{H})[\text{cm}^{-3}]$	gas type	main particles
1.	10 – 100	$10^3 - 10^6$	molecular cloud	H_2 , dust, CO, ...
2.	100 – 1000	≈ 10	diffuse atomic gas	H^0 , dust, C^+ , e^- , N^0 , O^0 , ...
3.	≈ 10000	$10 - 10^4$	H II-regions	H^+ , e^- , dust, X^{+i} , ...
4.	≈ 10000	≈ 0.1	diffuse, photo-ionized gas	H^+ , e^- , dust, X^{+i} , ...
5.	$\gtrsim 10^6$	$\approx 10^{-3}$	diffuse, collisionally ionized gas	H^+ , e^- , X^{+i} , ...

In the Milky Way disc the components 1 and 2 contribute about 90 % to the baryonic mass of the ISM, while the components 2 and 5 fill essentially the space in the disk.

2.4.1 The ISM in the solar neighborhood

The distribution of the ISM in the solar neighborhood is determined observationally by interstellar absorption lines in the spectra of nearby stars and the reddening of stellar colors by dust. Thanks to observations along many sight-lines to stars which are at different distances it is also possible to estimate the distances to the absorbing gas structures.

Important absorption lines for the ISM in the solar neighborhood are:

- | | | |
|-------|---------------------------|--|
| 1. | molecular gas | H_2 and CO in the far UV, CH in the visual |
| 2. | atomic gas | H I Lyman lines, C I, C II, Si II, O I in the far UV;
Ca II, Na I in the visual |
| 3.+4. | photo-ionized gas | C IV, Si III, Si IV in the far UV |
| 5. | collisionally ionized gas | O VI in the far UV |

Emission lines and continuum emission from dust are not suited for the investigation of the ISM in the immediate solar neighborhood ($d < 300$ pc). This emission is very diffuse because there are no high density regions within this distance. Further it is almost impossible to determine the distance of the diffuse emission.

For larger distances ($d > 300$ pc) there are some well defined high density regions, which produce emission with higher surface brightness. Usually, this emission can be associated with molecular clouds or H II regions with young stars, which allow the determination of the distance.

Distribution of gas and dust within 250 pc. The strongest interstellar absorption lines are the far-UV absorption of H_2 and HI $\text{Ly}\alpha$. These lines trace the most abundant element and therefore they provide very reliable and very sensitive results. However, the H_2 and HI line absorptions require satellite observations and a hot $\gtrsim 20'000$ K background star. Observationally less demanding are dust extinction measurements using photometric data or absorption line data of Na I or Ca II in the optical spectra of A to K-type stars. An example of such absorption line data for the Na I line at 589.6 nm is shown in slide 2-22.

Local bubble. In the immediate neighborhood of the sun there is a low-density bubble which extends to about 60 pc. The observations for this local bubble show that

- the hydrogen column density up to about 60 pc is
 $N(\text{H}) = N(\text{HI}) + 2 \cdot N(\text{H}_2) \sim 1 - 5 \cdot 10^{18} \text{ cm}^{-2}$
 $\rightarrow$ mean density $\sim 0.01 - 0.1 \text{ cm}^{-3}$,
- the local HI -gas shows a systematic expansion of about 30 – 50 km/s,
- the O VI -absorption line increases with distance.

The average particle density in the Milky Way disk is more like

$$\bar{\rho}_{\text{ISM}} \approx 1 \text{ cm}^{-3}.$$

This indicates, that the sun is located in a hot bubble with a density far below (factor 10-100) the mean density in the disc.

An important conclusion for Galactic astronomy is:

- Hot bubbles, like the one around the sun, are quite frequent in the Milky Way disc. They are associated with supernova explosions. The supernova interpretation indicates that the local bubble has an age of $\approx 10^7 - 10^8$ years.

The dense clouds around 150 pc. The sun is surrounded by several dense interstellar clouds and star forming regions at distances of about ≈ 150 pc. The location of these clouds is traced with measurements of the dust extinction and the absorption by the $\text{Na I } \lambda\lambda 589.0, 589.6$ nm resonance lines along the sight lines towards nearby stars. Slides 2-23 and 2-24 show inversion maps for the dust extinction and the Na I line absorption perpendicular to the galactic disk and for the Gould belt in the galactic plane. In Slide 2-23 the maps for the dust extinction and Na I are compared. Essentially the same structures are seen in both maps. The Gould belt is a local disk ring structure of young stars and star forming regions, which is inclined by $\pm 18^\circ$ with respect to the disk plane. The young stars are ideal background targets for accurate measurements of the Na I line absorption and the dust extinction.

The maps in the Slides 2-22 and 2-23 show:

- the gas is predominantly distributed in the Milky Way disc,
- molecular clouds with high density are e.g. located in the direction of the galactic center (Sco-Oph), or towards the anticenter (Tau). In these regions star formation takes place.
- A “tunnel” of low density gas extends through the disk and gives us clear views towards the North and South galactic poles. This low density region is filled with hot, atomic and ionized gas.

This separation between well localized dense clouds, containing cold molecular clouds, and diffusely distributed atomic hydrogen in low density bubbles is quite typical for the entire Milky Way disc.

Dust is present in the diffuse, low density regions and in the dense clouds. The interstellar extinction by dust is therefore a good measure for the mass column density along the line of sight. Observations provide a good empirical relation between dust reddening E_{B-V} and hydrogen column density $N(\text{H})$:

$$N(\text{H}) [\text{cm}^{-2}] \approx 5.8 \cdot 10^{21} E_{B-V} [\text{mag}].$$

2.4.2 Global distribution of the ISM in the Galaxy

Well suited for the investigation of the global distribution of the ISM in the Milky Way are line and continuum emissions in spectral regions with little interstellar absorption. These are observations in the radio range, in the far IR, hard X-rays and gamma rays.

Important emission features for the different gas components are:

- | | | |
|-------|---------------------------|---|
| 1. | molecular clouds | CO-lines, IR-dust emission, γ -rays from the π^0 -decay |
| 2. | atomic Gas | H I 21 cm line, fine structure lines (e.g. C II), IR-dust emission, γ -rays (π^0 -decay) |
| 3.+4. | photo-ionized gas | H I-recombination lines (near-IR, radio range), bremsstrahlung (radio-range), collisionally excited lines |
| 5. | collisionally ionized gas | X-ray radiation (bremsstrahlung and X-ray line emission) |

The observations of emission lines provide one important advantage, when compared to continuum emission; from lines one can also measure the **radial velocity** v_r of the emission region.

Distribution perpendicular to the disc. The distribution of the interstellar matter can be derived from maps showing the emission of the different components, like the H I map shown in Slide 2–3 or the dust emission map in Slide 2–5.

The following table gives estimates of the disk scale height β (using $D(z) = D_0 e^{-|z|/\beta}$) for different components in the direction perpendicular to the disc:

- | | | |
|----|---------------------------|---|
| 1. | molecular clouds | $\beta \approx 30$ pc |
| 2. | atomic Gas (H I) | $\beta \approx 180$ pc |
| 3. | photo-ionized gas | difficult to determine
$\beta > \beta_{\text{Pulsar}} > 200$ pc based on dispersion measurements (not discussed yet) for pulsars |
| 4. | H II-regions | β like molecular clouds |
| 5. | collisionally ionized gas | $b > 250$ pc, hot gas extends far into the galactic halo |

2.4.3 Galactic rotation curve from line observations

Line emission regions along a line of sight in the Milky Way disc have different radial velocities due to the galactic rotation curve. In Sect. 2.3.7, we have derived the general

formula 2.27 for the radial velocity for a given galactic longitude ℓ :

$$v_r = (\Omega - \Omega_0)R_0 \sin \ell.$$

An emission line observations shows therefore many radial velocity components v_i depending on the angular velocity $\Omega(R_i)$ of the emitting clouds along the line of sight (see Fig. 2.13)

Figure 2.13: Illustration of the measured radial velocities of emission line clouds located along a sight line with fixed $\ell \approx 20^\circ - 70^\circ$.

Usually, it is not possible to derive an accurate distance to the cloud and derive its galactocentric distance R_i .

However, for longitudes $0^\circ < \ell < 90^\circ$ or $360^\circ > \ell > 270^\circ$ there exists a maximum (respectively minimum) radial velocity $v_{\max}$ for the point with the smallest galactocentric distance along the line of sight $R_{\min} = R_0 \sin \ell$. There, we see the galactic motion exactly along the radial velocity direction component and there is

$$v_{\max}(\ell) = (\Omega(R_{\min}) - \Omega_0)R_0 \sin \ell.$$

This maximum radial velocity will only occur for a differentially rotating galaxy for which Ω is steadily increasing with decreasing radius R (or $\Omega \cdot R \approx \text{const}$). This is the typical case for the disks in spiral galaxies but not for the central bulge region. Therefore it is possible to determine from the maximum velocity $v_{\max}(\ell)$ the Galactic rotation curve for $R_{\text{bulge}} < R < R_0$.

$\ell - v_r$ -maps: Galactic longitude – radial velocity. The radial velocity and the distribution of the gas in the Milky Way can be plotted in a diagram for the galactic longitude and the radial velocity. Slide 2–25 shows these maps for H I and CO. In the exercises, we will derive the positions of different galactocentric rings (inner ring $R < R_0$, outer ring $R > R_0$) in this diagram.

From the $\ell - v_r$ -maps of H I and CO one can deduce:

- the rotation curve of the Milky Way is essentially flat
 $v_r(R) = R\Omega(R) \approx \text{const.}$ in the range $R \approx 3 - 7$ kpc,

- the H I-gas extends from about 3 kpc out to about 15 kpc from the galactic center (outermost rim ~ 18 kpc),
- the CO molecular clouds are mainly located within a broad ring extending from about 3 kpc out to 8 kpc.

The $v_{\max}$ -method works best in the longitude range $\ell \approx 20^\circ - 70^\circ$ (respectively $\ell \approx 340^\circ - 290^\circ$). This yields the galactic rotation curve from about 3 kpc to 7 kpc. Inside 3 kpc is the galactic bulge and the assumption of a differentially rotating system is not valid. Between $\ell \gtrsim 90^\circ \lesssim 270^\circ$ the method does not work, because there is no maximum radial velocity point.

Galactic rotation curve for $R > R_0$. The studies of the stellar dynamics in the solar neighborhood prove that the Galactic rotation curve is also flat near the solar cycle $R \approx R_0$. According to the Oort's constant there is:

$$\left(\frac{d\Theta}{dR}\right)_{R_0} \approx -(A + B) \approx -3(\pm 2) \frac{\text{km}}{\text{s kpc}}.$$

For $R > R_0$ the rotation curve $\Omega(R)$ can only be derived if the distance of a H I or a CO cloud with measured radial velocity can be determined. This can be achieved, if there are bright young stars, such as Cepheids, which can be associated with a CO cloud. The distance follows from the brightness of the Cepheid, the period-luminosity relation, and the correction for the interstellar extinction. From the distance d and the longitude ℓ follows R , so that $\Omega(R)$ can be derived from the measured radial velocity of the associated cloud.

2.4.4 H I and CO observations in other galaxies

The H I – 21 cm and the CO 2.3 cm lines are ideal for the investigation of the general distribution of the interstellar matter in disk galaxies. The line observations provide for resolved galaxies intensity and radial velocity maps. The maps provide also rotation curves using the $v_r(R)$ -profile along the major axis. For very distant, unresolved galaxies, one can measure the H I or CO velocity profile. This is sufficient for deriving the disk rotation velocities v_{rot} , if the inclination i of the disk can be determined from a resolved, e.g. optical image. The information which can be extracted from interstellar line observation is shown schematically in Slide 2–26, while Slide 2–27 and 2–28 give some examples for real data. Many nearby disk galaxies have been imaged in H I. CO data are still quite rare.

Sensitivity: A “deep” H I-map can reveal H I-gas with a column density down to $N(\text{H I}) \approx 10^{19} \text{ cm}^{-2}$. This corresponds to a mean surface density of $0.1 \text{ M}_\odot/\text{pc}^2$ per spatial resolution element.

The mean H I-surface density in disk galaxies is typically $1 - 4 \text{ M}_\odot/\text{pc}^2$. It seems that this is a kind of self-regulated value. If the surface density is larger than this value, then the disk becomes optically thick for ionizing UV-radiation and atomic hydrogen H I is transformed into molecular H_2 , because the radiative dissociation is strongly reduced.

In the centers of disk galaxies, the intensity maps show usually a minimum. Obviously, the density of atomic and molecular interstellar gas is strongly reduced in the bulges of disk galaxies.

In many galaxies the distribution of H I-gas can be traced to much larger distances than the visible star light. Edge-on galaxies show also the vertical distribution of H I which is usually broader than for the stars. On the other side, the CO is strongly concentrated to the disk mid-plane. Thus, the distribution of the interstellar matter in the Milky Way is typical for disk galaxies.

In external galaxies it is quite easy to recognize strong asymmetries in the distribution of the gas. Often the edge-on disks show warps, a tilt of the outermost disk ring regions with respect to the central disk. This phenomenon seems also to be present in the Milky Way.

H I-rotation curves. The main motion component of all spiral galaxies is the disk rotation. The motion of the H I of CO gas is always very smooth and deviates typically less than 20 km/s from the general rotation. The measured radial velocity at a given radius R and azimuthal angle ϕ (angle with respect to the line of nodes of the disk and the sky) is:

$$v_r(R, \phi, i) = v_{\text{sys}} + V(R) \cdot \cos \phi \cdot \sin i ,$$

where v_{sys} is the systemic radial velocity of the galaxy and i the inclination of the disk. The intrinsic rotation curve $V(R)$ follows then from the measurements $v_r(R, \phi, i)$ with a cut through the major axis ($\phi = 0^\circ, 180^\circ$) of the galaxy and a radial velocity correction for $\sin i$. Usually it is assumed that the studied galaxy is rotationally symmetric.

The obtained $V(R)$ -rotation curves are flat ($V(R) \approx \text{const.}$) out to large distances for essentially all disk galaxies. This cannot be explained with the mass distribution of the stars and the interstellar medium. In all these galaxies, there must be an additional mass component which extends far into the halo.

Chapter 3

Galactic dynamics

Galactic dynamics can be divided into different regimes. First, there is the motion of the gas and the stars in the overall galactic potential. On these large scales the stars behave in a first approximation like test masses in a smooth potential for which dynamical gravitational interactions (near encounters with other stars) can be neglected. The stellar dynamics and the motion of the gas can be used to constrain the potential and the corresponding density distribution of the different galactic components: the stars in the bulge and the disk, the gas located in disk, and the extended dark matter halo.

Second, on small scales the motion of a star is determined by the gravitational potential of many stars in a smooth “background” potential. Depending on the case, it must be distinguished whether the dynamics of a star is strongly affected by individual dynamical interactions with other stars or not. In this context it is important to consider the difference between collisionless systems and systems with collisions.

In this chapter we describe first simple models for smooth gravitational potentials, the associated density distributions and the expected motion parameters and time scales. Then we consider relaxation (collision) time scales and discuss the impact of collisions on the dynamics.

3.1 Potential theory

In this section we describe the force field for a smooth distribution of mass. There exist simple but powerful analytic formula, which give a lot of insight on the motion of test particles in a smooth potential. In particular, we will discuss how the density structure of the Milky Way can be modelled. The description of this section follows the corresponding chapter in the book “Galactic Dynamics” from Binney and Tremaine.

3.1.1 Basic equations for the potential theory

The force $\vec{F}(\vec{x})$ at position $\vec{x}$ on a star with mass m_S is generated by the space distribution of mass $\rho(\vec{x}')$:

$$\vec{F}(\vec{x}) = m_S \vec{g}(\vec{x}) = m_S G \int \frac{\vec{x}' - \vec{x}}{|\vec{x}' - \vec{x}|^3} \rho(\vec{x}') d^3 \vec{x}' . \quad (3.1)$$

$\vec{g}(\vec{x})$ is the vector gravitational field, the force per unit mass or the gravitational acceleration.

The gravitational potential $\Phi(\vec{x})$ is defined by

$$\Phi(\vec{x}) = -G \int \frac{\rho(\vec{x}')}{|\vec{x}' - \vec{x}|} d^3\vec{x}', \quad (3.2)$$

which is the integral of the mass distribution weighted by the inverse distance to the point $\vec{x}$. The gradient for the inverse distance is

$$\vec{\text{grad}}_x \left(\frac{1}{|\vec{x}' - \vec{x}|} \right) = \frac{\vec{x}' - \vec{x}}{|\vec{x}' - \vec{x}|^3}$$

and therefore the gravitational field $\vec{g}(\vec{x})$ can be expressed by the gravitational potential according to

$$\vec{g}(\vec{x}) = -\vec{\text{grad}}_x \Phi(\vec{x}) = \vec{\text{grad}}_x \left(G \int \frac{\rho(\vec{x}')}{|\vec{x}' - \vec{x}|} d^3\vec{x}' \right). \quad (3.3)$$

The potential $\Phi(\vec{x})$ is very useful because it is a scalar field which can be described and analyzed based on equipotential surfaces. Φ contains the same information as the vector gravitational field $\vec{g}(\vec{x})$ and the acceleration $\vec{g}(\vec{x})$ follows from the gradient of the potential.

The potential energy of a system follows from an estimate of the expected change in potential energy if a small additional mass is added to the system with potential $\Phi(\vec{x})$. If a small increment of density $\delta\rho(\vec{x})$ is added then the change in potential energy is:

$$\delta E_{\text{pot}} = \int \delta\rho(\vec{x}) \Phi(\vec{x}) d^3\vec{x}. \quad (3.4)$$

3.1.2 Newton's theorems

Let's start with the simple case of spherical systems to get familiar with the mathematical procedures. Spherical systems are particularly simple because of Newton's theorems.

First theorem of Newton. A body inside a spherical shell experiences no net gravitational force from that shell.

Second theorem of Newton. A body outside a spherical shell experiences a gravitational force equal to the force of a mass point in the center of the shell with the mass of the shell.

Figure 3.1 illustrates the proof of the first theorem. A point P inside the shell is attracted equally strong by opposite shell section “seen” under the same solid angle $d\Omega$. This is obvious for radial sight lines through the center of the shell because the areas (with surface mass $m_{1,2}$) of the opposite regions are proportional to the distances squared $d_{1,2}^2$ from point P . Thus there is $F_1 = m_1/d_1^2 = m_2/d_2^2 = F_2$

This is also valid for an arbitrary “sight” line (full line) because the tilt angles θ_1 and θ_2 between the tangential surfaces and the cone center lines are equal on both sides. Therefore the surface area defined by the solid angle cones are proportional to $d_1^2/\cos\theta_1$ and $d_2^2/\cos\theta_2$ and the attraction from the opposite sides is also equal.

Figure 3.1: Figures for the proof of Newton's first theorem (left) and Newton's second theorem (right).

Inside the shell the gravitational potential is constant because the gravitational force is zero

$$\vec{\text{grad}}_x \Phi = -g = 0.$$

The gravitational potential in the shell can be easily calculated for the central point from a radial form of Eq. 3.2 (see also Eq. 3.9)

$$\Phi(0) = -\frac{GM}{R}, \quad (3.5)$$

where $M = 4\pi\rho(r)dr$ is the total mass of a shell with thickness dr and R is the shell radius.

For the proof of Newton's second theorem a trick according to Fig. 3.1 with a special configuration of points p_1 , p_2 , q_1 and q_2 is needed. We consider two concentric shells with radius R_1 and R_2 and equal mass $M_1 = M_2$. Then one can write the potential for a point p_2 on the outer shell by a surface area region $\delta\Omega$ at point q_1 of the inner shell

$$\delta\Phi(\vec{p}_2) = -\frac{GM}{|\vec{p}_2 - \vec{q}_1|} \frac{\delta\Omega}{4\pi}.$$

This potential is equal to the potential for the point p_1 on the inner shell by a surface area region of the outer shell with the same angular dimensions $\delta\Omega$ at point q_2 .

$$\delta\Phi(\vec{p}_1) = -\frac{GM}{|\vec{p}_1 - \vec{q}_2|} \frac{\delta\Omega}{4\pi}.$$

Thus, there is $\delta\Phi(\vec{p}_2) = \delta\Phi(\vec{p}_1)$ because $|\vec{p}_2 - \vec{q}_1| = |\vec{p}_1 - \vec{q}_2|$ (symmetry) and the summation yields then that the potential due to the entire inner and outer shells are equal

$$\Phi_{\text{shell}_1}(\vec{p}_2) = \Phi_{\text{shell}_2}(\vec{p}_1).$$

We know $\Phi_{\text{shell}_2}(\vec{p}_1) = -GM/R_2$ (from Eq. 3.5) and therefore this is also the result for $\Phi_{\text{shell}_1}(\vec{p}_2)$ for the potential of a point at a radius $R = R_2$ outside a shell with $R_1 < R$ and mass M

$$\Phi_{\text{shell}_1}(R) = -\frac{GM}{R}. \quad (3.6)$$

This outside potential of a spherical shell is equal to the potential of a point with the same mass located at the center.

3.1.3 Equations for spherical systems

Simple equations can be derived for spherical systems using Newton's theorems.

The gravitational force of a spherical density distribution $\rho(r')$ on a star m_S at radius r is determined by the mass $M(r)$ interior to r

$$\vec{F}(r) = m_S \vec{g}(r) = -G \frac{m_S M(r)}{r^2} \vec{e}_r, \quad (3.7)$$

where

$$M(r) = 4\pi \int_0^r \rho(r') r'^2 dr. \quad (3.8)$$

The gravitational potential of a spherical system is the sum of the potentials of spherical mass shells $dM(r) = 4\pi\rho(r)r^2dr$ with $r' < r$ (located inside r):

$$\Phi_{r' < r}(r) = -\frac{G}{r} \int_0^r dM(r')$$

and the mass shells at $r' > r$ (located outside r):

$$\Phi_{r' > r}(r) = -G \int_r^\infty \frac{dM(r')}{r'},$$

or written alternatively

$$\Phi(r) = -4\pi G \left[\frac{1}{r} \int_0^r \rho(r') r'^2 dr + \int_r^\infty \rho(r') r' dr \right]. \quad (3.9)$$

The circular speed $v_c(r)$, which is the speed of a test particle with negligible mass m_S in a circular orbit at radius r , is an important parameter for the characterization of the gravitational potential. The circular speed follows from the equilibrium $F_g(r) = -F_c(r)$ of gravitational force and centrifugal force $F_c = m_S v_c^2/r$:

$$v_c^2(r) = r g(r) = r \frac{d\Phi}{dr} = \frac{GM(r)}{r}. \quad (3.10)$$

This can also be expressed with angular velocity

$$\Omega(r) = \frac{v_c(r)}{r} = \sqrt{\frac{GM(r)}{r^3}}.$$

The potential energy of a spherical system can be calculated from the incremental potential energy formula 3.4. For a spherical system this can be expressed as a change in potential energy due to the small addition of density in a shell at radius r :

$$\delta E_{\text{pot}}(r) = 4\pi r^2 \delta\rho(r) \Phi(r),$$

If we build up a whole spherical mass distribution from inside out by such small spherical mass (density) shell increments then the final potential energy is obtained by integration:

$$E_{\text{pot}} = - \int_0^\infty 4\pi r^2 \rho(r) \frac{GM(r)}{r} dr = -4\pi G \int_0^\infty r \rho(r) M(r) dr. \quad (3.11)$$

3.1.4 Simple spherical cases and characteristic parameters

Potential of a point mass. This is a very simple case which is often referred as a Keplerian potential. For a point mass there is

$$\Phi(r) = -\frac{GM}{r}, \quad \text{and} \quad v_c(r) = \sqrt{\frac{GM}{r}}. \quad (3.12)$$

The potential energy of a point is $-\infty$ (or not defined).

Potential of a homogeneous sphere. Inside a homogeneous sphere with constant ρ there is:

$$M(r) = \frac{4}{3}\pi r^3 \rho. \quad (3.13)$$

The circular velocity increases linearly with radius

$$v_c(r) = \sqrt{\frac{GM(r)}{r}} = \sqrt{\frac{4\pi G\rho}{3}} r. \quad (3.14)$$

The orbital period is then defined by the density ρ

$$T = \frac{2\pi r}{v_c} = \sqrt{\frac{3\pi}{G\rho}} \quad (3.15)$$

The potential energy of a homogeneous sphere with radius a , density ρ and total mass $M = (4/3)\pi G\rho a^3$ follows from Eq. 3.11:

$$E_{\text{pot}} = -4\pi G\rho \int_0^a r M(r) dr = -\frac{16\pi^2 G\rho^2}{3} \int_0^a r^4 dr = -\frac{16}{15}\pi^2 G\rho^2 a^5 = -\frac{3}{5} \frac{GM^2}{a}. \quad (3.16)$$

The gravitational potential of a homogeneous sphere with radius a is

$$\Phi(r) = -2\pi G\rho \left(a^2 - \frac{1}{3}r^2\right) \quad \text{for} \quad r < a, \quad (3.17)$$

$$\Phi(r) = -\frac{GM}{r} \quad \text{for} \quad r > a. \quad (3.18)$$

Gravitational radius. The size of a system is sometimes characterized by the gravitational radius which is defined as ratio between mass squared divided by the total gravitational (potential) energy:

$$r_g = \frac{GM^2}{|E_{\text{pot}}|}. \quad (3.19)$$

For a homogeneous sphere, where $E_{\text{pot}} = -(3/5)GM^2/a$ the corresponding gravitational radius is $r_g = (5/3)a$. The gravitational radius can be a convenient quantity for the definition of the size of systems which have no sharp boundary (e.g. stellar cluster).

The dynamical time scale. The homogeneous sphere is a useful model for an estimate of the dynamical time scale of a system.

If a mass is released from rest in a gravitational field of a homogeneous sphere then its equation of motion is given by the gravitational acceleration

$$g(r) = \frac{d^2 r}{dt^2} = -\frac{d\Phi(r)}{dr} = -\frac{GM(r)}{r^2} = -\frac{4\pi G\rho}{3}r.$$

This is the equation of motion for a harmonic oscillator ($\ddot{x} = -\omega^2 x$) with oscillation period $T = (2\pi/\omega) = \sqrt{3\pi/G\rho}$. This is the same time as is required for a full circular orbit (Eq. 3.15).

Thus, there is for a homogeneous sphere not only a unique circular orbital period but also a unique free fall time t_{ff} , which is the time it takes for any particle released at rest to fall into the center. This time is

$$t_{\text{ff}} = \frac{T}{4} = \sqrt{\frac{3\pi}{16G\rho}} = 0.767 (G\rho)^{-1/2}$$

The dynamical time scale is defined as

$$t_{\text{dyn}} = (G\rho)^{-1/2}. \quad (3.20)$$

This quantity is of the same order as the free-fall time, the crossing time or the orbital time for a particle. According to our definition there is:

$$t_{\text{dyn}} = 1.3 t_{\text{ff}} = 0.65 t_{\text{cross}} = 0.33 t_{\text{orbit}}$$

The dynamical time scale is also a good parameter for the characterization of systems with not too extreme density gradients, as long as ρ is replaced by the mean density $\bar{\rho}$ inside the location of the particle.

$$t_{\text{dyn}} \approx (G\bar{\rho})^{-1/2}.$$

This relation is therefore used for the characterization of systems like open clusters, globular clusters, bulges of galaxies, or clusters of galaxies.

The Plummer model. Plummer proposed in 1911 a spherical density model with a “soft edge” which can be described by a simple gravitational potential

$$\Phi(r) = -\frac{GM}{\sqrt{r^2 + b^2}}. \quad (3.21)$$

The corresponding density can be described by

$$\rho(r) = \frac{3M}{4\pi} \frac{b^2}{(r^2 + b^2)^{5/2}}. \quad (3.22)$$

Thus, there is a density distribution like for a homogeneous sphere for $r < b$ without a sharp edge but with a steep density fall off like $\propto r^{-5}$.

3.1.5 Spherical power law density models

Many galaxies have luminosity profiles which can be fitted with power law profiles. Therefore it seems useful to investigate spherical potentials for density distributions which can be described by a power law of the form

$$\rho(r) = \rho_0 \left(\frac{r_0}{r} \right)^\alpha. \quad (3.23)$$

The mass inside r is then

$$M(r) = 4\pi \int_0^r \rho(r') r'^2 dr' = 4\pi \rho_0 r_0^\alpha \int_0^r r'^{2-\alpha} dr' = 4\pi \rho_0 r_0^\alpha \frac{r^{3-\alpha}}{3-\alpha}.$$

We consider only $\alpha < 3$, because only for such cases the mass interior to r is finite. On the other side the mass $M(r)$ diverges for $r \rightarrow \infty$ at large radii if $\alpha < 3$. The models are still useful because according to Newton's first theorem the spherical mass shells outside r do not affect the gravitational forces and dynamics inside r .

Thus, we can derive the circular velocity v_c for the power law models and obtain

$$v_c^2(r) = \frac{GM(r)}{r} = 4\pi G \rho_0 r_0^\alpha \frac{r^{2-\alpha}}{3-\alpha}. \quad (3.24)$$

This is a very interesting formula which can be used for the interpretation of the flat rotation curves observed in disk galaxies out to very large radii. The circular velocity $v_c(r)$ is constant if $\alpha \approx 2$ or for a dark matter density distribution which behaves at radii $\gtrsim 10$ kpc like

$$\rho_{\text{dm}}(r) \propto \left(\frac{1}{r} \right)^2.$$

Two-power density models. A spherical density model combining two power laws, one approximating the flatter central region and one approximating a steeper density fall-off at larger radius provides more modelling possibilities. Well studied is a analytical parameterization for which the density is described by

$$\rho(r) = \frac{\rho_0}{(r/a)^\alpha (1+r/a)^{\beta-\alpha}} = \frac{\rho_0}{(r/a)^\alpha + (r/a)^\beta} \quad (3.25)$$

where a is a scaling radius. The α parameter is $\alpha < 3$ to avoid that the mass at small radius goes to infinity and $\beta \geq 3$ so that the total mass remains finite for large radius.

The following cases are simple and popular solutions:

- Hernquist model with $\alpha = 1$ and $\beta = 4$; this yields

$$\rho(r) \propto \frac{1}{(r/a)(1+r/a)^3}, \quad \Phi(r) \propto \frac{GM}{a+r} \quad v_c(r) = \frac{\sqrt{GMa}}{b+r}.$$

- Jaffe model with $\alpha = 2$ and $\beta = 4$,

$$\rho(r) \propto \frac{1}{(r/a)^2(1+r/a)^2}, \quad \Phi(r) \propto \frac{GM}{a} \ln(1+a/r), \quad v_c(r) = \sqrt{\frac{GM}{b+r}}.$$

- Navarro, Frenk and White or NFW model with $\alpha = 1$ and $\beta = 3$.

$$\rho(r) \propto \frac{1}{(r/a)(1+r/a)^2}, \quad \Phi(r) \propto GM \frac{\ln(1+r/a)}{r/a}.$$

3.1.6 Potentials for flattened systems

Potential of a “Toomre” disk. A simple potential for a disk was introduced by Kuzmin in 1956 and independently by Toomre in 1963. The disk potential can be described by

$$\Phi(R, z) = -\frac{GM}{\sqrt{R^2 + (|a| + |z|)^2}}. \quad (3.26)$$

According to Fig. 3.2 the potential at point $(R, -z)$ is equal to a potential generated by a mass M located at the point $(0, a)$ or for points above the disk by a mass located at $(0, -a)$.

Such a potential can be generated by a razor-thin disk with the surface density distribution

$$\Sigma(R) = \frac{aM}{2\pi(R^2 + a^2)^{3/2}}. \quad (3.27)$$

The central surface density at $R = 0$ is $M/2\pi a^2$, while the surface density drops for large R like $\Sigma(R) \approx aM/R^3$. The constant a is just a scale parameter which indicates where the surface density changes from constant to a step gradient.

Figure 3.2: Illustration of the parameters for Toomre’s disk.

A hybrid model between Toomre’s disk and the Plummer sphere We can now generalize the disk model to include also a matter distribution in z -direction. This can be achieved with a parameterization of the potential according to

$$\Phi(R, z) = -\frac{GM}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}. \quad (3.28)$$

This potential has two extreme cases:

- for $b = 0$ the potential of a thin Toomre’s disk is obtained,
- for $a = 0$ and using $R^2 + z^2 = r^2$ yields the spherical Plummer potential.

Depending on the selection of the parameters a and b one can create a family of potentials covering density distributions from a thin disk to a sphere. The corresponding mass distributions for this types of potentials are described by Miyamoto and Nagai

$$\rho(R, z) = \left(\frac{b^2 M}{4\pi}\right) \frac{aR^2 + (a + 3\sqrt{z^2 + b^2})(a + \sqrt{z^2 + b^2})^2}{[R^2 + (a + \sqrt{z^2 + b^2})^2]^{5/2}(z^2 + b^2)^{3/2}} \quad (3.29)$$

Slide 3–1 shows contour plots of this density distribution for a few parameter cases. The case $b/a \approx 0.2$ is at least qualitatively a quite good representation for a disk galaxy, while $b/a \approx 1$ resembles a S0 galaxy (e.g. Sombrero galaxy).

Potential of spheroids. Many astronomical systems are spheroidals, flattened spheres, because of the presence of angular momentum. The evaluation of potentials for spheroids in general is very difficult, because we have to consider the 2D or 3D density distribution of the system.

An important simplification is possible if we consider, homoeoids, thin concentrically nested spheroidal shells. These shells are similar to the spherical shell used for spherical systems.

One homoeid shell is bound by an inner surface and an outer surface described by

$$\frac{R^2}{a^2} + \frac{z^2}{b^2} = 1 \quad \text{and} \quad \frac{R^2}{a^2} + \frac{z^2}{b^2} = (1 + \delta\beta)^2,$$

respectively. The perpendicular distance between the two surfaces varies with position. This happens in such a way that Newton's first theorem can be generalized to spheroidal (ellipsoidal) shells.

Newton's third theorem. A mass that is inside a homoeid experiences no net gravitational force from the homoeoids.

The potential theory of spheroids was further developed in order to model with high precision the potential of the Milky Way and other galaxies. Important for these models is Newton's third theorem and the theory of multipole expansions for the gravitational potential. This theory is not discussed in this lecture. Some of the important results are:

- many potentials for flattened (oblate) spheroid and triaxial ellipsoids have been derived and applied to galaxy bulges, bars, and elliptical galaxies,
- potentials of exponential galactic disks are successfully described by strongly flattened spheroid using Newton's third theorem,
- potentials for non-axisymmetric disks can be calculated using Bessel functions, and special potential functions are used for the description of logarithmic spiral structures.

3.1.7 The potential of the Milky Way

In this subsection the potential of the Milky Way is described. In particular the density distributions of the main mass components are given: the bulge, the disk with different distributions for the stars and the interstellar gas, and the dark halo. The described model is only partly derived from studies of the dynamical properties of the Milky Way. A lot of information on the mass distribution is also derived from photometric studies. In this description the Milky Way is an axisymmetric system given in cylindrical coordinates R and z . The model picked for this description has the parameters of Model I in the book of Binney & Tremaine. This is a Milky Way model with a relatively small disk but all parameters are compatible with the available observations. Slide 3–2 shows the equipotentials for this model as well as the different components and Slide 3–3 illustrates the corresponding circular velocities $v_c(r)$.

The central bulge. The bulge can be described by a oblate, spheroidal power law model which is truncated at an outer radius r_b :

$$\rho_b(R, z) = \rho_{b0} \left(\frac{m}{a_b} \right)^{\alpha_b} e^{-m^2/r_b^2}, \quad (3.30)$$

with

$$m = \sqrt{R^2 + z^2/q_b^2}.$$

The parameters describe:

- $\rho_{b0} = 0.43 \text{ M}_\odot/\text{pc}^3$ is the density normalization for the bulge
- $a_b = 1 \text{ kpc}$ is the size normalization of the bulge,
- $q_b = 0.6$ describes the bulge flattening,
- $\alpha_b = -1.8$ is the power law index for the density distribution,
- r_b is the cut-off radius for the bulge.

The galactic disk. The Milky way disk consists of the stellar disk and a gas disks.

The stellar disk is described by an exponential fall-off in radial direction and two exponential laws for the vertical direction, one for the thin disk and one for the thick disks. The used formula is

$$\rho_s(R, z) = \Sigma_S e^{-R/R_S} \left(\frac{\alpha_0}{2z_0} e^{-|z|/z_0} + \frac{\alpha_1}{2z_1} e^{-|z|/z_1} \right). \quad (3.31)$$

The parameters describe:

- $\Sigma_S \approx 1500 \text{ M}_\odot/\text{pc}^2$ is the central surface density of the stellar disk which is not well known except for the solar radius R_0 . At R_0 the surface density of the stars is about $35 \text{ M}_\odot/\text{pc}^2$ to which the thick disk contributes about $3 \text{ M}_\odot/\text{pc}^2$.
- $R_S = 2.5 \text{ kpc}$ is the disk scale length,
- $\alpha_0 = 0.9$ and $\alpha_1 = 0.1$ are the relative normalizations of the thin and thick disk,
- $z_0 = 0.3 \text{ kpc}$ is the vertical scale height of the thin disk,
- $z_1 = 1 \text{ kpc}$ is the scale height of the thick disk.

The radial distribution of the interstellar gas disk is also described with an exponential law, but with a much larger scale length than for the stars. In addition, there is a hole with a radius of about 4 kpc in the center which is considered with an exponential cut-off. The vertical density distribution of the gas is much narrower than for the stars:

$$\rho_g(R, z) = \Sigma_g e^{-R/R_g} e^{-R_m/R} \frac{1}{2z_g} e^{-|z|/z_g}. \quad (3.32)$$

where the parameters are:

- $\Sigma_g \approx 500 \text{ M}_\odot/\text{pc}^2$, the surface density of the gas in the disk is not well known except for R_0 where the surface density is about $\Sigma_g(R_0) \approx 12 \text{ M}_\odot/\text{pc}^2$
- $R_g = 4 \text{ kpc}$ is the disk scale length for the gas (twice the value when compared to the stellar disk),
- $R_m = 4 \text{ kpc}$ is the radius of the inner hole,
- $z_g = 80 \text{ pc}$ is the scale height of the gas disk.

The dark halo. The dark halo can be described by an extension of the spherical two-power-law model to an oblate geometry.

$$\rho_h(R, z) = \frac{\rho_{h0}}{(m/a_h)^{\alpha_h} (1 + m/a_h)^{\beta_h - \alpha_h}} \quad (3.33)$$

where the flattening is described like for the bulge

$$m = \sqrt{R^2 + z^2/q_h^2}.$$

The parameters describe:

- $\rho_{h0} = 0.71 \text{M}_\odot/\text{pc}^3$ is the density normalization for the halo,
- $a_h = 3.8 \text{ kpc}$ is the size normalization for the halo,
- $q_h = 0.8$ is a guess for the possible flattening of the dark halo,
- $\alpha_h = 2.0$ and $\beta_h = 3$ are the power law indices for the halo density distribution.

3.2 The motion of stars in spherical potentials

This section discusses the orbits of individual stars in a static, spherical potential. Spherical potentials serve again as simple cases for the description of general principles.

3.2.1 Orbits in a static spherical potential

Spherical potentials describe very well globular cluster but less well flattened or triaxial systems like galaxies. Nonetheless the solutions for spherical potentials serve as very important guide for more complicated gravitational fields.

In a centrally directed gravitational field the position vector of a star is

$$\vec{r} = r\vec{e}_r$$

The motion of a star with a mass m_S in a spherical potential is defined by the radially directed gravitational force

$$\vec{F}(r) = m_S \frac{d^2 \vec{r}}{dt^2} = m_S g(r) \vec{e}_r.$$

Further, we know that the angular momentum in a static spherical system is constant

$$\vec{L} = m_s \vec{r} \times \frac{d\vec{r}}{dt} = \text{const.}$$

This implies that the stars moves in a plane. For this reason we can use plane polar coordinates.

Lagrange function. We introduce the Lagrange-function, which is a more general formulation for the equations of motions. The Lagrange-function for a star in free space is in Cartesian coordinates

$$\mathcal{L} = \frac{m_S}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2),$$

and in polar coordinates

$$\mathcal{L} = \frac{m_S}{2}(\dot{r}^2 + r^2\dot{\phi}^2 + \dot{z}^2).$$

The Lagrange-function for a mass m_S in a spherical potential $\Phi(r)$ can then be written as

$$\mathcal{L} = \frac{m_S}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - m_S\Phi(r). \quad (3.34)$$

We use polar coordinates because we can align the spherical coordinate system always with the orbital plane (where $\theta = 0$ and $\dot{z} = 0$).

Equation of motion. The equations of motions follow from the derivatives of the Lagrange equation

$$0 = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{r}} - \frac{\partial \mathcal{L}}{\partial r} = m_S \ddot{r} - m_S r \dot{\phi}^2 + m_S \frac{d\Phi}{dr}, \quad (3.35)$$

$$0 = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\partial \mathcal{L}}{\partial \phi} = \frac{d}{dt}(m_S r^2 \dot{\phi}). \quad (3.36)$$

The second equation is the formulation of the angular momentum conservation in polar coordinates

$$L = m_S r^2 \dot{\phi} = \text{const.}$$

With the angular momentum equation we can substitute the time derivative by the angle derivative

$$\frac{d}{dt} = \frac{L^2}{m_S r^2} \frac{d}{d\phi},$$

and this yields the equation of motion in the following form:

$$\frac{L^2}{r^2} \frac{d}{d\phi} \left(\frac{1}{r^2} \frac{dr}{d\phi} \right) = - \frac{d\Phi}{dr}. \quad (3.37)$$

With the substitution $u = 1/r$ a simplified form for the equation of motion is obtained:

$$\frac{du^2}{d\phi^2} + u = \frac{1}{L^2 u^2} \frac{d\Phi}{dr}(1/u). \quad (3.38)$$

Energy equation. We can write for a mass in a central potential the following energy equation

$$E_{\text{tot}} = \frac{m_S \dot{r}^2}{2} + \frac{L^2}{2m_S r^2} + m_S \Phi(r). \quad (3.39)$$

This provides very convenient formula for the motion of particles in a centrally symmetric gravitational field.

Further we can use for a stationary gravitational potential the **virial theorem**

$$2E_{\text{kin}} + E_{\text{pot}} = 0,$$

where $\Phi(r) = E_{\text{pot}}$ for the star in the central potential.

Effective Potential. The energy equation (3.39) shows that the radial motion can be described as 1-dimensional motion in an effective radial potential of the form

$$\Phi_{\text{eff}}(r) = \Phi(r) + \frac{L^2}{2m_S r^2}. \quad (3.40)$$

This potential goes, except for the case $L = 0$, for $r \rightarrow 0$ to infinity and for $r \rightarrow \infty$ from negative values to zero.

The potential has for small radii a centrifugal barrier, if $L \neq 0$. The r -values, where the total energy is equal to the effective potential energy, define the radial range of motion:

$$\frac{m\dot{r}^2}{2} = E - \Phi_{\text{eff}}. \quad (3.41)$$

The borders of this range are defined by the radius where the radial kinetic energy is zero or where $\dot{r} = 0$. At these points the total energy is equal to the effective potential energy. For bound orbits and $L \neq 0$ this equation has two roots r_1 and r_2 which are called the pericenter and apocenter distances, respectively.

Figure 3.3: Radial dependence of the effective potential energy for potentials with different angular momentum.

The different curves in Fig. 3.3 illustrate what happens if the total energy or the angular momentum is changed in the system. A change in angular momentum is equal to a jump to a different effective potential energy curve and a change in energy enhances or lowers the eccentricity. A dynamical interaction between two stars changes typically both, the total energy and the angular momentum.

The radial dependence of the effective potential energy is similar for essentially all gravitating systems. For small separation there is the centrifugal force barrier, in the intermediate range is the minimum of the potential energy, and for large separations the effective potential energy approaches zero.

3.2.2 Radial and azimuthal velocity component.

The motion in r and ϕ can be derived from the energy equation. The equation for the radial velocity component is

$$\dot{r} = \frac{dr}{dt} = \sqrt{\frac{2}{m_S} [E - \Phi(r)] - \frac{L^2}{m_S^2 r^2}}, \quad (3.42)$$

with the time dependence

$$t(r) = \int \frac{dt}{dr} dr = \int \frac{dr}{\sqrt{\frac{2}{m_S} [E - \Phi(r)] - \frac{L^2}{m_S^2 r^2}}} + \text{const.}, \quad (3.43)$$

and using the definition for the angular momentum $L = m_S r^2 \dot{\phi}$ or $d\phi = L/m_S r^2 dt$ yields the equation for the azimuthal velocity component (using $d\phi/dr = (d\phi/dt)(dt/dr)$)

$$\phi(r) = \int \frac{d\phi}{dr} dr = \int \frac{\frac{L}{r^2} dr}{\sqrt{2m_S [E - \Phi(r)] - \frac{L^2}{r^2}}} + \text{const.} \quad (3.44)$$

Figure 3.4: Typical orbit of a star in a spherical potential.

The radial period T_r is the time required for the star m_S to travel from apocenter to pericenter and back. This is:

$$T_r = 2 \int_{r_1}^{r_2} \frac{dr}{\sqrt{\frac{2}{m_S} [E - \Phi_{\text{eff}}(r)]}}. \quad (3.45)$$

Similarly, one can derive the azimuthal angle increase $\Delta\phi$ from pericenter to apocenter and back, which is

$$\Delta\phi = 2L \int_{r_1}^{r_2} \frac{dr}{r^2 \sqrt{\frac{2}{m_S} [E - \Phi_{\text{eff}}(r)]}}.$$

The azimuthal period is then

$$T_\phi = \frac{2\pi}{\Delta\phi} T_r, \quad (3.46)$$

or the mean azimuthal speed is equal to $2\pi/T_\phi$. The orbit will only be closed if $2\pi/\Delta\phi$ is a rational number, what is typically not the case except for the potential of a point source and a homogeneous sphere. The star moves therefore in general on a rosette around the center of the spherical potential (Fig. 3.4).

3.2.3 Motion in a Kepler potential

Effective potential. The effective potential energy for a point source is

$$\Phi_{\text{eff}}(r) = -\frac{GM}{r} + \frac{L^2}{2m_s r^2}. \quad (3.47)$$

The equation

$$\frac{d\Phi_{\text{eff}}(r)}{dr} = \frac{GM}{r^2} - \frac{L^2}{m_s r^3} = 0$$

provides the radius of the minimum

$$r_{\text{min}} = \frac{L^2}{GMm_s} \quad (3.48)$$

and the corresponding minimum effective potential energy

$$\min(\Phi_{\text{eff}}(r)) = -\frac{G^2 M^2 m_s}{2L^2}. \quad (3.49)$$

The total energy is for a given angular momentum equal or larger to

$$E_{\text{tot}} \geq \frac{L^2}{2m_s r_{\text{min}}^2} + \Phi(r_{\text{min}}) = \frac{G^2 M^2 m_s}{2L^2} - \frac{G^2 M^2 m_s}{L^2} = -\frac{G^2 M^2 m_s}{2L^2}.$$

For $E_{\text{tot}} = \min(\Phi_{\text{eff}}(r))$ we have a circular orbit with no radial motion component. In this case the angular momentum energy term is half the potential energy term. This is as predicted by the virial theorem for a system in gravitational equilibrium:

$$2E_{\text{kin}} + E_{\text{pot}} = 0.$$

The circular orbit is a minimum energy orbit for an object with a given angular momentum moving in a spherical potential.

Orbital periodicities in a Kepler potential. The motion in a Kepler potential can be derived from the equation of motion described in Equation 3.38.

We know from the first and third Kepler law that the orbits are closed:

$$T_r = T_\phi$$

and that the orbital period or radial oscillation period is

$$T_r^2 = 2\pi \frac{a^3}{GM}.$$

The Keplerian motion has the following properties:

- the mass m_s moves on closed ellipses with the point source in one focal point,
- according to the angular momentum conservation, the azimuthal velocity during an orbit behaves like

$$\frac{d\phi(r)}{dt} \propto 1/r.$$

3.2.4 Motion in the potential of a homogeneous sphere

According to Section the potential at $r < a$ inside a sphere with radius a is

$$\Phi(r) = -2\pi G\rho a^2 + \frac{2\pi G\rho}{3}r^2 = \frac{\omega^2}{2}r^2 + \text{const.},$$

with $\omega^2 = 4\pi G\rho/3$. The equation of motion $m_S\ddot{r} = m_S(d\Phi/dr)$ can be written in Cartesian coordinates $x = r \cos\phi$ and $y = r \sin\phi$:

$$\ddot{x} = -\omega^2 x, \quad \ddot{y} = -\omega^2 y, \quad (3.50)$$

with the solutions:

$$x = a_x \cos(\omega t + \delta_x), \quad y = a_y \cos(\omega t + \delta_y). \quad (3.51)$$

where a_x, a_y, δ_x and δ_y are arbitrary constants. The motion has the following properties:

- x and y have the oscillation period $T_r = 2\pi/\omega$,
- the oscillation phase in the x and y directions are independent,
- the mass m_S moves on closed ellipses which are centered on the center of the sphere $r = 0$,
- the radial period is half the orbital period, or an object completes two in-and-out oscillations during an orbital period:

$$T_r = \frac{1}{2}T_\phi. \quad (3.52)$$

If the x - and y -oscillations are in phase, then the motion corresponds to a swing from one side of the center to the other side and back along a straight line with a full oscillation period identical to the orbital period. However, for a radial coordinate system this corresponds to two full oscillation $r_{\max} - 0 - r_{\max} - 0 - r_{\max}$.

Figure 3.5: Qualitative illustration of the ellipse shape of a mass in a Kepler potential and a mass inside a homogeneous sphere.

Figure 3.5 illustrates the fundamental difference between the orbits in a homogeneous sphere and around a point source. All smooths potentials create orbits which have typically about two radial oscillation per azimuthal period.

3.3 Motion in axisymmetric potentials

Stars moving in the equatorial plane of an axisymmetric potential behave like stars in a spherical potential, because one can always find a spherical gravitational potential which induces the same gravitational force on the stars in the equatorial plane as the axisymmetric potential. For this reason the orbits discussed in the previous chapter for spherical potentials apply also for the stars in the equatorial plane of an axisymmetric potential.

The motion of the stars located in or near the equatorial plane is an important problem for the investigation of disk galaxies.

3.3.1 Motion in the meridional plane

We assume that the potential is symmetric with respect to the plane $z = 0$. Then we can write the Lagrange equation with the following terms

$$\mathcal{L} = \frac{m_S}{2}(\dot{R}^2 + R^2\dot{\phi}^2 + \dot{z}^2) - m_S\Phi(R, z)$$

The 3-dimensional motion of a star in an axisymmetric potential can be reduced to a 2-dimensional motion of a star in the R - z -plane, the meridional plane.

The equation of motion in this co-rotating plane are:

$$\ddot{R} = -\frac{\partial\Phi_{\text{eff}}(R, z)}{\partial R}, \quad \ddot{z} = -\frac{\partial\Phi_{\text{eff}}(R, z)}{\partial z}, \quad (3.53)$$

where the effective potential is

$$\Phi_{\text{eff}}(R, z) = \Phi(R, z) + \frac{L_z^2}{2m_S R^2} \quad (3.54)$$

Similar to the spherical case we can write the total energy equation, but now with an R and a z term for the kinetic energy

$$E_{\text{tot}} = \frac{1}{2m_S}(p_R^2 + p_z^2) + \Phi_{\text{eff}}(R, z). \quad (3.55)$$

The kinetic energy of motion in the R - z -plane is

$$\frac{1}{2m_S}(p_R^2 + p_z^2) = E_{\text{tot}} - \Phi_{\text{eff}}(R, z).$$

Orbits in the meridional plane are restricted to the area $E_{\text{tot}} > \Phi_{\text{eff}}(R, z)$ and one can define contour lines or the zero velocity curves in the meridional plane where the kinetic energy term is instantaneously zero

$$\Phi_{\text{eff}}(R, z) = E_{\text{tot}}.$$

The minimum of Φ_{eff} is in the equatorial plane $z = 0$ and the radial value follows from

$$0 = \frac{\partial\Phi_{\text{eff}}}{\partial R} = \frac{\partial\Phi}{\partial R} - \frac{L_z^2}{m_S R^3}$$

This yields the radius for a circular orbit with angular speed $\dot{\phi}$ which is identical to the radius of the minimum of the effective potential. At this radial point, which is called the **guiding-center radius** R_g , there is

$$\left(\frac{\partial\Phi}{\partial R}\right)_{(R_g,0)} = \frac{L_z^2}{m_S R^3} = m_S R_g \dot{\phi}^2,$$

($L_z = m_S R^2 \dot{\phi}$). This is the condition for a circular orbit with angular speed $\dot{\phi}$ for a mass located at the radius R_g , which is at the minimum of the effective potential.

Example. Slide 3-4 shows as example the contour plot and orbits for the effective potential

$$\Phi_{\text{eff}}(R, z) = \frac{v_0}{2m_S} \ln\left(\frac{R^2 + z^2}{q^2}\right) + \frac{L_z^2}{2m_S R^2},$$

for $v_0 = 1$, $L_z = 0.2$ and axial ratio $q = 0.9$ and 0.5 . This represents the effective potential for an oblate, spheroidal mass distribution like a central bulge of a disk galaxy, an elliptical galaxy, or a dark matter halo with a constant circular velocity speed $v_c = \text{const.}$ The effective potential energy rises strongly near $R = 0$ because of the “centrifugal barrier” for the given angular momentum L_z .

The equations (3.53) for the relative motion in a co-rotating frame must be integrated numerically. Slide 3-4 shows the calculated motion. The given results are for stars in the same potential, same energy and same angular momentum but they still differ significantly. This problem is called the third integral problem and it is linked in this case to the precession of the angular momentum vector in a flattened potential.

3.3.2 Nearly circular orbits: epicycle approximation

In disk galaxies many stars are on nearly circular orbits. For this case we can simplify the equation of motion in the co-rotating system (Eq. 3.53)

$$\ddot{R} = -\frac{\partial\Phi_{\text{eff}}(R, z)}{\partial R}, \quad \ddot{z} = -\frac{\partial\Phi_{\text{eff}}(R, z)}{\partial z},$$

with a linearization of the corresponding effective potential at $R = R_g$ and $z = 0$. We introduce x as new variable in the radial direction

$$x = R - R_g$$

The effective potential in Eq. 3.54 can then be written as Taylor expansion:

$$\Phi_{\text{eff}} = \Phi_{\text{eff}}(R_g, 0) + \frac{1}{2} \left(\frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2} \right)_{(R_g, 0)} x^2 + \frac{1}{2} \left(\frac{\partial^2 \Phi_{\text{eff}}}{\partial z^2} \right)_{(R_g, 0)} z^2 + O(xz^2) + \dots \quad (3.56)$$

The first order terms are zero because $\Phi_{\text{eff}}(R_g, 0)$ is at a minimum. One can introduce abbreviations for the second derivatives (curvature in the effective potential):

$$\kappa^2(R_g) = \left(\frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2} \right)_{(R_g, 0)}, \quad \text{and} \quad \nu^2(R_g) = \left(\frac{\partial^2 \Phi_{\text{eff}}}{\partial z^2} \right)_{(R_g, 0)}.$$

This approximation, which is called the **epicycle approximation**, yields very simple, harmonic, equations of motions for the radial x and vertical z directions:

$$\ddot{x} = -\kappa^2 x, \quad \ddot{z} = -\nu^2 z. \quad (3.57)$$

The two time scales are called:

- the epicycle or radial frequency κ ,
- the vertical frequency ν .

These frequencies can be evaluated using Eq. 3.54 for the effective potential in a co-rotating system

$$\Phi_{\text{eff}}(R, z) = \Phi(R, z) + \frac{L_z^2}{2m_S R^2}.$$

This yields for the vertical frequency the simple relation

$$\nu^2(R_g) = \left(\frac{\partial^2 \Phi}{\partial z^2} \right)_{(R_g, 0)} \quad (3.58)$$

Solution for the epicycle frequency. There are two terms for the epicycle frequency κ , a potential energy term and an angular momentum term (follows from $\partial^2/\partial R^2(L_z^2/2m_S R^2)$)

$$\kappa^2(R_g) = \left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{(R_g, 0)} + \frac{3L_z^2}{m_S R^4}. \quad (3.59)$$

We can now use the “global” angular velocity dependence for the circular motion at R_g which is (using also $v_c^2 = R(\partial\Phi/\partial R) = L_z^2/m_S^2 R^2$)

$$\Omega^2(R) = \frac{v_c^2(R)}{R^2} = \frac{1}{R} \left(\frac{\partial \Phi}{\partial R} \right)_{(R_g, 0)} = \frac{L_z^2}{m_S^2 R^4}$$

With this relation we can rewrite the equation for the epicycle frequency in terms of global, “galactic”, quantities:

$$\kappa^2(R_g) = \left(R \frac{d\Omega^2}{dR} + 4\Omega^2 \right)_{R_g} \quad (3.60)$$

using $d^2\Phi/dR^2 = d/dR(R\Omega^2) = \Omega^2 + R(d\Omega^2/dR)$. This relates the epicycle frequency to the radial dependence of the angular velocity $d\Omega^2(R)/dR$.

Comparison of epicycle period and orbital period. We can now compare the epicycle period T_r with the azimuthal orbital period T_{orb} which are simply:

$$T_r = \frac{2\pi}{\kappa} \quad \text{and} \quad T_\phi = \frac{2\pi}{\Omega}.$$

There are three useful approximate cases for a comparison between orbital frequency and epicycle frequency:

- Near the center of galaxies the circular speed v_c increases linearly and $\Omega(R)$ is essentially constant and therefore $d\Omega^2/dR \approx 0$. In this case there is

$$\kappa^2(R_g) \approx 4\Omega^2 \quad \text{or} \quad \kappa \approx 2\Omega,$$

This corresponds to the limiting case of a homogeneous sphere where the epicycle frequency is twice the orbital frequency, or the radial period is half the orbital period $T_r = T_\phi/2$.

- At large radii from the center the circular velocity falls off like (but usually less rapid) the Kepler law. For a Kepler law there is $\Omega \approx R^{-3/2}$ (using $R(d\Omega^2(R)/dR) = -3\Omega^2$). This limit implies

$$\kappa^2(R_g) \gtrsim \Omega^2 \quad \text{or} \quad \kappa \gtrsim \Omega.$$

Thus we have the case where the radial period and orbital periods are equal or $T_r = T_\phi$. This is as expected for a closed Keplerian orbit.

- At most points in a typical disk galaxy the circular velocity is constant or $\Omega \propto R^{-1}$. For this case the formula for the epicycle frequency is

$$\kappa^2 = 2\Omega^2 \quad \text{or} \quad \kappa \approx 1.4\Omega.$$

This indicates that in a disk the stars oscillate radially with a frequency of roughly 1.4 times the orbital frequency.

Application for the solar neighborhood. The third case, for intermediate separations, can be evaluated in detail for the solar neighborhood. As described in Chapter 2, we know quite well the Oort's constants A and B from the measurement of the radial and tangential velocities of stars in the solar neighborhood. We used the following formula for the Oort's constants:

$$A = \frac{1}{2} \left[\frac{\Theta_0}{R_0} - \left(\frac{d\Theta}{dR} \right)_{R_0} \right] \quad \text{and} \quad B = -\frac{1}{2} \left[\frac{\Theta_0}{R_0} + \left(\frac{d\Theta}{dR} \right)_{R_0} \right].$$

With $\Theta = R\Omega$ and $d\Theta/dR = d/dR(R\Omega) = \Omega + R(d\Omega/dR)$ we can write:

$$A = -\frac{1}{2}R\frac{d\Omega}{dR} \quad \text{and} \quad B = -\left(\Omega + \frac{1}{2}R\frac{d\Omega}{dR} \right)$$

The circular angular velocity is $\Omega = A - B$ while the epicycle frequency is

$$\kappa^2 = -4B(A - B) = -4B\Omega$$

or $\kappa^2/\Omega^2 = -4B/\Omega$, which yield the ratio between epicycle frequency and orbital period for the solar neighborhood

$$\frac{\kappa_0}{\Omega_0} = 2\sqrt{\frac{-B}{A - B}} \approx 1.3 \pm 0.1. \quad (3.61)$$

The result is obtained for the typical values for the Oort's constants $A \approx +15$ km/(s kpc) and $B \approx -12$ km/(s kpc). This means that the sun makes about 1.3 oscillations in radial directions within one orbit around the galactic center.

3.3.3 Density waves and resonances in disks

In the previous subsection we have introduced the following quantities for stars with almost circular orbits in disk galaxies:

- T_r : the epicycle or radial period for the in-and-out motion in radial direction,
- $\Delta\phi$: the azimuthal angle increase during the epicycle period,
- $\Omega_r = 2\pi/T_r$: the radial oscillation frequency,
- $\Omega_\phi = \Delta\phi/T_r$: the corresponding azimuthal oscillation frequency,
- $\Omega = 2\pi/T$: the orbital frequency or orbital angular velocity where T is the time for a full orbit around the galaxy center.

We now describe the motion of the stars in a frame which is rotating with some special angular velocity. The following quantities are defined in this system:

- Ω_P : angular velocity (or pattern speed) for the selected rotating frame,
- $\phi_p = \phi - \Omega_P t$: the azimuthal angle in the rotating reference system which changes with time,
- $\Delta\phi_p = \Delta\phi - \Omega_P T_r$: the azimuthal angle increase in the rotating system for one epicycle period.

One can always choose an angular velocity Ω_p for a rotating coordinate system in which the orbits are closed or $\Delta\phi_p/T_r = (n/m)\Omega_r$. This follows from the definition of $\Delta\phi_p$

$$\Omega_p = \frac{\Delta\phi}{T_r} - \frac{n}{m}\Omega_r. \quad (3.62)$$

For orbits close to circular orbits we can approximate $\Delta\phi/T_r = \Omega_\phi \approx \Omega$ and $\kappa \approx \Omega_r$ and write

$$\Omega_p = \Omega - \frac{n}{m}\kappa. \quad (3.63)$$

Figure 3.6 illustrates the appearance of an orbit with $\kappa/\Omega_r \approx 1.5$ in rotating frames with different m and n .

Figure 3.6: Closed orbits with different n and m in a rotating system.

In general, $\Omega - n\kappa/m$ is a function of radius, and no unique pattern speed Ω_p can be defined to close the orbits at all radii. Slide 3-5 shows curves for $\Omega - n\kappa/m$ for the Milky Way (model 1).

However, it was first noticed by Lindblad that the curve for $\Omega - \kappa/2$ is relatively constant for a wide range of galactic radii. A constant curve $\Omega - \kappa/2$ would mean that in a frame rotating at Ω_p all orbits would be ellipses, which are nested for a broad range of R . They would look like the ellipses shown in Slide 3-6. If stars move predominantly along these ellipses then they would create a **bar-like pattern**, which is stationary in a rotating frame. In a fixed frame this would then look like a **density wave** rotating with a pattern speed Ω_p .

In a real galaxy $\Omega - n\kappa/m$ depends on radius. Therefore, independent of the selected Ω_p , most orbits will not be exactly closed. The orientations at different radii will drift at slightly different speeds, and the pattern will twist, and might look like a spiral pattern (see Slide 3-6). This type of kinematic density waves, produce a non-axisymmetric disk pattern and an exact calculation of stellar orbits needs to take this into account.

Resonances. Resonances in a rotating disk occur if the circular frequency Ω and/or the epicycle frequency κ are in phase for different radii. The following resonances are distinguished:

- **Corotation resonance:** In this case the pattern frequency is equal to the circular frequency

$$\Omega_p = \Omega.$$

This resonance can be induced by an interaction with another galaxy or a strong density asymmetry in the disk.

- **Lindblad resonances:** They occur if the difference between the pattern and the orbital frequency is an integer of the epicycle frequency.

$$m(\Omega - \Omega_p) = \pm\kappa.$$

In this case a star encounters a weak perturbation in the potential in phase with its radial in-and-out motion. For this reason already a weak perturbation can build up a strong effect within several cycles. One distinguishes between inner $m(\Omega - \Omega_p) = +\kappa$ and outer $m(\Omega - \Omega_p) = -\kappa$ Lindblad resonances.

In a spiral galaxy the Lindblad resonances define the inner and outer boundaries of the spiral pattern.

- the inner Lindblad resonance is where the spiral structure starts. The stars are on elliptical orbits centered on the galactic center and they move faster than the disturbing spiral pattern. This happens in the Milky Way around 3 kpc.
- the outer Lindblad resonance is at the outer borders of the galaxy, where the spiral pattern ends. There, the orbits are again elliptical with the center in the middle. The disturbing spiral pattern is faster than the stars.

3.4 Two-body interactions and system relaxation

Up to now we have assumed that collisions, ie the interaction between individual stars, can be neglected. This is a reasonable assumption for galactic dynamics. We discuss now cases where such collisions between stars play an important role.

A star within a more or less homogeneous distribution of other stars “feels” the gravitational force of all these stars. The force $F = \Sigma F_i$ of all stars i in a given solid angle (see Fig. 3.7) behaves as follows:

- the force induced by an individual star is $F_i \sim 1/r_i^2$ and decreases with distance,
- the volume and therefore the number of stars in a fractional distance interval, e.g. $[r - r/2, r + r/2]$ increases like $\sim r^3$,
- the total force on the sample star is dominated by the more distant stars.

Figure 3.7: On the force induced by near and distant stars in a homogeneous distribution.

Therefore it is reasonable to assume that stars are smoothly accelerated by the force field that is generated by the Milky Way as a whole. In the following we investigate more quantitatively this simplification and consider cases where this approximation is no more valid.

3.4.1 Two-body interaction

We consider an individual star, called the **the subject star**, and investigate how much its velocity is disturbed by encounters with other stars, called two-body interactions, during the crossing through a system like a galaxy, or a star cluster. Thus we calculate the expected deflection of the trajectory of the subject star from the path it would have in the smooth overall potential. For our estimate we assume that all stars have the same mass m_S .

The velocity deflection $\delta\vec{v}$ induced by a two-body interaction can be simplified, if we consider only weak (distant) encounters which introduce small velocity deflections $|\delta\vec{v}|/v \ll 1$. Further it is assumed that the field star is stationary during the encounter. The velocity deflection follows then the perpendicular force $F_\perp$ of the field star induced onto the subject star which is moving with velocity v along an essentially straight line with impact parameter b (Fig. 3.8).

Figure 3.8: Geometry for an estimate of the deflection by a star-star interaction.

If both stars have the same mass, then the perpendicular force induced on the subject star is:

$$F_{\perp} \approx \frac{Gm_S^2}{b^2 + x^2} \cos\theta = \frac{Gm_S^2 b}{(b^2 + x^2)^{3/2}}, \quad (3.64)$$

using the trigonometric relation $\cos\theta = b/r = b/\sqrt{b^2 + x^2}$. The coordinate along the trajectory x can be expressed by the time and the velocity of the subject star $x = v \cdot t$ so that

$$F_{\perp} \approx \frac{Gm_S^2}{b} \left[1 + \left(\frac{vt}{b} \right)^2 \right]^{-3/2}.$$

According to Newton's law the acceleration or change in velocity $\dot{\vec{v}} = \vec{F}/m_S$ is the time integral of the acting force, or

$$\delta v \approx \int_{-\infty}^{+\infty} F_{\perp} dt = \frac{Gm_S}{b^2} \int_{-\infty}^{+\infty} \frac{1}{[1 + (vt/b)^2]^{-3/2}} dt \quad (3.65)$$

The integral is equal to $2b/v$ and the deflection is

$$\delta v \approx \frac{2Gm_S}{bv}. \quad (3.66)$$

This equation can be interpreted as follows:

- δv is proportional to the acceleration at closest approach Gm_S/b^2 times a characteristic duration of the acceleration $2b/v$,
- the derived approximative value is only valid for $\delta v \ll v$ or for an impact parameter larger than

$$b \gg Gm_S/v^2 = 900 \text{AU} \frac{(m_S/M_{\odot})}{(v/(\text{km/s}))^2}.$$

As next step we estimate the number of encounters of the subject star in a stellar system for the impact parameter range $[b, b + db]$. We use an estimate for the surface density of field stars

$$\Sigma_{\text{stars}} \approx \frac{N}{\pi R^2},$$

where N is the total number of stars and R the radius of the considered system, e.g. the stellar cluster or galaxy. The subject star will have during one crossing of the system the following number of encounters

$$\delta n = \frac{N}{\pi R^2} 2\pi b db = \frac{2N}{R^2} b db. \quad (3.67)$$

with impact parameter between b and $b + db$. Each such encounter produces a deflection $\delta \vec{v}$ to the subject star, but these deflections are randomly oriented and their mean will be zero. But the mean-square change will not be zero and after one crossing. The squared velocity deflection (change) for an impact parameter intervall db will be:

$$\Sigma \delta v^2 db \approx \delta v^2 \delta n db = \left(\frac{2Gm_S}{bv} \right)^2 \frac{2N}{R^2} b db. \quad (3.68)$$

Now, we have to take into account all impact parameters by integrating from $b_{\min}$ to $b_{\max}$

$$\Delta v^2 = \int_{b_{\min}}^{b_{\max}} \Sigma \delta v^2 db = 8N \left(\frac{Gm_S}{Rv} \right)^2 \ln b \Big|_{b_{\min}}^{b_{\max}}. \quad (3.69)$$

The logarithm term can be written as

$$\ln \Lambda = \ln b_{\max} - \ln b_{\min}.$$

The maximum impact parameter is of the order $b_{\max} \approx R$, the smallest, where the small deflection approximation is still valid, is $b_{\min} \approx 2Gm_S/v^2$. These are only approximate values with an uncertainty of a factor of a few. For this reason we can write

$$\ln \Lambda = \ln \left(\frac{R}{b_{\min}} \right) + \ln(\epsilon_1/\epsilon_2).$$

In most systems $R \gg b_{\min}$ and the typical ratio is $R/b_{\min} \gg 10^4$ while the uncertainty term ϵ_1/ϵ_2 is much smaller, of the order of a few. This term can therefore be neglected with respect to the first term. Thus, the parameter Λ is approximately

$$\Lambda \approx \frac{Rv^2}{2Gm_S} \approx N,$$

where we already used the next approximation for the typical stellar velocity v .

The encounters between the subject star and the field stars produce a diffusion of the star's velocity which is different from an acceleration induced by a smooth, large scale potential. This velocity diffusion is called **two-body relaxation**, because it is the result of a large number of mostly weak two-body interactions.

The typical speed v of the a field star can be approximated by the circular velocity of a star at radius R (at the edge) of the system

$$v^2 \approx \frac{GNm_S}{R}.$$

Equation 3.69 can be simplified with this velocity to

$$\frac{\Delta v^2}{v^2} \approx \frac{8 \ln \Lambda}{N}, \quad (3.70)$$

The subject stars makes typically many crossings until the velocity $\vec{v}$ changes by roughly Δv^2 . The number of crossings n_{relax} required for a change of the velocity by a value comparable to v is then

$$n_{\text{relax}} \approx \frac{N}{8 \ln \Lambda} \approx \frac{N}{8 \ln N}.$$

3.4.2 Relaxation time

The relaxation time is defined as $t_{\text{relax}} = n_{\text{relax}} t_{\text{cross}}$, where the crossing time is $t_{\text{cross}} = R/v$. Using all the approximations from above we can express the relaxation time by the number of stars and the crossing time

$$t_{\text{relax}} \approx \frac{0.1N}{\ln N} t_{\text{cross}}. \quad (3.71)$$

Thus the relaxation time exceeds the crossing time in a self-gravitating system for $N \gtrsim 40$.

After the relaxation time the orbit of a (subject) star is changed significantly by all the small kicks induced by other stars, so that its velocity is now different than from the what one would expect in a smooth potential.

Table 3.1: Typical characteristic parameters for stellar systems

system	R	N	t_{cross}	t_{relax}	t_{lifetime}
clusters of galaxies	1 Mpc	1000	1 Gyr	14 Gyr	10 Gyr
galaxies	10 kpc	10^{11}	100 Myr	$\gg 100$ Gyr	10 Gyr
central pc of galaxies	1 pc	10^6	10^4 yr	100 Myr	10 Gyr
globular clusters	10 pc	10^5	10^5 yr	100 Myr	10 Gyr
open clusters	10 pc	100	1 Myr	10 Myr	100 Myr

Table 3.1 gives typical numbers for the different stellar systems using these approximations. The numbers show that the relaxation times are extremely long for galaxies. Therefore they can be treated as collision-less systems. Important to notice is, that the dynamics of stars in galaxies preserve at least partly information from past events.

On the other side there are the stellar clusters. Globular clusters relax in about 100 Myr and open clusters on a very short timescale of about 1 Myr. Their dynamics is dominated by relaxation and the star motions are rapidly randomized. For this reason it is often not possible to extract the past history of clusters from dynamical studies.

3.4.3 The dynamical evolution of stellar clusters

Galactic stellar clusters have a very short relaxation time. A disturbance of their dynamics is therefore rapidly randomized. In addition, galactic stellar clusters are also quite fragile and dissolve rapidly, typically within about 300 Myr. On the other side there are the globular clusters which have survived more than 10 Gyr. For these systems we have only

little information about their formation history and all signatures from the formation process in the stellar dynamics has been washed out.

Some important dynamical processes for the evolution of stellar cluster can be inferred from their current properties:

For globular clusters, we know that

- they have a long life time of 10 Gyr or more,
- they have typically $N \approx 10^5$ to 10^6 stars,
- many globular clusters show a dense core and a low density halo,
- there are often “hard binary systems” in the center.

For galactic open cluster, we know that

- they have typically 100 - 1000 star members,
- they dissolve in about 300 Myr,
- they show often a mass segregation with more massive stars in the center and lower mass stars further out.

In the following we discuss a few processes in stellar dynamics which influence the evolution of stellar clusters.

Cluster formation. We discuss the formation of a stellar cluster, considering a very young population of N stars which is still embedded in the gas cloud out of which the stars were formed. We define the total embedded stellar mass of the cluster M_{ecl} and use m_S as mean stellar mass

$$m_S = \frac{M_{\text{ecl}}}{N}.$$

Further we can define a fractional star-formation efficiency ϵ , the fraction of the total mass of the initial gas cloud M_{cloud} which ends up in newly formed stars

$$\epsilon = \frac{M_{\text{ecl}}}{M_{\text{ecl}} + M_{\text{gas}}}.$$

Here M_{gas} is the gas left over from the star-formation process ($M_{\text{cloud}} = M_{\text{ecl}} + M_{\text{gas}}$). Usually it is very difficult to determine observationally the mass of the remaining gas after the end of the star formation process. For this reason the existing “typical” fractional star formation efficiency parameter is very uncertain. A value in the range

$$0.2 \lesssim \epsilon \lesssim 0.4$$

is often quoted. This means that less than half of the mass of a collapsing cloud ends up in stars. This is a strong hint that the star formation in a collapsing cloud is terminated by the newly formed stars: this is called **feedback effect** in star formation. The following energetic processes can be responsible for the termination of star formation:

- the production of turbulence by the outflows from circumstellar disk around newly forming stars,
- photoionization and heating by the energetic radiation produced by the gas accretion processes of protostellar sources or the UV radiation from the hot photosphere of young, massive stars,
- shocks created by the stellar winds of young stars,
- shocks from supernova explosions of very massive, short lived stars.

Feedback: instantaneous gas removal. We can estimate what happens if there is an embedded cluster of protostars where the gas is removed in a short time by energetic stellar processes.

We assume that the embedded, proto-stellar cluster is in a dynamical equilibrium state what is a reasonable assumption for a 10 Myr young cluster (see Table 3.1). Then the total energy (or binding energy) is

$$E_{\text{ecl}} = -\frac{GM_{\text{init}}^2}{r_{\text{init}}} + \frac{1}{2}M_{\text{init}}\sigma_{\text{init}}^2 \quad (3.72)$$

where σ_{init} is the velocity dispersion which can be written for a virialized system $E_{\text{pot}} + 2E_{\text{kin}} = 0$ as

$$\sigma^2 = \frac{GM}{r}. \quad (3.73)$$

A virialized systems relates also the binding energy and the potential energy

$$E = -\frac{1}{2}E_{\text{pot}}.$$

The initial mass is $M_{\text{init}} = M_{\text{ecl}} + M_{\text{gas}}$ and this quantity can be the same as the total mass of the collapsing cloud $M_{\text{init}} = M_{\text{cloud}}$. The formalism is also valid for later stages where already some gas is lost, so that $M_{\text{init}} < M_{\text{cloud}}$ and $M_{\text{gas}}(t) < M_{\text{gas}}(t=0)$.

If energetic processes remove instantaneously the gas then the total mass of the cluster is changed from M_{init} to $M_{\text{after}} = M_{\text{ecl}}$ which includes only the total mass of the stars. The instantaneous gas removal does not change instantaneously the radial distribution r_{init} and the kinetic motion σ_{init} of the stars. The total binding energy of the cluster immediately after the gas removal is then

$$E_{\text{after}} = -\frac{GM_{\text{after}}^2}{r_{\text{init}}} + \frac{1}{2}M_{\text{after}}\sigma_{\text{init}}^2 = -\frac{1}{2}\frac{GM_{\text{after}}^2}{r_{\text{init}}}. \quad (3.74)$$

The cluster evolves now with a timescale of the order of the relaxation time scale to a new equilibrium state. If we assume that the mass $M_{\text{cl}} = M_{\text{after}}$ and energy $E_{\text{cl}} = E_{\text{after}}$ are conserved during this phase then the new equilibrium state can be described by a new radius r_{cl} and a new velocity dispersion σ_{cl}

$$E_{\text{cl}} = -\frac{GM_{\text{cl}}^2}{r_{\text{cl}}} + \frac{1}{2}M_{\text{cl}}\sigma_{\text{cl}}^2. \quad (3.75)$$

The resulting cluster radius follows from $E_{\text{cl}} = E_{\text{after}}$ where

$$E_{\text{after}} = -\frac{GM_{\text{after}}}{r_{\text{init}}}\left(M_{\text{after}} - \frac{1}{2}M_{\text{init}}\right)$$

and with the relations from above $M_{\text{init}} = M_{\text{cl}} + M_{\text{gas}}$ we obtain

$$\frac{r_{\text{cl}}}{r_{\text{init}}} = \frac{1}{2}\frac{M_{\text{cl}}}{M_{\text{cl}} - M_{\text{init}}/2} = \frac{M_{\text{cl}}}{M_{\text{cl}} - M_{\text{gas}}}. \quad (3.76)$$

This equation implies that the cluster radius goes to ∞ , or becomes unbound

- for $M_{\text{gas}} \rightarrow M_{\text{cl}}$, or if the removed gas contains equal or more mass than the stellar mass of the cluster,
- this is potentially the case for inefficient star formation when the fractional star formation efficiency $\epsilon \leq 0.5$ is low and a lot of gas is still present in the newly formed clusters.

Since, the inferred fractional star formation rate is low $\epsilon < 0.5$, and there are many gas-less clusters observed there must be alternatives to the instantaneous gas removal model. Instantaneous gas removal would lead to the destruction of many galactic clusters, but this did not happen for all the known stellar clusters in our Milky Way.

Feedback: continuous removal of gas. One can talk of a continuous mass loss if the time scale for gas removal is much longer than the relaxation time or the cluster crossing time:

$$\tau_{\text{gas}} \gg \tau_{\text{cross}} .$$

In this case the cluster adjusts its dynamics continuously according to the virial equilibrium. The increase of the cluster radius can then be described as a result of small (infinitesimal) mass removals:

$$\frac{r_{\text{init}} + \delta r}{r_{\text{init}}} = \frac{M_{\text{init}} - \delta M_{\text{gas}}}{M_{\text{init}} - \delta M_{\text{gas}} - \delta M_{\text{gas}}} . \quad (3.77)$$

which can be written as

$$\frac{r + dr}{r} = \frac{M - dM}{M - 2dM} ,$$

We search now for the formula for the relative radius increase of the cluster because of a small mass loss. Useful formulae are obtained by rearranging

$$\left(\frac{dr}{r} + 1\right)(M - 2dM) = M - dM \quad \text{or} \quad \frac{dr}{r}(M - 2dM) = -(M - 2dM) + M - dM = dM .$$

Since the radius increases for a reduced mass we can write

$$\frac{dr}{r} \frac{M - 2dM}{M} = -\frac{dM}{M} \quad \text{or} \quad \frac{dr}{r} = -\frac{dM}{M} \left(\frac{M}{M - 2dM}\right)$$

For slow mass loss, there is $|dM| \ll M$. and we can approximate

$$\frac{dr}{r} \approx -\frac{dM}{M} . \quad (3.78)$$

Integration yields $\ln(r_{\text{cl}}/r_{\text{init}}) = -\ln(M_{\text{cl}}/M_{\text{init}})$ or

$$\frac{r_{\text{cl}}}{r_{\text{init}}} = \frac{M_{\text{init}}}{M_{\text{cl}}} = \frac{M_{\text{ecl}} + M_{\text{gas}}}{M_{\text{cl}}} = \frac{1}{\epsilon} . \quad (3.79)$$

If the mass-loss is slow, then one can have a low fractional star formation efficiency (say 0.2) and loose a lot of gas (80 %) from the initial cloud mass and still end up with a bound cluster. The radius of the cluster expands like $1/\epsilon$. For example, if 80 % of the mass is lost by a continuous gas removal then the initial radius of the cluster expands by a factor of 5.

The conclusion is that with a slow mass loss, which allows a continuous re-virialization of the cluster dynamics, the mass loss causes less expansion and a more likely survival of a cluster compared to an instantaneous mass loss.

Mass segregation and core-formation. A cluster contains stars with a range of masses. The interactions of stars in a cluster induces, like in the kinetic gas theory, an evolution towards **equipartition**:

- in two-body interactions, the more massive stars transfer a significant amount of their large kinetic energy to less massive stars i , until $m_1 v_1 \approx m_i v_i$,
- this leads in a self-gravitating star clusters to a mass segregation, the more massive stars have less specific (per unit mass) kinetic energy and sink towards the cluster center, while less massive stars gain kinetic energy and diffuse outwards to larger radii.

The concentration of massive stars towards the center would just continue and lead to a core collapse. A relatively small number of massive stars concentrate in a very compact cluster core while the halo expands. This evolution would lead to a singularity if hard binaries would not counteract to this process.

Compact binary stars. Binary stars can transfer a lot of energy to a dense stellar system by dynamic interactions. We consider here only a simple energy argument.

A virialized system has a binding energy of

$$E_{\text{cl}} \approx -\frac{GM^2}{R_{\text{cl}}} \approx -\frac{G N^2 m_S^2}{R_{\text{cl}}}.$$

We can compare this to the binding energy of a binary star which is

$$E_{\text{bin}} \approx -\frac{G m_S^2}{a}$$

where a is the orbital separation (semi-major axis).

If the binary is sufficiently compact then its binding energy (negative total energy) is equal to the total binding energy of the entire cluster. The corresponding binary separation is

$$a_{\text{eq}} \approx \frac{R_{\text{cl}}}{N^2}. \quad (3.80)$$

This separation corresponds to

- $a_{\text{eq}} \approx 2$ AU for a open cluster with 1000 stars and a radius of 10 pc,
- $a_{\text{eq}} \approx 10^{-4}$ AU (or $0.1 R_{\odot}$) for a globular cluster with 10^5 stars and 10 pc radius.

This comparison shows that compact binaries, also called hard binaries, can stabilize a stellar cluster against collapse. Interaction of a hard binary star with a single star can transfer orbital energy of the binary to the third star, which gains then kinetic energy and moves outward in the cluster. This interaction reduces of course the separation and the total energy of the binary. However, a compact binary can have more binding energy than an entire open cluster. Such binary star interactions act against the cluster core collapse due to two-body interactions and equipartition. Of course, the binaries become more and more compact with time and they may even merge. This scenario can also explain the presence of the blue stragglers in the HR-diagram of clusters.

In globular cluster several hard binaries are required to stabilize the system. With X-ray observations such hard binaries were indeed found in several globular clusters. There

are cases with about 10 or even more such binaries in one globular cluster. These X-ray binaries have characteristics of low mass X-ray binaries, which are composed of a neutron star and a companion, often a white dwarf, in a very compact orbit with an orbital period of about an hour. Thus the orbital separation is indeed very small, of the order 10^{-3} AU, or even less. Several such binaries are capable to stabilize a globular cluster against collapse of the compact core.

Evaporation. The stars in the cluster halo can escape from a cluster if the encounters with other stars transfer enough energy so that they can escape from the system. For this a star must reach a velocity above the **escape speed** $v_e(r)$ or its total energy must become positive:

$$E_{\text{kin}} + E_{\text{pot}} = \frac{1}{2}m_S v^2(r) + m_S \Phi(r) > 0.$$

or

$$v(r) > v_e(r) = \sqrt{2\Phi(r)}.$$

This can be generalized to an expression $v_e^2(\vec{x}) = -2\Phi(\vec{x})$ so that we can write a general mean-squared escape velocity for a system with a density $\rho(\vec{x})$ according to

$$\langle v_e^2 \rangle = \frac{\int \rho(\vec{x}) v_e^2(\vec{x}) d\vec{x}}{\int \rho(\vec{x}) d\vec{x}} = -2 \frac{\int \rho(\vec{x}) \Phi(\vec{x}) d\vec{x}}{M} = -4 \frac{E_{\text{pot}}}{M}$$

According to the virial theorem $2E_{\text{kin}} + E_{\text{pot}} = 0$, where E_{kin} is the total kinetic energy $M\langle v^2 \rangle/2$, the root mean squared (rms) escape speed is just twice the rms speed:

$$\langle v_e^2 \rangle = 2\langle v^2 \rangle.$$

We may assume that the velocity distribution behaves in a collisionally dominated system ($t > t_{\text{relax}}$ like a Maxwellian distribution, where a fraction of about $\gamma = 0.7\%$ of particles have a velocity which is $v > 2\langle v \rangle$. Thus we can assume that the two-body interaction removes about a fraction γ of stars by evaporation every relaxation time:

$$\frac{dN}{dt} = -\frac{\gamma N}{t_{\text{relax}}} = -\frac{N}{t_{\text{evap}}}.$$

Thus the **evaporation time** is of the order

$$t_{\text{evap}} = \frac{t_{\text{relax}}}{\gamma} \approx 140 t_{\text{relax}}.$$

Thus any system with an age comparable to $\tau \approx t_{\text{relax}}$ will have lost a substantial fraction of its stars. If we use the characteristic relaxation time scale for open cluster $t_{\text{relax}} \approx 10$ Myr then we obtain an evaporation time scale of the order 1.5 Gyr. This is of the same order of magnitude, although a bit higher, than the estimated typical age of stellar cluster of about $t \approx 0.3$ Gyr.

Most likely, there exist additional processes, which accelerates the evaporation of open clusters in the galactic disk. A possible process is the gravitational interaction of clusters with molecular clouds which enhances the stellar velocity dispersion in the cluster and shortens the evaporation time scale.

Chapter 4

Physics of the interstellar medium

Components of the interstellar medium

The description of the interstellar medium requires the consideration of several physical components: different forms of baryonic matter, the magnetic fields, and the radiation fields.

baryonic matter	gas	molecular gas
		atomic gas
		ionized gas
	dust	small, solid particles $\lesssim 1 \mu\text{m}$ (smoke)
	cosmic rays	relativistic particles

radiation field

magnetic field

Thus, the interstellar medium is a complicated physical system with properties that depend on the:

- mutual interaction of the different components of the interstellar medium,
- interaction of the interstellar medium with stars.

4.1 Gas

4.1.1 Description of a gas in thermodynamic equilibrium

The following formula are valid for a gas in thermodynamic equilibrium.

Temperature and kinetic motion of the particles. The mean kinetic energy of gas particles is given by the temperature of the gas according to:

$$\langle \frac{1}{2}mv^2 \rangle = \frac{3}{2}kT \quad (4.1)$$

For different particles there is (equipartition): $\langle m_1v_1^2/2 \rangle = \langle m_2v_2^2/2 \rangle$

The Maxwell-Boltzmann velocity distribution (Fig. 4.1) for the particles is :

$$f_v(T) = \frac{n(v) dv}{n} = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-mv^2/2kT} 4\pi v^2 dv \quad (4.2)$$

Maximum of $n(v)$: $v_T = \sqrt{2kT/m}$; mean value: $\sqrt{\langle v^2 \rangle} = \sqrt{3kT/m}$
 Examples: $v_T(\text{H}, 10^4 \text{ K}) = 12.9 \text{ km/s}$, $v_T(\text{e}^-, 10^4 \text{ K}) = 550 \text{ km/s}$.

Figure 4.1: Maxwell-Boltzmann velocity distribution.

Boltzmann equation for the level population of atoms and molecules:

$$\frac{N_i(X^n)}{N_1(X^{+n})} = \frac{g_i}{g_1} e^{-E_i/kT}.$$

In equilibrium the level population of atoms and molecules have a “Boltzmann” distribution, expressed here as population of level i with statistical weight g_i and excitation energy E_i of an ion X^n relative to the population of the ground state $i = 1$ of that ion.

Saha equation for the ionization degree: The Saha equation describes the gas or plasma ionization degree in thermodynamic equilibrium

$$\frac{N_e N_1(X^{n+1})}{N_1(X^n)} = 2 \frac{g_1(X^{n+1})}{g_1(X^n)} \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} e^{-\chi/kT}.$$

The Saha equation is given for the ground states of two consecutive ionization states X^n and X^{n+1} with statistical weight $g_1(X^n)$ and $g_1(X^{n+1})$. χ is the energy required to ionize X^n from the ground state.

Planck function for the radiation field: The radiation intensity in a volume element in thermodynamic equilibrium can be described by the Planck equation for the intensity distribution of a perfect black body

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}. \quad (4.3)$$

Detailed balance: For a gas in thermodynamic equilibrium there exists a detailed balance of microscopic processes, in the sense that the rates for a given process are equal to the rates for the inverse process. Examples for microscopic processes and inverse processes for neutral or ionized atoms are:

collisional excitation $\longleftrightarrow$ collisional de-excitation

line absorption $\longleftrightarrow$ spontaneous and stimulated line emission
 collisional ionization $\longleftrightarrow$ 3-body recombination
 photo-ionization $\longleftrightarrow$ radiative recombination

4.1.2 Description of the diffuse gas

The gas in the interstellar medium is far from a thermodynamic equilibrium. Therefore the gas properties cannot be simply described by the temperature T . For this reason the temperature equilibrium has to be evaluated considering individual heating and cooling processes, which depend for example on the radiation field, the gas temperature, the level excitation, and the ionization degree.

Radiation field. Essentially everywhere in the Universe the gas temperature T_{gas} is higher than the temperature T_{rad} of the black-body radiation from the 3 K micro-wave background which dominates the global radiation field. On the other hand there are various types of other radiation sources (e.g. thermal radiation of dust, stars, galaxies, ...) which can be important locally. The radiation from these discrete sources is usually strongly diluted and the energy distribution may depart strongly from a black-body curve. Thus, there is essentially everywhere:

$$T_{\text{rad}} \neq T_{\text{gas}}. \quad (4.4)$$

The radiation field may be described by a diluted Planck-function

$$F_{\nu} = W \cdot B_{\nu}(T_{\text{rad}}) \quad \text{with e.g.} \quad W < 10^{-10},$$

and for many application the radiation field can even be neglected.

Particle densities. In the disk of spiral galaxies, a rough average of the mean proton or mean baryon density is of the order

$$n_b \approx 1 \text{ cm}^{-3},$$

while a dense interstellar cloud may reach a density of $n_b \approx 10^{+6} \text{ cm}^{-3}$. The density is only $n_b \approx 10^{-3} \text{ cm}^{-3}$ in hot bubbles and in the galactic halo the value approaches the mean density of baryons in the universe, which is $n_b \approx 10^{-7} \text{ cm}^{-3}$.

Thus, there exist the following dominant density regimes for baryons:

stars	$n_p \gtrsim 10^{+20} \text{ cm}^{-3}$
diffuse matter in the galactic disk	$n_p \approx 10^{-3} - 10^{+7} \text{ cm}^{-3}$
galactic halo	$n_p < 10^{-3} \text{ cm}^{-3}$

The Universe is made up, except for an extremely small fraction ($\epsilon \ll 10^{-24}$), of space filled with diffuse matter having a very low baryon density. The density and pressure of the interstellar medium is typically lower than what is reachable in the best vacuum chambers in the laboratory.

Velocity distribution of particles. Fortunately, the Maxwell velocity distribution $f_v(T_{\text{gas}})$ is still a good approximation for the diffuse gas. This simplifies very much all calculations, because the kinetic motion of the particles is defined at a given point by a single parameter; the gas temperature T_{gas} or for ionized gas the electron temperature T_e .

Why is the Maxwell velocity distribution valid?

This is not obvious when considering the mean free path length of a particle $\langle \ell \rangle$ and the mean time $\langle t_\ell \rangle$ between two collisions:

$$\langle \ell \rangle \approx 1/n\sigma \quad \langle t_\ell \rangle \approx \langle \ell \rangle / v_T \quad (4.5)$$

Atomic cross sections are of the order of $\sigma = \pi r_B^2$, where $r_B = 0.53 \text{ \AA}$ is the Bohr radius; thus $\sigma \approx 1 \text{ \AA}^2$ ($\text{\AA} = 10^{-8} \text{ cm}$).

Example: The mean free path and the mean time between two collisions for an electron in the diffuse interstellar gas in the Milky Way (particle density of $n = 1 \text{ cm}^{-3}$ and temperature $T_{\text{gas}} = 10000 \text{ K}$ or $v_e = 550 \text{ km/s}$) are on the order:

$$\langle \ell \rangle \approx 10^{16} \text{ cm} = 670 \text{ AU} \quad \langle t_\ell \rangle \approx 2 \cdot 10^8 \text{ s} = 6 \text{ yr}$$

Note that the kinetic velocity v_T of protons and ions is much lower, and therefore they undergo much less frequently interactions with other particles, apart from many interaction with electrons.

The Maxwell velocity distribution is only valid, because:

- the typical structures (clouds) are larger than $\langle \ell \rangle$,
- the typical time scale for temporal variations is longer than $\langle t_\ell \rangle$,
- the predominant processes for the interstellar gas are the collisions between electrons and electrons, electrons and protons, and electrons and hydrogen or helium atoms, which are (essentially) all elastic collisions. Therefore the kinetic energy is well exchanged and randomized between the particles.

Level population for atoms and molecules: In general the Boltzmann equation for the energy level population is not valid in the interstellar medium because the radiation field is strongly diluted and the radiative transition rates are far from a detailed balance.

The Boltzmann equation for the level population may still be valid for cases where the collisional transitions rates are much higher than all radiative transitions rates (Fig. 4.2). This occurs in high density clouds (many transitions) and for low lying levels of many atoms, ions, and molecules with only slow downward (spontaneous) transitions. Important examples are:

- fine-structure levels of the ground state in many atoms and ions,
- hyperfine-structure level of H I,
- rotational levels of the ground state of H₂,
- rotational levels of molecules in dense molecular clouds.

For these cases, the level population is defined by the gas temperature T_{gas} .

Figure 4.2: Illustration of the dominant processes for the population of energy level of atoms and molecules.

Apart from these special cases the level population of an atom has to be calculated from equilibrium equations which take the individual transition processes (collisional and radiative) into account.

4.1.3 Ionization

The Saha equation is not valid for the gas in the interstellar medium and it must be distinguished between two ionization regimes, photoionization and collisional ionization.

Photoionization equilibrium. Hot stars emit a lot of energetic photons ($h\nu > 13.6$ eV) which are capable to ionize the surrounding hydrogen gas. The ionization degree at a given location can be described by the following equilibrium (rates per volume element and time interval):

number of photo-ionizations = number of radiative recombinations.

For hydrogen this can be written as:

$$N_{\text{H}^0} \int_{\nu_0}^{\infty} \frac{4\pi I_{\nu}}{h\nu} a_{\nu}(\text{H}) d\nu = N_e N_p \alpha(\text{H}, T). \quad (4.6)$$

The number of ionization depends on

$$\Gamma_{\nu} = \int_{\nu_0}^{\infty} \frac{4\pi I_{\nu}}{h\nu} d\nu, \quad (4.7)$$

which is the flux of ionizing photons $\nu > \nu_0$ in [photons/cm²s] ($\nu_0 = 3.3 \cdot 10^{15}$ Hz, equivalent to a photon energy of 13.6eV). Γ_{ν} dilutes with distance d from the photon source like $\propto 1/d^2$ and may be further reduced by absorptions.

$a_{\nu}(\text{H})$ is the photoionization cross section for hydrogen, given for $\nu > \nu_0$ by

$$a_{\nu}(\text{H}) \approx 6.3 \cdot 10^{-18} \text{ cm}^2 \cdot (\nu_0/\nu)^3.$$

Figure 4.3: Photoionization cross section for H^0 , He^0 and He^+ .

The number of radiative recombinations is described by the densities of electrons and protons and the recombination coefficient for hydrogen:

$$\alpha_B(H, T) = 2.6 \cdot 10^{-13} (T/10^4 K)^{-0.7} \text{ cm}^3 \text{ s}^{-1}$$

Photo-ionized nebulae have always a temperature on the order 10'000 K for reasons which will be discussed later in connection with the cooling curve.

For rough estimates on the ionization degree, the number of ionization can be simplified to $\approx N_{H^0} \bar{a}(H) \Gamma$, using $\bar{a}(H) = 2.6 \cdot 10^{-18} \text{ cm}^2$ and $\alpha_B(H, 10'000 \text{ K})$. This yields the following approximation for the ionization degree

$$\frac{N_p}{N_{H^0}} \approx 10^{-5} \frac{\Gamma}{N_e}, [\text{cm}^{-1} \text{ s}] \quad (4.8)$$

which is given by the ratio between the flux of ionizing photons and the electron density N_e . The term Γ/N_e is called the ionization parameter.

Equilibrium for collisional ionization. If diffuse gas is hot enough for collisional ionization then the ionization degree is given by the rates of the two following processes:

number of collisional ionizations = number of radiative recombinations.

This is equivalent to the rate equation for the ions X^m and X^{m+1} :

$$N_e N(X^m) \gamma_e(X^m, T) = N_e N(X^{m+1}) \alpha(X^m, T)$$

$\gamma_e(X^m, T)$: ionization coefficient for the ionization by electrons

$\alpha(X^m, T)$: recombination coefficient

Collisional ionization and radiative recombination are both proportional to the electron density and the equilibrium depends only on the ionization and recombination coefficients. Therefore, one can assume for a gas in an equilibrium state (not rapidly changing with

time):

The ionization degree of collisionally ionized gas is a function of temperature:

$$\frac{N(X^{m+1})}{N(X^m)} = \text{funct}(T) = \frac{\gamma_e(T)}{\alpha(T)}$$

For example the ionization degree of hydrogen is $N(H^+)/N(H) \approx 0.003, 0.09, 1.2, 14,$ and 83 for electron temperatures of $T_e = 10'000 \text{ K}, 12'500 \text{ K}, 15'800 \text{ K}, 20'000 \text{ K}$ and $25'100 \text{ K}$ respectively. A good diagnostic tool for the determination of the ionization degree of a hot, collisionally ionized gas are the emission lines from different ionization states of Fe (see Slide 4–1).

Recombination time scale: The equilibrium for collisional ionization or the photo-ionization equilibrium requires a constant input of energy. For collisionally ionized regions this energy is provided usually by shock fronts due to gas moving supersonically. In photo-ionized regions this is the ionizing radiation. If this energy sources stops then the gas will recombine within a typical time scale of:

$$t_{\text{rec}} \approx \frac{N_p}{N_e N_p \alpha(H, T)} \approx \frac{4 \cdot 10^{12} \text{sec}}{N_e [\text{cm}^{-3}]}$$

4.1.4 H II-regions

We assume that all ionizing photons from a hot star are absorbed by a surrounding nebula. Such an ionized nebula is called radiation bounded.

For such a nebula we can formulate a **“global” photo-ionization equilibrium**, where the emission rate of ionizing photons $\nu > \nu_0$ is equal to the number of recombinations in the entire nebula. For a spherically symmetric, homogeneous nebula this can be written as:

$$\int_{\nu_0}^{\infty} \frac{L_{\nu}}{h\nu} d\nu = Q(H^0) = \frac{4\pi}{3} r_s^3 N_e N_p \alpha_B. \quad (4.9)$$

$Q(H^0)$: emitted, ionizing photons [photons/s]

r_s : radius of the ionized nebula

$N_H = N_e = N_p$: electron density (= proton density for a pure hydrogen nebula)

α_B : recombination coefficient for all recombinations into excited levels of H I; radiative recombinations to the ground state produce again an ionizing photon and are not counted for the ionization equilibrium.

The radial extension of the nebula is called the Strömgren-radius r_s , which follows from the previous equation:

$$r_s = \left(\frac{3}{4\pi \alpha_B} \frac{Q(H^0)}{N_e^2} \right)^{1/3}. \quad (4.10)$$

The mass of the ionized matter in the ionized hydrogen nebula is (m_p = proton mass):

$$m_s = \frac{m_p}{\alpha_B} \frac{Q(H^0)}{N_e}$$

Table 4.1 lists Strömgren-radii for different types of hot and massive main sequence stars, adopting a density of $N_p = N_e = 100 \text{ cm}^{-3}$ and a temperature of $T = 7500 \text{ K}$.

Table 4.1: Parameters for a spherical Strömgren nebula ($N_e = 100 \text{ cm}^{-3}$, $T_e = 7500 \text{ K}$).

star	M_V	$T^* [K]$	$\log Q(\text{H}^0)$	$r_s [\text{pc}]$
O5	-5.6	48000	49.67	5.0
O7	-5.4	35000	48.84	2.6
B0	-4.4	30000	47.67	1.1

Photoionization of helium. Essentially the same formalism as for the hydrogen applies also for helium. If the light source emits sufficiently hard photons, then a He^{+2} -zone forms close to the source until the photons with energy $> 54\text{eV}$ are absorbed. Then follows a surrounding He^+ region by the ionization of He^0 by photons with energy $> 24.6\text{eV}$ and further out the He^0 -region (see Slide 4-2).

The H I-Strömgren radius remains practically unchanged when He-ionization is included, because each photon that ionizes He I or He II will produce also at least one recombination photon which is capable to ionize hydrogen.

Photo-ionization of heavy elements. For the calculation of the ionization structure of the heavy elements additional processes (e.g. charge exchange) and the diffuse radiation field has to be considered. Numerical calculations are required for accurate estimates. Thereby the ionization structure depends quite importantly on the ionization structure of helium, because the location of the He^{+2} , He^+ , and He^0 -zones defines the radiation spectrum in the nebula, which is strongly changed by the photoionization thresholds and the diffuse emission produced by the recombination of helium atoms and ions.

Qualitatively, there results for the heavy elements always an ionization stratification, with the more highly ionized species close to the radiation source and lower ionized species further out (Slide 4-2). The highest ionization stage present in the nebula gives a qualitative indication of the spectral distribution ($\approx$ radiation temperature) of the radiation from the ionizing source.

4.2 Dust

Interstellar dust is made of small solid particles with radius $a < 1 \mu\text{m}$ similar in size to cigarette smoke particles (nano-particles). Interstellar dust can be observed in different ways:

- dark clouds (e.g. the coal-sack region),
- extinction and reddening, mainly in the galactic disk,
- light polarization of back-ground sources,
- strong IR emission,
- scattering of light.

4.2.1 Extinction, reddening and interstellar polarization

Extinction and reddening. The extinction (absorption and scattering) of light from “background” sources depends strongly on wavelength. Short wavelengths (UV-light) are very strongly absorbed and scattered (see extinction curve). For this reason the extinction causes a reddening of the colors of “background” sources in the visual band. The extinction curve has a strong maximum around 220 nm. This can produce in the far-UV continuum of stars a strong absorption minimum. In the visual the extinction curve is smooth and it is approximately a straight line in the extinction A_λ vs. $1/\lambda$ plot (Fig. 4.4). This is equivalent to an extinction cross section which behaves like $\kappa(\lambda) \approx 1/\lambda$.

Figure 4.4: Mean extinction curve A_λ/E_{B-V} ; the normalization in the visual is $A_V = 3.1 \cdot E_{B-V}$.

The interstellar reddening of an object by dust along the line of sight is often described by the color excess [units in magnitudes] for the filters B (blue) and V (visual = green/yellow):

$$E_{B-V} = A_B - A_V = (B - V) - (B - V)_0, \quad (4.11)$$

which is equivalent to the difference between the measured color $(B - V)$ and the initial (intrinsic) color $(B - V)_0$ of an object. The intrinsic color of a star can for example be determined from its spectral type.

The relation between the extinction (reduction of the brightness of an object) and the color excess is:

$$A_V \approx 3.1 \cdot E_{B-V}. \quad (4.12)$$

This relation holds for the dust in most regions of the Milky Way disk. For some special star forming regions clear deviations from this relation are observed. This points to the fact that the properties of the dust particles are there different from “normal”.

Polarization. The absorption by dust particles introduces a linear polarization of the light from the “background” source. The polarization curve $p(\lambda)$ has a broad maximum around 5500 Å with half the maximum value around 12000 Å and 2600 Å (Fig. 4.5). Typically, the polarization p is several % for a reddening of $E_{B-V} = 1$ mag ($p \leq 9 \cdot E_{B-V}$ %/mag).

Figure 4.5: Wavelength dependence of the interstellar polarization.

The polarization is due to a preferred orientation of the anisotropic (oblate and prolate) dust particles in the Galactic magnetic field. The elongated dust particles are forced by magnetic torques to rotate with their rotation axis parallel to the magnetic field lines. Thus, the orientation of the particles is predominantly perpendicular to the magnetic field lines. For light with wavelengths on similar scales as the particle dimensions the absorption will be stronger for waves with an E-vector oriented parallel to the elongated particle. It results a polarization $p_Q = (I_{\perp} - I_{\parallel}) / (I_{\perp} + I_{\parallel})$ parallel to the magnetic field. The fact, that the polarization is strongest in the visual region indicates that particles with a size (diameter) of about 500 nm are most efficient for producing the interstellar polarization.

The measurements of the interstellar polarization to many stars in Milky Way reveal that the galactic magnetic field is aligned with the galactic disk in roughly azimuthal orientation (Slide 4–3). Such measurements of the interstellar polarization direction are very important for the investigation of the large and small scale magnetic field structure in the Galaxy.

4.2.2 Particle properties

Particle size. The extinction and polarization properties of the interstellar dust provide important information on the size of dust particles (see Fig. 4.6):

- Large particles with radii $a \gg \lambda$ absorb and scatter the light (UV-vis-IR domain) in a wavelength-independent way. Thus, the extinction is proportional to the cross section of the particle $\kappa(\lambda \ll a) \approx \pi a^2$.
- Very small particles $a \ll \lambda$ scatter light according to the Rayleigh-scattering laws with a cross section proportional to $\kappa(\lambda \gg a) \propto \lambda^{-4}$.
- The extinction curve is compatible with an average absorption cross section proportional to $\kappa(\lambda) \propto \lambda^{-1}$. This indicates that there exists a broad distribution of particle sizes in the range $a \approx 0.01 - 1 \mu\text{m}$ with a power law of roughly $n_S(a) \propto a^{-3}$ (detailed fits yield a power law index of -3.5) for the size distribution of the interstellar particles).
- A large fraction of particles with sizes $a \approx 0.3 \mu\text{m}$ must be anisotropic and well aligned perpendicular to the interstellar magnetic field in order to produce the observed maximum around $0.55 \mu\text{m}$ in the interstellar polarization curve.

Figure 4.6: Dust extinction for different particles sizes.

Dust particle density. The average density of the interstellar dust particles in the Milky Way disk can be estimated from the observed mean extinction. This extinction is roughly 1 mag/kpc (V-band) or an optical depth of about $\tau \approx 1/\text{kpc}$. The most efficient absorbers in the V-band are the particles with diameters $2a \approx \lambda = 0.55 \mu\text{m}$. We can use for the particles along the line of sight the cross section: $\pi a^2 \approx 2 \cdot 10^{-9} \text{cm}^2$. It follows from

$$\tau = \pi a^2 \cdot n_S \cdot \text{kpc} \approx 1 \quad (4.13)$$

the density of particles with sizes around $a = 0.2$ to $0.3 \mu\text{m}$ of about $n_S \approx 1.5 \cdot 10^{-13} \text{cm}^{-3}$ (this corresponds to 150 particles per km^3). This is a very small dust particle density when compared to hydrogen $n_H \approx 1 \text{cm}^{-3}$. Despite this, these dust particles are dominating the extinction in the visual region.

The average gas to dust mass ratio is about 160 (± 60) in the Milky Way disk. Thus, about one third of the mass of the heavy elements is bound in dust particles (assuming a metallicity of $Z = 0.02$). Dust and gas are quite well mixed in the interstellar medium and therefore there exists an empirical relation between extinction and hydrogen column density:

$$N_H \approx 6 \cdot 10^{21} E_{B-V} \text{ mag}^{-1} \text{ cm}^{-2}. \quad (4.14)$$

Composition. The main components in interstellar space, H and He, form no solid particles for the existing temperatures (> 5 K). For this reason, the main components of the dust are heavy elements. For the composition of dust particles one has to distinguish between two types of elements:

- elements, which easily condense in dust particles (refractory elements), e.g.: Al, Si, Mg, Ca, Cr, Ti, Fe and Ni.
- elements, which are not (noble gases) or not easily bound in dust particles, e.g.: Ne, Ar, N, O, S, Zn.

The abundance of different dust particle types can be inferred from spectroscopic signatures. The observations indicate for the Milky Way disk:

mass	particle type	examples
60 %	silicates	quartz SiO_2 , silicates $(\text{Mg,Fe})[\text{SiO}_4]$
20 %	organic molecules	carbon-polymers
12 %	graphite	
4 %	amorphous carbon	
1 %	“PAHs”	poly-aromatic-hydrocarbons, e.g. benzol

Ices of different kinds, e.g. from water H_2O , methane CH_4 , and ammonium NH_3 may condense in dense and cold molecular clouds as mantle around a dust nucleus.

Examples for spectroscopic signatures from dust particles are emission or absorption band at the following wavelengths (see also Slide 4–4):

silicates:	9.7, 18 μm	graphite:	2200 $\text{\AA}$ (?)
ice (H_2O):	3.1 μm ,	PAH:	3.3, 6.2, 7.7, 8.7, 11.3 μm

4.2.3 Temperature and emission of the dust particles

The temperature of the dust particles depends strongly on the radiation field, and therefore on the environment. The dust emission from the galactic discs has typically a spectral energy distribution corresponding to a black body radiation temperature of 10 – 30 K.

The dust temperature can also be significantly higher $\approx 100 - 1000$ K, for example in regions with a strong UV-visual radiation field as expected near bright stars. Above 1000 K the dust particle sublimate. The ice-mantles sublimate already for dust particle temperatures of about 100 K.

The dust particles absorb very efficiently visual and UV-radiation because the particle sizes are comparable to these wavelengths (see above). The emission of radiation in the

far-IR (around $100\ \mu\text{m}$) is not efficient, because the emitting particle is much smaller than the emitted wavelength (a dipole antenna with a length l is also not efficient in emitting radiation with $\lambda \gg l$). For this reason, each absorbed UV-visual photon enhances immediately the temperature of the absorbing particle. Then it takes some time ($\sim$ seconds) until the particle has cooled down again by the emission of many far-IR photons (Fig. 4.7). The absorption of an energetic photon by a small particles yields a high particle temperature because the heat capacity is small. Thereafter, the cooling time is relatively long, because of the small size (inefficient emission) of the particle. Thus, the ice mantels of small particles sublime first. The spectral energy distribution of the “thermal” emission of a large volume of dust particles corresponds to the black-body radiation with a temperature corresponding to the mean temperature of the dust particles.

Figure 4.7: Temperature as function of time for large (left) and small (right) dust particles in the radiation field of a hot star

IR-galaxies. In many galaxies the interstellar dust hides large regions with embedded sources, like clusters of young, massive stars or an active galactic nucleus. These sources emit a lot of radiation in the visual and UV wavelength range which is first absorbed by the surrounding dust and then re-radiated by the dust as black-body radiation in the far-IR spectral region (Slide 4–4). For this reason these galaxies emit most of their energy around $50 - 100\ \mu\text{m}$. Galaxies, which emit much more radiation in the IR than in the UV-visual are called IR-galaxies. The brightest galaxies of this type (so-called ULIGs = ultra-luminous infrared galaxies) emit more than $10^{12}L_{\odot}$ in the IR. They belong to the brightest galaxies in the Universe.

4.2.4 Evolution of the interstellar dust

Dust particles form in slow and dense stellar winds. They may also form and grow in dense clouds. Various processes erode, modify and destroy the dust particles and this “processing” homogenizes the particle properties in a galaxy.

Condensation and grows. There are two main regimes where dust particles may form or grow:

- stellar winds from cool stars,
- dense molecular clouds.

The gas temperature in dense stellar winds from cool stars drops rapidly with distance from the star. As soon as the temperature is below the dust condensation temperature $T < T_{\text{cond}} \approx 1000$ K dust particles will form. Because of the formation of dust, the stellar wind becomes optically thick in the visual (and UV) range so that momentum from the radiation field is transferred to momentum of the optical thick gas/dust (radiation pressure). The chemical composition of the dust particles formed depends on the chemical abundances in the stellar wind. In cool stars the most abundant molecule in the atmosphere besides H_2 is CO. For oxygen-rich stars ($\text{O} > \text{C}$; M-type stars) all carbon is blocked in CO and therefore the most abundant dust particles will be silicates $(\text{Fe,Mg})\text{SiO}_x$. In carbon-rich stars ($\text{C} > \text{O}$; C-stars) all oxygen is blocked in CO and the carbon rich particles like SiC, amorphous carbon, graphite, PAHs, etc. are formed.

Dust particles can also form and evolve in molecular clouds if the density is high enough. In particular the dust particles are reprocessed and therefore homogenized. Small particles can grow, and if the temperature is low enough then ice-mantels (H_2O , NH_3 or CO_2 , CH_4) may condense around the dust

Erosion and destruction. The most important processes for the erosion and destruction of dust particles are:

- sublimation,
- absorption of high energy radiation,
- collision with fast moving (thermal) gas particles,
- collision with other dust particles.

Dust particles erode via evaporation of single atoms or molecules. This process is gradual and starts to be significant for temperatures of $T \gtrsim 30$ K. The evaporation is also enhanced if the particle is in a strong UV-visual radiation field. Single, energetic photons may be able to evaporate small particles, because of the relatively strong temperature rise after a photon absorption.

The absorption of an UV photon ($\lambda \lesssim 2000$ Å) may excite an atom or molecule of the dust particle, followed by the ejection of a component.

Collisions with thermal ions can strip off single or several atoms or molecules from the particle. This process is certainly very important and efficient in regions with high gas temperatures $T \gtrsim 10^5$ K. For this reasons the dust particles cannot survive in dense collisionally ionized gas, like supernova remnants.

Collisions between dust particles with large relative velocities $\gtrsim 1$ km/s can lead to the melting and evaporation of both particles. This process is of importance in dense clouds.

4.3 Magnetic fields

Signatures from galactic magnetic fields, e.g. synchrotron emission, interstellar polarization, or Faraday-rotation, can be observed in the Milky Way and many other galaxies.

The large scale magnetic fields in disc-galaxies are aligned with the disc and follow often the spiral structure (Slide 4–5). The origin of the magnetic fields in the Universe is unclear. If there are seed fields present, then they can be enhanced in disc galaxies due to the differential rotation.

Small scale structures of the magnetic field observed in the Milky Way are often connected with high density regions (molecular clouds) or strong dynamic effects, e.g. connected to H II regions and supernova remnants. The magnetic fields are important for the gas motion, because charged particles can essentially only move along the magnetic field lines and not perpendicular to them. In addition, the magnetic fields determine the motion of the relativistic particles (electrons and cosmic rays) in the interstellar medium.

The average magnetic field in the Milky Way has a strength of about $2 \mu\text{G}$ ($1 \text{ G} = 10^{-4} \text{ T}$; magnetic flux density). The field strengths in H II-region can be about 10 times stronger and in molecular clouds even 100 times stronger.

Charge drift velocities in the magnetic field: The magnetic fields in the Milky Way require, that there exists a differential motion between the charged particles (electrons and ions), i.e. there must exist electric currents. For a field strength of $\approx \mu\text{G}$ and a ionization of 1 % (fraction of charged particles) one can write according to the first Maxwell law $\text{rot} \vec{B} = \mu_0 \vec{j}$ the following relation between the typical length scale $L = 100 \text{ pc}$ and the drift velocity v :

$$B/L \approx \mu_0 \cdot n_p \cdot e \cdot v.$$

($B = 1 \mu\text{G}$, $n_p = 0.01 n_{\text{H}}$, $e = 1.6 \cdot 10^{-19} \text{ C}$ and $\mu_0 = 1.26 \cdot 10^{-6} \text{ A s V}^{-1} \text{ m}^{-1}$). This yields a differential drift velocity of the charges on the order $v = 10^{-6} \text{ cm/s}$. It seems obvious that such drift velocities are possible in the interstellar medium.

Temporal evolution of magnetic fields. Existing magnetic field have a very long life time. The magnetic field can be reduced, if the charges collide so that their relative drift velocities are reduced. Thus, the currents $\vec{j}$ decay if there exists an electric resistivity η in the medium. However, η is extremely small in the interstellar medium, and $\eta = 0$ is a very good approximation. The result is that **the magnetic fields are frozen in** for the interstellar plasma and the magnetic fields behave in the following way:

- the fields move with the plasma,
- the field strength is roughly proportional to the plasma density,
- the magnetic fields pressure $\propto B^2$ acts like a gas pressure,
- the B-field is stabilizing dense clouds against collapse.

The magnetic field may drift out of the highly neutral gas in molecular clouds through a process that is called ambipolar diffusion. Neutral clouds without magnetic fields can easily collapse and form stars.

The differential rotation $\Omega(R)$ in disc galaxies enhances the magnetic field in azimuthal direction. The fields in radial direction are enhanced by the α -effect in small-scale turbulences. Small scale field motions get due to the Coriolis force in the rotating disk a predominant rotation direction, so that the radial field components are enhanced. After a few disc rotations the field strength saturates because the enhancement by the $\Omega - \alpha$ -dynamo is compensated by the field dissipation in magnetic reconnections.

4.4 Radiation field

The radiation field in the interstellar and intergalactic medium depends strongly on the location. Often there exists a bright star which dominates the radiation field. Dust may attenuate strongly for some places the visual, UV and soft ($E < 1$ keV) X-ray radiation. Important is also the distribution of neutral hydrogen, because H I blocks efficiently the ionizing far-UV radiation.

It is difficult to define an average radiation field for the ISM / IGM gas. One possibility is to take the radiation field at the position of the sun. This is quite a reasonable approach, because the sun is not in a particular region of the Milky Way. Important components of the diffuse radiation are the cosmic micro-wave background, the diffuse thermal IR radiation from the dust in the galactic disc and the stellar light from the stars in the Milky Way. Slide 4–6 shows the spectral distribution of the diffuse radiation field in the solar neighborhood.

The radiation field has an important effect on the properties of the interstellar gas. The absorption of (energetic) photons can cause the following changes:

- heating of dust particles, the evaporation of the ice mantles and the dust cores,
- photo-dissociation of molecules,
- photo-ionization of atoms and ions as discussed in Sect. 4.1.3.

The gas is heated and additional charged particles are created in all these processes so that also the gas pressure is enhanced. Most important is the ionization of H I, because the optical depths for the ionizing far UV radiation and the electron density depend critically on the $N(\text{H}^+)/N(\text{H}^0)$ -ratio.

4.5 Cosmic rays

Cosmic rays are high energy (relativistic) particles in the interstellar medium with

$$E \gg m_0 c^2 \quad (4.15)$$

($m_0 c^2$: 0.51 MeV for e^- ; 938 MeV for p^+).

4.5.1 Properties of the cosmic rays

Energy distribution. The energy distribution of the ions (p^+ , α , atomic nuclei) can be described by a power law:

$$J(> E) \sim E^{-q} \quad \text{with} \quad q \approx 1.7 - 2.1, \quad (4.16)$$

where $J(> E)$ is the flux of all particles with energy $> E$ (Slide 4–7). The highest measured energies are about $3 \cdot 10^{20}$ eV (≈ 50 Joule). However, such events are very rare $\approx 1 \text{ km}^{-2} \text{ yr}^{-1}$. Note, that the particle energies reached at CERN LHC are of the order 10 TeV, that is 10^{13} eV. Of course the LHC produces these energies in large quantities.

Observations of cosmic rays. Cosmic rays are detected with particle detectors and Cherenkov-telescopes, which measure essentially the products of a collision with a particle in the Earth atmosphere. For the determination of the ion abundances of the initial particles, one has to bring detectors into space or at least into the stratosphere.

The highly relativistic particles which penetrate into the Earth atmosphere produce in collisions with N and O nuclei a particle shower, a cascade of hadronic particles, mainly pions ($\pi^\pm, \pi^0$), but also nucleons (p,n), anti-nucleons ($\bar{p}, \bar{n}$), kaons and hyperons, which collide again with N- and O-nuclei. The unstable particles decay via weak interaction and they produce electrons, positrons, muons, neutrinos, and photons. The photons can also produce matter-antimatter pairs.

The particles in the shower are often relativistic and move faster than the speed of light in air $v > c/n$ (n : refractive index of air) and they produce therefore a Cherenkov-light-cone. Cherenkov telescopes on the ground (Slide 4–8) and particle detectors on the ground (hadrons and charged particles) or underground (e.g. neutrinos) provide then information on the direction and energy of the initial cosmic ray particle.

Elemental abundances. The elemental abundance of the cosmic rays is rather similar to the solar abundance with two important differences (Slide 4–9). The light elements Li, Be, B, which are very rare in the sun and the rather rare heavy elements Sc, Ti, V have a strongly enhanced abundance, similar to the abundance of the next elements in the periodic system of elements. The explanation is that the abundance minima are filled in by the spallation or fission of heavy elements, in particular of C and Fe. Because the particles travel with relativistic speed through interstellar space, they encounter on their path H and He nuclei, mainly in dense molecular clouds, and they lose in collisions protons and α -particles in a kind of erosion process. These collisions produce also π -mesons and other particles. Observationally important is the following decay of π^0 -mesons which produces γ -rays with > 100 MeV which can be observed with detectors on satellites.

Cosmic rays are an important heating source for cold (≈ 30 K), dense molecular clouds, which are dust-shielded from the radiation of stars. Similar to the Earth atmospheres, a shower of energetic particles are created by an interaction with a cosmic ray particle. This leads then to a temperature enhancement in the cloud.

Relativistic electrons. There exists also a component of relativistic electrons in the cosmic rays which is however much weaker. The flux is about 100-times less than for protons. The observed energy regime for electrons is in the range 2 MeV – 1000 GeV. The electrons produce due to their relativistic motion in the galactic magnetic field **Synchrotron-radiation**, which can be observed easily with radio telescope.

4.5.2 Motion in the magnetic field

The motion of the charged cosmic ray particles depends on the terrestrial, interplanetary, and interstellar magnetic field. For the velocity component $v_\perp$ perpendicular to the mag-

netic field B the motion is controlled by the equilibrium of Lorentz force F_L and centrifugal force F_Z :

$$F_L = \frac{e}{c} v_{\perp} \cdot B = F_Z = m\omega^2 r_c = \frac{mv_{\perp}^2}{r_c}. \quad (4.17)$$

Thus, the particle move along circles with the cyclotron radius r_c (momentum: $p = mv_{\perp}$)

$$r_c = p \frac{c}{eB}. \quad (4.18)$$

For relativistic particles there is $p = \gamma mv$, $E = \gamma mc^2$ with the Lorentz factor $\gamma = (1 - (v/c)^2)^{-1/2}$. This yields the relativistic cyclotron radius

$$r_c = \frac{E}{eB} \quad \text{oder} \quad r_c[\text{pc}] = 1.08 \cdot 10^{-6} \frac{E[\text{GeV}]}{B[\mu\text{G}]} \quad (4.19)$$

($e = 4.8 \cdot 10^{-10} \text{ g}^{1/2} \text{ cm}^{3/2} \text{ s}^{-1}$ and $1 \text{ G} = \text{g}^{1/2} \text{ cm}^{-1/2} \text{ s}^{-1}$).

The distribution of the directions of the cosmic rays is essentially isotropic due to the deviation of the particle motion in the galactic magnetic field (for small energies also the interplanetary and terrestrial magnetic fields are important). The cyclotron radius is of the order of the Galaxy $r_c \approx 10^5 \text{ pc}$ for very high energies $E \approx 10^{11} \text{ GeV}$. Particles with such energies move along a straight line and their direction of origin can be determined. On the other hand they can also escape easily from the galactic magnetic field into the intergalactic medium.

4.5.3 The origin of the cosmic rays

The decay of the π^0 -mesons, which are created by collisions of the cosmic rays with interstellar matter, can be measured as diffuse γ -radiation in the galactic disc tracing the dense molecular clouds. This indicates that the cosmic rays are not a local phenomenon, but that they exist throughout the entire galactic disc.

The observed spallation (e.g. the overabundance of Li, Be, B) requires, that the relativistic particles pass typically through a column density of matter of about $1/\sigma \approx 5 \text{ g cm}^{-2}$ before they reach us. Based on this, the following estimates can be made:

- mean density (ISM) $n_{\text{H}} = 1 \text{ cm}^{-3} \rightarrow \rho = n_{\text{H}} m_{\text{H}} = 1.7 \cdot 10^{-24} \text{ g cm}^{-3}$
- travel distance $3 \cdot 10^{24} \text{ cm} = 1 \text{ Mpc} \approx 3 \cdot 10^6 \text{ Lyr}$
- travel time $\approx 3 \cdot 10^6 \text{ years}$

The very high particle energies of the cosmic rays are most likely produced in magnetized shock-fronts. The particles are in these shocks mirrored back and forth (in and out) of a fast moving gas flows having a speed of ($\Delta v \approx 10000 \text{ km/s}$). Each time the particle is mirrored it is accelerated by Δv . Such shock fronts are produced by supernova explosions and pulsar winds. The highest energy particles may originate from extra-galactic sources, for example quasars where shock fronts in relativistic jets are responsible for the acceleration.

4.6 Radiation processes

In astronomy the most important source of information is the observations of the electromagnetic radiation. The flux of the observed radiation

$$\vec{F}(x, y, \lambda, t)$$

can be determined as function of the following parameters:

- coordinates: x, y (right ascension, declination) $\rightarrow$ intensity images,
- wavelength $\lambda \rightarrow$ spectral energy distribution,
- time $t \rightarrow$ light curves,
- polarization, which is the orientation of the electric vector of the electro-magnetic wave; thus the flux is a vector quantity $\vec{F}$.

Physical properties of astronomical objects can be derived from this information. However, this requires a good knowledge of the physics of the radiations processes which take place in the interstellar medium. Important radiation process are discussed in this chapter.

4.6.1 Radiation transport

The **radiative transfer equation** for a sight line describes the change of the radiation energy dI (or intensity) along the optical path ds by contributions from emission processes and the weakening of the intensity by absorption processes:

$$dI_\nu = \epsilon_\nu ds - \kappa_\nu I_\nu ds \quad (4.20)$$

$I_\nu =$ spectral intensity	$I(\vec{r}, \vec{n}, \nu, t)$	$\text{erg cm}^{-2} \text{s}^{-1} \text{Hz}^{-1} \text{sr}^{-1}$
$\epsilon_\nu =$ emission coefficient	$\epsilon(\vec{r}, \vec{n}, \nu, t)$	$\text{erg cm}^{-3} \text{s}^{-1} \text{Hz}^{-1} \text{sr}^{-1}$
$\kappa_\nu =$ absorption coefficient	$\kappa(\vec{r}, \nu, t)$	cm^{-1}

The absorption coefficient $\kappa = \sigma \cdot n$ includes the cross section per particle $\sigma [\text{cm}^2]$ and the particle density $n [\text{cm}^{-3}]$. The geometric dilution $\sim 1/d^2$ of the radiation energy coming from a source is taken into account by the solid angle dependence $[\text{sr}^{-1}]$.

The radiation transfer equation is a first order differential equation:

$$\frac{dI_\nu}{ds} = \epsilon_\nu - \kappa_\nu I_\nu. \quad (4.21)$$

Optical depth and source function. The transfer equation takes a particularly simple form if we use (except for $\kappa_\nu \approx 0$) the so-called optical depth $d\tau_\nu = \kappa_\nu ds$ and the source function $S_\nu = \epsilon_\nu / \kappa_\nu$. The source function is often a more convenient physical quantity than the emission coefficient, especially if the emission at a given point depends strongly on the absorption.

The optical depth is the absorption coefficient integrated along the optical path from x_0 to x :

$$\tau_\nu(x) = \int_{x_0}^x \kappa_\nu(s) ds. \quad (4.22)$$

The point x_0 is arbitrary (e.g. the location of the source or the observer), and it sets the zero point for the optical depth scale. A medium is called to be:

- optically thick or opaque for $\tau_\nu > 1$,
- optically thin or transparent for $\tau_\nu < 1$.

The photon **mean free path** ℓ_ν is defined by:

$$\langle \tau_\nu \rangle = \kappa_\nu \ell_\nu = 1 \quad \text{or} \quad \ell_\nu = \frac{1}{\kappa_\nu} = \frac{1}{N\sigma_\nu}. \quad (4.23)$$

The mean free path is just the reciprocal of the absorption coefficient for a homogeneous medium.

When using optical depth and source function then the transfer equation can be written as follows:

$$\frac{dI_\nu}{d\tau_\nu} = \frac{\epsilon_\nu}{\kappa_\nu} - I_\nu = S_\nu - I_\nu. \quad (4.24)$$

Integration gives the **formal solution**,

$$I_\nu e^{-\tau_\nu} = \int S_\nu e^{-\tau_\nu} d\tau \quad (4.25)$$

or expressed with the start and end values for the optical depths:

$$I_\nu(\tau_2) = I_\nu(\tau_1) e^{-(\tau_2-\tau_1)} + \int_{\tau_1}^{\tau_2} S_\nu(\tau) e^{-(\tau-\tau_1)} d\tau.$$

Often one can adopt $\tau_1 = 0$ and $I_\nu(\tau_1) = I_\nu(0)$:

$$I_\nu(\tau_\nu) = I_\nu(0) e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau) e^{-\tau} d\tau \quad (4.26)$$

Simple, but **very important special cases** for the description of the interstellar medium are:

- **only emission** of optically thin, diffuse gas without a background source ($\kappa_\nu = 0$; $I_\nu(0) = 0$):

$$I_\nu = \int \epsilon_\nu ds \quad (4.27)$$

- **only absorption** ($\epsilon_\nu = 0$) of radiation of a background source $I_\nu(0)$:

$$I_\nu = I_\nu(0) e^{-\tau_\nu} \quad (4.28)$$

4.7 Spectral lines: bound-bound radiation processes

The line emissivity ϵ^ℓ of an atom or molecule for a radiative decay of an upper level n to a lower level m is described by:

$$\epsilon^\ell = \int \epsilon_\nu^\ell d\nu = \frac{1}{4\pi} h\nu_{nm} A_{nm} N_n \quad \text{with} \quad \epsilon_\nu^\ell = \epsilon^\ell \Psi(\nu) \quad (4.29)$$

- N_n : density of particles in level n [cm^{-3}]
- A_{nm} : decay rate or transition probability for this transition [s^{-1}] (Einstein A-coefficient)
- $h\nu_{nm}$: energy for the radiated photon
- $1/4\pi$: per steradian
- $\Psi(\nu)$: normalized line profile function

$\Psi(\nu)$ describes the strong frequency dependence of the emission coefficient ϵ_ν^ℓ (and absorption coefficient κ_ν^ℓ). This includes the line profile due to the intrinsic line width or the natural line profile, the Doppler-broadening due to the kinetic motion of the particles (Gauss-function), and the Doppler-structure of the line due to large scale motions of the emitting gas.

The line absorption depends on the intensity:

$$\int \kappa_\nu^\ell I_\nu d\nu = \frac{1}{4\pi} h\nu_{nm} I(\nu_{mn}) (N_m B_{mn} - N_n B_{nm}) \quad (4.30)$$

- $g_m B_{mn} = g_n B_{nm}$: Einstein B-coefficients
- g_m, g_n : statistical weights for the levels N_m, N_n (there is $g_m = (2J_m + 1)$)

This gives the line integrated absorption coefficient:

$$\kappa^\ell = \frac{1}{4\pi} h\nu_{nm} (N_m B_{mn} - N_n B_{nm}) = \frac{1}{4\pi} h\nu_{nm} N_m B_{mn} \left(1 - \frac{g_m}{g_n} \frac{N_n}{N_m}\right), \quad (4.31)$$

where the frequency dependence is again described by the normalized profile function $\kappa_\nu^\ell = \kappa^\ell \Psi(\nu)$. If $N_n/g_n > N_m/g_m$, which is equivalent to an inversion of the level population (over-population with respect to the Boltzmann-distribution), then the line absorption coefficient κ_ν^ℓ becomes negative, and the radiation is amplified by **stimulated emission** like in a laser.

The relations for the Einstein coefficients A_{nm} and B_{nm} follow from the requirement of detailed balance in thermodynamic equilibrium:

$$A_{nm} = \frac{2h\nu^3}{c^2} B_{nm} \quad \text{and} \quad B_{nm} = \frac{g_m}{g_n} B_{mn} \quad (4.32)$$

Detailed balance requires that the transition rates for radiative processes between two levels (say 1 and 2) are equal:

$$N_1 B_{12} B_{\nu_{12}}(T) = N_2 A_{21} + N_2 B_{21} B_{\nu_{12}}(T).$$

where we have the processes: absorption equals spontaneous and induced emission. Solving for the Planck function and using the Boltzmann equation $N_1/N_2 = (g_1/g_2) e^{h\nu_{12}/kT}$ gives:

$$B_{\nu_{12}}(T) = \frac{N_2 A_{21}}{N_1 B_{12} - N_2 B_{21}} = \frac{A_{21}/B_{21}}{(N_1/N_2)(B_{12}/B_{21}) - 1} = \frac{A_{21}/B_{21}}{g_1 B_{12}/g_2 B_{21} (e^{h\nu_{12}/kT})}.$$

This gives only the Planck function with the relations for the Einstein coefficients given above.

Important case: optically thin emission line:

$$I_\nu = \int \epsilon_\nu ds = \frac{1}{4\pi} h\nu_{nm} \Psi(\nu) A_{nm} \int N_n ds \quad (4.33)$$

The measured column density $\int N_n ds$ is always an average value for the observed solid angle. An accurate determination of the column density requires that the internal structure of the emission region is spatially resolved.

4.7.1 Rate equations for the level population

N_n is the population of level n which is defined by all the transition rates which populate this level $\sum_m N_m (R_{mn} + C_{mn})$ and the rates which depopulate this level $N_n \sum_m (R_{nm} + C_{nm})$. In an equilibrium state there is $dN/dt = 0$:

$$\frac{dN_n}{dt} = \sum_m N_m (R_{mn} + C_{mn}) - N_n \sum_m (R_{nm} + C_{nm}) = 0 \quad (4.34)$$

Rates for radiative transitions are given by R_{nm} and R_{mn} per time interval [s^{-1}] ($E_n > E_m$):

$$\begin{aligned} - R_{nm} &= A_{nm} + B_{nm} I_\nu && \text{spontaneous and induced line emission} \\ - R_{mn} &= B_{mn} U_\nu && \text{line absorption} \end{aligned}$$

The transition rates for spontaneous emission depends on the type of transition. In astronomy it is distinguished between allowed transitions, inter-combination or semi-forbidden transitions, and forbidden transitions. In atomic physics the terms, electric dipole transitions, magnetic dipole transitions, and multipole (usually quadrupole) transitions are used.

Typical transition rates A are:

- $A \approx 10^8 s^{-1}$: allowed transitions (electric dipole)
- $A \approx 10^2 s^{-1}$: semi-forbidden transitions (electric dipole with spin-flip)
- $A \approx 10^{-2} s^{-1}$: forbidden transitions (magn. dipole and electric quadrupole)
- $A \approx 10^{-5} s^{-1}$: forbidden fine structure transitions
- $A \approx 10^{-15} s^{-1}$: forbidden hyperfine structure transitions (e.g. H I)

Selection rules for dipole transitions:

For one electron atoms the selection rules are:

- $\Delta l = \pm 1$ (this includes a parity change for one electron systems)
- $\Delta m = 0, \pm 1$.

For many electron systems the selection rules are:

- parity change
- $\Delta S = 0$
- $\Delta L = 0, \pm 1$
- $\Delta J = 0, \pm 1$ except $J = 0$ to $J = 0$

In higher multipole transitions the spin may change (semi-forbidden transitions) or no parity change may be required (magnetic dipole or electric quadrupole transitions). There is no parity change in all transitions between states with the same electron orbit configurations like transitions between fine-structure levels or hyperfine-structure levels.

Rates for collisional transitions are described by C_{nm} and C_{mn} (in $[s^{-1}]$). It has to be distinguished between collisional deexcitation $n \rightarrow m$ and collisional excitation $m \rightarrow n$ ($E_n > E_m$):

$$C_{nm} = N_s Q_{nm} \quad \text{und} \quad C_{mn} = N_s Q_{mn} \quad (4.35)$$

where N_s is the density of the colliding particles (often e^- , p^+ , H, H_2 , etc.), and Q_{nm} , Q_{mn} are the collision rates $[cm^3 s^{-1}]$.

For collisional transitions in atoms by e^- between level n and m ($\Delta E_{nm} = h\nu_{nm} = E_n - E_m$, $n > m$) there is

$$C_{nm} = N_e \frac{1}{g_n} \frac{8.63 \cdot 10^{-6} cm^3 s^{-1}}{\sqrt{T [K]}} \Omega_{nm} \quad \text{collisional deexcitation} \quad (4.36)$$

$$C_{mn} = N_e \frac{1}{g_m} \frac{8.63 \cdot 10^{-6} cm^3 s^{-1}}{\sqrt{T [K]}} \Omega_{mn} e^{-\Delta E_{nm}/kT} \quad \text{collisional excitation} \quad (4.37)$$

$\Omega_{nm} = \Omega_{mn}$ are the collision strengths. Typical values for the collision strengths are of the order $\approx 0.1 - 10$, with usually only a small temperature dependence. It follows the following relation for the opposite collisional processes between level m and n :

$$C_{mn} = \frac{g_n}{g_m} C_{nm} e^{-\Delta E_{nm}/kT}. \quad (4.38)$$

4.7.2 Collisionally excited lines

2-level atom. The rate equation for a 2-level atom (or molecule) is, if we consider only collisional processes due to free electrons e^- (reasonable assumption for an ionized gas):

$$N_1 (B_{12}U_\nu + N_e Q_{12}) = N_2 (B_{21}U_\nu + A_{21} + N_e Q_{21}) \quad (4.39)$$

This equation becomes even more simplified if we neglect absorption and stimulated emission. This is a good approximation for the interstellar medium because of the weak radiation field. Typical dilution factors are $W \ll 10^{-10}$, and there is $B_{12}U_\nu \ll N_e Q_{12}$ and $B_{21}U_\nu \ll A_{21} + N_e Q_{21}$. This gives the following simple but very useful rate equation for a 2-level atom:

$$N_1 N_e Q_{12} = N_2 (A_{21} + N_e Q_{21}), \quad (4.40)$$

and with $Q_{12} = Q_{21} (g_2/g_1) e^{-h\nu/kT}$ ($h\nu = \Delta E_{21}$) follows the **level population ratio for a 2-level atom**:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \frac{N_e Q_{21} e^{-h\nu/kT}}{A_{21} + N_e Q_{21}} = \frac{g_2}{g_1} \frac{e^{-h\nu/kT}}{A_{21}/N_e Q_{21} + 1} \quad (4.41)$$

Low density regime: The spontaneous emission is much faster than the collisional deexcitation $N_e Q_{21} \ll A_{21}$ and it follows for the level population and the line emissivity ($\epsilon^\ell = (1/4\pi) h\nu_{nm} A_{nm} N_n$):

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \frac{N_e Q_{21}}{A_{21}} e^{-h\nu/kT} \quad \text{and} \quad \epsilon_{21} = \frac{1}{4\pi} h\nu_{21} N_e N_1 \frac{g_2}{g_1} Q_{21} e^{-h\nu_{21}/kT} \quad (4.42)$$

The line emission coefficient is then proportional to the particle density squared $\epsilon^\ell \propto N_1 N_e$ (Fig. 4.8).

High density regime: Collisions are frequent and the collisional de-excitation is much faster than the spontaneous emission $N_e Q_{21} \gg A_{21}$. In this case the level population depends only on the collisions and we obtain the Boltzmann level distribution

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} e^{-h\nu/kT} \quad \text{and} \quad \epsilon_{21} = \frac{1}{4\pi} h\nu_{21} N_1 \frac{g_2}{g_1} A_{21} e^{-h\nu_{21}/kT} \quad (4.43)$$

The line emission is proportional to the density $\epsilon^\ell \propto N_1$ (Fig. 4.8).

Figure 4.8: Schematic behavior for emissivity per electron for a collisionally excited line as function of the density (N_k =critical density).

The **critical density** $N_k = A_{21}/Q_{21}$ for a line transition defines the border line between the high and low density regimes. When considering “real” multi-level atoms, then the critical density is a quantity of a particular level x , for which one has to evaluate the ratio between all line transitions and all de-populating collisional transitions

$$N_{x,k} = \sum_n \frac{A_{xn}}{Q_{xn}}. \quad (4.44)$$

Collisional rates due to electrons in warm (photoionized) gas $T \approx 10^4$ K are on the order $Q_{21} \approx 10^{-7} \text{ cm}^3 \text{ s}^{-1}$. Rough estimates for the critical densities N_k for different types of line transitions in a photoionized nebula ($T \approx 10^4$ K, electron collisions dominate) are as follows

– allowed transitions	$A \approx 10^8 \text{ s}^{-1}$	$N_k \approx 10^{15} \text{ cm}^{-3}$
– inter-combination lines	$A \approx 10^2 \text{ s}^{-1}$	$N_k \approx 10^9 \text{ cm}^{-3}$
– forbidden transitions	$A \approx 10^{-2} \text{ s}^{-1}$	$N_k \approx 10^5 \text{ cm}^{-3}$
– forbidden fine structure lines	$A \approx 10^{-5} \text{ s}^{-1}$	$N_k \approx 10^2 \text{ cm}^{-3}$

Important example: H I – 21 cm line (1420 MHz)

The ground level of atomic hydrogen H I has due to the non-zero nuclear spin a hyperfine splitting, between the parallel and anti-parallel configurations of the spins of the proton and the electron.

energy level diagram H I
1²S

The parallel configuration (quantum number for the hyperfine-structure $f=1$, statistical weight $g=2f+1=3$) is energetically slightly higher ($\Delta E = h\nu \approx 10^{-5}$ eV) than the anti-parallel configuration ($f=0$, $g=1$).

The transition from $f = 1$ to $f = 0$ has the following properties:

- the transition rate for spontaneous emission is extremely small $A_{21} = 3 \cdot 10^{-15} \text{ s}^{-1}$ (decay time 10^7 year!),
- The collision frequency with other particles in the cold ($T = 100$ K), diffuse $N_{\text{H}} \approx 1 \text{ cm}^{-3}$, partly neutral interstellar medium (electron collision dominant) is on the order $N_{\text{H}} Q_{12} \approx 10^{-9} \text{ s}^{-1}$ (an H-atom gets a kick about every 60 years).

→ **collisions define the level population** and therefore the H I hyperfine structure level population is essentially always and everywhere in the interstellar medium in the high density regime:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} e^{-h\nu/kT} \quad \text{where} \quad h\nu/kT < 10^{-5} \quad \text{for} \quad T > 10 \text{ K} \quad (4.45)$$

$$\text{result : } N_2 = 3 N_1 = \frac{3}{4} N_{\text{H}^0} \quad \text{and} \quad \epsilon_{21 \text{ cm}} = \frac{1}{4\pi} h\nu A_{21} \frac{3}{4} N_{\text{H}^0} \quad (4.46)$$

Thus, 75 % of the atomic hydrogen in the Universe will be in the excited state $f = 1$ of the two hyperfine structure levels. Because H is so abundant it is possible to observe the decay $f=1 \rightarrow f=0$, despite the very long lifetime (small transition rate) of the excited level. In practice, the H I 21 cm line observations belong to the most important diagnostic tool for the investigation of cool, diffuse gas in the Milky Way and other galaxies. Thereby the measured surface brightness yields directly the column density along the line of sight $\int N_{\text{H}^0} ds$.

Temperature and density determinations. We consider a 3-level atom (or molecule) with the following simplifications:

- no absorption or stimulated emission (induced radiation transitions),
- no transitions between level 2 and 3 ($N_1 \gg N_2, N_3$).

Then there is:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \frac{e^{-h\nu_{21}/kT}}{A_{21}/N_e Q_{21} + 1} \quad \frac{N_3}{N_1} = \frac{g_3}{g_1} \frac{e^{-h\nu_{31}/kT}}{A_{31}/N_e Q_{31} + 1} \quad (4.47)$$

$$\rightarrow \frac{N_3}{N_2} = \frac{g_3}{g_2} \frac{A_{21}/N_e Q_{21} + 1}{A_{31}/N_e Q_{31} + 1} e^{-h(\nu_{31}-\nu_{21})/kT} \quad (4.48)$$

If the line emissivities of two collisional excited lines differ in their temperature or density dependence, then they can be used for determining T_e and N_e in the emitting gas. For this we use the relation between level population ratio N_3/N_2 and line emissivity ratio (with $\epsilon_{nm} = (1/4\pi) h\nu_{nm} A_{nm} N_n$):

$$\frac{\epsilon_{31}^\ell}{\epsilon_{21}^\ell} = \frac{\nu_{31} A_{31}}{\nu_{21} A_{21}} \frac{N_3}{N_2} \quad (4.49)$$

Density determination:

- ideal for $\nu_{31} - \nu_{21} \approx 0$, $\rightarrow T_e$ -dependence small,
- gas density N_e between $N_{k,2}$ and $N_{k,3}$,
- critical densities for the two transitions differ $N_{k,2} \neq N_{k,3}$,
- low and high density regimes: (with simplification $e^{-h(\nu_{31}-\nu_{21})/kT} = 1$)
 - low density: $N_e \ll N_{k,2}, N_{k,3}$ (or: $A_{nm}/N_e Q_{nm} \gg 1$)

$$\frac{N_3}{N_2} = \frac{g_3}{g_2} \frac{Q_{31}/A_{31}}{Q_{21}/A_{21}} \rightarrow \frac{\epsilon_3^\ell}{\epsilon_2^\ell} = \frac{g_3 Q_{31}}{g_2 Q_{21}} = \frac{\Omega_{13}}{\Omega_{12}} \quad (4.50)$$

- high density: $N_e \gg N_{k,2}, N_{k,3}$ (or: $A_{nm}/N_e Q_{nm} \ll 1$)

$$\frac{N_3}{N_2} = \frac{g_3}{g_2} \rightarrow \frac{\epsilon_3^\ell}{\epsilon_2^\ell} = \frac{g_3 A_{31}}{g_2 A_{21}} \quad (4.51)$$

- Slide 4–10 illustrates the density determination using the [O II] and [S II] lines.

Figure 4.9: Schematic illustration for the density determination (left) using two emission lines with essentially identical excitation energy (right) but with different critical density

Temperature determination:

- ideal for the temperature determination is a 3-level atom with a large difference in the excitation energies $\nu_{31} - \nu_{21}$, so that the temperature dependence for the line ratio becomes large,
- the gas density should be for both transitions either in the low or high density regime, $N_e \ll N_{k,2}, N_{k,3}$ or the high density regime $N_e \gg N_{k,2}, N_{k,3}$, so that the line ratio is not strongly density dependent. For these conditions, there is:

$$\frac{N_3}{N_2} = \text{const} \cdot e^{-h(\nu_{31}-\nu_{21})/kT}. \quad (4.52)$$

- The constant *const.* is for the low density regime:

$$\text{for } N_e \rightarrow 0 \quad \text{const} = \frac{g_3}{g_2} \frac{Q_{31}/A_{31}}{Q_{21}/A_{21}}, \quad (4.53)$$

- and for the high density regime

$$\text{for } N_e \rightarrow \infty \quad \text{const} = \frac{g_3}{g_2} \quad (4.54)$$

Figure 4.10: Schematic illustration for the temperature determination (left) using two transitions with strongly different excitation states (right) and similar critical densities.

Slide 4–11 illustrates the temperature determination for the [O III] lines. This is an ideal case because level $n = 3$ decays to $n = 2$ and produces a line in the same wavelength range as the decay $2 \rightarrow 1$. For “real” emission line ratios there exists often a simultaneous dependence on temperature and density. In addition, the derived values are only valid for one ion and they represent some sort of average for the observed (usually inhomogeneous) emission line region of that line. For this reason, the T_e and N_e determination for a nebula should be based on all diagnostic lines available. Slide 4–12 shows the emission line spectrum for the Orion nebula, a typical H II region.

4.7.3 Collisionally excited molecular lines

We consider here only the very important molecules H_2 and CO as examples for the excitation of molecules in the interstellar medium. There exist the following transitions between different energy states of molecules:

- rotational transitions: $J' - J''$
energy levels of a rotator: $E_J \propto J(J+1)$; $\Delta E_J \approx 0.01$ eV
- (ro)-vibrational transitions: $\nu' - \nu''$
energy levels of a harmonic oscillator: $E_\nu \propto \nu + 1/2$; $\Delta E_\nu \approx 0.3$ eV
- electronic transitions: $n'L' - n''L''$
energy levels $E_n \propto -1/n^2$; $\Delta E_n \approx 10$ eV

Selection rules for the angular momentum change of a molecule due to allowed (dipole) transitions are:

$\Delta J = \pm 1$ or $\Delta L = \pm 1, 0$ (but not $L' = 0 - L'' = 0$).

schematic energy level diagram for a 2-atomic molecule:

transition	H_2	CO
rotational transitions		
dipole-transitions		
$J = 1 \rightarrow 0$	—	2.60 mm, $A = 7.2 \cdot 10^{-8} \text{s}^{-1}$
$J = 2 \rightarrow 1$	—	1.30 mm, $A = 6.9 \cdot 10^{-7} \text{s}^{-1}$
$J = 3 \rightarrow 2$	—	0.65 mm, $A = 2.5 \cdot 10^{-6} \text{s}^{-1}$
quadrupole transitions		
$J = 2 \rightarrow 0$	28.2 μm , $A = 2.9 \cdot 10^{-11} \text{s}^{-1}$	—
$J = 3 \rightarrow 1$	17.6 μm , $A = 4.8 \cdot 10^{-10} \text{s}^{-1}$	—
ro-vibrational transitions		
dipole transitions		
$\nu = 1 \rightarrow 0, J = 1 \rightarrow 0$	—	4.7 μm ,
$\nu = 2 \rightarrow 0, J = 1 \rightarrow 0$	—	2.3 μm ,
quadrupole transitions		
$\nu = 1 \rightarrow 0, J = 2 \rightarrow 0$	2.12 μm ,	
$\nu = 2 \rightarrow 0, J = 2 \rightarrow 0$	μm ,	
electronic transitions		
e.g. $^1\Sigma_u^+ \rightarrow ^1\Sigma_g^+$	UV (≈ 100 nm) $A \approx 10^7 \text{s}^{-1}$	UV (≈ 100 nm) $A \approx 10^7 \text{s}^{-1}$

4.7.4 Recombination lines: excitation through recombination

A recombination process involves the collision of a free electron with an ion X^{i+1} forming together an ion X^i . After this process, the ion (electron) may be in the ground state X_g^i or in the excited state X_n^i . The excited state can then decay to lower states by an emission of a line photon. Alternatively it may also be possible, but very unlikely for the low density in the interstellar medium, that the excited state is either re-ionized or deexcited by a collision, before a recombination line photon is emitted. Slides 4–13 and 4–14 show the levels and transitions of H I and He I, which are excited by recombination. The excited states X_n^i and corresponding lines of an ion, which are populated by recombination, can usually be distinguished from collisional excited lines.

There are two main recombination processes:

radiative recombination: $X^{i+1} + e^- \rightarrow X^i + h\nu$

This is the predominant process in the interstellar medium.

3-body recombination: $X^{i+1} + e^- + e^- \rightarrow X^i + e^-$

This process is only important in high density gas, like stellar atmospheres, and can be neglected for the interstellar gas, because there are two electrons involved.

Emissivity for recombination lines: The emissivity for recombination lines depends on the recombination rate for radiative recombination, which can be described by

$$\epsilon^\ell = \frac{1}{4\pi} h\nu_{nm} \alpha_{nm}^{\text{eff}} N_e N(X^{i+1}). \quad (4.55)$$

The emissivity per volume element is proportional to the density squared. The effective recombination coefficient for a recombination line $\alpha_{nm}^{\text{eff}}(T_e, N_e)$ (units $[\text{cm}^3/\text{s}]$) considers the population of the upper level n through the following processes:

- recombination directly into the level n
- cascades into level n from higher levels, which are populated by recombinations
- collisional transitions into level n from other level, which were also populated by recombinations.

The **temperature dependence** of the recombination lines behaves roughly like

$$\alpha_{nm}^{\text{eff}}(T_e) \propto 1/T. \quad (4.56)$$

This can be explained by the fact, that slow electrons have a higher chance to be captured by an ion.

Hydrogen and helium recombination lines The strongest recombination lines from diffuse gas regions are the lines from H I. In the visual range are the Balmer-transitions (transitions $n - 2$), in the near-IR the Paschen ($n - 3$), e.g. visible in the Orion spectrum in Slide 4–12, and Brackett lines ($n - 4$) and in the far-UV the Lyman lines ($n - 1$).

The He I and He II recombination lines are significantly weaker than the H I lines. One important factor is the abundance of helium which is typically 10 times lower. In addition, the He I energy levels are not degenerate for levels with different orbital angular momentum and further there are different levels for the singlet (electron spins anti-parallel) and the

triplet states (spins parallel). Due to this there are many more but rather weak He I-lines in the spectrum (Slides 4–13 and 4–14).

Recombination lines from heavy elements are in astronomical objects again much weaker than the H I lines mainly due to the much lower abundance of these elements.

4.7.5 Absorption lines

Absorption lines are a very important source of information for the investigation of the interstellar medium. The atoms and ions in the diffuse gas are predominately in the ground state, because of the low density in the ground state. For this reason the interstellar gas produces essentially only absorptions from resonance lines. These are the absorptions by allowed transitions from the ground states of atoms and ions.

the strongest absorption lines are the resonance lines of atomic and molecular hydrogen H I and H₂ in the far-UV between 912Å and 1215Å (see Slides 4–15 to 4–17).

For the heavy elements the strongest lines are often the doublet-transitions ²S–²P of the isoelectronic sequences of Li, Na, and K. Some important absorption lines are:

Li-sequence	Na-sequence	K-sequence
C IV $\lambda\lambda 1548, 1551$	Na I $\lambda\lambda 5990, 5996$	Ca II $\lambda\lambda 3934, 3968$
N V $\lambda\lambda 1239, 1243$	Mg II $\lambda\lambda 2786, 2803$	
O VI $\lambda\lambda 1032, 1038$	Al III $\lambda\lambda 1855, 1863$	
	Si IV $\lambda\lambda 1394, 1403$	

Line strength. The strength of the line absorption coefficient κ^ℓ is defined by atomic parameters and the volume density of the absorbing atom (or ion) N_m in state m :

$$\kappa^\ell = \int \kappa_\nu^\ell d\nu = \frac{1}{c} h\nu_{nm} N_m B_{mn}. \quad (4.57)$$

The contribution from the stimulated emission ($-N_n B_{nm}$) can be neglected for most cases. The important point in this equation is the fact that the strength of the line absorption coefficient is proportional to the density of the absorbing particle.

The absorption coefficient is often expressed with oscillator strength f_{mn} , a description which comes from classical electrodynamics:

$$\kappa^\ell = \frac{\pi e^2}{m_e c} f_{mn} N_m. \quad (4.58)$$

The relation between oscillator strength and Einstein B coefficient is:

$$f_{mn} = (m_e h\nu / \pi e^2) B_{mn}.$$

The total line absorption follows through the integration of all particles along the line of sight and considering the frequency dependence of the line profile:

$$\int \int I_\nu \kappa_\nu^\ell(N_m) d\nu ds \quad (4.59)$$

It is often difficult to determine from the observed line absorption the column density $\int N_m ds$. This problem exists because the absorptions saturate, so that the absorption line depths in the spectrum is far from a linear relationship to the column density. For a meaningful interpretation a detailed analysis of the line structure is required.

Line profile structure

The line absorption depends usually strongly on frequency (or wavelength). Different effects play a role for the line structure:

The natural line profile. The natural line profile describes the line structure of a transition with frequency ν_0 of an atom, ion or molecule at rest. The line profile can be described by the Lorentz profile:

$$\psi_L(\nu) = \frac{1}{\pi} \frac{\Gamma/4\pi}{(\nu - \nu_0)^2 + (\Gamma/4\pi)^2} \quad (4.60)$$

Γ is the transition rate, $1/\Gamma$ the life time of the two levels. The natural line width for allowed transitions is of the order $\Delta\lambda_n \approx 10^{-4}\text{\AA}$ (wavelength independent). The natural line width is extremely small when compared to the Doppler effect caused by the kinetic and dynamic motion of the gas. For this reason a pure Lorentz profile is rarely used in astrophysics.

The Doppler profile. The Doppler profile is used for absorption lines which are weak or have an intermediate strength. The Doppler profile takes the kinetic motion of the absorbing particles in the gas cloud into account:

$$\psi_D(\nu) = \frac{1}{\Delta\nu_D\sqrt{\pi}} e^{-(\nu-\nu_0)^2/\Delta\nu_D^2} \quad (4.61)$$

The structure of the Doppler profile is a Gauss curve. $\Delta\nu_D$ is the Doppler width which follows from the velocity dispersion σ_v of the absorbing particle

$$\Delta\nu_D = \frac{\sigma_v}{c}\nu_0 \quad (4.62)$$

due to the kinetic velocity as defined by the Maxwell-Boltzmann velocity distribution. For a given temperature the Doppler width is:

$$\sigma_v = \left(\frac{2kT}{m}\right)^{1/2} = 12.9 \left(\frac{T[\text{K}]/10^4}{A}\right)^{1/2} \text{ km/s}. \quad (4.63)$$

Sometimes, the turbulent motion of the gas is included in this Doppler profiles.

The Voigt profile. The Voigt profile must be used for very strong absorption lines. This profile considers the fact, that the line wings defined by $(\nu - \nu_0) > \Delta\nu_D$ decrease faster for the Doppler profile than for the Lorentz profile. The Doppler profile falls off exponentially, but only quadratically for the Lorentz-profile. The Voigt profile is simply a more general line profile description which folds together the Lorentz and the Doppler profile:

$$\psi_V(\nu) = \frac{1}{\Delta\nu_D\sqrt{\pi}} \frac{\Gamma}{4\pi^2} \int_{-\infty}^{\infty} \frac{e^{-(\Delta\nu)^2/\Delta\nu_D^2}}{(\nu - \nu_0 - \Delta\nu)^2 + (\Gamma/4\pi)^2} d(\Delta\nu) \quad (4.64)$$

Multiple components. Often different components of a line absorption are observed which are displaced in the spectrum. This can be due to clouds located at different

distances and having different radial velocities. It is often hard to find out from absorption line observations which component is closer to the observer.

Spectroscopic observations:

A high spectral resolution is required to measure with sufficient accuracy the structure of interstellar absorption lines. The high spectral resolution requires that the background source is bright and has intrinsically no narrow lines. Thus it is only possible to probe with an absorption line analysis only certain line of sights towards well suited background objects, which are:

- bright, hot stars with broad lines (fast rotators) for interstellar absorptions,
- bright quasars for intergalactic absorptions.

Line equivalent width. A very basic quantity for the characterization of an absorption line is the equivalent width W_λ . W_λ measures the strength of a line absorption in the spectrum I_λ . W_λ measures the area in the spectrum between the normalized flux $I_\lambda^n = I_\lambda/I_{\text{cont}}$ and the normalized continuum flux $I_{\text{cont}} = 1$ (area in units of Å or nm). Expressed as mathematical formula:

$$W_\lambda = \int_{\text{Linie}} (1 - I_\lambda^n) d\lambda. \quad (4.65)$$

The absorption depth at a given wavelength (or frequency) is defined by the optical depth τ :

$$W_\lambda = \int (1 - I_\lambda^n) d\lambda = \frac{\lambda^2}{c} \int (1 - e^{-\tau_\nu}) d\nu \quad (4.66)$$

while the optical depth is the absorption coefficient integrated along the line of sight:

$$\tau_\nu = \int \kappa_\nu^\ell ds \quad (4.67)$$

Curve of growth

Weak lines. For weak absorption lines $I_\lambda^n \gtrsim 0.8$ the approximation $1 - e^{-\tau} \approx \tau$ is applicable and the equivalent width is:

$$W_\lambda \approx \frac{\lambda^2}{c} \int \int \kappa_\nu^\ell d\nu ds = \frac{\lambda^2}{c} \frac{\pi e^2}{m_e c} f_{mn} \int N_m ds. \quad (4.68)$$

The curve of growth is for weak absorption lines in the “linear regime”. This means that each contribution to the line absorption (column density) produces an enhancement of the equivalent width independent of the wavelength of the absorption. This is equivalent to the statement that the equivalent width is proportional to the column density for weak lines (Fig. 4.11). The following relationship for λ and W_λ expressed in Å can be used:

$$\int N_m ds = \frac{1.13 \cdot 10^{20}}{\lambda^2[\text{Å}] f_{nm}} \cdot W_\lambda[\text{Å}] \quad (4.69)$$

Saturated lines. Stronger lines saturate in the Doppler core and the line can only grow in the Doppler wings if there are more absorbing particles. The equivalent width changes not much if the column density is enhanced because of the exponential decrease of the absorption coefficient in the line wings. The equivalent width approaches a limiting value which is proportional to the line width of the Doppler profile.

In this regime, one has to consider in the line analysis, that a displaced wavelength component, e.g. due to a cloud with different radial velocity, can produce a significant contribution to the equivalent width, while an additional component in the line center has no effect on the line profile.

Figure 4.11: Schematic illustration of the curve of growth for absorption lines.

Damped absorption lines. In very strong absorption lines the damping wings become visible which grow with increasing column density. These lines are called damped absorption lines. The damped profiles are due to the natural line profile for which the absorption coefficient decreases only quadratically with distance from the line center. Although the line wings of the natural line profile are very weak, they are in the far line wing still stronger than the exponentially decreasing Doppler wings and become visible for very strong absorption lines. Lines in which the damping wings dominate are in the regime where the equivalent width is proportional to the square of the column density:

$$\int N_m ds \propto W_\lambda^2 \quad (4.70)$$

Example: H I Ly α :

$\lambda = 1215 \text{ \AA}$, $f_{\text{Ly}\alpha} = 0.41$

– for unsaturated (optically thin) lines ($W_\lambda \ll 0.3 \text{ \AA}$) there is:

$$\int N(\text{H I}) ds = 1.8 \cdot 10^{14} \text{ cm}^{-2} \cdot W_\lambda[\text{\AA}], \quad (4.71)$$

– for damped absorptions ($W_\lambda \gg 1 \text{ \AA}$) there is:

$$\int N(\text{H I}) ds = 1.9 \cdot 10^{18} \text{ cm}^{-2} \cdot (W_\lambda[\text{\AA}])^2 \quad (4.72)$$

Example: H₂-molecular absorptions

H₂ is a special case: Only the singlet states $^1\Sigma_g$ with anti-parallel spins for the electrons are stable. In addition the Pauli principle requires that the quantum states of the H₂ systems are anti-symmetric (non-exchangeable).

→ there are two types of molecular hydrogen H₂ depending on the relative orientation of the nuclear spins (see Fig. 4.12):

- **para-H₂**: nuclear spin antiparallel, J even, statistical weight $J(J+1)$
- **ortho-H₂**: nuclear spin parallel, J odd, statistical weight $3J(J+1)$

For this reason there are no allowed dipole-transitions between the different rotation and vibrational states of the ground level. A dipole-transition requires $\Delta J = \pm 1$, but such a transition would also require a nuclear spin flip which is not possible.

Figure 4.12: Energy level diagram for molecular hydrogen.

Electronic transitions from the ground state are possible, because the symmetry requirement (Pauli principle) does not apply if the principle quantum numbers of the two electrons are different. e.g. $^1\Sigma_g - ^1\Sigma_u$ with $\nu' (= 0) - \nu''$ und $J' - J'' (= J' \pm 1)$

These electronic transitions produce very strong H₂ Lyman- und Werner bands in the far UV.

Temperature determination using H₂ absorption lines. Collisional processes dominate the level population for the lowest states of H₂, because the radiative transitions are forbidden between the rotational states. The population of the level $N_{J,\nu} = 0$ are therefore given by the Boltzmann-equation:

$$N_J \propto g_J e^{-E_J/kT} \quad (4.73)$$

The observations of the strength of H₂ far-UV lines yields therefore a good estimate for the temperature in molecular clouds.

4.8 Free-bound and free-free radiation processes

4.8.1 Recombination continuum

The electron, which is captured in a radiative recombination process, emits a photon in particular wavelength regions, which are characteristic for the recombining atom or ion. This radiation is well visible for the H I recombination. The recombination continuum may also be seen for He I and He II in high quality spectra.

The energy of the emitted photon is defined by the kinetic energy (relative to the recombining ion) of the capture electron and the energy difference between the ionization energy (usually set to zero) and the (negative) energy of the bound state into which the electron is captured initially:

$$h\nu = \frac{1}{2}m_e v_e^2 - \chi_n \quad (4.74)$$

For hydrogen this is $\chi_n = -Ry/n^2$. Recombination into level n produce according to this equation photons with an energy of at least $-\chi_n$ or more. This produces characteristic discontinuities in the spectrum of photoionized regions. The strongest case is the Balmer jump at 3648 Å.

Figure 4.13: Wavelength dependence of the Recombination continua for a hot and a cold emission nebula.

The emissivity for the recombination continuum can be calculated from the following formula:

$$j_\nu = \frac{1}{4\pi} N_e N(X^{+m+1}) \sum_{n,L} \int_{v_e} v_e \sigma_{n,L}(X^{m-1}, v_e) f(v_e, T_e) h\nu(v_e) dv_e. \quad (4.75)$$

The meaning of the different terms are:

- $1/4\pi$: considers emission in all directions
- $N_e N(X^{+m+1})$: density of the particles involved
- $\sum_{n,L}$: summation over all levels
- $\int_{v_e} f(v_e, T_e) dv_e$: integration for a Maxwell distribution of electrons
- v_e : number of interaction is proportional to the electron velocity
- $\sigma_{n,L}(X^{+m}, v_e)$: cross section for the recombination into level n, L .
In general σ is large for small v_e , thus slow electrons are more frequently captured.
- $h\nu(v_e)$: energy of the emitted photon

Temperature dependence: The intensity jump and the gradient of the recombination continuum depend on the gas temperature. For low temperature the average kinetic energy

of the captured electron is lower and more photons are emitted with an energy just above the “jump” energy. A high temperature gas emits more photons significantly above the jump, so that intensity jump and the gradient are smaller (Fig. 4.13).

low temperature $\rightarrow$ steep continuum and relatively strong jump

high temperature $\rightarrow$ flat continuum and relatively small jump

4.8.2 Photoionization or photo-electric absorption

In a photo-ionization process a photon “is pulling out” an electron from an atom or ion:



The liberated electron has after the process a kinetic energy which is equal to that part of the photon energy, which was beyond the ionization energy χ_{ion} :

$$e^-(E_{\text{kin}}) = h\nu - \chi_{\text{ion}}. \quad (4.77)$$

This extra energy is in photo-ionized regions the most important energy source for the heating of the gas.

Absorption cross section. The photo-ionization cross section a_ν is zero $a_\nu = 0$ for photon energies below the ionization energy of a given atomic state. The cross section has a maximum value at the ionization energy and for higher energies the cross section decreases typically like (see Fig. 4.3):

$$a_\nu \propto \nu^{-3}. \quad (4.78)$$

The H I ionization edge . The H I ionization edge at 912 Å, or 13.6 eV is most important for the interaction between the radiation field and the ISM / IGM. If radiation above 13.6 eV is present, then the gas can be ionized and become transparent for ionizing radiation. If no radiation $E > 13.6 \text{ eV}$ is present then the gas becomes neutral and opaque for the ionizing radiation.

The H I ionization edge defines further two types of neutral elements. Elements which have an ionization edge below 13.6 eV and can be ionized by UV-photons with $h\nu < 13.6 \text{ eV}$. These elements are also in the neutral H I-regions often ionized, e.g.:

– C II, Mg II, Si II, Ca II, Fe II.

Atoms with $\chi_{\text{ion}} \geq 13.6 \text{ eV}$ are neutral when hydrogen is neutral, e.g.:

– He I, N I, O I, Ne I.

The absorption cross section of hydrogen and helium are small for high photon energies $h\nu > 100 \text{ eV}$, in the soft (= low energy) X-ray range. For these energies one has to consider also photoelectric absorptions from heavy element despite their low abundance. In many electron atoms, X-ray photons can be absorbed efficiently by inner shell electrons (K- or L-shell). This produces discontinuities in the photoelectric absorption cross sections. The more abundant heavy elements dominate the interstellar absorption in the soft X-ray range due to K- and L-shell electron absorptions (Slide 4–18).

The averaged photo-electric absorption in the soft X-ray range is quite universal for **neutral** interstellar gas. The strength of the X-ray absorptions is essentially identical for atomic or molecular gas, or for neutral gas with dust particles.

The photo-ionization cross section is of course strongly reduced for highly ionized gas. Due to this, soft X-rays can also be observed for extra-galactic sources for sight lines perpendicular to the Milky Way disk.

4.9 Free-free radiation processes or bremsstrahlung

4.9.1 Radiation from accelerated charges

Whenever a charged particle is accelerated or decelerated it emits electromagnetic radiation. If this radiation is created by the interaction of fast electrons with atomic nuclei then it is called bremsstrahlung. In atomic physics this process is called free-free emission because the radiation corresponds to transitions between unbound states in the field of a nucleus.

The following scheme illustrates the origin of the radiation from an accelerated charged particle (from M.S. Longair, High Energy Astrophysics). The field lines are shown for a particle that suffers a small acceleration Δv . The electric field lines inside a sphere with radius $r = ct$ already “know” that the charge has moved, while the field lines outside this sphere have still the configuration from before the kick. In a shell with thickness $c\Delta t$ there must be an electric field component in i_ϕ or tangential direction. This “pulse” of transverse electromagnetic field propagates away from the charge with speed of light and represents the radiation from the accelerated charge (Slide 4–19).

The total power emitted from a single accelerated charge q is given by Larmor’s formula:

$$P = \frac{2q^2|\dot{\vec{v}}|^2}{3c^3}. \quad (4.79)$$

This formula is valid for any form of acceleration (including charges moving in magnetic fields).

The emitted radiation from an accelerated particle has the following properties:

- the radiated energy is proportional to $P \propto q^2|\dot{\vec{v}}|^2$,
- the radiated energy has an angle dependence like a dipole $dP/d\Omega \propto \sin^2\theta$ where θ is the polar angle with respect to the acceleration vector ,
- the radiation is polarized with electric field vector parallel to the acceleration vector

Radiation spectrum. The spectrum of the emitted radiation depends on the time variation of the electric field. A regularly oscillating field (e.g. from a bound electron, or from a rotating or vibrating molecule) produces a line at a given wavelength or frequency. The frequency spread of this line, or the natural line width, is defined by the energy uncertainty principle $\Delta E \Delta t > h/2\pi$ or if we insert the energy of the emitted photon $E = h\nu$:

$$\Delta\nu \Delta t > \frac{1}{2\pi}. \quad (4.80)$$

If a charge, say an electron, is accelerated in the electric field of an ion then the radiation pulse is extremely short. Lets assume an electron with a relative speed of $v_T(10 \text{ K}) = 550 \text{ km/s}$ is accelerated by an an atom with a dimension of $1 \text{ \AA}$, then the pulse duration is on the order 10^{-16} sec . This implies that the frequency (or energy) of the emitted photon is essentially unconstrained. For this reason, the free-free radiation is essentially frequency independent.

4.9.2 Thermal bremsstrahlung

Bremsstrahlung or free-free emission is produced by Coulomb collisions of electrons e^- with ions (p^+ , He^+ , etc.). In these collisions, charged particles are strongly accelerated so that radiation pulses are created. Most efficient is the acceleration of electrons in the field of ions. The frequency spectrum of a short pulse (Fourier transformation of a delta-function) is broad band. Thus the resulting spectrum I_ν is essentially flat for low frequencies ν and it has an exponential cut-off at the high frequency end. The exponential cut-off is defined by the kinetic energy distribution of the electron, which is for a thermal gas defined by the Maxwell velocity distribution, which has also an exponential cut-off. Essentially, there can be no photons emitted with an energy higher than the kinetic energy of the accelerated particle.

The emission coefficient has the following temperature and density dependence:

$$j_\nu \propto \frac{N_i N_e}{\sqrt{T}} e^{-h\nu/kT}. \quad (4.81)$$

The exact formula is:

$$j_\nu = 5.44 \cdot 10^{-39} g_{\text{ff}} z_i^2 \frac{N_i N_e}{\sqrt{T}} e^{-h\nu/kT} \text{erg cm}^{-3} \text{s}^{-1} \text{Hz}^{-1}, \quad (4.82)$$

where $g_{\text{ff}}(T_e, z_i, \nu) \approx 1 - 2$ is the Gaunt-factor, a quantum-mechanical correction factor to the classical formula, z_i is the charge for the ion ($= 1$ for a hydrogen nebula).

The characteristic energy or wavelength for the exponential cut-off is given by $-h\nu/kT = 1$. It is:

- for $T = 10^4$ K at $\lambda = 1.4 \mu\text{m}$ (warm photo-ionized gas)
- for $T = 10^7$ K at $\lambda = 14 \text{\AA} \approx 0.9 \text{keV}$ (hot collisionally ionized gas)

In total the energy radiated by bremsstrahlung is obtained through integration over all frequencies:

$$\epsilon_{\text{ff}} \propto N_i N_e \sqrt{T}, \quad (4.83)$$

or exactly: $\epsilon_{\text{ff}} = 1.43 \cdot 10^{-27} z_i^2 \langle g_{\text{ff}} \rangle N_i N_e \sqrt{T} \text{erg cm}^{-3} \text{s}^{-1}$ where N_i , N_e are particle densities per cm^3 and T_e the electron temperature. Free-free emission dominates the cooling of collisionally ionized gas if $T_e > 10^6$ K.

Free-free absorption coefficient

A gas which emits free-free radiation becomes for frequencies low enough $\nu < \nu_0$ optically thick. This fact follows from the Kirchhoff law, which defines the Planck radiation $B_\nu(T)$ as the maximum source function for a thermal gas with temperature T :

$$S_\nu = \frac{j_\nu}{\kappa_\nu} \leq B_\nu(T) \quad (4.84)$$

The emissivity j_ν for the free-free radiation is for low frequencies essentially frequency-independent, while the Planck radiation decreases for $\nu \rightarrow 0$ (with $e^{h\nu/kT} \rightarrow 1 + h\nu/kT$) like:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \rightarrow \frac{2\nu^2 kT}{c^2} \quad (4.85)$$

According to the Kirchhoff law $\kappa_\nu = j_\nu/B_\nu$ the free-free absorption coefficient is:

$$\kappa_\nu = j_\nu \frac{c^2}{2\nu^2 kT} \propto \frac{N_i N_e}{T^{3/2}} \frac{1}{\nu^2} \quad (4.86)$$

Thus, for high frequencies $\nu \rightarrow \infty$ the free-free absorption becomes rapidly very small.

Figure 4.14: Wavelength dependence of the radio continuum from an ionized nebula.

Result (Fig 4.14):

- for high frequencies the emission is optically thin and the observed radiation flux is:

$$I(\nu) \propto \int j_\nu ds \propto \int N_i N_e \frac{e^{-h\nu/kT}}{\sqrt{T}} ds \quad (4.87)$$

- for low frequencies the free-free emission is optically thick and the radiation flux is defined by the black-body radiation with a temperature of T , thus:

$$I(\nu) \propto \nu^2 T. \quad (4.88)$$

Approximation for relativistic bremsstrahlung. In collisionally ionized hot gas $T > 10^8$ K, as observed in rich clusters of galaxies, the thermal electrons may reach relativistic velocities $v_T = 55'000$ km/s. For such gas a relativistic correction is necessary. A simple formulation of this effect for the total bremsstrahlung emission is:

$$\epsilon_{\text{ff}} = 1.43 \cdot 10^{-27} z_i^2 \langle g_{\text{ff}} \rangle N_i N_e \sqrt{T} (1 + 4.4 \cdot 10^{-10} T[K]) \text{ erg cm}^{-3} \text{ s}^{-1} \quad (4.89)$$

The term in brackets is the relativistic correction which is only relevant for very high temperatures.

4.10 Compton and Thomson scattering

In Compton scattering energy and momentum of a photon is transferred to the scattering particle, usually an electron. This process is for the electron gas not so important, because the hard radiation field is too weak in the interstellar space to produce any additional gas heating via this process.

More important is the inverse effect, **inverse Compton scattering**. In this process a moving electron transfers kinetic energy to the scattered photon so that the energy distribution of the radiation is changed. This process provides an important diagnostic tool for the investigation of the electron energies (or gas temperature) in hot gas. This is particularly important for fully ionized gas, where no or hardly any atomic emission lines are emitted.

Thomson scattering: For low energies, $h\nu \ll m_e c^2$ (or $E_\gamma \ll 511$ keV) one can use the classical Thomson scattering cross section for photon scattering. In this case the scattering is essentially elastic in the frame of the electron.

The total cross section is:

$$\sigma_e = \frac{8\pi}{3} r_0^2 = 5.56 \cdot 10^{-25} \text{ cm}^2 \quad (4.90)$$

where $r_0 = e^2/m_e c^2$ is the classical electron radius.

- The scattering cross section has a angle dependence like $d\sigma/d\Omega = r_0^2/2 (1 + \sin^2 \theta)$. Thus the scattering favors forward and backward scatterings.
- The scattered radiation is linearly polarized even for unpolarized incoming radiation. The polarization is 100 % for right angle scatterings, with an orientation perpendicular to the scattering plane. The polarization degree is given by:

$$\Pi = \frac{1 - \cos^2 \theta}{1 + \cos^2 \theta} \quad (4.91)$$

Thomson scattering is a dipole-type scattering process and the polarization can be understood like the emission of an oscillating particle which was disturbed (accelerated) by an oscillating radiation field.

Thomson scattering can form a significant wavelength independent opacity source for astrophysical plasmas.

Compton scattering: In Compton scattering (photon energy $h\nu \approx m_e c^2$) energy and momentum is transferred from the photon to the electron (assumed to be at rest). The wavelength change $\lambda_2 - \lambda_1$ for the photon in a Compton scattering (e.g. Tipler) follows from the conservation of energy and momentum in an inelastic collision.

$$\lambda_2 - \lambda_1 = \frac{h}{m_e c} (1 - \cos \theta) . \quad (4.92)$$

This is equivalent to a relative photon energy loss of:

$$\frac{\epsilon_1 - \epsilon_2}{\epsilon_2} = \frac{\epsilon_2}{m_e c^2} (1 - \cos \theta) \quad (4.93)$$

When averages are taken over the scattering angle θ then the net loss for the photon field, or the energy increase for the electron gas is:

$$\left\langle \frac{\Delta\epsilon}{\epsilon} \right\rangle = \frac{h\nu}{m_e c^2} \quad (4.94)$$

Inverse Compton scattering: Electrons with kinetic motion can also transfer energy to photons via the Doppler effect. However, for a cold gas with slowly moving electrons, there are equal rates for “positive” and “negative” Doppler shifts. In a hot gas, where the electrons move fast (relativistically) there exists a second order effect (the fast moving electrons see more photons in the direction of motion), which leads to an enhancement of the average photon energy via electron scattering (inverse Compton scattering).

Without going into details the mean amplification of photon energies per scattering is

$$\left\langle \frac{\Delta\epsilon}{\epsilon} \right\rangle = \frac{4}{3} \left(\frac{v}{c} \right)^2 = \frac{4kT_e}{m_e c^2}, \quad (4.95)$$

where $\langle m_e v^2 \rangle / 2 = 3kT_e / 2$ was used for the kinetic motion of the electrons.

Comptonization: As a result we get the equation which describes the energy exchange between the radiation field and the electron gas through Compton collisions. The energy change of the radiation field is:

$$\frac{\Delta\epsilon}{\epsilon} = -\frac{h\nu}{m_e c^2} + \frac{4kT_e}{m_e c^2} \quad (4.96)$$

This equation defines the conditions under which energy is transferred to and from the photon field:

- if $h\nu = 4kT_e$, then there is no energy transfer
- if $h\nu > 4kT_e$, then energy is transferred from a hard radiation field to the cool gas
- if $h\nu < 4kT_e$, then energy is transferred from hot gas to the radiation field.

Due to the large distance to high energy sources (AGN, X-ray binary stars) the hard radiation field is always strongly diluted. Therefore the first and second cases are not important for the interstellar medium.

4.11 Temperature equilibrium

The gas in the interstellar space is far from a thermodynamic equilibrium. For this reason the equilibrium temperature at a given position depends on various **heating** and **cooling** processes. Detailed computations are required to estimate the equilibrium temperature.

4.11.1 Heating function H for neutral and photo-ionized gas

Important processes for the gas heating in molecular clouds are collisions by cosmic rays (relativistic particles) and the photo-dissociation of molecules. Photoionization by radiation from stars and other UV sources is the dominant heating process for the atomic, diffuse gas and the photo-ionized gas.

Shocks due to supersonic gas motions can in addition heat the gas to high temperatures. In shocks the dynamic energy of a gas cloud is converted into kinetic (internal) energy of the gas. Supersonic gas flows are produced by stellar winds and supernova explosions. The heating by shocks depends strongly on time and the location and is therefore difficult to describe accurately.

Important heating processes are:

- photo-ionization: $h\nu + X^m \rightarrow X^{m+1} + e^-(E_{\text{kin}})$.

The heating per volume element is given by the number of ionizations multiplied by the extra photon energy above the ionization threshold:

$$H = N_{\text{H}^0} \int_{\nu_0}^{\infty} \Gamma_{\nu} h(\nu - \nu_0) a_{\nu} d\nu. \quad (4.97)$$

- photo-dissociation: $h\nu + XY \rightarrow X(E_{\text{kin}}) + Y(E'_{\text{kin}})$
- photo-electric absorption by dust: $h\nu + \text{dust} \rightarrow \text{dust}' + e^-(E_{\text{kin}})$
- collisions with cosmic ray particles: $P_{\text{cr}} + X \rightarrow P_{\text{cr}} + Y_1(E_{\text{kin}}) + \dots + e^-(E_{\text{kin}}) + \dots$

A heating process produces particles with a kinetic energy ($\gg 3kT/2$) and therefore it contributes to the heating of the gas. In each microscopic heating process, one particle takes part. The energy originates from remote sources (e.g. stellar radiation or relativistic particles), which is converted into kinetic energy of the gas:

the heating per unit volume (cm^3) is proportional to the particle density n :

$$\text{heating} = n \cdot H \quad (4.98)$$

As a first approximation the heating H does not depend on gas parameters, like T or n . However, H depends on the intensity of the radiation field or of the cosmic rays.

4.11.2 Cooling of the gas

The cooling of the gas is mainly due to line emission. Bremsstrahlung (free-free radiation) is the dominant cooling process for gas with very high temperatures $T > 10^6$ K. Important cooling processes are:

- collisionally excited lines: $e^-(E_{\text{kin}}) + X_g^m \rightarrow X_i^m \rightarrow X_g^m + h\nu$
- Bremsstrahlung: $e^-(E_{\text{kin}}) + X^m \rightarrow e^-(E'_{\text{kin}}) + h\nu$

Less important gas cooling processes are the thermal emission by dust particles (contributes in particular in molecular clouds to the cooling) and thermal conduction (in regions with strong temperature gradients like shocks).

The basic process for the cooling by **line emission** is, that an atom or molecule is put into an excited state by a collision with another gas particle (e.g. by an electron), from where it returns to the ground state through the emission of a photon.

kinetic energy of the gas $\rightarrow$ inner energy of the particle
 $\rightarrow$ emission of a line photon ($h\nu$)

Bremsstrahlung is emitted by charged gas particles which are accelerated or decelerated by collisions with other gas particles. Also in this process kinetic energy of the gas is transformed into radiation energy.

Radiation energy is produced in all important gas cooling processes by the collision of **two** particles, thus:

the cooling is proportional to the particle density squared n^2 :

$$\text{cooling} = n^2 \cdot \Lambda(T) \quad (4.99)$$

4.11.3 The cooling function $\Lambda(T)$

The cooling of the gas depends strongly on the temperature of the gas. For this reason the efficiency of the gas cooling is described by the cooling function $\Lambda(T)$. In addition there exists also some dependence of the cooling on the elemental abundances which are important in special cases (e.g. early universe or supernova remnants). Since the elemental abundances are rather homogeneous in the Universe the abundance effects can often be neglected.

An efficient cooling requires:

- an abundant particle (e.g. hydrogen H, C, N, O, CO),
- with an excited state having an excitation energy χ within the range of the kinetic energy of the gas particles, thus $\chi \approx kT$,
- and an excited state with a decay time shorter than the typical time interval to the next collision which may de-excite collisionally the particle (this would convert the excitation energy back to kinetic energy).

Order of magnitude values for collisional rates γ are:

- $\gamma \approx 10^{-11} \text{cm}^3 \text{s}^{-1}$ for collisions between neutral particles
- $\gamma \approx 10^{-9} \text{cm}^3 \text{s}^{-1}$ for collisions between a charged and a neutral particle
- $\gamma \approx 10^{-7} \text{cm}^3 \text{s}^{-1}$ for collisions between charged particles

The collisions per second [s^{-1}] and particle are $n \cdot \gamma$.

Important emission lines for the gas cooling.**T < 1'000 K**

The most abundant particle in molecular clouds is H₂. But the gas cooling by H₂ is very inefficient due to the symmetry of this particle and there exists no fast decay from excited rotational levels of H₂ ($A_{2 \rightarrow 0} = 3 \cdot 10^{-11} \text{s}^{-1}$). For this reason the main process for the cooling in molecular clouds is the emission of photons by rotational transitions of CO (see Slide 4–20). The lowest transitions are:

$$\begin{aligned} \text{CO } 2.6 \text{ mm}, J = 1 \rightarrow 0, & \quad A \approx 10^{-2} \text{s}^{-1} \\ \text{CO } 1.3 \text{ mm}, J = 2 \rightarrow 1 & \quad A \approx 10^{-2} \text{s}^{-1} \end{aligned}$$

The cooling of cold, atomic gas is mainly due to lines from fine structure transitions emitted in the far IR, e.g.:

$$\begin{aligned} \text{O I } 63.2 \text{ } \mu\text{m}, {}^3\text{P}, J = 1 \rightarrow 2, & \quad A = 9 \cdot 10^{-5} \text{s}^{-1} \\ \text{O I } 145.5 \text{ } \mu\text{m}, {}^3\text{P}, J = 0 \rightarrow 1, & \quad A = 2 \cdot 10^{-5} \text{s}^{-1} \\ \text{C II } 157.7 \text{ } \mu\text{m}, {}^2\text{P}, J = 3/2 \rightarrow 1/2, & \quad A = 2 \cdot 10^{-6} \text{s}^{-1} \end{aligned}$$

T = 1'000 – 30'000 K

Gas with neutral hydrogen H I can cool through the excitation of H I and the emission of Lyman lines e.g.:

$$\text{H I Ly}\alpha \text{ } \lambda 1215 \text{ \AA}, n = 2 \rightarrow 1, \quad A = 5 \cdot 10^8 \text{s}^{-1}$$

This process is only efficient for gas with high temperature because the excitation energy is rather high for the first excited state $n = 2$: $\chi = 10.6 \text{eV} = 1.7 \cdot 10^{-11} \text{erg} \rightarrow e^{-\chi/kT} \approx e^{-10^5 \text{K}/T}$.

Often, hydrogen is highly ionized and therefore the cooling through neutral hydrogen can be very in-efficient. Efficient for the cooling are different nebular lines from ions which are abundant in ionized nebulae (Slide 4–21). Dependent on the ionization degree of the gas the following lines are important coolants:

$$\begin{aligned} \text{C III]} & \quad [1907], 1909 \text{ \AA}, {}^3\text{P}^o \rightarrow {}^1\text{S}, & \quad A \approx [0.01], 100 \text{s}^{-1} \\ \text{C IV} & \quad 1548, 1551 \text{ \AA}, {}^2\text{P}^o \rightarrow {}^2\text{S}, & \quad A \approx 3 \cdot 10^8 \text{s}^{-1} \\ \text{[N II]} & \quad 6548, 6583 \text{ \AA}, {}^3\text{S} \rightarrow {}^1\text{D}, & \quad A \approx 10^{-3} \text{s}^{-1} \\ \text{[O II]} & \quad 3726, 3728 \text{ \AA}, {}^4\text{S} \rightarrow {}^2\text{D}, & \quad A \approx 10^{-4} \text{s}^{-1} \\ \text{[O III]} & \quad 4959, 5007 \text{ \AA}, {}^3\text{S} \rightarrow {}^1\text{D}, & \quad A \approx 10^{-2} \text{s}^{-1} \\ \text{O VI} & \quad 1032, 1038 \text{ \AA}, {}^2\text{P}^o \rightarrow {}^2\text{S}, & \quad A \approx 4 \cdot 10^8 \text{s}^{-1} \\ \text{[S II]} & \quad 6716, 6731 \text{ \AA}, {}^4\text{S} \rightarrow {}^2\text{D}, & \quad A \approx 10^{-3} \text{s}^{-1} \end{aligned}$$

T = 30'000 – 10⁷ K

Hot (collisionally ionized) gas emits many lines from different, highly ionized atoms. Strong lines are e.g. from ions of the H and He iso-electronic sequences (Slide 4–22), like O VII and O VIII or from the many ionization states of iron (Fe X – Fe XXVI).

T > 10⁶ K

Bremsstrahlung contributes always to the cooling of an ionized gas. At very high temperature, essentially all atoms are fully ionized and line radiation is no more possible. Bremsstrahlung is for this case the dominating gas cooling process. The cooling, equivalent to the radiation emitted by Bremsstrahlung is proportional to $\sqrt{T}$. For fully ionized gas with solar abundances the emitted luminosity per volume element is given by

$$L_{\text{BS}} = 2 \cdot 10^{-27} n_e^2 \sqrt{T} \text{ erg cm}^{-3} \text{s}^{-1}. \quad (4.100)$$

Table 4.2: Summary of the most important heating and cooling processes.

gas type	heating	cooling
molecular clouds	cosmic rays	molecular lines, CO, H ₂ O
cold, neutral gas	UV radiation (stars)	fine structure lines, C II, O I
warm, neutral gas	UV radiation (stars, AGN)	Ly α , nebular lines, [O I], [S II]
photo-ionized gas	UV radiation (stars, AGN)	Ly α , [O II], [N II], [O III]
collisionally ionized gas	shocks	X-ray lines, bremsstrahlung

Strongly simplified, it can be said that the temperature equilibrium for the diffuse gas in the interstellar medium is determined by:

$$n^2 \cdot \Lambda(T) = n \cdot H \quad \text{and} \quad n \cdot \Lambda(T) = H \approx \text{const.} \quad (4.101)$$

The “cosmic” cooling curve. The gas cooling processes are always the same for diffuse gas and one can describe the cooling with an universal cooling curve (Fig. 4.15). This curve illustrates the energy loss by gas cooling processes and it is given in units of [energy cm³/s]. The cooling curves has a major maximum around 10⁵ K where the cooling by atomic lines is most efficient and a smaller bump around 300-1000 K where molecules and atomic fine structure lines are efficient. There is a minimum around 10⁷ K where all atoms are fully ionized so that no line emission is possible.

Figure 4.15: Schematic illustration of the cooling curve.

Table 4.3: Characteristic cooling time scales.

gas type	n [cm ⁻³]	T [K]	Λ [erg cm ³ s ⁻¹]	τ_{th}
molecular gas	10 ⁵	50	10 ⁻²⁶	$7 \cdot 10^6 \text{ s} \approx 80 \text{ d}$
diffuse, ionized gas	10 ⁻¹	10 ⁴	10 ⁻²⁴	$1.4 \cdot 10^{13} \text{ s} \approx 4 \cdot 10^5 \text{ J}$
H II-region	10 ³	10 ⁴	10 ⁻²⁴	$1.4 \cdot 10^9 \text{ s} \approx 40 \text{ J}$
SN-remnant	10	10 ⁶	$2 \cdot 10^{-23}$	$7 \cdot 10^{11} \text{ s} \approx 2 \cdot 10^4 \text{ J}$
coll.ionized gas	10 ⁻³	10 ⁷	10 ⁻²³	see exercise

4.11.4 Cooling time scale

The cooling time scale for the diffuse gas, which is the time required for the cooling of a gas cloud if the heating is switched off, can be roughly estimated from the cooling function according to:

$$\tau_{\text{th}} = \frac{U}{n^2 \Lambda(T)} \approx \frac{nkT}{n^2 \Lambda(T)} = \frac{kT}{n \Lambda(T)} \quad (4.102)$$

(U: kinetic energy (inner energy) of the gas in cm⁻³). As first approximation one may approximate $\Lambda(T) \propto T$. Thus the cooling time scale behaves (very roughly) like $\tau_{\text{th}} \approx 1/n$. Thus:

high density gas cools rapidly, while diffuse, low density gas cools slowly.

Figure 4.16: The spezific cooling curve $\Lambda(T)/T$.

4.11.5 Equilibrium temperatures.

For the Milky Way disk it can be assumed that there exists a very rough pressure equilibrium for the diffuse Gas. Thus a hydrostatic stratification of the gas can be assumed in the direction perpendicular to the disk. In addition we can adopt the (simplified) law for the temperature equilibrium: $n \cdot \Lambda(T) = H \approx \text{const.}$ Based on this we obtain the following, very rough relation between gas pressure, temperature and cooling function:

$$p = nkT = kT \frac{H}{\Lambda(T)} = \text{const.} \quad (4.103)$$

The “specific” cooling function $\Lambda(T)/T \propto 1/p$ is drawn in Fig. 4.16. There exist for a given gas pressure different intersections with the “specific” cooling curve $\Lambda(T)/T$. At these intersections the gas temperature could be in an equilibrium state. However, the temperature equilibrium is only stable for intersections where the gradient of the “specific” cooling curve is positive. For intersections with curve sections having a negative gradient the equilibrium is not stable.

stable equilibrium: $(d(\Lambda(T)/T)/dT > 0)$,

if the temperature is slightly disturbed then the temperature will go back to the equilibrium point:

- increase of $T \rightarrow$ increase of $\Lambda(T)/T \rightarrow$ more cooling
- decrease of $T \rightarrow$ decrease of $\Lambda(T)/T \rightarrow$ less cooling

unstable equilibrium: $(d(\Lambda(T)/T)/dT < 0)$,

after a small temperature disturbance the temperature T will drift away from the equilibrium point:

- increase of $T \rightarrow$ decrease of $\Lambda(T)/T \rightarrow$ less cooling
- decrease of $T \rightarrow$ increase of $\Lambda(T)/T \rightarrow$ additional cooling

The “specific” cooling curve has two stable temperature regimes with a positive gradient. Due to this, there exist two predominant temperatures for the interstellar gas:

cold gas $T < 100$ K and warm gas $T \approx 10000$ K

Hot $T > 10^5$ K gas cannot exist in a stable temperature equilibrium (theoretically). But the cooling time scale for hot gas is often so long (because of the low density), that it can survive for a very long time. Diffuse, hot gas $T > 10^6$ K is therefore the third type of interstellar gas which is frequently present.

The observed parameters of the dominant interstellar components in the Milky Way can be plotted in a density-temperature diagram (Fig. 4.17).

Figure 4.17: Parameters for dominant interstellar components.

The diagram in Fig. 4.17 illustrates the following:

- the gas exists predominately in 3 temperature regimes (cold, warm and hot)
- there exists, very roughly, a pressure equilibrium ($n \cdot T \approx 1000 \text{ K/cm}^3$) for the diffuse gas in the Milky Way

4.12 Dynamics of the interstellar gas

The interstellar matter is not static but moves always and everywhere around. This gas motions is caused by various processes which induce typical gas velocities as follows:

- galactic rotation $v \approx 200$ km/s
- peculiar motion of galaxies $v \approx 500$ km/s
- speed of stellar winds $v \approx 20 - 3000$ km/s
- ejection velocity of supernova explosions $v \approx 10000$ km/s,
- outflows (broad absorption lines) from active galactic nuclei $v \approx 10000$ km/s.

The gas velocities are very often much larger than the sound velocity of the gas $v \gg v_s$. For this reason many important hydrodynamical processes in the interstellar medium are due to **supersonic flows** which produce non-linear effects, in particular **shocks**.

4.12.1 Basic equations for the gas dynamics

For a simple description of gas-dynamical processes in the interstellar medium one can often start for first useful estimates with strongly simplified equations based on the following assumptions:

- $\vec{B} = 0$ no magnetic field :
Neglecting magnetic fields simplifies the treatment of hydrodynamic processes enormously. However, neglecting magnetic fields can be a very critical choice because many hydrodynamic problems may not be understood without magnetic fields. Sometimes it is useful to assume at least that the magnetic field moves with the gas (it is frozen in) and that the field adds just another pressure term $\sim B^2/8\pi$.
- $\vec{E} = 0$ no electric fields :
This is a very good approximation because charged particles e^-, p^+ are abundant and they can move freely.
- viscosity $\eta = 0$:
A very good approximation due to the low density.
- no hydrodynamic coupling between matter and radiation field (no radiation pressure)

The equation of motion:

$$\rho \frac{d\vec{v}}{dt} = \rho \left(\frac{\partial \vec{v}}{\partial t} + \vec{v} \text{grad} \vec{v} \right) = -\text{grad} p - \rho \text{grad} \Phi, \quad (4.104)$$

and the equation for the conservation of mass:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \vec{v} \text{grad} \rho = -\rho \text{div} \vec{v} \quad (4.105)$$

where $\vec{v}$ is the gas velocity, p the pressure, ρ the density and Φ the gravitational potential. Further there is:

- d/dt : the time-derivative in the co-moving coordinate system
→ **Lagrange-system**,

- $\partial/\partial t$: the partial time-derivative for a fixed point in space
→ **Euler-system**.

The variables are:

velocity field $\vec{v}(x_i, t)$	density $\rho(x_i, t)$
gas pressure $p(x_i, t)$	gravitational potential $\Phi(x_i, t)$

These equations alone cannot be solved because there are more variables than equations. This means that some of the functions must be known, for example the gravitational potential Φ or the local energy balance which requires knowledge on the heating and cooling processes.

Energy conservation:

The energy conservation can in general not be expressed as local differential equation, because the heating depends on the interaction with the distant surroundings, e.g. the heating by the absorption of ionizing UV-radiation from stars or the heating by collisions with relativistic cosmic ray particles. Two useful simplifications for first order estimates are:

- **Adiabatic hydrodynamics:** It is assumed, that the energy is conserved locally. This means, that the heating and cooling of the gas is neglected. This is a quite reasonable simplifying assumption for regions with very low densities and long time scales for cooling (e.g. hot, collisionally ionized, diffuse gas).
- **Isothermal hydrodynamics:** In this case a constant temperature is adopted for the gas, e.g. 10'000 K for photo-ionized gas or 100 K for neutral gas. In this approach it is assumed, that the heating due to compression or the cooling due to expansion is compensated immediately by enhanced or reduced radiative cooling. Thus, there is no local energy conservation in this case. This approximation is useful for regions with high gas density where the radiative cooling is very efficient.

Gravitation

For the gravitation the Poisson equation is used:

$$\Delta\Phi = 4\pi G \rho(x_i, t) \quad (4.106)$$

The solution of this equation depends very much on the geometric scale for the distribution of the mass with respect to the size of the gas structure studied.

relatively simple:	The gas-dynamics is solved in a pre-defined and constant gravitational potential, e.g. the motion of the gas in the potential of a galaxy.
very difficult:	The gas-dynamics for a “self-gravitating gas” is very delicate, because $\text{grad}\Phi$ has a dynamic (non-linear) component and small inhomogeneities in the density distribution can grow to large gravitational instabilities for the gas. An important example for this non-linear behavior is the collapse of a gas cloud in a star forming region.

4.12.2 Shocks

The sound velocity v_s defines the propagation velocity of a pressure waves. If there exist supersonic flows $v > v_s$, then hydrodynamic effects occur at a given location without a “preceding warning”. This will produce strong discontinuities in the gas parameters, so-called **shocks**.

Sound velocity. The square of the sound velocity v_s is defined in compressible media by the density derivative $d/d\rho$ of the pressure p :

$$v_s^2 = \frac{dp}{d\rho} \quad (4.107)$$

The adiabatic sound velocity for adiabatic gas $p = (p_0/\rho_0) \cdot \rho^\gamma$ is given by (for an ideal gas):

$$v_s^2 = \frac{dp}{d\rho} = \gamma \frac{p_0}{\rho_0} \rho^{\gamma-1} = \gamma \frac{p}{\rho} = \frac{5}{3} \frac{kT}{m_T} \quad (4.108)$$

$\gamma = 5/3$ for ionized and atomic gas; $\gamma < 5/3$ else

m_T = mean particle mass (e.g. $m_T \approx 0.5 m_p$ for ionized H-gas)

The isothermal sound velocity for isothermal gas $p = \rho \cdot kT/m_T$:

$$v_s^2 = \frac{dp}{d\rho} = \frac{p}{\rho} = \frac{kT}{m_T} \quad (4.109)$$

The sound velocity is of the same order as the mean (mass weighted) kinetic velocity of the particles ($\sim$ protons in ionized gas). Rough estimates for the sound velocity are: $v_s \approx 1$, 10, and 100 km/s for temperature of $T = 100$, 10^4 , and 10^6 K, respectively.

Conservation laws for idealized shocks. The hydrodynamic conservation laws can be used for the description of basic properties of shocks, without studying the complicating processes taking place at the shock fronts. For this we consider one-dimensional flows with a shock front. The parameters p_1 , ρ_1 , and v_1 , stand for the pressure, density, and gas velocity in front of the shock front p_2 , ρ_2 , and v_2 after the shock front. The gas velocities v_1 and v_2 are expressed relative to the velocity of the shock front which is set equal to zero.

Figure 4.18: Illustration of shock parameters.

The so-called **Rankine-Hugoniot conditions** are:

- based on the mass conservation: $\rho_1 v_1 = \rho_2 v_2$
- based on the equation of motion: $p_1 + \rho_1 v_1^2 = p_2 + \rho_2 v_2^2$

In addition we have to consider also the energy budget. The energy budget depends on the treatment for the energy loss due to radiative cooling.

Isothermal shocks. The isothermal shock is a simple model case, in which one assumes that the temperature is identical before and after the shock:

$$T_1 = T_2 \quad (4.110)$$

This assumption requires an extremely efficient cooling, in order to radiate away (instead of heating up the gas) all the energy produced by the work due to the shock. The isothermal shock can be a useful approximation for shocks in high density gas, where the cooling is very efficient. It is also necessary that the gas is optically thin so that the energy can be radiated away.

With the isothermal sound velocity $v_s^2 = p/\rho$ the equations can be solved with the following algebra:

$$\underbrace{p_1}_{\rho_1 v_s^2} + \rho_1 v_1^2 = \underbrace{p_2}_{\rho_2 v_s^2} + \rho_2 v_2^2 \quad (4.111)$$

and:

$$v_s^2 (\rho_1 - \rho_2) = \underbrace{v_2^2 \rho_2}_{v_1^2 \rho_1^2 / \rho_2} - v_1^2 \rho_1 = v_1^2 \frac{\rho_1}{\rho_2} (\rho_1 - \rho_2). \quad (4.112)$$

The result is:

$$\frac{\rho_2}{\rho_1} = \frac{v_1^2}{v_s^2} = M^2, \quad (4.113)$$

where M is the Mach number for the gas flow in front of the shock (in the coordinate system of the shock). For an outside observer this is the Mach number for the shock velocity in the pre-shock medium. The compression or **the density jump in an isothermal shock is proportional to the square of the Mach number**. The Mach number for shocks in the interstellar medium is often very high, e.g. $M \approx 100 - 1000$ for supernovae. The compression is under isothermal conditions very high, on the order $\rho_2/\rho_1 = 10^4 - 10^6$.

Further there is:

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{v_1^2}{v_s^2} \rightarrow v_s^2 = v_1 \cdot v_2, \quad (4.114)$$

which is equivalent to the statement, that the velocity of the post-shock gas is smaller than the sound velocity in the coordinate system of the shock front.

An illustrative description for an isothermal shock is a snow-plough, which piles all material up and carries it away.

Adiabatic shocks. In an adiabatic shock the energy is conserved locally and the work done by the shock front is put into the heating of the post shock gas. The adiabatic shock is a good approximation for very thin gas, where the cooling is not efficient.

For adiabatic shocks we have another equation, besides the Rankine-Hugoniot conditions, which describes the total energy flow through the shock front:

$$v_2 \left(\frac{1}{2} \rho_2 v_2^2 + U_2 \right) - v_1 \left(\frac{1}{2} \rho_1 v_1^2 + U_1 \right) = v_1 p_1 - v_2 p_2 \quad (4.115)$$

where $\rho v^2/2$ is the kinetic energy of the gas, $U = p/(\gamma - 1)$ the inner energy of the gas, and the term on the right side is the work due to the pressure change at the shock front $d/dt(p \cdot A \Delta x) = d/dt(F_A \Delta x)$.

With the Rankine-Hugoniot conditions and a lot of algebra for the case $M \gg 1$ there is:

$$\frac{\rho_2}{\rho_1} \approx \frac{\gamma + 1}{\gamma - 1} \rightarrow \frac{\rho_2}{\rho_1} \approx 4 \quad \text{for} \quad \gamma = \frac{5}{3} \quad (4.116)$$

The density jump is a factor 4 for an adiabatic shock of an ideal gas.

The temperature after the shock follows from the gas equation and the equation of motion:

$$T_2 = \frac{m_H}{k} \frac{p_2}{\rho_2} \quad \text{and} \quad \frac{p_2}{\rho_2} = \underbrace{\frac{p_1}{\rho_2}}_{=0 \text{ für } p_1 \ll p_2} + \underbrace{\frac{\rho_1}{\rho_2}}_{1/4} v_1^2 - \underbrace{v_2^2}_{v_1^2 \rho_1^2 / \rho_2^2 = v_1^2 / 16} \quad (4.117)$$

This gives the result:

$$T_2 \approx \frac{3}{16} \frac{m_H}{k} v_1^2 = 1.4 \cdot 10^7 \text{ K} \left(\frac{v_1}{100 \text{ km/s}} \right)^2 \quad (4.118)$$

Thus the temperature of the post-shock gas in an adiabatic shock is typically on the order $10^6 - 10^8$ K.

More realistic shocks. Observations of shocks show a combination of both cases, the adiabatic and the isothermal shock. The temperature can reach near the shock front a very high temperature and the adiabatic approximation is not bad. Correspondingly one has then a density jump which is not far from the factor 4. Further away from the front in the post-shock region the gas has sufficient time to cool and it approaches more the parameters for gas in an isothermal shock. This means that the gas cools down and becomes quite dense and may be visible as so-called “radiative shock”.

Magnetic fields may also play a role in shocks. Especially, the B-field may be responsible for a magnetic pressure term which can be significant or even a dominant contribution to the total pressure in shocks which are dense and behave like isothermal shocks. Thus, the compression for shocks in dense gas with magnetic fields may be significantly smaller than in isothermal shocks due to the magnetic pressure.



Figure 4.19: Schematic structure for a more realistic shock model.

4.12.3 Example: supernova shells

The velocity and the kinetic energy of a supernova shell is immediately after the explosion enormous:

$$v_{\text{SN}} \approx 15000 \text{ km/s} \quad \text{und} \quad E_{\text{kin}} \approx 4 \cdot 10^{50} \text{ erg} \quad (4.119)$$

This corresponds to the radiation energy which is delivered by our Sun in $3.5 \cdot 10^9$ years.

first phase: free expansion

The first phase is characterized by:

- essentially a gas motion in free space (vacuum) $\rightarrow$ free expansion,
- last until the swept-up mass of the interstellar medium is comparable to the mass of the supernova shell.

An estimate on the swept-up mass may be based on the mean mass density in the Milky Way disc $\rho = 1.6 \cdot 10^{-24} \text{ g cm}^{-3}$ (corresponds to a particle density of $n_{\text{H}} = 1 \text{ cm}^{-3}$). A sphere of diffuse gas with a mass comparable to the supernova shell with a mass of $M_{\text{SN}} \approx 1 M_{\odot}$ has a radius of:

$$\frac{4\pi r^3}{3} \rho = M_{\text{SN}} \quad \rightarrow \quad r = 2 \text{ pc} \approx 6 \text{ Lj}. \quad (4.120)$$

The free expansion phase last with an expansion velocity of $15000 \text{ km/s} = c/20$ about:

$$t = \frac{r}{v} = 120 \text{ years}. \quad (4.121)$$

second phase: adiabatic shock

The density of the supernova-shell is in this phase still relatively small, because not much ISM has been swept-up. Due to the low density the radiative cooling is relatively unimportant and the shock can be approximated by adiabatic conditions:

- The velocity of the pre-shock gas is for a distant observer equal to zero. However, the velocity is relative to the shock front on the order $10'000 \text{ km/s}$.

- The post-shock gas moves with a velocity of 2500 km/s relative to the shock. For an outside observer the velocity of the post-shock gas is 7500 km/s.
- The density inside the shock is about 4 times higher than the density of the ISM, but the temperature is extremely high $> 10^7$ K. Thus in this phase a hot, tenuous bubble is formed.
- This phase lasts about 100-1000 years and the shell radius grows to $r = 1 - 10$ pc

Figure 4.20: Structure of a spherical, adiabatic shock from a supernova shell.

Third phase: isothermal shock

This last phase is characterized by the fact that the supernova shell has accumulated a lot of interstellar material and there will be an isothermal or radiative shock.

- the shock velocity has decreased to $\sim 100 - 3000$ km/s
- there is a huge density jump $\rho_2/\rho_1 = M^2 \approx 10^4 - 10^6$ like for a snow-plough.

The velocity evolution, or the deceleration of the supernova shell can be roughly described by the momentum conservations, considering a spherical shell of gas moving into a thin surrounding gas.

$$M_{\text{SN}} v_{\text{SN}} \approx (M_{\text{SN}} + m_{\text{ism}}) v_{\text{shock}} \quad \text{because} \quad m_{\text{ism}} \propto r^3 \quad \text{thus} \quad v_{\text{shock}} \propto \frac{1}{r^3}. \quad (4.122)$$

This equation describes very roughly the shock velocity v_{shock} and the radius of the supernova shell r which can both be used to make estimates on the age of the supernova shell.

Chapter 5

Star formation

Star formation is a very complicated process which depends on different physical processes. It is often unclear which process is dominant. In the end there seems to result a quite universal initial mass function (IMF) for the star formation. On the other side the star formation occurs only in dense, cold molecular clouds which are strongly concentrated to the midplane of the Milky Way disks. In this section a few topics in star formation are briefly discussed without going into much detail.

5.1 Molecular clouds.

Molecular clouds are overdense regions in the Milky way disk predominantly composed of molecular H_2 and CO and dust. Because they are dense, their dust and gas is self-shielding the cloud from stellar optical and UV-light from the outside. For this reason the molecular clouds can not be seen in the visual except for the fact that they obscure the background objects. The dark irregular bands of absorption in the Milky Way are due to these absorbing clouds. The best way to see molecular clouds are CO line observations, e.g. at $\lambda = 2.6$ mm, in the radio range (see Slides 5–1 and 5–2). The following types of molecular clouds are distinguished:

- Bok globules are small, isolated, gravitational bound molecular clouds of $\lesssim 100 M_\odot$ in which at most a few stars are born,
- molecular clouds have masses of $10^3 - 10^4 M_\odot$ distributed in irregular structures with dimensions of ≈ 10 pc consisting of clumps, filaments bubbles and containing usually hundreds of new-born stars,
- giant molecular clouds are just larger than normal molecular clouds with a total mass in the range $10^5 - 10^7 M_\odot$, dimensions up to 100 pc, and thousands of young stars.

Molecular clouds in the solar neighborhood. The sun resides in a hot (10^6 K), low density bubble with a diameter of ≈ 50 pc. The nearest star forming clouds are located at about 140 pc and because of their proximity they are important regions for detailed investigations of the star and planet formation process. Well studied regions are:

- The **Taurus molecular cloud** at a distance of about 140 pc is a large, about 30 pc wide, loose association of many molecular cores with a total mass of about $\approx 10^4 M_\odot$

and several hundred young stars (Slide 5–1). Because of its proximity there are many well known prototype objects, like T Tau or AB Aur in this star forming region.

- The ρ **Oph cloud** at a distance of 130 pc has a denser gas concentration than Taurus with a main core and several additional smaller clouds and about 500 young stars with an average age of about 0.2 Myr. The total gas mass is about $\approx 10^4 M_\odot$.
- the **Orion molecular cloud** complex has a distance of about 400 pc and a diameter of 30 pc. Orion is the nearest high mass star forming region with in total about 10'000 young stars with an age less than 15 Myr (Slides 5–2 and 5–3). The Orion molecular cloud complex includes the Orion nebula M42 (H II region), reflection nebulae, dark nebulae (Horsehead nebula), an OB associations mainly located in the Belt and Sword of the Orion constellation. The Orion nebula is ionized by the brightest star in the Trapezium cluster.

5.2 Elements of star formation

Stars form in dense, molecular clouds. If regions in clouds become dense enough then they may collapse and form under their own gravitational attraction a sphere which evolves into a star. The star formation process is very complex involving many different physical phenomena like the interaction of gas with radiation, hydrodynamics, magnetic fields, gas chemistry, dust grain evolution, gravitation and more.

Key parameters of the gas change strongly for the transition from a cloud to a star:

- the density of a cloud is enhanced from $\sim 10^{-20} \text{ g cm}^{-3}$ to $\sim 1 \text{ g cm}^{-3}$ in a star,
- the specific angular momentum (per unit mass) of the gas must be lowered from $\sim 10^{22} \text{ cm}^2 \text{ s}^{-1}$ to about $\sim 10^{20} \text{ cm}^2 \text{ s}^{-1}$ for a binary system or $\sim 10^{17} \text{ cm}^2 \text{ s}^{-1}$ for a single star with a planetary system,
- and the magnetic energy per unit mass must be lowered from about $\sim 10^{11} \text{ erg g}^{-1}$ to about $\sim 10 \text{ erg g}^{-1}$.

Thus, star formation means that the gas is strongly compressed by self-gravity, that it must loose essential all its angular momentum by fragmentation and magnetic breaking, and it must be strongly de-magnetized by processes like ambipolar diffusion.

Gravitational equilibrium and Jeans mass. Because molecular clouds have often a lifetime $\gg 100$ Myr, there must exist, besides an equilibrium for the temperature and the pressure, also a hydrostatic equilibrium. The virial theorem is valid for systems in a gravitational equilibrium.

$$2 E_{\text{kin}} + E_{\text{pot}} = 0 \quad (5.1)$$

If we consider a homogeneous (constant density) and isothermal cloud then we can write for the kinetic (or thermal) energy $E_{\text{kin}} = E_{\text{therm}} = 3kTM/2\mu$. This yields for the virial theorem:

$$2 \cdot \frac{3}{2} \frac{k}{\mu} T M - \frac{3}{5} \frac{G M^2}{R} = 0 \quad (5.2)$$

This can be rearranged into $kT/\mu = GM/5R$. The third power of this equation and inserting the mean density of a homogeneous sphere ($\rho = 3M/4\pi R^3$) provides an estimate

for the equilibrium density or equilibrium mass for a given gas temperature T . These quantities are called **Jeans-mass**

$$M_J = \left(\frac{375}{4\pi}\right)^{1/2} \left(\frac{k}{G\mu} T\right)^{3/2} \frac{1}{\rho^{1/2}}.$$

and **Jeans-density**

$$\rho_J = \frac{375}{4\pi} \left(\frac{k}{G\mu} T\right)^3 \frac{1}{M^2}$$

Example: The Jeans-density for $M = M_\odot$, $T = 10$ K and $\mu = 2.7$ is $\rho_J \approx 7 \cdot 10^{-19} \text{ g cm}^{-3}$ equivalent to a particle density (H_2) of $2 \cdot 10^5 \text{ cm}^{-3}$.

The Jeans mass gives for a fixed cloud temperature and density the mass required for being in gravitational / hydrodynamic equilibrium. The Jeans mass is smaller for cold or/and high density clouds. Similarly, the Jeans density describes for a given cloud mass and temperature the density which must be achieved to be in a gravitational equilibrium. The density can be rather low for high mass, cool clouds.

The Jeans-density and Jeans-mass are parameters for an interstellar cloud in an idealized gravitational equilibrium. For real clouds, there are other parameters which come into play and they may stabilize or de-stabilize the cloud. Under certain conditions, already a small disturbance could produce a collapse to a star or an expansion and diffusion of the cloud.

For a theoretical **closed box model** the cloud remains in a hydrostatic equilibrium. If the cloud is slightly compressed, then the liberated potential (or gravitational) energy is converted into thermal energy, which enhances the gas pressure and the system goes back into the equilibrium state.

Contraction by radiation. A contraction is possible if energy is radiated away. If contraction occurs then potential energy is converted into kinetic energy $\Delta E_{\text{kin}} = -\Delta E_{\text{pot}}$ and if part of this thermal (or kinetic) energy is radiated away then the system can find a more compact quasi-equilibrium configuration. According to the virial theorem half of the liberated potential energy must be radiated away, while the other half is converted into thermal energy

$$L_{\text{cloud}} = -\frac{1}{2}\Delta E_{\text{pot}} \quad \text{and} \quad \Delta E_{\text{kin}} = -\frac{1}{2}\Delta E_{\text{pot}}. \quad (5.3)$$

The contraction speed depends on the radiation or cooling time-scale:

$$\tau_{\text{cooling}} \approx \frac{3}{2}nkT \frac{1}{n^2\Lambda} = \frac{3}{2} \frac{kT}{n\Lambda}.$$

The cooling time scales becomes shorter during the collapse because the particle density increases steadily.

- the contraction is rapid in the **optically thin** case, because then the radiation can escape from the entire cloud volume,
- the contraction is slow if the cloud is **optically thick**, because the radiation can only escape from the surface.

The virial theorem requires that the cloud temperature raises during contraction if no radiation is emitted. Thus, contracting clouds heat up. But, because warmer gas emits more efficiently (as long as it is below $T < 1000$ K) for higher temperatures (see cooling curve in Sect. 4), the luminosity and therefore the loss of radiation energy of the contracting object becomes higher until the fast contraction changes into a slow quasi-static contraction when the cloud becomes optically thick.

Stabilization mechanisms must exist for self-gravitating clouds because else all existing clouds would collapse in a short timescale. Mechanisms which can stabilize a cloud against collapse are:

- cloud **heating** processes, like radiation from external stars, cosmic rays, magneto-hydrodynamic turbulence and waves, which enhance the gas temperature and the gas pressure so that the cloud expands,
- **angular momentum** conservations may inhibit collapse because of enhanced centrifugal forces for more compact and therefore more rapidly rotating clouds,
- **magnetic fields** are frozen into the plasma as long as there are ions in the molecular cloud so that the contraction enhances the magnetic pressure like $p_{\text{magn}} \propto B_0^2 / r_{\text{cloud}}^2$.

Star formation is complicated because so many different processes play a role and from observations it is often hard to get detailed information about the cloud geometry, heating processes, specific angular momentum, and magnetic properties of a gas. A few important aspects of star formation follow from the stabilizing processes discussed above.

Star formation feedback is the influence of new-born stars on their environment. Young stars have strong outflows and emit energetic radiation which both can heat the surrounding cloud and stop the star formation process. On the other hand, this heating produces over-pressurized bubbles, like the Orion nebula which expand and which may compress the adjacent gas and trigger the collapse of a cloud. Depending on the details positive or negative feedback occurs and there is strong observational evidence that both mechanisms happen. However, many aspects of the star formation feedback are still unclear.

Figure 5.1: Schematic illustration of the fragmentation process.

Fragmentation is linked to the Jeans mass. If a cloud contracts isothermally (loss of energy through radiation) then the density increases and the Jeans mass becomes smaller like $M_J \propto 1/\sqrt{\rho}$. Thus, a large contracting cloud can decay in smaller clouds so that many stars are formed simultaneously in a big cloud complex. Typically there are many low mass stars formed $M < 1 M_\odot$ but only a few high mass stars $M > 1 M_\odot$.

The specific angular momentum of the gas in a molecular cloud is very large when compared to a contracted proto-stellar clouds. Therefore the angular momentum barrier inhibits a global contraction of a cloud. However, if subunits can collapse into proto-stars then the global angular momentum with respect to the entire cloud is preserved as motion of the proto-stars around the center of gravity. The remaining specific angular momentum of the gas with respect to the individual proto-stellar cloud unit is then much smaller. The angular momentum barrier is a second important aspect in favor of cloud fragmentation and the quasi-simultaneous formation of many stars out of big molecular cloud.

Proto-stellar disks and binaries are a further result of the angular momentum conservation. A contracting pre-stellar cloud core needs still to get rid of angular momentum. Angular momentum transfer via magneto-hydrodynamic processes helps to transport angular momentum away from the contracting cloud.

Another option is the formation of a binary star or a circumstellar disk. Both are configurations which can “store” more angular momentum than a rapidly rotating star.

Figure 5.2: Schematic illustration of ambipolar diffusion.

Ambipolar diffusion can solve the problem of the magnetic field pressure. A contracting cloud with charged particles contracts also the galactic magnetic field and will therefore “feel” soon the magnetic pressure which acts against further contraction. The magnetic field can move out of a molecular cloud by the so-called ambipolar diffusion. Magnetic fields can diffuse out of neutral cloud cores, either because also the charged particles diffuse out or all charged particles (ions, electrons) form neutral atoms, molecules, and solids. This leaves in the end a compact, demagnetized, cloud core. The fact that stars form predominantly in dense, cool, neutral clouds could be due to the lack of magnetic pressure.

5.2.1 Time scale for contraction

A simple estimate for the time scale for a molecular cloud contraction can be derived for a spherically homogeneous cloud. It is assumed that a cloud loses suddenly the gas pressure support (e.g. due to cooling) and starts to contract under its own gravitation. According to Newton's theorems the gravitational acceleration at a radius r is given by the mass inside r . During the contraction phase (decreasing r) this “inner” mass remains constant but gets more and more concentrated:

$$\frac{d^2r}{dt^2} = v \frac{dv}{dr} = \frac{-GM_r}{r^2} = -\frac{4\pi Gr_0^3 \rho_0}{3r^2}, \quad (5.4)$$

where we used a trick based on the chain rule

$$\frac{d^2r}{dt^2} = \frac{d}{dt} \frac{dr}{dt} = \frac{dr}{dt} \frac{d}{dr} \frac{dr}{dt} = v \frac{dv}{dr}.$$

Integration yields

$$\int v dv = -\frac{4\pi Gr_0^3 \rho_0}{3} \int \frac{1}{r^2} dr \quad (5.5)$$

or

$$v^2 = +\frac{8\pi Gr_0^3 \rho_0}{3} \cdot \frac{1}{r} + \text{const.}$$

The integration constant follows from the start conditions $r = r_0$ and $v = 0$

$$\text{const.} = -\frac{8\pi Gr_0^2 \rho_0}{3}.$$

This yields then the formula for v^2 or for $v(r)$ adopting in addition a negative sign because of the inward motion:

$$v(r) = -\sqrt{\frac{8\pi Gr_0^2 \rho_0}{3} \left(\frac{r_0}{r} - 1 \right)}. \quad (5.6)$$

This formula indicates:

- the infall velocity is initially zero $v(r) = 0$,
- it increases first slowly for $0.5r_0 < r < r_0$,
- goes to $\rightarrow \infty$ for $r \rightarrow 0$, when the gas collapses into a singularity.

This description is only meaningful for a contraction by at most a few orders of magnitude in r . For a contraction to a singularity (e.g. black hole) other processes would counteract to the contraction.

The contraction or collapse time scale can be derived by an integration of the velocity from radius r_0 to 0

$$t_{\text{coll}} = \int dt = \int_{r_0}^0 \frac{dr}{v} = \sqrt{\frac{3}{8\pi Gr_0^2 \rho_0}} \int_0^{r_0} \frac{1}{\sqrt{r_0/r - 1}} dr. \quad (5.7)$$

The integral is equal to $\pi/2$ and the resulting collapse time scale is

$$t_{\text{coll}} = \sqrt{\frac{3\pi}{32G\rho_0}} \quad (5.8)$$

This is shorter but of the same order as the dynamical time scale or free-fall time scale for the motion of a particle in the gravitational potential of a (static) homogeneous sphere.

Important result of this collapse timescale are:

- an entire cloud has the same collapse timescale which is only defined by the mean density ρ_0 ,
- the timescale for the collapse of a hydrogen cloud is defined by the particle density n_H according to:

$$t = 5 \cdot 10^7 [\text{yr}] \frac{1}{\sqrt{n_H [\text{cm}^3]}}.$$

Thus, a dense cloud with $n_H = 10^6 \text{ cm}^3$ would collapse in just 50'000 years if the pressure support is switched off.

In a more realistic model one could assume that the cloud has a density profile which decreases with radius. For such a cloud the collapse time is not the same for small radius and larger radius. Because of the higher density the innermost regions would collapse faster than the outer regions.

Alternatively, it is also reasonable to assume that the cooling by radiation is more effective at the edge of the cloud at large radii. This means that the gas pressure decreases faster further out and the cloud would start to collapse from the outside. This would cause a collision between the infalling outer material with the stationary inner gas. This shock would brake the collapse and then again complicated physical processes will occur.

Thus, the collapse time derived above is just a rough estimate for the shortest possible collapse time. A more realistic cloud collapse will always last longer than this estimate.

5.3 Initial mass function

Collapsing interstellar clouds form stars in the mass range from 0.1 to 100 $M_\odot$. The **initial mass function (IMF)** describes the mass distribution for the formed stars. According to the classical work of Salpeter (1955), this distribution can be described for stars of about solar mass and above with a potential law of the form:

$$\frac{dN_S}{dM} \propto M^{-2.35} \quad \text{for } M > 0.5 M_\odot. \quad (5.9)$$

This relation is often given as a logarithmic power law of the form

$$\frac{dN_S}{d \log M} \propto M^{-1.35} \quad \text{because} \quad \frac{dN_S}{dM} = \frac{dN_S}{d \log M} \frac{d \log M}{dM} = \frac{1}{M} \frac{dN_S}{d \log M}.$$

This is equivalent to a linear fit with slope -1.35 in $\log M$ - $\log N_S$ diagram (Figure 5.3). This law indicates, that the number of newly formed stars with a mass between 1 and 2 $M_\odot$ is about 20 times larger than the stars with masses between 10 and 20 $M_\odot$. If we consider the gas mass of the molecular cloud, then about twice as much gas ends up in stars between 1 and 2 $M_\odot$ when compared to stars with masses between 10 and 20 $M_\odot$. The initial mass function seems to be valid for many regions in the Universe, for the star formation in small molecular clouds, larger cloud complexes, and the largest star forming regions in the local Universe. Up to now no star forming regions have been found for which the Salpeter IMF is a bad description.

For low mass stars the mass distribution shows a turn over. Since M-stars $M < 0.5 M_\odot$ have a main-sequence life time which is longer than the age of the universe we can just use as first approximation the frequency of stars with different spectral types as rough

description for the IMF of low mass stars. This distribution shows a maximum in the range of M3V to M5V stars (or stars with $M \approx 0.35 M_{\odot}$).

Figure 5.3: Schematic illustration of the initial mass function for stars.

Considering the complex physics involved in the star formation process it is surprising that the initial mass function is such an universal law which seems to be valid everywhere in the Universe. There must be one essential process which dominates the outcome of the stellar mass distribution. This could be the fragmentation process. Further it seems to be clear that there are different regimes of formation between stars and planets. The low frequency of substellar object in the mass range $0.01 - 0.1 M_{\odot}$ indicates that such objects are not easily formed via the normal star forming process, perhaps because the formation of small fragments or their survival in molecular clouds is rather unlikely. On the other side the planets are very frequent but seem to form predominantly around stars.

This indicates that there exists a bimodal formation mechanism of hydrostatic astronomical objects.

- stars are formed by the collapse and fragmentation of clouds,
- planets are the result of a formation process in circumstellar disks.

5.4 Proto-stars

There are different phases in the star formation process, from a collapsing cloud, to a pre-stellar core, to a proto-star, and a pre-main sequence star. Some of these phases have specific observational characteristics in the spectral energy distribution (SED). The SED show the signatures of the following components:

- the Planck-spectrum of the main energy source, with its characteristic peak flux wavelength indicating the temperature of the object,
- an infrared excess, if optical to near-IR radiation is absorbed by the circumstellar material and re-radiated at longer wavelength,
- an UV-visual excess because of energetic processes due to gas accretion onto the star,
- emission lines if the energetic processes are strong enough to dissociate and ionize gas.

Figure 5.4: Schematic illustration of the spectral energy distribution for the different types of young stellar objects.

According to the presence and characteristics of these features different types of young stellar objects are distinguished:

- **Class 0:** The SED peaks in the far-IR or sub-mm part of the spectrum near $100 \mu\text{m}$ (30 K), with no flux in the near-IR. These are the dense, pre-stellar cloud cores.
- **Class I:** They have a flat or rising SED from about $1 \mu\text{m}$ towards longer wavelengths indicating that a hot source ($\approx 1000 \text{ K}$) is still embedded in a cloud, so that most radiation from the young stellar object is absorbed and re-radiated as far-IR emission by the circumstellar dust.
- **Class II:** They have falling SED into the mid-IR and the underlying objects have the characteristics of so-called classical T Tauri stars or Herbig Ae/Be stars. They exhibit strong emission lines and often a strong UV excess from the accretion process. These are the systems with extended circumstellar disks which are strongly irradiated by the central proto-star.
- **Class III:** These are pre-main-sequence stars with little or no excess in the IR, but with still some weak emission lines due to gas accretion. One of the subgroups of this class are the weak-lined T Tauri stars.

Class II and Class III objects can be placed into the Hertzsprung-Russell diagram if the temperature and luminosity are corrected for the contribution from the accretion processes. Compared to normal, main-sequence stars the Class II and Class III objects are located above the main sequence. These object evolve then “down” to the main-sequence (Slide 5–4).

The **pre-main-sequence** time scale, which describes the quasi-static contraction of a young star follows from the Virial theorem. The Virial theorem requires that half of the potential energy gained by the gravitational contraction is radiated away as described by

$$\tau_{\text{KH}} \approx \frac{E_{\text{pot}}}{L} \approx \frac{GM^2}{RL}. \quad (5.10)$$

This time-scale is also called the Kelvin-Helmholtz timescale. For solar parameters there is $\tau_{\text{KH}} \approx 30 \text{ Myr}$. Pre-main sequence stars start as relatively large $\approx 3 R_{\odot}$ and luminous objects $\approx 10 L_{\odot}$ with correspondingly shorter time-scales.

Chapter 6

Milky Way formation and evolution

A description of the Milky Way formation and evolution requires the understanding of many different processes:

- the hydrodynamics of proto-galaxies,
- the transport of angular momentum,
- the incidence of gravitational instabilities,
- the star formation, nucleosynthesis and the chemical enrichment of the gas,
- the accretion history of the galaxy,
- the effects of a time-dependent gravitational potential,
- the origin, growth and effects of magnetic fields.

Many of these processes can only be described with significant uncertainties because we do not know well essential parameters. Nonetheless one can try to estimate the properties of certain processes with idealized models and comparisons with observations of the Milky Way and other galaxies.

6.1 Virial theorem and galaxy formation

Some basic constraints on the galaxy formation process can be gained from simple considerations of the virial theorem

$$E_{\text{kin}} = -\frac{1}{2}E_{\text{pot}}.$$

The standard picture is that galaxies form in growing dark matter concentrations. If gas cooling can be neglected then the virial temperature for the gas falling into the proto-galactic potential will have a temperature of

$$E_{\text{therm}} \approx \frac{M kT}{m_{\text{H}}} \approx \frac{GM^2}{R}.$$

This yields for typical quantities of a galaxy ($R = 10$ kpc, $M = 10^{11}M_{\odot}$), a virial temperature of

$$T_{\text{virial}} = \frac{GMm_{\text{H}}}{kR} \approx 5 \cdot 10^6 \text{ K} \frac{M/10^{11}M_{\odot}}{R/10 \text{ kpc}}.$$

The fact that the Milky Way disk is a significant component of the galaxy consisting of co-rotating stars and gas moving in an orbital equilibrium and a small velocity dispersion ($\sigma \approx 30$ km/s) around the galactic center implies the following:

- the gas must dissipate a lot of energy by collisional excitation of atoms and molecules and subsequent radiative emission in order to cool to low temperature,
- the gas has settled into a rotating disk by gas friction and angular momentum conservation,
- most disk stars share the same overall circular motion around the galactic center indicates that they were all born over many Gyr out of a cold rotating disk.

A very similar estimate based on the virial theorem can be made for the kinetic velocity dispersion σ of N collisionless particles (= stars) with mass m_S ($M = Nm_S$) falling into such a proto-galactic potential

$$E_{\text{kin}} = \frac{1}{2}Nm_S\sigma^2 \approx -\frac{1}{2}\frac{GM^2}{R}.$$

For the velocity dispersion of the stars it is not so important whether the mass falling into the potential consists only of stars or a mixture of stars and gas. The estimate for the resulting velocity dispersion is then for the galaxy parameters given above

$$\sigma \approx \sqrt{\frac{GM}{R}} \approx 200 \text{ km/s} \sqrt{\frac{M/10^{11}M_\odot}{R/10 \text{ kpc}}}.$$

This is about the velocity dispersion observed for the system of globular clusters and halo stars. Thus they have been formed far from the center of the potential and seem to have “fallen” into the system without much energy dissipation.

On the other side one can say that the stars in the central bulge of the Milky Way have much smaller kinetic velocities than expected for a population of infalling “collisionless” stars. This indicates that these stars were formed after a lot of energy dissipation from cold gas which had a more spherical distribution and less angular momentum than the disk.

Time-scales. One should also consider the typical time-scales for the mass infall and the energy dissipation for a proto-galaxy. The time-scale for the collapse of a gas cloud is derived in the section on star formation

$$t_{\text{coll}} \approx 5 \cdot 10^7 \text{ yr} \frac{1}{\sqrt{n_{\text{H}}}}.$$

If we consider that the dark matter mass is about a factor 5 higher than the baryonic matter mass, the collapse time scale is shorter by a factor of a few.

It can be assumed that the infalling gas was either neutral or photo-ionized as can be inferred from the intergalactic Ly α absorption lines in quasar spectra. Thus the infalling gas has a temperature of 10^4 K or less and the corresponding sound speed is about $c_s \approx 10$ km/s or less. Thus the infalling gas will produce shocks which results in significant gas heating. This needs to be radiated away by radiative cooling which happens on a time scale of

$$t_{\text{cool}} \approx \frac{nkT}{n^2\Lambda(T)}.$$

We have seen in Chapter 4 that the cooling is very roughly proportional to the gas temperature

$$\Lambda(T)/T \sim 10^{-28 \pm 1} \text{erg cm}^3 \text{s}^{-1} \text{K}^{-1}$$

with deviations of about one order of magnitude. Thus the cooling time scale is roughly

$$t_{\text{cool}} \approx 4 \cdot 10^4 \text{yr} \frac{1}{n_{\text{H}}}.$$

This is typically much shorter than the collapse time scale. This means that gas falling into the dark matter potential of a proto-galaxy is cooling fast, and can dissipate efficiently its energy. Thus, gas will settle rapidly in the center of the galaxy where it can cool and form stars.

6.2 Milky Way evolution and high redshift observation

From Galactic studies it is difficult to derive the temporal history of its formation and evolution. Stars can be used as age indicators but it is hard to associate individual stars with past evolutionary events.

With modern large telescopes it is possible to follow the evolution of galaxies with high redshift observations. The look-back times for galaxies at redshift $z = 0.5, 1, 2$, and 5 are about $\tau_{\text{b}} \approx 5, 8, 11$ and 13 Gyr. These high redshift galaxy observations trace mainly the phases with strong star formation because then the objects are bright. Galaxy surveys at different red-shifts show:

- galaxies like the Milky Way show for $\tau_{\text{b}} > 5$ Gyr a significantly higher star formation rate, which is about an order of magnitude enhanced for red-shifts between $z = 1$ and 3 when compared to galaxies in the local Universe,
- galaxies at $z > 1$ show less well defined spiral structures, but often very bright knots located near the galaxy center, which shine so bright because of very strong starburst events,
- the estimated stellar mass of objects which are expected to evolve into a disk galaxy like our Milky Way is typically about 1000 times smaller for $\tau_{\text{b}} \approx 13$ Gyr ($z = 5$), about 100 times smaller for $\tau_{\text{b}} \approx 10$ Gyr, and about 10 times smaller for $\tau_{\text{b}} \approx 7$ Gyr than the current mass of our Milky Way.

The general picture on the Evolution of the Milky Way galaxy which emerges from the study of distant galaxies can be described as follows:

- 10 Gyr ago, the Milky Way was continuously fed by streams of fresh gas and merging galaxies as it is observed for high redshift galaxies $z > 2$,
- the constant inflow of gas and minor mergers produce disk instabilities and disk contractions due to counter-rotating streams, inducing phases of violent star formation in the central region (< 2 kpc) which become “in the end” the central bulges,
- since about 5 Gyr up to about 2 Gyr the gas inflow is significantly reduced and the star formation in the central bulge region is suppressed by e.g. rapid gas consumption, the activity of the central supermassive black hole, and/or blow out of gas by stellar winds and supernovae, while the rotationally supported extended disk shows continuous star formation and grows (inside-out grows) but at a slow rate,

- since about 2 Gyr the disk structure is well established, gas inflow and minor mergers with other galaxies still takes place but at a much reduced rate (100 times less) than in the early Universe and the star formation is much reduced but steadily continuing in the disk outside the bulge.

The high redshift galaxy surveys provide an excellent overall picture of the evolution of galaxies of the type of our Milky Way. On the other side it is impossible to deduce from high redshift observations detailed information about the properties of the disk, bulge and halo of the Milky Way.

6.3 Gas infall and minor mergers today

Galaxy surveys demonstrate that the life of galaxies was much “wilder” in the past, with many gas accretion events and minor mergers with small galaxies. Similar events are still happening in the Milky Way today but at a lower rate.

6.3.1 Gas inflow

With H I radio observations many hydrogen clouds were detected in the galactic halo. These clouds are special because they show high radial velocities $|v_r| > 70$ km/s with respect to the local standard of rest. This property allows us to distinguish them from local H I high latitude clouds. Slide 6–1 shows a galactic map of these high velocity clouds.

The distance to the high velocity clouds is typically between 2 and 15 kpc and they are located up to 10 kpc above or below the galactic plane. The distance can be determined by measurements of the Ca II or Na I absorptions in the spectra of halo stars with known distances. From the column density of about $n_H = 10^{18} \text{cm}^{-3}$ and the typical cloud extensions the total mass in the high velocity clouds is estimated to be of the order $\approx 10^8 M_\odot$.

The overall gas infall rate is estimate to be at a level of about $\approx 1 M_\odot/\text{yr}$ or a bit less. This inflow of fresh gas keeps the star formation rate at a constant level (also about $1 M_\odot/\text{yr}$) in the Milky Way disk.

Besides the high velocity clouds, there are two other types of high galactic latitude H I clouds:

- The Magellanic stream which is made of H I gas stripped off from the two Magellanic clouds due to tidal interaction between SMC and LMC. The Magellanic stream is at a distance of about 55 kpc (like the LMC and SMC) and it contains a mass of about $\approx 10^9 M_\odot$.
- Intermediate velocity clouds ($|v_r| \approx 40$ km/s) which could be the results of the ejection of gas from the galactic disk by supernova explosions, which cooled and falls now back onto the disk.

6.3.2 Mergers with dwarf galaxies

There are about 10 galaxies with are located very close, within 100 kpc, of the Milky Way. This list includes the Magellanic Clouds, LMC at 50 kpc and SMC at 65 kpc, but also several dwarf spheroidals and dwarf elliptical galaxies (UMa dwarf galaxy I and II, Sgr dwarf, UMi dwarf, Sex dwarf, Scu dwarf, Dra dwarf and a few others). There is a quite

large probability that some of these galaxies will collide in the coming few Gyr with the Milky Way.

Actually the Sagittarius dwarf spheroidal (Sgr dSph) galaxy is currently colliding with the Milky Way. The collision takes place at a distance of about 20 kpc from the sun on the “other side” of the galactic center. Sgr dSph is a small galaxy with only about 10^7 population II stars ($[\text{Fe}/\text{H}] < -1.6$) and apparently no gas. The system moves through our Milky Way in a roughly polar orbit. It is elongated, most likely due to the strong tidal forces, and many stars were already lost by the galaxy and they spread along a stellar stream in the halo (Slide 6–2 and 6–3). The orbit of Sgr dSph extends to about 40 kpc with a period of the order 600 Myr. Several globular clusters are associated with the Sgr dSph galaxy and one of them, M54, could be the core of the galaxy.

This example shows that the interpretation of stellar properties of stars in the halo and the thick disk can be very complex. Such merger events have happened often in the past and in the end it is very difficult to find out for a given star or stellar group whether they were “born” in the Milky Way, or whether they were accreted by a merger event. Slide 6–3 illustrates with an artist impression the collision of the Sgr dSph with the Milky Way. A most impressive example (and therefore not representative) are the observations of edge-on galaxy NGC 5907 where the stellar stream created by the collision can be nicely observed with very deep observations (Slide 6–4).

6.4 The chemical evolution of the Milky Way

The metallicity of the stars and gas in the Milky Way is far from homogeneous and the distribution provides interesting information about the Milky Way evolution. Important features are:

- the old stars $\tau_{\text{age}} \gtrsim 10$ Gyr in the globular clusters and the halo are metal poor with a metallicity $[\text{Fe}/\text{H}] \lesssim -1$,
- the galactic bulge consists mainly of intermediate age and old stars $\tau_{\text{age}} \gtrsim 2$ Gyr, but hardly any young stars $\tau_{\text{age}} \lesssim 300$ Myr. The galactic bulge stars show typically a metallicity which is higher than the solar metallicity $[\text{Fe}/\text{H}] > 0$,
- the Galactic disk shows two metallicity gradients:
 - one as function of radius where the metallicity decreases from $[\text{Fe}/\text{H}] \gtrsim 0$ inside the sun ($R < R_0$) to $[\text{Fe}/\text{H}] \lesssim 0$ for the outer disk $R > R_0$,
 - one in vertical direction with older and lower metallicity stars located in a thick disk while the young, higher metallicity stars are located close to the disk mid-plane.

This metallicity distribution is in qualitative agreement with the evolution history derived from the galaxy surveys. However, it is possible to extract much more details from accurate abundance measurements and detailed modelling of the gas enrichment with heavy elements.

6.4.1 Nucleosynthesis and stellar yields

The stars produce heavy elements by nuclear fusion and some fraction of these products are expelled at the end of the stellar evolution in stellar winds and explosions. This

“lost” stellar material enriches the interstellar gas with heavy elements. This production of heavy elements depends on the mass of the star. This subsection on stellar yields follows the paper from A. Maeder (1992, *Astron. & Astrophys.* 264, 105). Calculations for the stellar yields are reproduced on Slide 6–5, 6–6 and 6–7. Slide 6–5 shows the end products resulting from the stellar nuclear burning for gas which has initially a metallicity of 0.02 (or 2 %).

There are three main components which define the resulting heavy element enrichment by a star of a certain star:

- the fraction of the initial mass which is in the end locked in the stellar remnants,
- the fraction of the gas which is lost by the star but which is not converted to heavy elements by nuclear processes,
- the fraction of the gas which is processed into heavy elements by nuclear processes and subsequently ejected to the interstellar medium by stellar winds or/and explosions.

We discuss the enrichment of the gas with heavy elements for three different masses according to Slide 6–5.

- **1 $M_{\odot}$:** low mass stars produce a white dwarf composed of C and O with a mass of about $0.5 M_{\odot}$. Most of the nuclear burning products from the H- and He-burning phase are concentrated in this stellar remnants. The star loses at the end of its evolution the outer envelope, but these outer layers were not enriched with the nuclear burning products and therefore low mass stars contribute very little to the chemical enrichment of the Milky Way.
- **6 $M_{\odot}$:** intermediate mass stars produce also a white dwarf but with a mass of about $1 M_{\odot}$. During the red-giant phase the outer convective layer penetrates deeply into the star and “dredges-up” the products of the hydrogen burning layer. For this reason the mass lost during the red giant phase is enriched in helium which enhances the He abundance of the interstellar medium.
- **40 $M_{\odot}$:** high mass stars are important for the enrichment of the interstellar matter with heavy elements. Objects with such high initial mass lose a lot of mass via stellar winds. First, during the blue supergiant phase, the lost mass is hardly enriched by heavy elements. But, during later evolutionary phases, as luminous blue variable or Wolf-Rayet star, the stellar wind is so strong ($\approx 10^{-5} M_{\odot}/\text{yr}$) that it peels off subsequent layers of the star and expels about $10 M_{\odot}$ of He and N rich layers (WN-phase) where previously the H-burning took place, and then about $5 M_{\odot}$ of C and O rich material (WC and WO-phase) from the He-burning zone. The star has only about $5 M_{\odot}$ left, when it explodes as SN Ib.

For stars with an initial mass in the range $10\text{--}25 M_{\odot}$ the enrichment is mainly due to SN II explosions, which produces predominately α -elements such as Ne, Mg, Si, S, and Ca. Stellar wind mass loss is much less important for the enrichment of the α -elements.

Metallicity dependence. The stellar yields are different for low metallicity stars (see Slide 6–6). The main reason is that the gas opacities (metal line absorptions and e^{-} -scattering) are much reduced for low metallicity gas so that the radiation driven stellar winds are much weaker. This means that a star with a given initial stellar mass loses

much less mass during its evolution and the star will be much more massive and has still a hydrogen rich envelope when the “final” supernova (type II) explosion occurs. This leads to some variation in the production rates of heavy elements.

Weighting with the IMF. The stellar yield shown in Slide 6–5 and 6–6 must be weighted with the initial mass function. There result the IMF-weighted stellar yields diagrams shown in Slide 6–7.

This diagram shows that the low and intermediate mass stars with $M < 8 M_{\odot}$ dominate strongly the enrichment of the Milky Way in He, while these objects have no impact for heavier elements (but see below the section on SN Ia). The more massive stars $M > 8 M_{\odot}$ are the dominant producers of all heavy elements and they contribute also to the He enrichment.

6.4.2 The role of SN Ia.

The previous section discussed the yields of heavy element from single star stellar evolution. However, there is a small group of strongly interacting binary stars which explode at the end of their evolution as supernova Ia and they have a significant impact on the abundances of the “iron peak” (Cr, Fe, Co, Ni, ...) elements in the interstellar medium. Most binary stars are not strongly interacting because they have wide orbits $a > 100$ AU or they loose predominantly unprocessed material. Such systems can be counted like single stars.

SN Ia explosions. A SN Ia is an explosions of a CO white dwarf. These explosions convert the carbon and oxygen in explosive thermonuclear processes into mainly ^{56}Ni and other “iron peak” elements and the whole white dwarf explodes without remaining compact remnant. This means that the fusion products (Ni, Fe, ...) are all distributed into the interstellar medium. This is unlike to the core collapse supernovae (all other supernova types) where the “final” Ni–Fe-core collapses into a neutron star or black hole. For this reason the SN Ia are a main contributor to the Fe abundance, while high mass stars produce mainly the “ α -elements” such as O, Ne, Mg, Si, S, Ca.

The models for a SN Ia predict that the explosion takes place when a CO white dwarf accretes matter and approaches the Chandrasekhar mass limit of $1.4 M_{\odot}$. This can happen in close binary stars where mass is transferred from a companion to the white dwarf or if two very close white dwarfs merge due to the loss of orbital energy by the emission of gravitational radiation. Several different types of progenitor system have been identified, but it is not clear yet, which are the best candidates for becoming in the end a SN Ia event. Nonetheless SN Ia are good standard candles for cosmological studies, because the explosions are at least in most cases due to a nuclear explosion of a rather well defined CO white dwarf “bomb” of $1.4 M_{\odot}$.

SN Ia abundance effect. SN Ia are produced by intermediate mass binary systems which live about 1 Gyr before they explode. This is much longer than the core-collapse supernovae of massive stars which explode after less than 100 Myr. For this reason one observes for metal poor stars in the Milky Way a strong α/Fe -element overabundance because during the first Gyr no SN Ia contributed to the iron-abundances. This effect is illustrated in Slide 6–8.

The $[\text{O}/\text{Fe}] - [\text{Fe}/\text{H}]$ plot (Slide 6–8) gives the abundance ratios of the elements relative to solar abundance ratios. The important features in this plot are:

- low metallicity stars $[\text{Fe}/\text{H}] < -1$ have typically an overabundance of 0.5 dex in $[\text{O}/\text{Fe}]$ (= factor 3) with respect to the sun,
- $[\text{Fe}/\text{H}] \approx -1$ corresponds roughly to the epoch where the Milky Way was about 1 Gyr old and where the SN Ia started to add significant amounts of Fe,
- since then, the $[\text{O}/\text{Fe}]$ ratio evolved steadily towards the solar value,
- the slope in the distribution of stars in the $[\text{O}/\text{Fe}] - [\text{Fe}/\text{H}]$ diagram for $[\text{Fe}/\text{H}] \gtrsim -1$ can be explained with a supernova ratio of SN Ia/SN II+Ib $\approx 1.5/1$ as indicated by the line in the diagram.

Similar plots were compiled for many different element ratios for investigations of the origin of individual elements (SN Ia or SN II ?).

6.4.3 Modelling the chemical evolution of the Milky Way

The chemical evolution of the Milky Way is often modelled with simple stellar population models which consider for a very basic first approximation at least the following processes (see also corresponding exercise):

- the star formation rate, which is for the Milky Way of the order $\approx 1 - 10 \text{ M}_\odot/\text{yr}$, what yields about 10^{11} stars in 10^{10} yr,
- the initial mass function which describes the mass distribution of the newly formed stars ($f(M) \propto M^{-2.35}$)
- stellar lifetimes which are roughly $t_{\text{age}} \propto M^{-3}$
- the stellar yields as function of mass as shown in Slides 6-5 to 6-7.
- a simple history for the description of the low metallicity mass accretion by using e.g. a closed box model $M_0 = 10^{11} \text{ M}_\odot$ or a continuous mass inflow rate in the range $\approx 1 - 10 \text{ M}_\odot/\text{yr}$.

This modelling can be elaborated in much more detail with the consideration of more realistic descriptions and including additional processes. The models can then be compared with the available observational data. However, often the modelling is ambiguous or the uncertainties in the data are too large for firm conclusions.

Despite this, several important results are based on the study of elemental abundances of the Milky Way. The models can constrain in particular:

- the chemical evolution of the stars in the halo, the bulge, and the thick and thin disk; halo stars are old, metal poor and a large fraction of them were accreted by infall or merger events, the bulge stars originate mainly from the early galactic evolution, while the disk stars were formed later.
- continuous gas inflow is important because there are much less low metallicity G-stars than expected from a closed box model. The existence of less gas in the beginning allowed for a faster rise of the metallicity early in the galactic history, explaining the low frequency of metal poor, 10 Gyr old, G-dwarfs,
- the outward metallicity drop in the Milky Way disk is simply a result of the inside-out growth of the Milky Way as observed directly from the changing average properties of galaxies with red shift.