PHYS3080 Content Notes

Ryan White s4499039

Semester 1, 2022

All content taken from *Schneider - Extragalactic Astronomy and Cosmology [2nd Ed]* unless otherwise stated. Some excerpts from *"An Introduction to Modern Astrophysics – Second Edition"* [Carrol, B & Ostlie, D] also. If you'd like my reading notes on PHYS2082 (which is referred to in these notes on occasion), or to tell me I've messed up somewhere in these notes, please contact me at ryan.white@uqconnect.edu.au

Full disclaimer in that I haven't been paid to make/distribute this, but I'm legally obligated to accept an offer of a coffee or two for my suffering in making this as pretty as I know how.

Worth noting is that these notes don't really take into account the information given specifically in the course's accompanying lectures and serve only as a "summary" of the textbook in the context of the course's aims. I'm sure you'll notice that some topics have been regrettably glossed over, while some others have comparatively too much detail. In any case, I hope you'll find this useful and should you see any errors you can shoot me an email.

Contents

1 Week 1 - Galaxy Introduction

Ch 1.1 - Introduction

Typical galaxies are divided primarily into two major classes: spiral galaxies (where the majority of stars are confined to a relatively thin disk) and elliptical galaxies (with no discernible structure). The diameter of luminous galaxies are always far less than the separation between them.
While some stars within galaxies may be as old as the galaxy itself, the life cycle of stars allows for the birth of new stars (in *stellar nurseries*) throughout denser regions of a galaxy.

Since the speed of light is finite (and minuscule over cosmological distances), galaxies nearer to the edge of the observable universe (which could be progenitors to galaxies similar to the Milky Way) can be observed in an effort to understand the galaxy formation process.

Tremendous advances in telescope aperture and optical detectors have allowed for angular resolution observations small enough to study such distant and (apparently) dim objects. Simultaneously, advances in computing allowed for the analyses of such large datasets so that experimental physicists could test theories about galaxy formation and behaviour.
One of the most profound observations in recent history is that *most* galaxies harbour a supermassive black hole in their center. Even more interesting is that the mass of these black holes was found to be closely related to the other properties of its parent galaxy, implying that the evolution of this black hole is linked to the behaviour of it's host galaxy.

Detailed studies of galactic groups (as well as individual galaxies) led to the result that the majority of mass within is not visible in the form of stars or gas; only about 10-20% of the mass consists of stars, gas and dust that is able to be observed by photon emission and absorption. The remaining fraction of mass is termed *'Dark Matter'*, which can only be inferred by its gravitational effects on visible matter.

The most important development in recent years is the establishment of the standard model of cosmology (the science of the universe as a whole). This details that the universe has a finite age ($t_0 = 13.8$Gyr), which initially evolved from a very dense and hot state. Echoes of the *Big Bang* are visible today in the form of the *Cosmic Microwave Background Radiation* (CMBR). Observations of the CMBR led to much of the knowledge of today, but revealed that only $\sim 4\%$ of the energy content of the universe can be accounted for by known matter (with about 25% of the universe consisting of dark matter). This means that about 70% of the universe is comprised of *vacuum energy*, or *dark energy*, which reveals its existence exclusively in its effect on the cosmic expansion.

Ch 1.2 - Overview

1.2.1 - Our Milky Way as a Galaxy

From the analysis of the distribution of stars and gas in the Milky Way, it was found that the Galaxy consists of several distinct components:

- A thin disk of stars and gas with a radius of about 20 kpc and a scale height of about 300pc
- a ~ 1 kpc thick disk, which contains a different, older stellar population compared to the thin disk
- a central bulge, similar to other spiral galaxies
- a nearly spherical halo which contains most of the globular clusters, some old stars, and gas with different densities and temperatures.

Star formation is currently taking place in the spiral arms, with mainly old stars towards the center (especially in the galactic bulge). The Galactic disk rotates with velocity $V(R)$ depending on the distance R from the center. Using a Newtonian approximation, this can be modelled as:

$$V_0 = \sqrt{\frac{G \cdot M(R_0)}{R_0}} \tag{1.1}$$

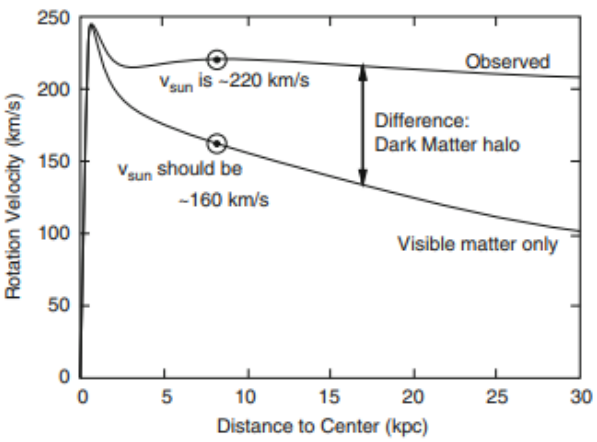


Fig. 1.7 Newtonian vs Actual Rotation Speed

The observed rotational velocity of the Sun around the Galactic center is significantly higher than would be expected from the Newtonian approximation (using the observed mass distribution). From the visible matter in stars, we would expect a rotational velocity of $\sim$ 160km/s, but $V_0 \sim$ 220km/s is observed. This discrepancy, coupled with the unexpected shape of the rotation curve $V(R)$ for larger distances R from the Galactic center, indicates that our Galaxy contains significantly more mass than is visible in the form of stars.
The main candidates for dark matter are weakly interacting elementary particles and (more unlikely) macroscopic objects which emit very little light.
There is cold gas falling into the Galactic disk and hot gas outflowing. Currently, the small neighbouring Sagittarius

dwarf galaxy is being torn apart in the tidal gravitational field of the Milky Way, and will completely merge in the (cosmically) near future.

Due to optical extinction from galactic dust, the Galactic center must be observed in longer wavelengths of light (infrared and radio). The view from our position inside the Milky Way makes it very difficult to measure and infer many properties of the Galaxy. Only by combining the study of the Milky Way with that of other galaxies can we hope to fully understand the physical nature of galaxies and their evolution.

1.2.2 - The World of Galaxies

In addition to the major classes of spiral and elliptical galaxies, there are additional classes such as irregular and dwarf galaxies, active galaxies, and starburst galaxies (which have an unusually high star formation rate). These classes differ not only from morphology, but also in their physical properties such as colour (indicative of stellar content), internal reddening (indicative of dust content), amount of interstellar gas, star-formation rate, etc. Galaxies of different morphologies have evolved in different ways.

Spiral galaxies are stellar systems in which active star formation is still taking place, whereas elliptical galaxies consist mainly of old stars, with star formation terminated long ago. S0 galaxies, an intermediate type, are characteristic of older stars with a galactic disk still present. Ellipticals and S0 galaxies together are called *early-type galaxies*, and spirals referred to as *late-type galaxies*.
Spiral galaxies contain a large amount of dark matter, with the visible matter embedded in a halo of dark matter.

Like with the interconnected characteristics of stars (mass-luminosity relations on a HR diagram), dynamical properties of galaxies are closely related to their luminosity. Also, the mass of the galactic center supermassive black hole (SMBH) is inherently linked to the velocity distribution of stars in elliptical galaxies and in the bulge of spirals.

Ch 3.1 - Classification

3.1.1 - Morphological Classification: The Hubble Sequence

Figure 3.2 shows the classification scheme defined by Hubble. According to this, three main types of galaxies exist:

- ***Elliptical Galaxies:*** (E's) are galaxies that have nearly elliptical isophotes without any clearly defined structure. They are subdivided according to their ellipticity:

  $$\epsilon \equiv 1 - b/a$$

  where a and b denote the semi-major and minor axes respectively. The upper limit of elliptical ellipticity is $\epsilon = 0.7$. En is commonly used to classify the ellipticals, where $n = 10\epsilon$ and E0 galaxies have circular isophotes.

- ***Spiral Galaxies:*** consist of a disk with spiral arm structure and a central bulge. They are divided into two subclasses:

  – *normal spirals* (S's)
  – *barred spirals* (SB's)

  In each of these subclasses, a sequence is defined that is ordered according to the brightness ratio of the bulge and the disk, denoted by a, ab, b, bc, c, cd, and d. Galaxies along this sequence are often referred to as early-type or late-type based on how prominent the spiral structure is respectively.

- ***Irregular Galaxies:*** (Irr's) are galaxies with only weak (Irr I) or no (Irr II) regular structure.

- ***S0 Galaxies*** are a transition between ellipticals and spirals; they are also called Lenticular on occasion. They contain a bulge and a large enveloping region of relatively unstructured brightness which often appears like a disk without spiral arms.

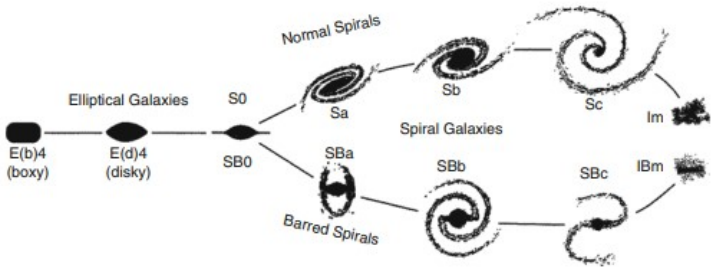


Fig. 3.2 Hubble Galaxy Classification

2 Week 2 - Galactic Structure and Dark Matter

Ch 2.3 - The Structure of the Galaxy

Roughly speaking, the Galaxy consists of the disk, the central bulge, and the Galactic halo (a roughly spherical distribution of stars and globular clusters that surround the disk). The disk, whose stars form the bulk of the visible band of the Milky Way, constitutes spiral arms similar to those observed in other galaxies. The Sun is situated in the disk, with a rough distance of $R_0 = 8.5$ kpc from the center of the Galaxy. The diameter of the galactic disk is about 50 kpc.

2.3.1 - The Galactic Disk: Distribution of Stars

One finds that the *scale-height, h,* of the thickness (z height) of each stellar component of the galaxy varies between populations of stars. In principle, three stellar components need to be distinguished:

1) The ***young thin disk***: contains the largest fraction of gas and dust in the Galaxy, and contains the bulk of present-day stellar formation. The youngest stars are found here, and has a scale height of $h_{\rm ytd} \sim 100$ pc.

2) The ***old thin disk***: thicker than the young disk, with a scale height of $h_{\rm otd} \sim 325$ pc.

3) The ***thick disk***: contributes only about 2% to the total mass density of the Galactic plane at $z = 0$, and has a scale height of $h_{\rm thick} \sim 1.5$ kpc.

Molecular gas, characterised by being cooler and denser than other (e.g. ionized) gas, has the smallest scale height $h_{\rm mol} \sim 65$ pc and is the main contributor to stellar formation. The younger a stellar population is, the smaller it's scale height.

Another notable characterization of stellar populations is their ***velocity dispersion*** (the amplitude of the components of their random motions). The random motion of the stars in the direction perpendicular to the disk is the reason for the finite thickness of the population. The larger the velocity dispersion σ_z perpendicular to the disk, the larger the scale-height h will be.
The Sun is a member of the young thin disk, and is located above the plain of the disk at $z \approx 30$ pc.

2.3.2 - The Galactic Disk: Chemical Composition and Age; Supernovae

Stellar Populations: The chemical composition of stars in the thin and thick disks differs, with a strong tendency that stars in the thin disk have a higher metallicity. In contrast, the stars in the Galactic halo and in the bulge have lower metallicities than those in the thin and thick disks. Population I (pop I) stars have a Solar-like metallicity (Z $\sim$ 0.02) and are mainly located in the thin disk, and population II stars (pop II) are metal-poor (Z $\sim$ 0.001) and found mainly in the thick disk, the halo and the bulge.

Metallicity and Supernovae: In a supernova explosion (SNe), most of the matter of the star is driven into the interstellar medium, enriching it with metals that were produced in the course of stellar evolution and death.

Classification of Supernovae: Based on their spectral properties, SNe are divided into several classes. Type 1 SNe don't show any Balmer lines of hydrogen, in contrast to the core-collapse Type II supernovae. SNe Ia show strong emission of SiII, whereas no SiII at all is visible in spectra of Type Ib,c. Type Ia SNe are found in all galaxies, whereas Ib,c and II are only found in spiral and irregular galaxies (where blue stars are common).

Core-Collapse Supernovae: SNe II and Ib,c are the final stages in the evolution of massive ($\gtrsim 8M_\odot$) stars. For more on the process of core-collapse supernovae, see my PHYS2082 Content Notes.
The major fraction of the binding energy released in the formation of the SN compact remnant is emitted in the form of neutrinos: about 3×10^{53} erg. Due to the high opacity in the stellar core during supernovae, even neutrinos are absorbed and scattered which results in part of their outward-momentum contributing to the explosion of the stellar envelope, which expands at $v \sim 10000$ km/s, with a kinetic energy of $E_{\rm kin} \sim 10^{51}$ erg. Of this, only about 10^{49} erg (about 1%) is converted into photons.
Due to the nuclear reaction chains taking place over a star's life, the majority of nuclei fused are composed of an even number of protons and neutrons. This is because of the prevalence of α-particles (or ^{4}He-nuclei comprised of two protons and two neutrons) in the nuclear chain reactions. Such elements are called ***α-elements***. The dominance of α-elements in the interstellar medium (and even in the Solar-system) is a clear indication of nuclear fusion occurring in the He-rich zones of stars.

Type Ia Supernovae: SNe Ia are most likely the explosions of white dwarfs (WDs) which exceed the Chandrasekhar mass limit, $M_{\rm Ch} \approx 1.44M_\odot$. In the case of mass accretion unto an existing WD by a main-sequence star (*single-degenerate* scenario), at about $M \approx 1.3M_\odot$, carbon burning will ignite the interior of the star, in the process transforming about half of the star into iron-group elements. The resulting explosion enriches the ISM with $\sim 0.6M_\odot$ of Fe and destroys the WD completely. A second, *double-degenerate*, scenario exists for the merger of, or accretion by two WDs.

Since initial conditions for single-degenerate Ia SNe are more or less homogeneous, they are good candidates for *standard candles*. This relation doesn't hold for double-degenerate Ia SNe, since the progenitor masses aren't strictly defined.

Since SNe Ia occur as the result of ignition of nova-remnant stars (and thus, much older), SNe Ia are visible in all galaxies. Contrast this to core-collapse supernovae, which are only visible in regions of relatively recent star formation (on account of the massive, and thus young stellar progenitors), and so are only typically visible in spiral and irregular galaxies.

In addition to SNe, the ISM can be enriched from stellar winds and planetary nebulae, particularly from stars with strong convection which transports newly fused material to the surface.

Age-Metallicity Relation: Assuming that at the beginning of its evolution the Milky Way had negligible metal content, the metallicity should be strongly related to the age of a stellar population. With each new stellar generation, more metals are produced and ejected into the ISM by winds and SNe. Stars that form later should thus have higher metallicities.
Indeed, one finds [Fe/H] $= -4.5$ for extremely old stars, and [Fe/H] $= 1$ for very young stars. However, this relation isn't very tight since the frequency of SNe Ia isn't well defined and stellar winds can create large [Fe/H] inhomogeneities locally. Another, preferred measure of metallicity is [O/H], because oxygen (an α-element)

is produced and ejected mainly in core-collapse supernovae which happen just $\sim 10^7$yr after stellar formation, which, on cosmological timescales is practically instantaneous.

Origin of the Thick Disk: Two explanations arise for the origin of the thick disk:

1) Star formation started earlier and/or ceased earlier in the thick disk compared to the thin disk, or
2) Stars that originally formed in the thin disk migrated into where the thick disk is now

The second explanation is more favoured, mainly due to inhomogeneities in the gravitational field throughout the galaxy caused by spiral arms, massive molecular clouds and other stars. Stellar orbits are perturbed by such fluctuations, in turn gaining a random velocity component perpendicular to the galactic plane. In other words, the velocity dispersion σ_z of a stellar population grows in time, and the scale height increases. In contrast to stars, gas keeps its narrow distribution about the galactic plane due to internal friction.

Another, origin is likely, with the two explanations potentially working simultaneously; the stars of the thick disk were formed outside the Milky Way and became constituents later through accretion of satellite galaxies. This is especially likely in galaxies with a thick disk that rotates about the center of the galaxy in a way significantly inconsistent with that of the thin disk (e.g. rotating in the opposite direction).

Mass-to-Light Ratio: The mass-to-luminosity ratio in the thin disk is

$$\boxed{\frac{M}{L_B} \approx 3\frac{M_\odot}{L_\odot} \text{ in thin disk}} \tag{2.37}$$

The M/L ratio in the thick disk is higher, as expected from an older and dimmer stellar population. The relative contribution of the thick disk to the stellar budget of the Milky Way is very uncertain due to the difficulty in attribution of individual stars to the thin vs thick disks. In any case, the thick disk contributes little to the total luminosity of the Milky Way. Spiral galaxies generally carry a value of about 4 solar units for the M/L ratio.

2.3.3 - The Galactic Disk: Dust and Gas

Spatial Distribution: The spiral structure of spiral galaxies is delineated by very young and luminous objects like O/B stars and HII-regions. The molecular clouds contract under their own gravity to form new stars inside the spiral arms, and as such the spiral arms are much less prominent in red light.

Open Clusters: Star formation in molecular clouds leads to the formation of open star clusters, since stars are born in groups. The mass of open clusters is heavily dependent on the progenitor gas cloud mass, and typically ranges from $100M_\odot$ to $10^4M_\odot$. Stars within open clusters have very small velocity dispersion and orbit more or less cohesively about the center of galaxies.

Most open clusters are younger than 300Myr and found within $\sim$ 50pc of the Galactic plane. The reason that we only see few old ($\geq$ 1Gyr) open clusters is that they are only weakly gravitationally bound and are subject to tidal gravitational forces and gravitational perturbations.

Observing Gas in the Galaxy: The 21cm line emission of neutral atomic hydrogen, HI, is mainly used to observe gas in the Milky Way since the Galaxy is optically thin in this wavelength. In most cases, it is assumed that the ratio of HI to CO (the second most common molecule after H_2) is constant throughout the universe and so the prevalence of CO can be used to infer the amount of HI present.

Distribution of Dust: Galactic dust is primarily observed in two ways:

1) The extinction it causes, quantified mainly by stellar reddening, and
2) The thermal radiation by the dust, primarily around temperatures between 17 and 21K.

Gas and dust are primarily found in the spiral arms where they serve as raw material for star formation. The gas mass in the Milky Way is less than 10% of the stellar mass.

Phases of the Interstellar Medium: Since molecules are easily destroyed by photons from hot stars, molecular gas needs to be shielded from the interstellar radiation field which is provided by the dust embedded in the gas. The range of temperatures in the molecular gas phase extends from $\sim$ 10K to about 70K, with characteristic densities of 100 particles per cm^3.

A second prominent phase is the warm interstellar gas, with characteristic temperatures on the order of thousands of Kelvin. Depending on T, the fraction of atoms which are ionized can range from 0.01 to 1. This gas can be heated by hydrodynamical processes or by photoionization (e.g. in gas nearby to a hot star). The best known examples for this gas are the HII regions around hot stars and planetary nebulae.

2.3.4 - Cosmic Rays

The Magnetic Field of the Galaxy: The properties of the Galaxy's magnetic field can be analyzed using a variety of methods, including but not limited to:

- *Polarization of stellar light:* The light of distant stars is partially polarized, with the degree of polarization being strongly related to the extinction of the star. The light scattered by dust particles is partially linearly polarized, with the direction of polarization depending on the alignment of the dust grains relative to the magnetic field. If the dust grains' orientation were random, the superposition of the scattered radiation over all of the dust particles would cancel out to net zero polarization, but this is not what is seen. As such, the observed polarization indicates the projected direction of the Galactic magnetic field.
- *The Zeeman effect:* The energy levels in an atom change if the atom is placed in a magnetic field. In fact, the observed 21cm transition line of neutral hydrogen is split in a magnetic field, where the amplitude of the split is proportional to the field strength.

Through the above, and also observations of synchrotron radiation and Faraday rotation, a magnetic field is known to exist in the Milky Way (primarily within the spiral arms) with a strength of about 4×10^{-6} G.

Cosmic Ray Acceleration and Confinement: For particle energies below 10^{15}eV, supernova remnants interacting with the ISM are the most likely origins. Accelerated particles propagate through the galaxy along complicated helical tracks (due to the present magnetic field), and so aren't able to be linked to a specific location. This magnetic field also confines the accelerated particles to within the field (as opposed to propagating outwards, away from the source) for on the order of 10s of Myr before they eventually diffuse out.
For higher energy particles, SN remnants are suggested to be the source with the most energetic, **Ultra-high energy cosmic rays**, likely being from active galactic nuclei (AGN).

2.3.5 - The Galactic Bulge

Owing to the strong extinction in the disk, the bulge is best observed in the IR. The scale-height of the bulge is ~ 400 pc.

Stellar Age Distribution in the Bulge: The stars in the bulge cover a large range in metallicity, $-1 \lesssim$ [Fe/H] $\gtrsim +0.6$, with a mean of about 0.3 (about twice that of the Sun). The metallicity also changes as a function of distance from the center, with more distant stars having a smaller value of [Fe/H]. The high metallicity suggests either that the stars of the bulge formed rather late (according to the age-metallicity relation), or that it is an old population with very intense star formation activities at an early cosmic epoch.
With the help of spectroscopic observations, one finds that the ratio of [Mg/Fe] in the bulge stars is much higher than that of the thin disk and is close to the thick disk, which implies that the enrichment must have occurred predominately by core-collapse supernovae as opposed to Type Ia SNe (further implying that the enrichment occurred long ago before Ia SNe were common). Hence, most of the bulge stars formed within a short time-scale of ~ 1 Gyr.
In the region of the bulge, one also finds stars that kinematically belong to the disk and the halo, as both extend to the inner region of the Milky Way.
The stellar mass-to-light ratio of the bulge is

$$\boxed{\frac{M}{L_B} \approx 5 \frac{M_\odot}{L_\odot} \text{ in the bulge}} \tag{2.44}$$

which is notably larger than the thick and thin disks.

2.3.6 - The Stellar Halo

The visible halo of our Galaxy consists of about 150 globular clusters and field stars with a high velocity component perpendicular to the Galactic plane. Most globular clusters are at a distance of $r \lesssim 35$ kpc, and field stars have been found out to $r \sim 50$ kpc, hence one assumes a characteristic value of $r_{\text{halo}} \sim 50$ kpc for the extent of the visible halo. Star counts indicate that the halo is an oblate spheroid (with the larger radii parallel with the galactic plane), with an axis ratio of the smallest axis to the longer one being $q \sim 0.6$.

The stellar halo is highly structured, with several over- and underdensities of stars. Some of the overdense streams of stars have been linked to the disruption of a low-mass satellite galaxy by tidal gravitational forces, most notably the Sagittarius dwarf spheroidal (Sgr dSph).

Tidal disruption: For a sufficiently large system of gravitationally bound particles (such as a gas cloud, star cluster, etc) moving inside a gravitational field, particles further away from the parent gravitational force experience a lesser pull from gravity compared to the system center of mass. Conversely, particles closer to the source (relative to the system center of mass) experience a stronger force. Hence, in the rest frame of the moving system, there is a net acceleration of the particles away from the center due to tidal gravitational forces.
If the net force for the particles in the outer part of the system is directed outwards (and is stronger than the force binding them to the system), these particles will be removed from the system in a process called ***tidal stripping***.

On its orbit through the Galaxy, a satellite galaxy or star cluster will experience a tidal force which varies with time. The 180°-symmetry of the tidal force mentioned before leads to the occurrence of two almost symmetric tidal tails, one moving slightly faster than the cluster (*leading tail*) and another slightly slower (*trailing tail*).

Ch 2.4 - Kinematics of the Galaxy

The Milky Way rotates clockwise when viewed from above, i.e. from the North Galactic Pole (NGP).

2.4.1 - Determination of the Velocity of the Sun

Local Standard of Rest: To link local measurements to the Galactic coordinate system (R, θ, z), the *local standard of rest* (LSR) is defined. It is a fictitious rest-frame such that a point that is located today at the position of the Sun moves along a perfectly circular orbit in the plane of the Galactic disk. The velocity components in the LSR are then

$$\boxed{U_{\rm LSR} \equiv 0 \qquad V_{\rm LSR} \equiv 0 \qquad W_{\rm LSR} \equiv 0} \tag{2.49}$$

where velocity components in the galactic coordinate system are defined as

$$U \coloneqq \frac{dR}{dt} \qquad V \coloneqq R\frac{d\theta}{dt} \qquad W \coloneqq \frac{dz}{dt} \tag{2.48}$$

Although the LSR changes over time, the time-scale of this change is so large (orbital period is $\sim 2.3 \times 10^7$ yr) that this effect is negligible.

Peculiar Velocity: The velocity of an object relative to the LSR is called its peculiar velocity. It is denoted by $\boldsymbol{v}$, and its components are given by

$$\boxed{\begin{aligned}\boldsymbol{v} &\equiv (u,\, v,\, w) \\ &= (U - U_{\rm LSR},\, V - V_{\rm LSR},\, W - W_{\rm LSR}) \\ &= (U,\, V - V_0,\, W)\end{aligned}} \tag{2.50}$$

The velocity of the Sun relative to the LSR is denoted by $\boldsymbol{v}_\odot$. If the relative velocity of a star with respect to the Sun is known, then the peculiar velocity of said star is

$$\boldsymbol{v} = \boldsymbol{v}_\odot + \Delta\boldsymbol{v} \tag{2.52}$$

Peculiar Velocity of the Sun: The Sun is currently moving inwards, upwards, and faster than it would on a circular orbit at its location, described by

$$\boxed{\boldsymbol{v}_\odot = (-10,\, 5,\, 7)\ \text{km/s}} \tag{2.55}$$

Velocity Dispersion of Stars: For young stars (A-type and more massive), the velocity dispersion relative to the LSR is small. For older K giants it is larger, and larger still for old, metal-poor red dwarf stars. There's a well defined velocity-metallicity relation, which, when combined with the age-metallicity relation, suggests that the oldest stars have the highest peculiar velocities.

2.4.2 - The Rotation Curve of the Galaxy

Galactic Rotation Curve for $R < R_0$; Tangent Point Method: With the tangent point method, applied to the 21cm line of neutral hydrogen or to radio emission lines of molecular gas, the rotation curve of the Galaxy inside the Solar orbit ($R < R_0$) can be measured.

Rotation Curve for $R > R_0$: Measuring $V(R)$ for $R > R_0$ requires measuring v_r for objects whose distance can be determined directly. It turns out that the rotation curve does not decline outwards as one would expect from the distribution of visible matter in the Milky Way. The observed total mass within $R = 60$ kpc is about $(4.0 \pm 0.7) \times 10^{11} M_\odot$, and so in order to get the almost constant rotational velocity of the Galaxy, much more matter has to be present than is observed in gas and stars.

The Milky Way contains, besides stars and gas, an additional component of matter that dominates the mass at $R \gtrsim R_0$. Its presence is known only by its gravitational effect, since it has not been observed directly yet, neither in emission nor in absorption. Hence, it is called *dark matter*. To be more specific, the evidence for dark matter is especially strong in spiral galaxies, which are embedded in a *dark matter halo*.

2.4.3 - The Gravitational Potential of the Galaxy

We have little direct indications about the spatial extent of the dark matter halo, and it is largely unknown whether it is approximately spherical or deviates significantly from sphericity (being either oblate or prolate).

The Nature of Dark Matter is thus far unknown. In principle, two different kinds of dark matter candidates are viable:

- *Astrophysical dark matter*, consisting of compact objects which are too faintly visible to be observed. Such objects are referred to as 'Massive Compact Halo Objects' (MACHOs).
- *Particle physics dark matter*, consisting of elementary particles which are yet to be discovered.

Ch 3.3 - Spiral Galaxies

3.3.1 - Trends in the Sequence of Spirals

Looking at the sequence of early-type to late-type spirals, a number of difference can be used for classification:

- a decreasing luminosity ratio of bulge and disk, with $L_{\rm bulge}/L_{\rm disk} \sim 0.3$ for Sa's and ~ 0.05 for Sc's
- an increasing opening angle of the spiral arms, from $\sim 6^{\rm o}$ for Sa's to $\sim 18^{\rm o}$ for Sc's
- an increasing brightness structure along the spiral arms: Sa's have a 'smooth' distribution of stars along the spiral arms while Sc's have resolved light distributions of bright knots of stars and HII regions

Compared to ellipticals, the spirals cover a distinctly smaller range in magnitude and mass: $-16 \gtrsim M_B \gtrsim -23$ and $10^9 M_\odot \lesssim M \lesssim 10^{12} M_\odot$ respectively.

3.3.2 - Brightness Profile

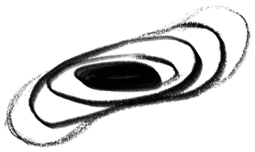
The light profile of the bulge of spirals is described by a de Vaucouleurs profile while the disk follows an exponential

brightness profile. Expressing these distributions of the surface brightness in $\mu \propto -2.5\log(I)$ measured in mag/arcsec^2,

$$\mu_{\rm bulge}(R) = \mu_e + 8.3268\left[\left(\frac{R}{R_e}\right)^{1/4} - 1\right] \tag{3.13}$$

$$\mu_{\rm disk}(R) = \mu_0 + 1.09\left(\frac{R}{h_R}\right) \tag{3.14}$$

where μ_e is the surface brightness at the effective radius R_e which is defined such that half of the bulge luminosity is emitted within R_e. The parameters μ_0 and h_R are the central surface brightness (of the disk) and scale-length of the disk respectively.



Disk Warping

Warps in Disks: The disks of galaxies are not always lying in a plane. In this case, the plane in which the orbit of stars and gas rotate about the center at a radius R changes its inclination with R. The origin of these warps are not well understood.
In some cases, the warps have found to be correlated with a significant tidal force from other galaxies in a group, but this has not been found in all cases of disk warping.

3.3.4 Rotation Curves and Dark Matter

The inclination angle (its orientation with respect to line of sight) of other galaxies is determined from the observed axis ration of the disk, assuming that disks are intrinsically axially symmetric.

The rotation curves of spirals do not decrease for $R \geq h_R$, as one would expect from the light distribution, but are basically flat. We therefore conclude that spirals are surrounded by a halo of dark matter. The density distribution of this dark matter halo can be derived from the rotation curves.

Assuming that M/L is independent of radius for a stellar population, a distribution of dark matter can be inferred from the discrepancy between $v^2_{\rm lum}$ (rotational speed due to luminous matter) and v^2 (observed rotational speed):

$$M_{\rm dark}(R) = \frac{R}{G}\left[v^2(R) - v^2_{\rm lum}(R)\right] \tag{3.17}$$

Correlations of Rotation Curves with Galaxy Properties The form and amplitude of the rotation curves of spirals are correlated with their luminosity and their Hubble type. The larger the luminosity of a spiral, the steeper the rise of $v(R)$ in the central region, and the larger the maximum rotation velocity $v_{\rm max}$. The shape (not amplitude) of different Hubble types rotation curves' is similar, despite different brightness profiles – further implying the existence of dark matter.

3 Week 3 - Spiral Structure and Scaling Relations

3.3.6 - Spiral Structure

The spiral arms are the bluest regions in spirals, containing young stars and HII regions. Because of this, the brightness contrast of spiral arms increases as the wavelength decreases.

Due to the differential rotation speeds and orbit sizes, the spiral structure wouldn't persist over cosmological timescales if they were a material structure of stars and gas. Instead, it is suspected that the spiral arms are a wave structure; quasi-stationary density waves with regions of higher density than the local disk environment.
As gas orbits about the center of the Galaxy and encounters the abnormally dense region, it is compressed which leads to increased star formation; this accounts for the blue colour of the spiral arms.

The initial generation of spiral arms may be induced by a non-axially symmetric perturbation of the gravitational potential in the galaxy, such as by a massive central bar or a companion galaxy. About 65% of spiral galaxies have a central bar which suggests this plays a key importance in spiral formation.

CO Ch 25.3 - Spiral Structure

The most majestic spiral galaxies, known as ***grand-design spirals*** usually have two very symmetric and well-defined arms. Galaxies that do not possess well-defined spiral arms that are traceable over a significant angular distance are called ***flocculent spirals***. Only about 10% of all spirals are considered grand-design galaxies, 60% multiple-arm galaxies, and the remaining 30% are flocculent galaxies.
Optical images of spirals are dominated by their arms, containing luminous O and B stars and HII regions. Since massive OB stars are short-lived relative to the characteristic rotation period of a galaxy, spiral structure corresponds to regions of active star formation. Dust and gas clouds are also more prevalent near the inner edges of the spiral arms.

When spirals are observed in red light, the arms become much broader and less pronounced, but still detectable. Since red stars are typically older and lower mass, this implies that the bulk of the galactic disk is dominated by older stars. Despite this, observations indicate that there is still an increase in the number density of older stars within the spiral arms.

25.3.1 - Trailing and Leading Spiral Arms

Although the general appearance of spiral galaxies suggest that their arms are *trailing* in rotation (meaning that the tips of the arms point in the opposite direction of rotation), verifying this isn't trivial. In almost all cases where a determination can be made, it appears that the spiral arms are trailing. There are some determined cases of leading arms,

however, and were suggested to be due to a tidal encounter with a retrograde moving object.

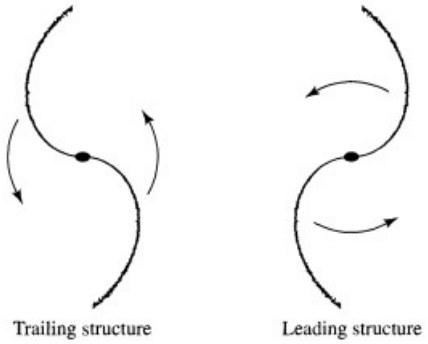


Fig. 25.20 Trailing and Leading Arm Structures

25.3.2 - The Winding Problem

One problem immediately arises when the nature of spiral structure is considered: *material arms* composes of identifiable stars and gas clouds would "wind up" on a timescale that is short compared to the age of the galaxy, visualised in Fig 25.21:

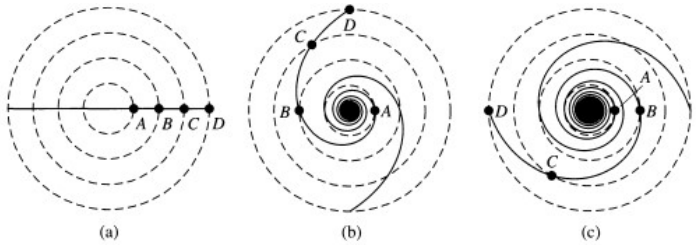


Fig. 25.21 Winding Problem for Material Arms

The winding effect would lead to a natural generation of spiral arms, becoming more tightly wound as time progresses. Obviously spirals generally don't display this behaviour, so another mechanism is at work.

25.3.3 - The Lin-Shu Density Wave Theory

It is theorised that spiral structure arises because of the presence of long-lived ***quasistatic density waves***, consisting of regions in the galactic disk where the mass density is 10-20% greater than average. Stars, dust, and gas clouds move through the density waves during their orbits around the galactic center.

When the spiral galaxy is viewed from a noninertial reference frame that is rotation with a specific angular speed known as the ***global pattern speed***, $\Omega_{\rm gp}$, the spiral pattern appears to be stationary as seen in Fig 25.22. At a specific distance from the galactic center, called the ***corotation radius*** $R_{\rm c}$, the stars and density waves will move together. In this noninertial frame, stars with $R < R_{\rm c}$ will appear to pass through the arms in one direction, and stars with $R > R_{\rm c}$ will appear to be moving in the opposite sense.

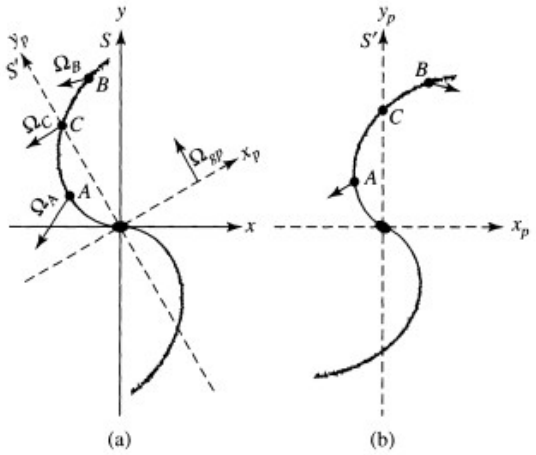


Fig. 25.22 Spiral Galaxy viewed from (a) Inertial Frame and (b) Noninertial Frame

As dust and gas clouds within the corotation radius overtake a density wave, they are compressed by the effects of the increase in local mass density, leading to star formation. Since this cloud collapse takes time, new stars will appear within the arm slightly "downstream" from dust and gas clouds at the edge of the wave.

Newly formed OB stars ionize the gas around them (HII regions) which results in the displayed red regions in spiral arms. Since their lifetimes are short, these stars die before they can leave the density wave, while less massive stars leave to populate the disk.

In principle, the density wave theory solves the Winding Problem proposed earlier.

Ch 3.4 - Scaling Relations

The properties of galaxies are characterized by a number of quantities including but not limited to: luminosity, size, mass, rotational velocity, velocity dispersion, colour, star-formation rate, etc. These properties are inherently linked, and the properties of isolated galaxies usually may be determined by just a few parameters from which the others follow.

3.4.1 - The Tully-Fisher Relation

The maximum rotation velocity of spirals is closely related to their luminosity, following the *Tully-Fisher Relation*,

$$\boxed{L \propto v_{\rm max}^{\alpha}} \tag{3.19}$$

where the power-law index is about $\alpha \sim 4$. That said, the index is dependent on wavelength with more massive, redder galaxies having a larger α.

Because of the close correlation of data-to-model, the luminosity of spirals (and hence the distance [by comparison of apparent magnitude and luminosity]) can be estimated by measuring the rotational velocity, which is independent of distance from the observer.

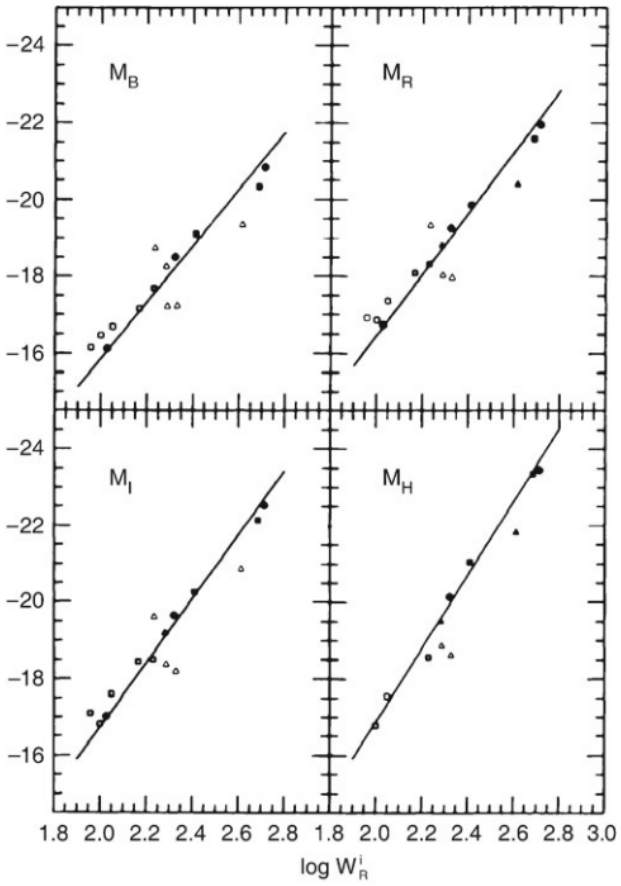


Fig. 3.27 The Tully-Fisher Relation for Galaxies in the local group; absolute magnitude on the y-axis, with log velocity dispersion on the x-axis.

Since the shapes of rotation curves for spirals seem to be very similar, the radial dependence of the ratio of luminous to dark matter may also be quite similar among spirals. The flat rotation curve implies

$$\boxed{M = \frac{v_{\max}^2 R}{G}} \tag{3.20}$$

where the value of R is the distance from the center of the galaxy chosen to be the extent of the 'flat' part of the rotation curve (i.e. where $v_{\rm rot}(R) \approx V_{\max}$). The exact value of R isn't particularly important since M is a function of radius and a the value of R used will give the mass within said radius.

Mass-to-Light Ratio of Spirals: We're unable to infer the total mass of a spiral since the extent of the dark halo is unknown. As such, M/L can only be measured within a fixed, luminous radius defined as R_{25} – the radius at which the surface brightness attains the value of 25 mag/arcsec^2 in the B-band. The spiral then follows the relation

$$\boxed{\log\left(\frac{R_{25}}{\rm kpc}\right) = -0.249 M_B - 4.00} \tag{3.23}$$

independently of the type of spiral. Within R_{25}, one finds $M/L_B = 6.2$ for Sa's, 4.5 for Sb's, and 2.6 for Sc's. This trend is expected since late type spirals contain more blue, luminous stars.

The Baryonic Tully-Fisher Relation: For spirals will a small $v_{\max} \lesssim 100$ km/s, the Tully-Fisher relation does not hold. Accounting for the mass of the gas in the galaxy (which is proportionally higher for smaller galaxies) gives a much tighter correlation:

$$\boxed{M_{\rm disk} = 2 \times 10^9 h^{-2} M_\odot \left(\frac{v_{\max}}{100 \text{ km/s}}\right)^4} \tag{3.24}$$

3.4.2 - The Faber-Jackson Relation

An analogous relation to the Tully-Fisher relation for elliptical galaxies shows that the velocity dispersion in the center of ellipticals scales with luminosity:

$$\boxed{L \propto \sigma_0^4} \quad \text{or} \quad \boxed{\log(\sigma_0) = -0.1 M_B + \text{const}} \tag{3.25}$$

3.4.3 - The Fundamental Plane

The *fundamental plane* relates the surface brightness and effective radius for elliptical galaxies whose velocity dispersion is too small for the Faber-Jackson relation to apply. It reads

$$\boxed{R_e \propto \langle I \rangle_e^{-0.83}} \tag{3.26}$$

where $\langle I \rangle_e$ is the average surface brightness within the effective radius such that

$$\boxed{L = 2\pi R_e^2 \langle I \rangle_e} \tag{3.27}$$

From this, an *approximate* relation between effective surface brightness and total luminosity is $\boxed{\langle I \rangle_e \propto L^{-1.5}}$
Hence, more luminous ellipticals have smaller surface brightnesses. By means of the Faber-Jackson relation, L is related to σ_0 and therefore, σ_0, $\langle I \rangle_e$, and R_e are all related to each other.
The distribution of elliptical galaxies in the three-dimensional parameter space (R_e, $\langle I \rangle_e$, σ_0) is located close to a plane defined by

$$\boxed{R_e \propto \sigma_0^{1.4} \langle I \rangle_e^{-0.85}} \tag{3.29}$$

$$\boxed{\log R_e = 0.34 \langle \mu \rangle_e + 1.4 \log \sigma_0 + \text{const}} \tag{3.30}$$

where $\langle \mu \rangle_e$ is the average surface brightness within R_e, measured in mag/arcsec^2. Equation 3.30 defines a plane in this three-dimensional parameter space that is known as the *fundamental plane*.
Like the Tully-Fisher relation, the fundamental plane is an important tool for distance estimations.

3.4.4 - The D_n-σ Relation

D_n is defined as the mean diameter of an ellipse within which the average surface brightness I_n corresponds to a value of 20.75 mag/arcsec^2 in the B-band. This gives

$$\boxed{D_n \propto \sigma_0^{1.4} I_e^{0.05}} \tag{3.34}$$

which implies that D_n is nearly independent of I_e and only really depends on σ_0. This relation describes the properties of ellipticals considerably better than the Faber-Jackson relation and, in contrast to the fundamental plane, it is a relation between only two observables.

3.4.5 - Summary: Properties of Galaxies on the Hubble Sequence

- Most luminous galaxies in the local universe fit onto the Hubble sequence; they are either ellipticals, spirals, or belong to the class of S0 galaxies.
- Ellipticals and spirals not only differ in morphology, but in composition, star formation rates and colour.
- Stars in spirals have very ordered motion [moving about the center in nearly circular orbits with relatively minimal velocity dispersion – referred to as 'dynamically cold']. In contrast, the stars in ellipticals have largely random motion – 'dynamically hot'.
- Some elliptical galaxies show signs of complex structure which infer past galaxy interaction. In contrast, spiral galaxies are thought to be largely undisturbed.
- The rotation curves of spiral galaxies are almost flat across large radii, not what would be expected from the visible mass distribution. This implies that there is some gravitational influence, *dark matter*, in which the galaxy is subject to.
- Both spirals and ellipticals follow scaling relations which connect their luminous properties with their dynamical properties.

4 Week 4 - Galactic Centers and Active Galaxies

Ch 2.6 - The Galactic Center

2.6.1 - Where is the Galactic Center?

Radio observations in the direction of the GC show a complex structure; a central disk of HI gas exists at radii from several 100 pc up to about 1kpc. Its rotational velocity yields an estimate of the enclosed mass $M(R)$ for $R \gtrsim 100$ pc. Furthermore, radio filaments are seen perpendicular to the galactic plane and many supernova remnants are seen. In addition, a fairly large number of globular clusters and gas nebulae are observed towards the central region.

The innermost 8 pc contain the radio source Sagittarius A (Sgr A), which itself consists of:

- A circumnuclear molecular ring shaped like a torus, extending between radii of 2 pc $\lesssim R \lesssim$ 8 pc and inclined by 20° relative to the Galactic disk. It has a sharp inner boundary and has a rotational velocity of ~ 110 km/s, nearly independent of R. It is evidence of an energetic event in the Galactic center's past.
- Sgr A* is a compact radio source close to the center of Sgr A West, with an extent smaller than about 1AU. Except for the emission in the mm and cm domain, Sgr A* is a weak source and is an excellent candidate for the center of the Milky Way (since other galaxies have compact radio sources in their center).

2.6.2 - The Central Star Cluster

Density Distribution: Observations show a compact star cluster centered on Sgr A*, with its density behaving like $\propto r^{-1.8}$ within the distance range 0.1 pc $\lesssim r \lesssim$ 1 pc. The number density of stars here is so large that close stellar encounters are common (about every $\sim 10^6$ yrs), and so the distribution of the stars is 'thermalized', which means that the local velocity distribution of the stars is the same everywhere. Most of the stars in the nuclear star cluster have an age $\gtrsim 1$ Gyr and are late-type giant stars.
Some young O and B stars are found in the central parsec, too, and are inferred to reside in one of two strongly-inclined, counter-rotating disks. The age of these stars is 6 ± 2 Myr; the same order of time between two strong encounters.
A strong radial dependence of velocity dispersion on radius is observed: σ increases towards smaller r. This indicates that the gravitational potential in which the stars are moving is generated not only by themselves, and so there is a central mass concentration in the star cluster.

The Origin of Very Massive Stars near the GC: The small ($\sim 10^8$) lifetime of the inner O and B stars implies that these stars were born close to the Galactic center. Both the strong tidal gravitational field of the central black hole and presumably strong magnetic field in the region prevent this 'standard' star formation picture, though. Perhaps dense gas infall in previous periods led to intense star formation. Other solutions, such as binary disruption and dynamical processes are also likely.

2.6.3 - A Black Hole in the Center of the Milky Way

The S-Star Cluster: There is a distribution of stars within ~ 1" of Sgr A* which is composed mainly of B-stars, known as the S-Star cluster. Some stars of this cluster have a proper motion well in excess of 1000km/s, up to $\sim$ 10000 km/s. From the observed kinematics of the stars, the enclosed mass $M(r)$ can be calculated. The corresponding analysis yields that $M(r)$ is basically constant over the range 0.01 pc $\lesssim r \lesssim$ 0.5 pc, indicating the presence of a point mass. A characteristic value obtained yields a distance to the Galactic center of $R_0 \approx 8.3$ kpc, which results in a black hole mass of Sgr A* of

$$\boxed{M = (4.3 \pm 0.4) \times 10^6 M_\odot} \tag{2.94}$$

Why a Black Hole?

- The energy for the central activity in quasars, radio galaxies, and other AGNs is produced by accretion of gas onto a supermassive black hole (SMBH). Thus we know that at least a sub-class of galaxies contains a central SMBH.
- Even if such an ultra-dense star cluster (with a central density of $\gtrsim 4\times10^{12}M_\odot/\mathrm{pc}^3$) was present, it could not be stable and would dissolve within $\sim 10^7$ years.
- Sgr A* itself has a proper motion of less than 20km/s – far smaller than the stars around it and is therefore the dynamical center of the Galaxy.

The gravitational effect of the black hole on the motion of stars and gas is constrained to the innermost region of the Milky Way. The gravitational field of the SMBH dominates the rotation curve of the Galaxy only for $R \lesssim 2$ pc. At larger radii, the presence of the SMBH is of no relevance for the rotation curve of the Milky Way.

2.6.6 - Hypervelocity Stars in the Galaxy

The inherent instability associated with N-body systems is the currently accepted source of observed *hypervelocity stars* at the outreaches of the Milky Way; young OB stars with velocities hundreds of km/s relative to the Galaxy and at distances of > 100 kpc. By three-body interactions, the orbit of two masses can become tighter and the corresponding excess energy is transferred to the third mass which may then become gravitationally unbound. As such, when a binary system gets close to the black hole, this three-body interaction can lead to the ejection of one of the two stars into an unbound orbit, whereas the other star gets bound to the black hole.
This acceleration of hypervelocity stars near the Galactic center is not the only accepted mechanism, and binary interactions with intermediate mass black holes in the center of globular clusters is also suggested as an acceleration mechanism.
As another source of acceleration, tightly bound binaries where one star undergoes a supernova explosion will accelerate the other star in the system to high velocities. These runaway stars typically have lower velocities than the aforementioned hypervelocity stars, however, and are typically found/produced near the Galactic disk.

Ch 5.1 - Active Galactic Nuclei (Introduction)

There are some galaxies which have a much broader energy spectrum than what would be expected of the superposition of all of the stars' Planck distributions within said galaxy. Some of these show significant emission in the full range from radio wavelengths to the X-ray and even gamma range. This emission originates mainly from a very small central region of such an ***active galaxy*** which is called the ***active galactic nucleus*** (*AGN*).
Some classes of AGNs, in particular the quasars, belong to the most luminous sources in the Universe and have been observed out the the highest measured redshifts, exceeding the luminosity of normal galaxies by a factor of a thousand. This luminosity originates from a region $r \lesssim 1$ pc.

5.1.2 - Fundamental Properties of Quasars

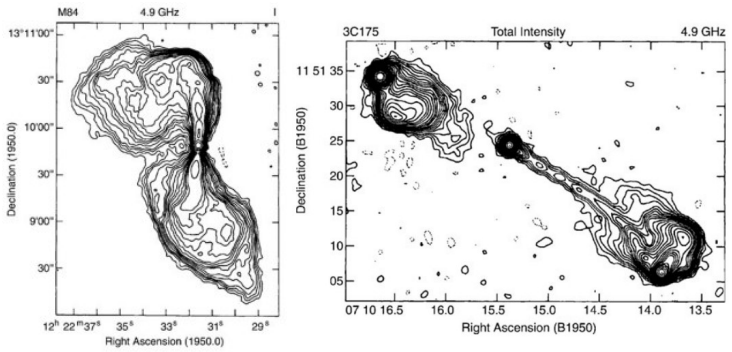
Quasars were discovered by identifying radio sources with point-like optical sources, although they emit at all wavelengths. The flux of the source varies at nearly all frequencies, where the variability time-scale differs among the objects and also depends on the wavelength. As a rule, it is found that the variability time-scale of the observed radiation is smaller, and its amplitude larger, at higher frequencies. The optical spectrum of quasars, especially at high redshift, is very blue (comparable in colour-index to hot white dwarfs).

5.1.3 - AGNs as Radio Sources: Synchrotron Radiation

The morphology of quasars and other AGNs in the radio regime depends on the observed frequency and can be very complex, consisting of several extended source components and one compact central one. In most cases, the extended component is observed as a double source in the form of two radio lobes. These lobes are often connected to the central core by jets, which are thin emission structures related to the energy transport from the core into the lobes. The observed length-scales of the radio source can reach up to 1 Mpc, and the extent of the central source corresponding to $r \lesssim 1$ pc.

Classification of Radio Sources: Extended radio sources are often divided into two class;

- *Fanaroff-Riley Type I* (FR I) are brightest close to the core, and the surface brightness decreases outwards. They typically have a luminosity of $L_\nu(1.4\mathrm{GHz}) \lesssim 10^{32}$ erg/s/Hz.
- *Fanaroff-Riley Type II* (FR II) sources increase in surface brightness outwards, and their luminosity is generally higher than that of FR I sources.



(a) FR I **(b)** FR II

Fig. 5.7 Comparison of Radio Source Types

FR II radio sources often have *jets*: extended linear structures that connect the compact core with a radio lobe. Jets often show internal structure such as knots and kinks. Worth noting is that the jets are not symmetric; often only one jet is observed and in most sources where two jets are found, one of them (the 'counter-jet') is much weaker than the other.

Synchrotron Radiation: AGN radiation in the radio is often linearly polarized, where the extended radio source may reach a degree of polarization up to 30% or more. The spectral form and high degree of polarization are interpreted such that the radio emission is produced by *synchrotron radiation* of relativistic electrons. Electrons in a magnetic field propagate along a helical path so that they are continually accelerated by the Lorentz force. Since accelerated charges emit electromagnetic radiation, this motion of the electrons leads to the emission of synchrotron radiation. These electrons must be highly relativistic, and to obtain particles at such high energies, very efficient processes of particle acceleration must occur in the inner regions of quasars.

It should be noted that cosmic ray particles of considerably higher energies have been observed, presumably produced in the shock fronts of supernova remnants. Thus, it is supposed that the energetic electrons in quasars and other AGNs are also produced by 'diffusive shock acceleration', where here the shock fronts are caused by hydrodynamical phenomena as opposed to supernovae. AGN outflow velocities are considerably higher than the speed of sound in the plasma, and so the conditions for the formation of shock fronts are satisfied.

Ch 3.8 - Black Holes in the Centers of Galaxies

It is generally accepted that the energy for the activity of AGNs is generated by accretion onto a black hole.

3.8.1 - The Search for Supermassive Black Holes

Even for the closest SMBH (Sgr A*), the Schwarzschild radius is significantly smaller than the achievable angular resolution of current telescopes.

The Radius of Influence: For a mass concentration of $M_\bullet$ in the center of a galaxy, with characteristic (star) velocity dispersion σ_v, distances smaller than

$$\boxed{r_{\rm BH} = \frac{GM_\bullet}{\sigma_v^2} \sim 0.4 \left(\frac{M_\bullet}{10^6 M_\odot}\right)\left(\frac{\sigma_v}{100{\rm km/s}}\right)^2 \ {\rm pc}} \quad (3.44)$$

the SMBH will significantly affect the kinematics of stars and gas in the galaxy.

Kinematic Evidence: The presence of a SMBH inside $r_{\rm BH}$ is revealed by an increase in the velocity dispersion and rotational velocity in the local neighbourhood ($r \gtrsim r_{\rm BH}$).

3.8.3 - Correlation Between SMBH Mass and Galaxy Properties

One finds that $M_\bullet$ is correlated with the absolute magnitude of the bulge component of the galaxy in which the SMBH is located. Here, the bulge component is either the bulge of a spiral (or S0) galaxy, or an elliptical galaxy as a whole. This correlation is described by

$$M_\bullet = 1.7 \times 10^9 M_\odot \left(\frac{L_V}{10^{11} L_{V\odot}}\right)^{1.11} \quad (3.46)$$

Instead of the bulge luminosity, one can also study the correlation of $M_\bullet$ with the mass of the bulge, for which the best power-law fit is

$$\boxed{M_\bullet = 2.9 \times 10^8 M_\odot \left(\frac{M_{\rm bulge}}{10^{11} M_\odot}\right)^{1.05}} \quad (3.47)$$

(where the trend is slightly more accurate than for the luminosity relation). Notice that the power is almost 1, and hence $M_\bullet \approx 3 \times 10^{-3} M_{\rm bulge}$ which gives an interesting correlation: about 0.3% of the baryon mass that was used to make the stellar population in the bulge of these galaxies was transformed into a central black hole.
An even tighter correlation exists between $M_\bullet$ and σ_v:

$$\boxed{M_\bullet = 2.1 \times 10^8 M_\odot \left(\frac{\sigma_v}{200{\rm km/s}}\right)^{5.64}} \quad (3.49)$$

Hence we conclude that galaxies with a bulge component host a supermassive black hole, whose mass is tightly correlated with the properties of the stellar component; in particular, the black hole amasses about 0.3% of the stellar mass in the bulge.

These correlations has to be linked to the fact that the spheroidal component of a galaxy evolves together with the SMBH.

Ch 5.2 - AGN Zoology

A wide range of objects are subsumed under the name AGN, all of which have in common strong non-thermal emission in the core of a *host galaxy*. The classification of AGNs described below can be confused; different classes refer to different appearances of AGNs but do not necessarily correspond to the physical nature of these sources.

Surrounding the central supermassive black hole is an accretion disk which emits the bulk part of the optical and UV continuum emission. The central region around the accretion disk is the source of most of the X-ray radiation. Gas clouds above and below the accretion disk are responsible for the broad emission lines. The appearance of the AGN depends on the relative inclination of the galactic disk; an edge on view will absorb much of the radiation emanating from the center.
The aforementioned ***emission lines*** correspond to the light

emitted from the accretion disk present around the central SMBH. 'Broad' emission lines correspond to fast moving gas close to the central SMBH where the high velocity broadens the characteristic lines. The gas responsible for the narrow emission lines is located at much larger distance from the black hole, and is never obscured by the gas around the nucleus. Since this gas is far from the black hole, it is characterized by slower rotational speeds and hence the comparatively narrow emission line.

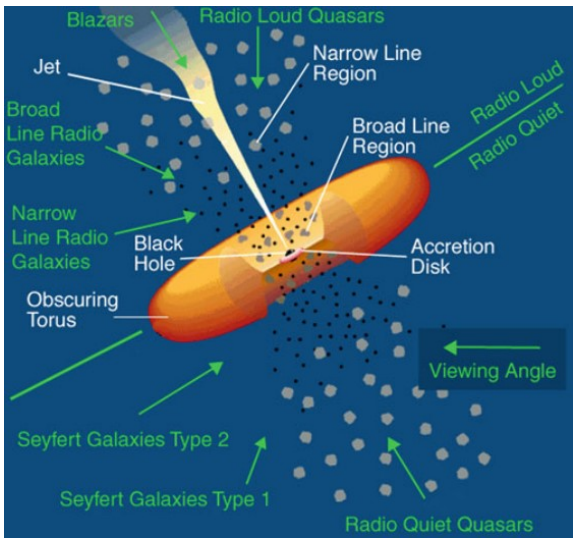


Fig. 5.12 Unification of AGN Types

The classification of AGNs is heavily dependent on the viewing angle (which dictates the type and magnitude of radiation received) as shown in Figure 5.12.

5.2.1 - QSOs

Apart from their radio properties (or lack thereof), ***radio-quiet quasars***, or quasi-stellar objects, QSOs, appear to be like quasars. Most of these sources are invisible in the radio domain of the spectrum, but their optical properties are virtually indistinguishable from those of quasars. In particular, they have a very blue optical energy distribution, strong and broad emission lines, and (in general) a high redshift.
In modern terminology, the expression QSO encompasses both the quasars and the radio-quiet QSOs. About 10 times more radio-quiet QSOs than quasars are thought to exist.

The QSOs are the most luminous AGNs, with their core luminosity sometimes being as high as $1000L^*$. They can outshine their host galaxy and appear point-like on optical images.

5.2.2 - Seyfert Galaxies

Seyfert galaxies are the AGNs that were first detected, and their nuclear luminosity is considerably lower than that of QSOs. On optical images, they are identified as spiral galaxies which have an extraordinarily bright core whose spectrum shows strong emission lines which are broader than typical velocities in galaxies. We distinguish between two types:

- *Seyfert 1* galaxies have both very broad and also narrower emission lines, where 'narrow' still means several hundred km/s and thus a significantly larger width than characteristic velocities (like rotational velocities) found in normal galaxies.
- *Seyfert 2* galaxies show only the narrower lines.

Intermediate galaxies also exist, for example a Seyfert 1.5 might show very broad lines but with a smaller ratio of broad-to-narrow line flux than in Seyfert 1 galaxies.
The optical spectrum of Seyfert 1 galaxies is very similar to that of QSOs, and both classes are often combined under the name ***Type 1 AGNs***. Except for the different core luminosity, no fundamental physical difference seems to exist.

5.2.3 - LINERs

The least luminous and by far most common type of AGNs are the *low-ionization nuclear emission-line regions* (LINERs). In fact, at least one third of all nearby galaxies contain a LINER in their core, characterized by emission lines from neutral atoms or ions with rather low ionization energies. The width of emission lines in LINERs is typically smaller than the narrow emission lines in Seyfert galaxies, and not much larger than the rotational velocity of the galaxy.

5.2.4 - Radio Galaxies

Radio galaxies are elliptical galaxies with an active nucleus. Similarly to Seyfert galaxies, radio galaxies are distinguished between those with and without broad emission lines: broad-line radio galaxies (BLRG) and narrow-line radio galaxies (NLRG), respectively. In principle, the two types of radio galaxies can be considered as radio-loud Seyfert 1 and 2 galaxies.
Besides the classification of radio galaxies into BLRG and NLRG with respect to the optical spectrum, they are distinguished according to their radio morphology; radio sources are divided into FR I and FR II sources.

5.2.5 - OVVs

One subclass or QSOs is characterized by the very strong and rapid variability of its optical radiation. The flux of these sources, which are known as *Optically Violently Variables* (OVVs), can vary by a significant fraction on time-scales of days. The light emitted by OVVs is typically several times more polarized than by typical QSOs also.

5.2.6 - BL Lac Objects

The class of AGNs called BL Lac objects (named after its prototypical source BL Lacertae) are characterized by very strongly varying radiation, like the OVVs, but without strong emission and absorption lines. As with OVVs, the optical radiation from BL Lacs are highly polarized.

	Normal Galaxy	Radio Galaxy	Seyfert Galaxy	Quasar	Blazar
Example	Milky Way	M87, Cygnus A	NGC 4151	3C273	BL Lac, 3C279
Galaxy Type	Spiral	Elliptical, Irregular	Spiral	Irregular	Elliptical?
$L_{\rm AGN}/L_\odot$	$< 10^4$	$10^6 - 10^8$	$10^8 - 10^{11}$	$10^{11} - 10^{14}$	$10^{11} - 10^{14}$
$M_{\rm BH}/M_\odot$	4×10^6	3×10^9	$10^6 - 10^9$	$10^6 - 10^9$	$10^6 - 10^9$
Radio Emission	Weak	Core, Jets, Lobes	Only $\approx 5\%$ Radio-loud	Only $\approx 5\%$ Radio-loud	Strong, Short-time Variable
X-Ray Emission	Weak	Strong	Strong	Strong	Strong
Gamma Emission	Weak	Weak	Medium	Strong	Strong

Table 5.1 Overview of the Classification of Active Galactic Nuclei

The optical luminosity of some BL Lacs varies by several magnitudes if observed over a long time period. In epochs of low luminosity, emission lines are sometimes observed and then a BL Lac appears like an OVV. For this reason, OVVs and BL Lacs are collectively called ***blazars***. All known blazars are radio sources. Besides their variability, blazars also show highly energetic and strongly variable γ-radiation.

Table 5.1 summarizes the fundamental properties of the different classes of AGNs.

Ch 5.3 - The Central Engine: A Black Hole

To justify the importance of black holes as AGN originators, first summarise some of the relevant observational facts of AGNs:

- The extent of some radio sources in AGNs may reach $\gtrsim 1$ Mpc, and hence a minimum lifetime of the source can be found to be $\tau \gtrsim 10^7$ yr.
- Luminous QSOs have a luminosity up to $L_{\rm bol} \sim 10^{47}$ erg/s. Assuming that the luminosity doesn't change substantially over the lifetime of a source, the total energy is on the order of 10^{61} erg.
- The luminosity of some AGNs varies by more than 50% on time-scales of a day, which puts the maximum spatial extent of such a source at about 1 lightday (otherwise it would violate causal properties of relativity).

5.3.2 - Accretion

The Principle of Accretion: Gas falling onto a compact object loses its potential energy, which is first converted into kinetic energy. Through the gas' inherent angular momentum and friction with other gas particles (and by the resulting momentum transfer), the gas will assemble in a disk oriented perpendicular to the direction of the angular momentum vector. The disk will rotate with the Kepler velocity, and since it rotates differentially (since it depends on r), the gas in the disk will be heated by internal friction. This friction causes a slight deceleration of the rotational velocity, whereby the gas slowly moves inwards. This is the energy source for heating the gas in the disk.

According to the virial theorem, half of the potential energy released is converted into kinetic energy; in the situation here, this is the rotational energy. The other half of the potential energy can be converted into internal energy.

Temperature Profile of a Geometrically Thin, Optically Thick Accretion Disk: When a mass falls further into a gravitational potential, some energy is released with half of the release being in the form of heat. In this case, the self-gravity of the accretion disk can be ignored, and the only gravitational attractor is that of the black hole. If the accretion disk is optically thick, the local emission corresponds to that of a black body. When accounting for the dissipation by friction and that some of the energy is advected inwards, the disk temperature can be expressed as

$$T(r) = \left(\frac{3GM_\bullet \dot{m}}{8\pi\sigma_{\rm SB} r^3}\right)^{1/4} = \left(\frac{3c^6 \dot{m}}{64\pi\sigma_{\rm SB} G^2 M_\bullet^2}\right)^{1/4} \left(\frac{r}{r_S}\right)^{-3/4} \quad (5.13 \text{ \& } 5.14)$$

Interpretation and Conclusions: The temperature in the accretion disk increases inwards $\propto r^{-3/4}$ as expected. To a first approximation, the total emission of the disk is a superposition of black bodies consisting of rings with different radii at different temperatures. Because of this, the resulting spectrum is much more broad than that of a Planck curve (as observed).

Most of the luminosity from a disk comes from the inner parts and so it depends critically on how far the disk extends inside the gravitational potential. Also, for any fixed ratio of radius to Schwarzschild radius, r/r_S, the temperature increases with the accretion rate $\dot{m}$. Furthermore, at fixed radius ratio, the temperature decreases with increasing $M_\bullet$ due to a reduction in tidal forces. This implies that the maximum temperature of the disk in an AGN is much lower than in accretion disks around compact objects of stellar mass (such as neutron stars and stellar mass black holes).

Radiatively Inefficient Accretion: In a system where the accretion rate $\dot{m}$ is low, the disk may be optically thin and the emission process of the heated gas can become inefficient. In this case, the gas cannot effectively cool and thermal energy is advected inwards together with the gas. Such a disk (called 'advection-dominated accretion flow', or ADAF) is rather inefficient in converting rest mass

into radiation, but accretion flow may be quite efficient in generating outflows, such that part of the accreted material is ejected in the form of jets. Hence, this accretion may be of importance for analysing radio galaxies.

5.3.4 - Further Arguments for SMBHs

From observational data, the shape of the average emission line can be well modeled by emission from an accretion disk around a black hole where the radiation originates from a region lying between ~ 3 and ~ 400 Schwarzschild radii.

The Spin of Black Holes: The spectral shape of the line is affected by the spin of the SMBH. According to General Relativity, there is a maximum spin a black hole can have, and the ratio of a black hole's spin to this maximal value is called the spin parameter $a_{\rm spin}$. It's been found that a large fraction of SMBHs have a spin parameter $a_{\rm spin} \gtrsim 0.9$, consistent with a black hole gaining most of its mass through accretion (since accretion transfers angular momentum, spinning it up). Conversely, black hole mass growth via mergers would likely cancel out much of the angular momentum (due to opposing angular momentum vectors), and so smaller values of $a_{\rm spin}$ are expected.

5.3.5 - The Eddington Luminosity

Radiation Force: The largest part of energy generation is in the innermost region of an accretion disk. This energy propagates outwards and can interact with infalling matter via absorption and scattering, transferring outwards momentum to the infalling matter; i.e. the infalling matter experiences an outwards-directed radiation force. In order for the mass to accrete onto the black hole, the radiation force must be smaller than the gravitational force.

Eddington Luminosity: The Eddington Luminosity for a black hole can be shown as

$$\boxed{L < L_{\rm edd} \coloneqq \frac{4\pi G c m_p}{\sigma_T} \approx 1.26 \times 10^{38} \left(\frac{M_\bullet}{M_\odot}\right) \text{erg/s}} \tag{5.25}$$

A Lower Limit on $M_\bullet$: For accretion to occur at all, we need $L < L_{\rm edd}$. Hence, the condition must hold:

$$\boxed{M_\bullet > M_{\rm edd} \coloneqq \frac{\sigma_T}{4\pi G c m_p} L \approx 8 \times 10^7 \left(\frac{L}{10^{46}\text{erg/s}}\right) M_\odot} \tag{5.26}$$

where m_p is the proton mass and σ_T is a constant related to *Thomson scattering*. Equation (5.26) shows that a lower limit for the mass of a black hole can be derived from its luminosity.
In the above definition of the Eddington luminosity, it was implicitly assumed that the emission of radiation is isotropic. In principle, luminosities exceeding the Eddington luminosity can be obtained if the emission is highly anisotropic.

Eddington Accretion Rate: If the conversion of infalling mass into energy takes place with an efficiency ϵ:

$$\epsilon = \eta = \frac{L}{\dot{m} c^2} \tag{5.15}$$

the accretion rate is then

$$\dot{m} = \frac{L}{\epsilon c^2} \equiv \frac{L}{L_{\rm edd}} \dot{m}_{\rm edd} \Rightarrow \dot{m}_{\rm edd} = \frac{L_{\rm edd}}{\epsilon c^2} \qquad (5.27 \to 5.29)$$

Growth Rate of the SMBH Mass: The Eddington accretion rate is the maximum accretion rate if isotropic emission is assumed. We can now estimate a characteristic time in which the mass of the SMBH will significantly increase:

$$t_{\rm evo} \coloneqq \frac{M_\bullet}{\dot{m}} \approx \epsilon \left(\frac{L}{L_{\rm edd}}\right)^{-1} 5 \times 10^8 \text{yr} \tag{5.30}$$

5 Week 5 - Quasars and Galaxy Clusters

Ch 5.4 - Components of an AGN

5.4.2 - The Broad Emission Lines

Characteristics of the Broad Line Region: The presence of very strong emission lines originating from AGN cannot be due to thermal line broadening, as the observed strength would imply temperatures in excess of 10^{10}K, at which point *all* atoms would be fully ionized and there would be no emission lines. Therefore, the observed line width is interpreted as Doppler broadening, where the gas emitting these lines then has large-scale velocities of order ~ 10000km/s.
For velocities this high, we expect the gas to be rotating in the vicinity of a SMBH. Consequently, characteristic velocities of $v \sim c/30$ occur at a distance of about $r \sim 500 r_S$.

The region in which the broad lines are produced is called the ***broad line region*** (BLR). From the ionization stages of the line-emitting elements in the BLR, a temperature can be estimated and is typically $T \sim 20000$K. Also, the emission measure (number of emitted line photons per unit time and volume element) can be calculated. Upon doing so, it's found that the emitting gas fills a volume far smaller than the total volume of the BLR in all cases and so the BLR is not homogeneously filled with gas. The gas in which the broad lines originate from fills only a small fraction ($\sim 10^{-7}$ to ~ 0.1) of the total volume of the BLR, and hence it must be concentrated in clouds.

Reverberation Mapping; Results: The picture of an inhomogeneous BLR is obtained which extends over a large range in r and which consists of different 'layers'. The various emission lines are emitted at different radii, because the ionization structure of the BLR varies with r; the higher the ionization energy of a transition, the smaller the corresponding radius r. Intuitively, the fact that lines of higher ionization energy are located closer to the central continuum source implies that they are also broader than low-ionization lines, according to the scaling $v \sim \sqrt{GM_\bullet/r}$.

The size of the region from which a particular line is emitted increases with increasing continuum luminosity. Essentially, the larger the L, the larger the extent of the BLR.

The nature of the clouds in the BLR is to some degree still unknown. Their small extent and high temperature imply that they should vaporize on very small time-scales unless somehow stabilized by replenishment, external pressure, by magnetic fields or perhaps by gravitation.

5.4.5 - The Host Galaxy

As the term 'active galactic nuclei' already implies, AGNs are considered the central engines of otherwise regular galaxies. From detailed images, AGNs' host galaxies are often heavily disturbed (e.g. by tidal interactions of nearby galaxies), and these disturbances are considered essential for the gas to overcome the angular momentum barrier and to flow towards the center of the galaxy to accrete onto the central SMBH. At the same time, such disturbances seem to increase the star formation rate enormously (since star burst galaxies are characterized by disturbances), and so AGNs are correlated with nuclear starbursts.
Today it seems established that the hosts of low-redshift luminous QSOs are predominantly massive and bulge-dominated galaxies. This is in good agreement with the fact that SMBH mass scales with the mass of the spheroidal component of the galaxies.

Binary QSOs: The connection between the activity of galaxies and the presence of close neighbours is also seen from the clustering properties of QSOs. The number of binary QSOs found is considerably larger than the expectation. In the case of close binary galaxies, both attain a perturbed gravitational potential and start to become active.

Ch 5.7 - Quasar Absorption Lines

Absorption lines from QSOs may be caused by absorbing material in the AGN itself or in its host galaxy, indicating an intrinsic origin. They may also be caused by intermediate gas between the observer and the AGN along the line of sight. Both of these situations occur, and the latter is an important probe into the gas makeup of the universe.
This intervening gas can either be in intergalactic space or being correlated with foreground galaxies, where the former implies metal-poor gas and the latter (comparably) metal-rich gas.

Classification of QSO Absorption Lines: The different absorption lines in QSOs are distinguished by classes according to their wavelength and width:

- *Metal Systems:* In general, these are narrow absorption lines with redshift of $0 < z_{\rm abs} < z_{\rm em}$ ($z_e m$ corresponding to redshift of the QSO source), caused by intervening matter along the line-of-sight. Normally, a metal system consists of many different lines all at the same redshift. From the line strength, the column density of the absorbing ions can be derived. These metal systems originate in either overdense regions of intergalactic space or they are associated with galaxy halos located along the line-of-sight.
- *Associated Metal Systems:* These systems have characteristics very similar to the aforementioned metal systems, but their redshift is $z_{\rm abs} \sim z_{\rm em}$. As such, they are interpreted as being related to the QSO itself, by its host galaxy or some other interaction. This is caused by the diffuse intergalactic distribution of gas.
- *Lyα Forest:* A large set of lines at $\lambda < (1 + z_{\rm em})1216$Å is interpreted to be Lyman-α absorption by hydrogen along the line-of-sight to the QSO.
- *Broad Absorption Lines:* For about 15% of QSOs, very broad absorption lines are found in the spectrum at redshifts slightly below $z_{\rm em}$. The lines show a profile which is typical for sources with outflowing material. The Doppler width of the lines in the *broad absorption line* (BAL) QSOs is a significant fraction of c. Since the redshift is slightly lower than that of the corresponding emission lines, the absorbing gas must be moving towards us.

Ch 6.1 - The Local Group

Galaxies are not uniformly distributed in space, but instead show a tendency to gather together in *galaxy groups*, and *clusters of galaxies*. The Milky Way is itself a member of a group called the Local Group, which implies that we are living in a locally overdense region of the universe.
The transition between groups and clusters is smooth. Roughly speaking, an accumulation of galaxies is called a group if it consists of $N \lesssim 50$ members within a sphere of diameter $D \lesssim 1.5h^{-1}$ Mpc, while clusters have $N \gtrsim 50$ and $D \gtrsim 1.5h^{-1}$ Mpc.

6.1.1 - Phenomenology

The Milky Way, M31 (Andromeda) and M33 (Triangulum) galaxies are the three most luminous members of the Local Group, and are all spiral galaxies. About 11 of the galactic members of the Local Group are Irregular galaxies (such as the Large and Small Magellanic Clouds), and the remaining members are all dwarf galaxies which are very small and faint. Many of these dwarf galaxies are in fact satellite galaxies of the more luminous members.
By strict classification, the Local Group is too spatially extended to classify as a group for it's mass and it is not dynamically relaxed. The bimodal distribution of galaxies in the Local Group instead suggests that two small galaxy groups – one centered on the Milky Way and another on M31 – are merging towards a common barycenter.

6.1.2 - Mass Estimate

A mass estimate of

$$M \sim 3 \times 10^{12} M_\odot \tag{6.5}$$

is obtained for the Local Group, which is far larger than the mass of the Milky Way and M31 (who themselves make up 90% of the observed mass) as observed in stars and gas. This mass estimate yields a mass-to-light ratio for the Local Group of $M/L \sim 70 M_\odot/L_\odot$, much larger than that of any known stellar population. This is therefore another indication of the presence of dark matter, since we can only see about 5% of the estimated mass in the Milky Way and Andromeda.

6.1.3 - Other Components of the Local Group

Tidal Streams: One of the most interesting galaxies in the Local Group is the *Sagittarius Dwarf Galaxy*, which was discovered in an analysis of the stellar kinematics in the direction of the Galactic bulge. It is located close to the Galactic plane at a distance of 16kpc from the Galactic center. This proximity implies that it must be experiencing strong tidal gravitational forces on its orbit around the Milky Way, and so it is being slowly disrupted. In fact, a relatively narrow band of stars have been found likely originating from the SgrD Galaxy in orbit of the Milky Way. With a low metallicity, it is very likely that they are stripped members. In addition, globular clusters were identified which presumably once belonged to the Sagittarius dwarf galaxy which were also removed by tidal forces.

The Neighbourhood of the Local Group: Most galaxies are members of a group. Many more dwarf galaxies exist than luminous galaxies, and dwarf galaxies are located preferentially in the vicinity of larger galaxies.

Ch 6.2 - Optical Cluster Searches

6.2.2 - Morphological Classification of Clusters

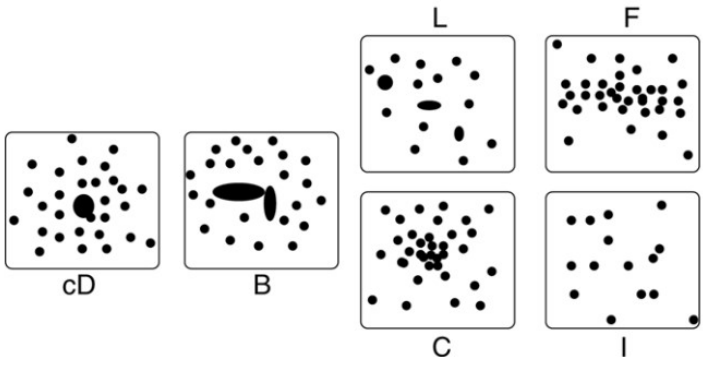


Fig. 6.10 Rough Morphological Classification of Clusters
cDs are those dominated by a central cD galaxy, Bs contain a pair of bright galaxies in the center. Ls are clusters with a nearly linear alignment of the dominant galaxies, Cs have a single core of galaxies, Fs are cluster with an oblate galaxy distribution, and Is are clusters with an irregular distribution. Regular clusters are on the left, whereas irregular clusters are on the right.

A rough classification can provide an idea of the state of a cluster, i.e., whether it is currently in dynamical equilibrium or whether it has been heavily disturbed by a merger process with another cluster. Therefore, one distinguishes in particular between regular and irregular cluster, and also those which are intermediate; the transition between classes is continuous. Regular clusters are 'compact' whereas irregular clusters are 'open'.
Regular clusters are found to be completely dominated by early-type galaxies, whereas irregular clusters have a fraction of spirals nearly as large as in the general distribution of field galaxies. Very often, regular clusters central galaxy density is very high. In contrast, irregular clusters are significantly less dense in the center and often show strong substructure which is rarely found in regular clusters.

6.2.3 - Galaxy Groups

Groups are the continuation of clusters towards fewer member galaxies and therefore presumably of lower mass, lower velocity dispersion and smaller extent, although the distinction is at least partially arbitrary.
Compared to clusters, the intergalactic gas in groups has a lower temperature, and possibly a lower metallicity.

6.2.4 - Modern Optical Cluster Catalogs

Colour-Magnitude Diagram: All early-type galaxies in a cluster have nearly the same colour, only weakly depending on luminosity; more massive (and so more luminous) ellipticals have somewhat higher metallicity, rendering the stellar emission slightly redder. The colour-magnitude diagrams of different clusters at the same redshift define a very similar red cluster sequence: early-type cluster galaxies with the same redshift and luminosity have virtually the same colour.

The fact that cluster galaxies at the same redshift all have roughly the same colour indicates that their stellar populations have very similar ages. In fact, the colour of cluster galaxies is compatible with their stellar populations being roughly the same age as the Universe at that particular redshift. This also provides an explanation for why the red cluster sequence is shifted towards *intrinsically* bluer colours at higher redshifts – there, the age of the Universe was smaller and thus the stellar population was younger. This effect is of particular importance at high redshifts.

Ch 6.3 - Light Distribution and Cluster Dynamics

6.3.2 - Dynamical Mass of Clusters

Consider the dynamical time-scale of clusters, defined as the time a typical galaxy needs to traverse the cluster once,

$$t_{\rm cross} \sim \frac{R_A}{\sigma_v} \sim 1.5 h^{-1} \times 10^9 \text{ yr} \qquad (6.19)$$

where, assuming a velocity dispersion of 1000km/s, gives a dynamical time-scale shorter than the age of the universe which implies that clusters are gravitationally bound.
As for the mass of the cluster at large, with defined velocity dispersion and cluster radius R_G, the mass can be calculated

as:

$$M = \frac{3\pi R_G \sigma_v^2}{2G} \simeq 1.1 \times 10^{15} M_\odot \left(\frac{\sigma_v}{1000\text{km/s}}\right)^2 \left(\frac{R_G}{1\text{Mpc}}\right) \tag{6.28}$$

This mass estimate doesn't depend on the mass of individual galaxies, but rather the gravitational potential of them all.

The 'Missing Mass' Problem in Clusters of Galaxies: With the mass M and the number N of galaxies, one can derive a characteristic mass $m = M/N$ for the luminous galaxies. Typically, $m \sim 10^{13} M_\odot$. Alternatively, M can be compared with the total optical luminosity of the cluster galaxies, and hence the mass-to-light ratio can be calculated. Typically,

$$\left(\frac{M}{L_{\text{tot}}}\right) \sim 300h \left(\frac{M_\odot}{L_\odot}\right) \tag{6.29}$$

This value exceeds the M/L ratio of early-type galaxies by at least a factor of 10. Hence, a major fraction of the mass in galaxy clusters consists of (non-baryonic) dark matter; *the stars visible in galaxies contribute less than about 5% to the total mass in clusters of galaxies.*

Ch 6.4 - Hot Gas in Galaxy Clusters

6.4.1 - General Properties of the X-Ray Radiation

Clusters of galaxies are the brightest extragalactic X-ray sources besides AGNs. Outside of the Milky Way, about 85% of X-ray sources are AGNs while the remaining 15% are galaxy clusters. In contrast to AGNs (for which X-ray emission is essentially point-like), cluster X-ray emission is extended. The fact that it is spatially extended implies that it does not originate from individual galaxies.

Continuum Radiation: The spectral energy distribution of the X-rays leads to the conclusion that the emission process is optically thin thermal free-free radiation from a hot gas which is collisionally ionized. This radiation is produced by the acceleration of electrons in the Coulomb field of protons and atomic nuclei. Since an accelerated electrically charged particle emits radiation, such scattering processes between electrons and protons in an ionized gas yields emission of photons. For galaxy clusters with mass in the range of $\sim 3 \times 10^{13} M_\odot$ to $\sim 10^{15} M_\odot$, a gas temperature in the range of 10^7 to 10^8K is obtained.

Line Emission: As a rule, hotter gas implies more ionized gas, which corresponds to weaker line emission. One finds that the mass in the intracluster gas is about five to ten times larger than the mass of the stars in the galaxies.

Morphology of the X-ray Emission: From the morphology of their X-ray emission, one can roughly distinguish between regular and irregular clusters. Regular clusters show a smooth brightness distribution, centered on the optical center of the cluster, with an outward decreasing surface brightness. In contrast, irregular galaxies may have several brightness maxima, often centered on cluster galaxies or subgroups of cluster galaxies.

Ch 6.6 - Clusters of Galaxies as Gravitational Lenses

6.6.1 - Luminous Arcs

By differential light deflection, the light beam of a distant source can be distorted in such a way by a foreground cluster that highly elongated arc-shaped images are produced. Hence, arcs are strongly distorted and highly magnified images of galaxies at high redshift. A single, distant galaxy can be represented by multiple arcs, and a single cluster can arc many distant galaxies (each with perhaps more than one arc).

Many clusters of galaxies in which arcs are observed are not dynamically relaxed, and are still undergoing dynamical evolution. They are often young systems with an age not much larger than t_{cross}, or systems whose equilibrium was disturbed by a recent merger.

The investigation of galaxy clusters with the gravitational lens method provides a third, completely independent method of determining cluster masses. It confirms that the mass of galaxy clusters significantly exceeds that of the visible matter in stars and in the intracluster gas. This result means that clusters of galaxies are dominated by dark matter.

6 Week 6 - Galaxy Interactions and Formation

CO Ch 26.1 - Interactions of Galaxies

Evidence of Interactions

Interactions between galaxies within clusters tend to increase the velocity dispersions of stars in the galaxies involved, possibly even destroying disk structures in late-type galaxies.
At least half of all disk galaxies display warped disks, possibly due to tidal interactions with satellite galaxies. Also, more than half of elliptical galaxies harbor discrete shells of stars which is evidence of past mergers.
Observations also suggest that hot, X-Ray emitting gas occupies much of the space between galaxies in rich clusters and has a mass at least equal to (and often exceeding) the stellar mass in the entire cluster. It's hypothesized that gravitational perturbations strips the gas from galaxies while still leaving the gas trapped within the gravitational potential well of the cluster.
Galaxy mergers will often initiate a burst of star formation, where the resulting stellar mass loss and supernovae will liberate gas from the galaxy over time.

Dynamical Friction

Given that stars are generally spread very far apart in galaxies, stellar collisions in galaxy mergers are very rare. Instead, interactions between stars will be gravitational. In this way, the merging galaxies will experience ***dynamical friction***, where as one galaxy moves through the other, stars are gravitationally pulled towards its path and hence slowing the merging galaxy via gravitational drag. In this case of merging galaxies, there will be an overdensity of stars behind the center of the merging galaxy which enact a net gravitational force that opposes the motion.

Not only are globular clusters affected by dynamical friction, but satellite galaxies too. Any giant galaxy will undoubtedly devour numerous satellite galaxies during its lifetime, with the process of satellite accretion having a variety of possible consequences. For instance, the counter-rotating cores observed in some elliptical galaxies may be explained by the merger of a satellite galaxy originally in a retrograde orbit. Mergers are also capable of producing the sort of featureless disks that are characteristic of Ir galaxies.

Rapid Encounters

Occasionally, galaxies will pass-by or interact with each other so quickly that their stars don't have time to respond. Even in the case where the two systems pass through each other, there is no significant dynamical friction. Since the stars (approximately) don't alter their positions during the interaction, the internal potential energy of each galaxy is unchanged over the event. However, the gravitational work that each galaxy performs on the other increases the internal kinetic energies of both galaxies in a random way.
One way for a galaxy to regain equilibrium after a sudden increase in kinetic energy is to convert the excess KE into an increased (less negative) gravitational potential energy, and so the galaxy expands slightly. Another way to return to equilibrium is for the most energetic components to carry the KE out away from the galaxy in a stream of stars and gas. In practice, both of these phenomena may occur.

For instance, a high speed, nearly head-on collision may produce a ***ring galaxy***.
Polar-ring galaxies and ***dust-lane ellipticals*** are normal galaxies that are orbited by rings of gas, dust, and stars that were stripped from other galaxies as they passed by or merged. Polar rings typically contain $10^9 M_\odot$ or more of gas and are found only around elliptical or S0 galaxies.

Modelling Interactions with N-Body Simulations

If the relative speed of two galaxies is substantially greater than the velocity dispersions of their stars, the collision will not result in a merger. Conversely, if the relative speed of the collision is less than the velocity dispersion, a merger is inevitable. The result is not trivial in the case that the relative speed is comparable to the velocity dispersion.

In general, only close, slow galactic encounters produce bridges and tidal tails that link the two galaxies after the interaction event. Two bulges tend to develop on opposite sides of one (or both) of the galaxies. Tidal stripping then pulls out streams of stars and gas as the two galaxies pirouette around one another. When conditions are right, the stars and gas torn from the near side will form an apparent bridge, while, because of angular momentum conservation, the material stripped from the far side moves off to form a curving tail.

There are two different spatial distributions of globular clusters around the Milky Way, with the inner clusters being old and varying little in age, and the outer population having a wider age spread. The inner clusters tend to have a preferential direction for their orbits about the Galaxy, while the outer clusters have randomly oriented orbits — suggesting that the outer clusters were captured via mergers/stripping events.

Starburst Galaxies

Strongly interacting galaxies tend to be bluer than isolated ones of the same genus. The excess blue light may be attributed to hot newborn stars formed via the strong tidal interactions between two or more galaxies. The light emitted by young stars (within dense gas clouds) at visible and ultraviolet wavelengths is absorbed by the clouds and reradiated in the infrared spectrum. These starburst galaxies are extremely bright in the infrared, with up to 98% of their total luminosity being in the infrared (compared to 30% for the Milky Way and only a few percent for some ellipticals).

While most common, observational evidence suggests that starbursts don't occur exclusively within the nuclei of galaxies. As galaxies merge, angular momentum in gas is lost and shock fronts compress the gas, leading to a burst of star formation.

Binary Supermassive Black Holes

An intriguing consequence of the merging of two large galaxies is the apparently inevitable formation of a binary system of supermassive black holes. The two central SMBHs in a merger migrate toward the center of the potential well of the combined system as a direct result of dynamical friction. Depending on their trajectories, the two black holes would probably enter into a well-separated binary system.
These SMBHs would likely have an extreme effect on the stars between them, ejecting stars and gas via 3-body interactions. If enough star ejection occurs, the angular momentum loss of the SMBHs would result in them spiraling closer together. Then, when sufficiently close, they would produce large amounts of gravitational radiation that would pull them closer together still via loss of energy and angular momentum. The final outcome would be the merger of the two black holes. This is a mechanism by which the central SMBH of a galaxy can grow extremely large.

Ch 10.1 - Galaxy Evolution Introduction

The evolution of structure in the Universe is seeded by density fluctuations. In the cold dark matter universe, small density structures formed first, which means that low-mass dark matter halos preceded those of higher mass. The baryonic gas in these halos is compressed and heated, the source of which being the potential energy. If the gas can cool by radiative processes, it can collapse into denser structures and eventually form stars.

The mass of halos grow either by merging processes of smaller-mass halos or by accreting surrounding matter through the filaments of the large-scale density field. If the gas within the halo can cool, it will eventually coalesce towards the center and form a disk galaxy with the disk perpendicular to the angular momentum vector. As soon as star formation sets in, it has a feedback on the gas: the most massive stars very quickly explode as supernovae, putting energy into the galactic gas and consequently heating it. This feedback then prevents all of the gas from turning into stars on a very-short timescale, providing a self-regulation mechanism of the star-formation rate. With the accretion of additional material as a result of the gravity of the surrounding dark matter halo, additional gas is added to the forming galaxy as raw material for further star formation.

Ch 10.4 - The Formation of Disk Galaxies

10.4.1 - The Contraction of Gas in Halos

A non-spherical overdensity of matter can attain angular momentum due to a torque caused by the tidal gravitational field in which the overdensity is located. Therefore, dark matter halos are born with a finite angular momentum.
When the gas in a halo cools, it collapses toward the center, thereby conserving its angular momentum. The gas therefore cannot collapse to an arbitrarily small region since the angular momentum barrier prevents this. Frictional forces in the gas drive the gas onto approximately circular orbits (depending on the symmetries in the halo) in a plane perpendicular to the angular momentum vector — it forms a flat disk. The gas in the disk is much denser than it would be in a spheroidal distribution and so the gas in the disk finds it easier to cool and form stars.

The Necessity for Dark Matter: The characteristic spin parameter of the forming halo is about 0.05, while the spin parameter of a self-gravitating, thin exponential disk can be calculated as $\lambda_d \approx 0.425$. Using equation (10.11), it can be found that the halo size required for the Milky Way to form from a *baryonic* halo would be about 700kpc. With a baryonic mass of $\sim 5 \times 10^{10} M_\odot$, the collapse time of such a halo from 700kpc to the current 10kpc of the visible disk would be about 4×10^{10} years, which is about three times the current age of the universe. Therefore, the Milky Way could not have formed as it is today solely from baryonic matter. Since we observe *many* old stars in the Galaxy, it must be very old and so the galaxy formation must have involved a dark matter halo.

Gas Collapse in a Dark Matter Halo: When accounting for dark matter mass, the necessary halo size would be about 70kpc, a factor of 10 less than that of the required size from a strictly baryonic halo. In this case, the baryonic gas collapse within a dark matter halo would occur on a timescale of $\sim 10^9$ years, and so these flat galactic disks can form sufficiently early in the cosmic evolution.

Ch 10.5 - Formation of Elliptical Galaxies

Properties of Ellipticals: Stars in ellipticals feature a high velocity dispersion, indicating that they were not formed inside a cool gas disk, or that the stellar distribution was subsequently heated very strongly.

Monolithic Collapse: A simple model capable of coherently describing the observational characteristics is that of the monolithic collapse, whereby the gas in a halo is nearly instantaneously transformed into stars, consuming most of the gas in the process. The process of star formation remains unexplained in this picture, and so ellipticals are believed to have been formed via the transformation of existing galaxies.
Most ellipticals are found in dense environments, like groups and clusters, and ellipticals congregate towards the center of these high-mass structures. Furthermore, ellipticals have complicated kinematics, often exhibiting small disks (sometimes counter-rotating) around the center, as well as shells and ripples which indicate a lively history.

10.5.1 - Merging of Halos and their Galaxies

When two halos merge to form one with larger mass, their baryonic components will be affected as well. Clearly, after two spiral galaxies collide, the resulting stellar distribution almost always doesn't resemble that of a spiral anymore.
There must now be distinction between the type of halo mergers:

- *Minor Mergers*, where the mass ratios of halos is large (typically in excess of 3:1)
- *Major Mergers*, where the two halo masses are similar.

Conditions for Merging: In order for a merger to happen, the collisional speed between the galaxies has to be of the same order, or smaller, than the intrinsic velocity dispersion. This implies that effective mergers of galaxies do not occur in massive clusters where the velocity dispersion of the galaxies of the cluster is considerably higher than the stellar velocity dispersion of individual galaxies. In contrast, groups of galaxies have both a high density of galaxies (making collisions probable) and a sufficiently low velocity dispersion to enable to merging of galaxies.

Minor Mergers: If a smaller galaxy merges with a massive one, the properties of the dominating galaxy are expected to change only marginally; the small galaxy will be embedded into the bigger halo, and survive as a satellite galaxy for a long while. Depending on the orbit of the

satellite galaxy, however, it may not survive forever due to tidal stripping and dynamical friction. Since the stars of satellite galaxies have small velocity dispersions, they are added as coherent 'streams' to the main galaxy over time.

Major Mergers and Morphological Transformations of Galaxies: In such 'major' mergers, the galaxies will change completely. The disks are very often destroyed due to a sudden increase in the disks' velocity dispersion which transforms the stellar distribution into a spheroid. Furthermore, the gas orbits are disrupted which may result in massive starbursts. By means of these perturbations, angular momentum loss can lead to gas infalling towards the central SMBH which leads to AGN activity. Due to the violence of interactions, gas and stars are ejected from the system, often visible as tidal tails.

Dry vs Wet Mergers: The situation is slightly more complex than the above text suggests. Galaxy mergers are often called 'wet' if the two progenitor galaxies are gas rich (and hence triggering starbursts), contrasted against 'dry' mergers where gas only plays a small role in the merger.
Wet mergers are often characterised by the dissipational properties of the gas, where the associated friction can lead to higher spatial densities than is possible for collisionless matter only. The gas can be driven towards the center of the merger remnant where it condenses to form stars. Wet mergers lead to considerably larger rotational velocities and central velocity dispersions than dry mergers.
As most ellipticals are seen as 'red-and-dead' in terms of colour and star formation rate, wet mergers must have happened early on in the cosmological epoch where there was a large concentration of gas in the progenitor galaxies. The resultant star formation from the merger would then have happened long ago, resulting in the old stars seen today.

The Resulting Scenario for the Formation of Ellipticals: Lower-mass normal ellipticals (not incl. dwarfs) are formed by wet major mergers of gas-rich (disk) galaxies at high redshift. Such mergers preferentially occur in overdense regions, which explains why mergers are so often found in groups and clusters (where clusters are mainly formed by merging of groups). In these dense environments, some of the ellipticals merge with other ellipticals, and these dry mergers lead to the formation of more massive galaxies with the characteristics of observed massive ellipticals.

7 Week 7/8 - Cosmology and Expansion Kinematics

Ch 4.1 - Cosmology Introduction and Fundamental Observations

Cosmological observations are difficult in general, simply because the majority of the universe is very far away from us. Distant light sources are very dim, and so our knowledge of the universe runs parallel with the development of more advanced telescopes.

The finite speed of light in a Euclidean space, in which we are located at the origin $r = 0$ today ($t = t_0$), implies that we can only observe points in spacetime for which $|r| = c(t_0 - t)$. This implies that we can't observe any point in spacetime, and the set of points which satisfy that relation are called our *backward light cone*.

4.1.1 - Fundamental Cosmological Observations

1) The sky is dark at night (Olbers' paradox)
2) Averaged over large angular scales, faint galaxies are uniformly distributed across the sky.
3) With the exception of a very few nearby galaxies, a redshift is observed in the spectra of galaxies — most galaxies are moving away from us, with their velocity increasing linearly with distance.
4) In nearly all baryonic cosmic objects, the mass fraction of helium is 25-30%.
5) The oldest star clusters in our Galaxy have an age of ~ 12 Gyr.
6) A microwave radiation (CMBR) is observed, reaching us from all directions. This radiation is isotropic except for very small fluctuations with amplitude $\sim 10^{-5}$.
7) The spectrum of the CMBR corresponds, within very small error bars, to that of a perfect blackbody of a temperature $T_0 = 2.728 \pm 0.004$ K.
8) The number counts of radi sources at high Galactic latitude does *not* follow the simple law $N(> S) \propto S^{-3/2}$

4.1.2 - Simple Conclusions

Start with the assumption of an infinite, on-average homogeneous, Euclidean, static universe.

Olbers' paradox (1): In this assumed universe, the sky would be uncomfortably bright. With these assumptions, there would be a uniform spatial density of stars throughout the whole universe; in a spherical shell twice as far away, there would be four times as many stars (and so on). This implies that the universe would be filled with equally dense volumes of stars, and so the sky would be filled with starlight of temperature on the order of 10^4 K or above. This is obviously not the case.

Source Counts (8): Consider a population of light sources with a luminosity function that is constant in space and time. In a spherical shell, we would expect the counts of sources with flux $F > S$ to follow $N(> S) \propto S^{-3/2}$. This is in contradiction to observations.

From these two contradictions — Olbers' paradox and the non-Euclidean source counts — we conclude that at least one of the underlying assumptions of this universe must be wrong; our Universe cannot be all four of Euclidean, homogeneous, infinite, and static. The Hubble flow, i.e., the redshift of galaxies, indicates that the assumption of a static

Universe is wrong.

The Age of Globular Clusters (5): requires that the Universe is at least 12 Gyr old, as it cannot be younger than the objects it contains. The age estimates for globular clusters yield values which are very close to the ***Hubble time*** $H_0^{-1} = 9.78h^{-1}$ Gyr. This suggests that the Hubble expansion may be directly linked to the evolution of the universe.

The apparently isotropic **distribution of galaxies (2)** when averaged over large scales, and the **CMBR isotropy (6)** suggest that the Universe is isotropic over large angular scales. If we assume, in addition, that our place in the cosmos is not privileged over any other place, then the assumption of isotropy around us implies that the Universe appears isotropic as seen from any other place. The homogeneity of the Universe follows immediately from the isotropy about every location.
The combined assumption of homogeneity and isotropy of the Universe is also known as the ***cosmological principle***.

The assumption of homogeneity breaks down at small scales, of course, due to galaxies, galactic clusters, galactic superclusters, etc. Large scale structures have been found that extend over $\sim 100h^{-1}$ Mpc (where $h = H_0$ in units of 100km/s/Mpc), but no indication of structures much larger than this. This length-scale can be compared to a characteristic length of the universe. This ***Hubble radius***, the characteristic length-scale of the observable Universe, is found by

$$R_H \coloneqq \frac{c}{H_0} = 2998h^{-1} \text{ Mpc} \tag{4.3}$$

The Hubble volume found from a sphere of Hubble radius can contain a very large number of structures of size $\sim 100H^{-1}$ Mpc, so it still makes sense to assume an on-average homogeneous cosmological model.

Ch 4.2 - An Expanding Universe

Only gravitational forces and electromagnetic forces can act over large distance. Since cosmic matter is electrically neutral on average, electromagnetic forces do not play any significant role on large scales; gravity has to be considered as the driving force in cosmology.

4.2.2 - Kinematics of the Universe

Comoving Coordinates: Consider a homogeneous sphere which may be radially expanding (or contracting), requiring that the density $\rho(t)$ be spatially homogeneous. The density may vary in time due to expansion or contraction. For some coordinate, $\boldsymbol{x}$, in this radially expanding sphere with radius from the origin of r at $t = t_0$, the radius at time t of this coordinate is

$$r(t) = a(t)\boldsymbol{x} \tag{4.4}$$

Since $\boldsymbol{x}$ and r both have the dimension of a length, $a(t)$ is dimensionless and only depends on time and is spatially constant. The function $a(t)$ is called the ***cosmic scale factor***, with $a(t_0) = 1$. Particles which move according to (4.4) are called comoving particles.

Expansion Rate: The velocity of a comoving particle is obtained by

$$\boldsymbol{v}(r,t) = \dot{a}\boldsymbol{x} = \frac{\dot{a}}{a}r \equiv H(t)r \tag{4.6}$$

where the ***expansion rate*** is

$$\boxed{H(t) \coloneqq \frac{\dot{a}}{a}} \tag{4.7}$$

Equation 4.6 yields a result strikingly similar to the Hubble law

$$\boxed{v = H_0 D} \tag{4.9}$$

where D is the distance to an object. These three equations express the fact that any observer expanding with the sphere will observe an isotropic velocity field that follows the Hubble law. Since we are observing an expansion today, we have $H_0 > 0$ and $\dot{a}(t_0) > 0$.

4.2.3 - Dynamics of the Expansion

Equation of Motion: From conservation of mass, one finds the acceleration of the scale factor can be calculated by

$$\ddot{a}(t) = \frac{\ddot{r}(t)}{x} = -\frac{4\pi G}{3}\rho(t)a(t) \tag{4.13}$$

where $\rho(t)$ is the mass density of the universe at some time. It's important to note that this equation does not depend on x, and so the dynamics of expansion are determined solely by matter density.

Conservation of Energy: Another way to describe the dynamics of the expanding shell is based on the law of energy conservation: the sum of kinetic and potential energy is constant in time.
The constant K is proportional to the total energy of a comoving particle, and thus the history of the expansion depends on K. The sign of K characterizes the qualitative behaviour of the cosmic expansion history:

- If $K < 0$, da/dt remains positive for all times, and the universe expands forever.
- If $K = 0$, da/dt is still positive for all times, but approaches 0 as $t \to \infty$. Therefore the universe will expand forever but to a diminishing extent.
- If $K > 0$, the expansion will turn into a contraction and such a universe will re-collapse.

In the special case of $K = 0$, the universe has a current density called the ***critical density***, with

$$\rho_{\text{cr}} \coloneqq \frac{3H_0^2}{8\pi G} = 1.88 \times 10^{-29} h^2 \text{ g/cm}^3 \tag{4.15}$$

This is a characteristic density of the current universe. Define the ***density parameter*** as

$$\boxed{\Omega_0 \coloneqq \frac{\rho_0}{\rho_{\rm cr}}} \tag{4.16}$$

where $K > 0$ corresponds to $\Omega_0 > 1$, and $K < 0$ corresponds to $\Omega_0 < 1$.

The matter which is visible as stars contributes only a small fraction to the density of our universe, $\Omega_* \lesssim 0.01$. However, it is already known that dark matter plays an important role in the dynamics on smaller scales, and so it can dominate the value of Ω_0 too, at least in principle.

4.2.4 - Modifications due to General Relativity

The Newtonian approach contains nearly all essential aspects of homogeneous and isotropic world models. However, the image of an expanding sphere needs to be revised — this picture implies that a 'center' of the universe exists. Our Universe neither has a center, nor is it expanding away from a privileged point.
General Relativity modifies the Newtonian model in several aspects:

- Via Special Relativity, mass and energy are equivalent. This implies that it is not only the matter density that contributes to the equations of motion, and radiation fields like the CMB have an energy density which affects the expansion equations.
- The field equation of GR as originally formulated by Einstein did not permit a solution which corresponds to a homogeneous, isotropic, and static cosmos. He modified his equations by introducing an additional term (the cosmological constant) which gave a solution for a static universe, in which he falsely believed in.
- The interpretation of the expansions is changed completely: it is not the particles or the observers that are expanding away from each other, nor is the Universe and expanding sphere. Instead, it is space itself that expands. In particular, the redshift is not Doppler redshift, but it itself a property of expanding spacetimes.

4.2.5 - The Components of Matter in the Universe

The essential components of matter in our Universe can be described as pressure-free matter, radiation, and vacuum energy.

Pressure-Free Matter: In cosmology, a substance with $P \ll \rho c^2$ is denote as (pressure-free) matter, also called cosmological dust. We approximate $P_m = 0$, where the index 'm' is for matter. The constituents of the (pressure-free) matter move with velocities much smaller than c.

Radiation: If the above condition is not satisfied (thus if the thermal velocities are no longer negligible compared to the speed of light), then the pressure will also no longer be small compared to ρc^2. In the limiting case that the thermal velocity equals the speed of light, we denote this component as 'radiation'. One example of course is electromagnetic radiation. Even particles of finite mass can have a thermal velocity very close to c if the thermal energy of the particles is much larger than the rest mass energy, i.e. $k_B T \gg mc^2$.

Vacuum Energy: The cosmological constant is equivalent to a finite vacuum energy density. The energy density of the vacuum is constant in space and time, and has a negative pressure. The vacuum energy density is given by

$$\boxed{\rho_{\rm v} = \frac{\Lambda}{8\pi G}} \tag{4.23}$$

4.2.6 - "Derivation" of the Expansion Equation

When accounting for pressure, the equation of motion of the universe expansion is

$$\boxed{\ddot{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) a} \tag{4.22}$$

4.2.7 - Discussion of the Expansion Equations

For matter, radiation and vacuum density, we obtained:

$$\begin{aligned} &\rho_m(t) = \rho_{m,0} a^{-3}(t); \quad \rho_r(t) = \rho_{r,0} a^{-4}(t); \\ &\rho_{\rm v}(t) = \rho_{\rm v} = \text{ constant} \end{aligned} \tag{4.24}$$

where the index '0' indicates the current time, $t = t_0$. As for the matter, the number density of photons changes $\propto a^{-3}$ because the number of photons in a comoving volume is unchanged. However, photons are redshifted by the cosmic expansion. Their wavelength changes proportionally to a, and so the energy of a photon changes as a^{-1} due to cosmic expansion. With these two together, the photon energy density changes as $\propto a^{-4}$.
We define the dimensionless density parameters for matter, radiation, and vacuum as

$$\boxed{\Omega_m = \frac{\rho_{m,0}}{\rho_{\rm cr}}; \quad \Omega_r = \frac{\rho_{r,0}}{\rho_{\rm cr}}; \quad \Omega_\Lambda = \frac{\rho_{\rm v}}{\rho_{\rm cr}} = \frac{\Lambda}{3H_0^2}} \tag{4.25}$$

so that $\Omega_0 = \Omega_m + \Omega_r + \Omega_\Lambda$.
The matter densities of galaxies (including their dark matter halos) correspond to $\Omega_m \gtrsim 0.02$, and so provides a lower limit of Ω_m. Studies of galaxy clusters yield a lower limit of $\Omega_m \gtrsim 0.1$, and across the universe shows that $\Omega_m \sim 0.3$.

Photons are not the only contributors to the radiation energy density. In addition, there are neutrinos from the early cosmic epoch which add to the density parameter of radiation. Today, the energy density of radiation in the Universe can be neglected when compared to that of matter.

The constant K is obtained from the density parameters, mainly those of matter and vacuum since $\Omega_r \ll \Omega_m$,

and has the dimension of (length)$^{-2}$. In the context of GR, K is interpreted as the curvature scalar of the universe today, or more precisely, the homogeneous, isotropic three-dimensional space at time $t = t_0$ has a curvature K. Depending on the sign of K, we can distinguish 3 cases:

- If $K = 0$, the three-dimensional space for any fixed time t is Euclidean, i.e. flat.
- If $K > 0$, the curvature radius of the spherical 3-space is given as $1/\sqrt{K}$, where a 2D analogy would be the surface of a sphere.
- If $K < 0$, the space is called hyperbolic, the 2D analogy would be the surface of a saddle.

Hence, GR provides a relation between the curvature of space and the density of the universe. If the universe has a simple topology, it is finite in the case of $K > 0$, whereas it is infinite if $K \leq 0$. However, in both cases it has no boundary.

We finally obtain the expansion equation:

$$\begin{aligned} H^2(t) &= \left(\frac{\dot{a}}{a}\right)^2 \\ &= \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3} \qquad (4.18) \\ &= H_0^2\left[\frac{\Omega_r}{a^4(t)} + \frac{\Omega_m}{a^3(t)} + \frac{(1-\Omega_m-\Omega_\Lambda)}{a^2(t)} + \Omega_\Lambda\right] \\ &\equiv H_0^2 E^2(t) \qquad (4.33) \end{aligned}$$

where the dimensionless Hubble function, $E(t) = H(t)/H_0$ was defined in the last step.

Paper: *Cosmological Constraints on Dark Energy* — Tamara Davis

Form observations of type Ia supernovae at distance, the universe has been found to be expanding at an accelerating rate. The cause of this is due to some unobserved ***dark energy***. A naive approximation of vacuum energy (via quantum physics) to explain this dark energy results in a value 54 orders of magnitude smaller than what is observed. This expansion in the form of dark energy is uniform throughout space.

The standard cosmological model, ΛCDM, consists of approximately 30% matter and 70% cosmological constant, Λ. Most of the matter is cold (non-relativistic) and dark (does not interact electromagnetically), so is known as cold dark matter (CDM). It is expanding at about $H_0 = 70$km s^{-1}Mpc^{-1}, and redshift is related to the scale factor a by equation (4.41) later on in the document.

How the expansion rate changes with time is determined by the gravitational effect of the components of the universe, such as the matter density, cosmological constant density, and radiation density, denoted by: Ω_m, Ω_Λ, Ω_r. Observations show that the universe is close to flat, and so $\Omega_m + \Omega_\Lambda + \Omega_r \sim 1$.
The strength of cosmological constraints come from the complementarity of many different types of observations. In particular, the cosmic microwave background (CMB) is very important, and most of the other observables give strong constraints only when combined with some information from the CMB.

The number counts of galaxies, the predictions of inflation along with the measurements of the CMB, and the ages of the oldest stars, were the fertile ground into which the discovery of the acceleration of the expansion of the universe fell.

Peculiar velocities refer to any motion that is not just the smooth flow of the expansion of the universe. They arise because galaxies tend to fall towards each other and form clusters and filaments. These motions shift the redshifts away from what they would be if the galaxies were flowing perfectly with the expansion, and also shift the luminosities of supernovae slightly due to Doppler beaming. The contribution of peculiar velocities is dominant at low redshift, and comes in the form

$$(1 + z_0) = (1 + \overline{z})(1 + z_p) \qquad (1)$$

where the observed redshift, z_0, gets contributions from the cosmological redshift, $\overline{z}$, and the peculiar redshift, z_p. The peculiar redshift can be calculated from a velocity by

$$z_p = \gamma - 1 = \frac{1}{\sqrt{1 - v^2/c^2}} - 1$$

The second discussed feature that would cause scatter about the magnitude-redshift relation for Type Ia supernovae is *gravitational lensing*. Supernovae behind overdense regions of space (dense in the sense of matter) should be brighter, on average, than those found behind empty regions.

5 Cosmic Microwave Background

When the universe was still hot and dense enough to be plasma, there were no atoms to put absorption or emission lines in spectra. As such, the CMB remains to this day the most perfect black-body spectrum ever measured.
The tiny, $T \sim 10^{-5}$K fluctuations seen in the CMB act as seeds for the structure of the universe, where overdensities in the CMB eventually collapsed under gravity to form the galaxies, filaments, and clusters that make up the cosmic web today.

The components of the early universe interacted in various ways. Dark matter essentially ignored everything else from the get-go, and simply collapsed under its own gravity without any opposing pressure forces. Normal matter has to deal with very high pressures, and thus initial density fluctuations propagated as sound waves. The very high pressure was primarily generated by interactions with photons, whose mean-free-path was so short as to make the protons, electrons, and photons act as a single fluid. The CMB was

emitted when the universe had expanded and cooled enough for electrons to stick to protons to form hydrogen, and with the electrons no longer free, the mean-free-path for the photons was essentially infinite. Shortly after this time, the sound waves froze as the pressure supporting the waves dropped dramatically when the photon pressure disappeared.

When plotting a power spectrum of the CMB temperature fluctuations, one finds that there is a preferred separation of about 1 degree where hot spots (and cold spots) are correlated. Physically, this corresponds to the distance that sound waves could have travelled between the birth of the universe and the time when the CMB was emitted (the surface of last scattering). The second CMB peak corresponds to the pattern of rarefractions on the scale where a sound wave has had time to compress and rarefy once before last scattering. The third peak corresponds to the second compression, and so on. The spacing and height of these peaks were determined by several effects, on of the dominant ones being the ratio of baryonic matter to dark matter, another being the damping effect of radiation.

One of the most important measurements from the CMB gives an indication of the curvature of space. We know the wavelength of the first acoustic peak theoretically from our calculations of sound waves in the early universe, and so it may be used as a standard ruler. By measuring how large it appears, we can measure how far away it is. The apparent size, however, is heavily dependent on the curvature of the universe. A negatively curved universe would mean that the angular size of a standard ruler would appear smaller than in a positively curved universe. The position of the smaller-scale peaks allows for a more accurate model, and so the position of the first peak in the CMB gives a strong indication of curvature which tells us whether the universe is close to the critical density (which would make the universe flat).

Since the CMB is such a uniform background radiation, it also serves as a beautiful backlight for revealing matter distribution in the foreground. The gas and galaxies it passes through en route to our telescopes interact with the CMB and change it in a variety of ways.

8 Week 9 - Expansion Dynamics

Ch 4.3 - Consequences of the Friedmann Expansion

4.3.1 - The Necessity of a Big Bang

From the Hubble law, we conclude that $\dot{a}(t_0) > 0$, i.e. a is currently an increasing function of time.

Classification of the Model: Which of these alternatives describes our universe depends on the density parameters.

- If $\Lambda = 0$, then $H^2 > 0$ for all $a \leq 1$, whereas the behaviour for $a > 1$ depends on Ω_m:
  - If $\Omega_m \leq 1$ (or $K \leq 0$, respectively), $H^2 > 0$ for all a: the universe will expand for all times.
  - If $\Omega_m > 1$ $(K > 0)$, H^2 will vanish for $a = a_{\max} = \Omega_m/(\Omega_m - 1)$. The universe will have its maximum expansion when the scale factor is $a_{\max}$ and will collapse thereafter.
- In the presence of a cosmological constant $\Lambda > 0$, the discussion becomes more complicated; the sign of K is not sufficient to predict the qualitative expansion behaviour. One finds, for $\Lambda \geq 0$:
  - If $\Omega_m < 1$, the universe will expand for all $a > 1$
  - For $\Omega_m > 1$, the future behaviour of $a(t)$ depends on Ω_Λ: if Ω_Λ is sufficiently small, a value $a_{\max}$ exists at which the expansion comes to a halt and reverse. Conversely, if Ω_Λ is large enough, the universe will expand forever.
  - If $\Omega_\Lambda < 1$, then $H^2 > 0$ for all $a \leq 1$.
  - If $\Omega_\Lambda > 1$, it is in principle possible that $H^2 = 0$ for an $a = a_{\min} < 1$. Such models (where a minimum value for a existed in the past — so called bouncing models) can be excluded by observations.

With the exception of the last case, we come to the conclusion that a must have attained the value $a = 0$ at some point in the past. At this instant, the "size of the universe" formally vanished. As $a \to 0$, both matter and radiation densities diverge so that the density in this state must have been singular. The epoch at which $a = 0$ and the evolution away from this state is called ***The Big Bang***. It is useful to define this epoch as the origin of time, so that t is identified with the age of the universe, the time since the Big Bang.

Only a negative pressure can cause an accelerated expansion — more precisely, $P < -\rho c^2/3$ (from equation (4.22)) is needed for $\ddot{a} > 0$. We know that our universe is currently undergoing an accelerated expansion, and thus that the cosmological constant differs significantly from 0. If Ω_Λ is sufficiently large, the universe's expansion accelerates.

Age of the Universe: The age of the universe at a given scale factor a follows from

$$\begin{aligned} \int dt &= \int \left(\frac{da}{dt}\right)^{-1} da \\ &= \int \frac{1}{aH}\, da \\ \Rightarrow t(a) &= \frac{1}{H_0}\int_0^a \left[x^{-2}\Omega_r + x^{-1}\Omega_m + \right. \\ &\qquad \left. (1 - \Omega_m - \Omega_\Lambda) + x^2\Omega_\Lambda\right]^{-1/2} dx \end{aligned} \tag{4.36}$$

To obtain the current age of the universe, t_0, (4.36) is calculated with $a = 1$.
The values of the cosmological parameters are now quite well known. They are, approximately,

$$\boxed{\Omega_m \sim 0.3; \quad \Omega_\Lambda \sim 0.7; \quad h \sim 0.7} \tag{4.37}$$

4.3.2 - Redshift

The relation between the redshift of an object and the scale factor a is

$$\boxed{1 + z = \frac{1}{a}} \tag{4.41}$$

Cosmic Microwave Background: If a Planck distribution of photons had been established at an earlier time, it will persist during cosmic expansion. The CMB is such a blackbody radiation, with a current temperature of $T_0 = T_{\rm CMB} \approx 2.73$K. It is meaningful to consider the temperature of the CMB as the 'temperature of our universe' which is a function of redshift:

$$\boxed{T(z) = T_0(1 + z) = T_0\, a^{-1}} \tag{4.45}$$

and so the universe was hotter in the past than it is today.

Interpretation of Cosmological Redshift: The redshift results from the fact that during the expansion of the universe, the energy of the photons decreases in proportion to $1/a$, which is the reason, together with the decreasing proper number density, that $\rho_r(a) \propto a^{-4}$.

The Necessity of a Big Bang: One gap in our argument that led to the necessity of a Big Bang still remains, namely the possibility that at some time in the past, $\dot{a} = 0$ occurred, i.e. that the universe underwent a transition from a contracting to an expanding state. This is possible only if $\Omega_\Lambda > 1$, and if the matter density parameter is sufficiently small. In this case, a attained a minimum value in the past, which is dependent on both Ω_m and Ω_Λ. For instance, for $\Omega_m > 0.1$ (as in our universe), the value is $a_{\rm min} \gtrsim 0.3$. A minimum value for a implies a maximum redshift $z_{\rm max} = 1/a_{\rm min} - 1 \lesssim 2.4$. However, we have observed quasars and galaxies with $z > 6$ and the density parameter is known to be $\Omega_m > 0.1$, and so such a model without a Big Bang can be safely excluded.

4.3.3 - Distances in Cosmology

Because of the monotonic behaviour of the corresponding functions, each of a, t, and z provide the means to sort objects according to their distance. An object at redshift z_2 is more distant than one at $z_1 < z_2$ such that light from a source at z_2 may be absorbed by gas in an object at redshift z_1, but not vice versa. The more distant a source is from us, the longer the light takes to reach us, the earlier it was emitted, the smaller a was at emission, and the larger z is. Since z is the only observable of these parameters, distance in extragalactic astronomy are nearly always expressed in terms of redshift.

In a *static Euclidean space*, the separation between two points is unambiguously defined, and several prescriptions exist for measuring a distance. Examples include typical flux/luminosity relations, and standard geometric formulae. In a non-Euclidean, or expanding/contracting space-time (our universe), this is no longer the case. The equivalence of different distance measures is only ensured in Euclidean space, and does not hold in a curved spacetime. In cosmology, the same measuring prescriptions as in Euclidean space are used for defining distances, but the different definitions lead to different results. The two most important definitions of distance are:

- ***Angular-Diameter Distance:*** Consider a source of radius R observed to cover a solid angle ω. The angular-diameter distance is defined as

$$\boxed{D_A(z) = \sqrt{\frac{\pi R^2}{\omega}}} \tag{4.49}$$

- ***Luminosity Distance:*** Consider a source with bolometric luminosity L and flux F, and define its luminosity distance as

$$\boxed{D_L(z) = \sqrt{\frac{L}{4\pi F}}} \tag{4.50}$$

The two distances agree locally, and on small scales the curvature of spacetime is not noticeable. There exists a general relation between angular-diameter distance and luminosity distance:

$$D_L(z) = (1 + z)^2\, D_A(z) \tag{4.52}$$

In the low-redshift universe ($z \ll 1$), the differences between different distance definitions are very small.

9 Week 10 - Misc Lecture Material

10 Week 11 - Big Bang Nucleosynthesis and Inflation

Ch 4.4 - Thermal History of the Universe

Since $T \propto (1+z)$, our Universe was hotter at earlier times. As such, we might expect energetic processes like nuclear fusion to have taken place in the early Universe. A few comments should serve as preparation for the content in this section:

- Temperature and energy may be converted into each other since $k_B T$ has the dimension of energy.
- Statistical physics and thermodynamics of elementary particles are described by quantum mechanics. A distinction has to be made between ***bosons*** (particles of integer spin, e.g. the photon), and ***fermions*** (particles of half-integer spin, e.g. electrons, protons, neutrinos, etc).
- If particles are in thermodynamic and chemical equilibrium, their number density and energy distribution are specified solely by the temperature.

4.4.2 - Expansion in the Radiation-Dominated Phase

The radiation density behaves like $\rho_r \propto T^4$. Since $T \propto 1/a$, $\rho_r \propto a^{-4}$ and so radiation dominates in the early universe when $a \ll 1$.

4.4.3 - Decoupling of Neutrinos

We start our consideration of the universe at a temperature of $T \approx 10^{12}$K, which corresponds to ~ 100 MeV. This energy can be compared to the rest mass of various particles:

$$\begin{aligned} \text{protons: } & m_p = 983.3 \text{ MeV}/c^2 \\ \text{neutrons: } & m_n = 939.6 \text{ MeV}/c^2 \\ \text{electrons: } & m_e = 511 \text{ keV}/c^2 \\ \text{muons: } & m_\mu = 140 \text{ MeV}/c^2 \end{aligned}$$

Therefore, protons and neutrons (the baryons) are too heavy to be produced at the temperature considered, and so all baryons that exist today must have been present at this early time. The production of muon pairs at this temperature is also not efficient because the temperature is not sufficiently high. Hence, at $T \approx 10^{12}$K, electrons, positrons, photons and neutrinos would have been present, all contributing to the radiation density ρ_r since they were relativistic.

Like baryons, WIMPS are non-relativistic at the energies considered in the early universe and so neither of them effectively contribute to the energy density at these times.

From about one second after the big bang, neutrinos began propagating without considerable interaction and so maintained their thermal distribution up to the present day with a temperature decreasing as $T \propto a^{-1}$. This prediction suggests that those neutrinos are still in the universe today at a temperature of 1.9K, however are indetectable due to their low cross section.

4.4.4 - Pair Annihilation

At temperatures smaller than $\sim 5 \times 10^9$ (or $k_B T \sim 500$keV), electron-positron pairs can no longer be produced efficiently, however their pair annihilation continues and the density of e^+e^- pairs decreases rapidly. Pair annihilation injects additional energy into the photon gas as a result of the kinetic and rest mass energy of the destroyed electrons/positrons. This changes the energy distribution of photons, which continues to be a Planck distribution but now with a modified temperature relative to that it would have had without annihilation. The neutrinos, already decoupled at this stage, don't benefit from the additional energy. This means that the photon temperature now exceeds that of the neutrinos. From the thermodynamics of this process, the change in photon temperature is computed as

$$\boxed{aT = \left(\frac{11}{4}\right)^{1/3} aT_\nu} \tag{4.62}$$

where T_ν is the neutrino temperature. This temperature ratio is preserved afterwards so that neutrinos have a temperature lower than that of photons by about $(11/4)^{1/3} \simeq 1.4$ until the present epoch.

Before pair annihilation, there were about as many electrons and positrons as there were photons. After annihilation, nearly all electrons were converted into photons, but some remained as there was a very small excess of electrons over positrons to compensate for the positive electrical charge density of the protons. Therefore, the number density of electrons that survive the pair annihilation was exactly the same as the number density of protons, so that the Universe remained electrically neutral.

4.4.5 - Primordial Nucleosynthesis

Protons and neutrons can fuse to form atomic nuclei if the temperature and density of the plasma are sufficiently high. In the interior of stars, these conditions for nuclear fusion are obviously provided. The high temperatures in the early phases of the universe suggest that atomic nuclei may also have formed then.

Proton-to-Neutron Abundance Ratio: Since free neutrons decay with a time-scale of $\tau_n = 881$s, the early neutrons must have been bound to atomic nuclei quite quickly to explain the number of them we see in the present day. At the time of neutrino decoupling, the ratio of protons to neutrons was about $n_n/n_p \approx 1/3$.

Deuterium Formation: The simplest compound nucleus is that of deuterium (D), consisting of a proton and a neutron and formed in the reaction

$$\mathrm{p} + \mathrm{n} \rightarrow \mathrm{D} + \gamma$$

At the early stages of the universe, there are a very large number of highly energetic photons which break apart deuterium atoms and so a deuterium abundance only becomes appreciable when $k_B T \ll E_b$ (where E_b is the binding energy of D). The temperature at which the deuterium formation rate exceeds the photon dissociation rate at about $T_D \approx 8 \times 10^8$K, corresponding to $t \sim 3$ minutes. Up to then, a fraction of the neutrons decayed, yielding a proton-neutron ratio at T_D of $n_n/n_p \approx 1/7$.
After that time, deuterium formation begins rapidly and virtually all neutrons become bound in D. Once the deuterium density has become appreciable, helium (He^4) forms, which is a nucleus of high binding energy and therefore cannot be destroyed by photo-dissociation. Except for a small remaining fraction, all deuterium is quickly transformed into He^4.

Helium Abundance: The number density of helium nuclei can now be calculated, since virtually all neutrons present are bound in He^4. The mass fraction Y of He^4 of

the baryon density follows,

$$\boxed{Y = \frac{4n_{\mathrm{He}}}{4n_{\mathrm{He}} + n_{\mathrm{H}}} \approx 1/4} \tag{4.67}$$

Therefore, about 1/4 of the baryonic mass in the universe should be in the form of He^4. This is a robust prediction of Big Bang models and is in excellent agreement with observations.
The helium content in the universe of course changes later by nuclear fusion in stars, but this contribution is expected to be an order of magnitude smaller than from the Big Bang Nucleosynthesis (BBN).

4.4.6 - WIMPs as Dark Matter Particles

There is a wide variety of evidence for the existence of dark matter, from scales of individual galaxies (rotation curves of spirals), clusters (velocity dispersion of galaxies, X-ray temperature, lensing), to cosmological scales, where the baryon density as inferred from the BBN is lower than the lower bound on the total matter density. The MACHO experiments rule out astronomical objects as the dominant contribution to dark matter, at least in the halo of milky way. In addition, the fact that the mass density in the universe is ~ 6 times higher than the baryonic density precludes any 'normal' astronomical objects as the main constituent of dark matter. Therefore, the solution must lie in particle physics.

Constraints on the Dark Matter Particle: Since dark matter particles are 'dark', they must be electrically neutral in order to not interact electromagnetically. Furthermore, the particle must be stable, or at least have a lifetime much longer than the age of the universe. The only known neutral particles in the Standard Model are the neutrinos and the neutron. However, the neutron is baryonic (with its density in the early universe well constrained by BBN), and the free neutron is unstable. The neutrinos in principle could be good candidates, but would be hot dark matter and as such would lead to a large-scale structure in the universe that would be different to the one we observe. In particular, they would be too 'hot' to cluster on scales of galaxies, with their thermal velocity exceeding that of the escape velocity of galaxy halos.

4.4.7 - Recombination

About 3 minutes after the Big Bang, BBN comes to an end. At this time, the universe has a temperature of roughly $T \sim 8 \times 10^8$K and consists of photons, protons, helium nuclei, traces of other light elements, and electrons. In addition, there are neutrinos that dominate, together with photons, the energy density and thus also the expansion rate. Except for the neutrinos and (probably) WIMPs, all particle species have the same temperature which is established by interactions of charged particles with the photons (which resemble a sort of heat bath).

After some time, pressureless matter (i.e. the non-relativistic dust) begins to dominate the cosmic energy density and thus the expansion rate. This dominates the expansion behaviour until either the curvature term or, if this is zero or very small, the Λ-term starts to dominate.

After further cooling, the free electrons can combine with the nuclei and form neutral atoms. This process is called *recombination* (even though the particles had never combined in the first place and this represents the first transition to a neutral state). The recombination of electrons and nuclei is in competition with the ionization of neutral atoms by energetic photons (photoionization), whereas collisional ionization can be disregarded completely. Because photons are so much more numerous than electrons, cooling has to proceed to well below the ionization temperature (corresponding to the binding energy of an electron in hydrogen), before neutral atoms become abundant. This happens for the same reasons that apply in the context of deuterium formation. The ionization energy of hydrogen corresponds to a temperature of $T > 10^5$K, but T has to first decrease to ~ 3000K before the ionization fraction falls considerably below 1.

Photons can propagate from $z \sim 1000$ (the '*last-scattering surface*') until the present day essentially without any interaction with matter, provided the wavelength is larger than 121.6nm. The absorption cross section for smaller wavelength photons is large with neutral atoms, and so, disregarding these, we conclude that photons present after recombination have been able to propagate without further interactions until the present epoch. Before recombination, the photons followed a Planck spectrum, and as was discussed earlier, will remain a Planck spectrum after recombination, only with temperature changing. Thus, these photons from the early universe are still observable today, redshifted into the microwave regime of the electromagnetic spectrum in what is the Cosmic Microwave Background.

Although only the recombination of hydrogen was discussed, helium has a higher ionization energy and so recombines earlier than hydrogen. Although recombination defines a rather sharp transition, we receive photons from a recombination layer of finite thickness of about $\Delta z \sim 60$.

The gas in the intergalactic medium at lower redshift is highly ionized. If this were not the case, we wouldn't be able to observe any UV photons from sources at high redshift. Sources with redshifts $z > 6$ have been observed, and so the universe at at least the epoch corresponding to this redshift must have been nearly fully ionized. This means that at some time between $z \sim 1000$ and $z \sim 6$, a reionization of the intergalactic medium must have occurred, presumably by a first generation of stars or by the first AGNs. The results from new satellites suggest that this occurred at $z \sim 10$.

11 Week 12 - Constraints on Dark Energy and CMB History

Ch 4.5 - Achievements and Problems of the Standard Model

4.5.1 - Achievements

The standard model of the Friedmann-Lemaitre universe described as been extremely successful in numerous ways:

- It predicts that gas which has not been subject to much chemical processing should have a helium content of $\sim 25\%$. This is in great agreement with observations.
- It predicts that sources of lower redshift are closer to us than sources of higher redshift. Therefore, modulo any peculiar velocities, the absorption of radiation from sources at high redshift must happen at smaller redshifts. Not a single counter-example has yet been found.
- It predicts the existence of a microwave background, which was indeed found.
- It correctly predicts the number of neutrino families.

The Friedmann-Lemaitre universe is excellent in that it is inherently falsifiable. If any of the above points were not found in observations, the whole model would be disproven.

4.5.2 - Problems

Despite the achievements of the standard model, there are some aspects which require further consideration.

The Horizon Problem: Since no signal can travel faster than light, the CMB radiation from two directions separated by more than about one degree originates in regions that were not in causal contact before recombination, i.e. the time when the CMB photons interacted with matter the last time. Therefore, these two regions have never been able to exchange information, for example about their temperature. Nevertheless, their temperature is the same, as seen from the high degree of isotropy of the CMB (which shows relative fluctuations of only $\Delta T/T \sim 10^{-5}$).

The Flatness Problem: For the total density parameter to be of order unity today, it must have been extremely close to 1 at earlier times, which means that a very precise 'fine tuning' of this parameter was necessary.

The above also suggests that the curvature term has never had a dominant effect on the universe.

4.5.3 - Extension of the Standard Model: Inflation

For all of the cosmological parameters to be the value that they are today (and with them, the universe not collapsing/rapidly expanding so that we see it in its current state today) is unlikely from a random standpoint. Different parameters could have led to the universe recollapsing billions of years ago or expanding so rapidly that no structure could be formed. These parameters come down to the initial conditions of the universe just after the Big Bang. The only answer to why these parameters are the values that they are today lies in some processes that must have taken place even earlier in the universe, and so the initial conditions of the normal Friedmann-Lemaitre expansion have a physical origin. Cosmologists believe they have found such a physical reason in the inflationary model.

Inflation: Physical laws and properties of elementary particles are well known up to energies of ~ 100 GeV because they have been experimentally tested in particle accelerators. For higher energies, particles and their interactions are unknown. This means that the history of the Universe can only be considered secure up to this 100 GeV energy. The extrapolation to earlier times, up to the Big Bang, is considerably less certain. From particle physics, we expect new phenomena to occur at an energy scale of the Grand Unified Theories (GUTs), at about 10^{14} GeV, corresponding to $t \sim 10^{-34}$s.
In the inflationary scenario, it is presumed that at very early times the vacuum energy density was much higher than today, and so it dominated the Hubble expansion. Because of this, there would have been exponential expansion (inflationary phase), but only for a brief period. We assume that a phase transition took place in which the vacuum energy density is transformed into normal matter and radiation (a process called reheating), which ends the inflationary phase and after which the normal Friedmann evolution of the Universe begins.

Inflation solves the Horizon Problem: During inflation, the Hubble parameter is constant and so the horizon may become arbitrarily large in the inflationary phase (dependent on the duration of the exponential expansion). According to this scenario, the whole universe visible today was in causal contact prior to inflation, so that the homogeneity of the physical conditions at recombination, and with it the nearly perfect isotropy of the CMB, is provided by causal processes.

Inflation solves the Flatness Problem: Due to the tremendous expansion, any initial curvature is straightened out. Formally, during the inflationary phase we have

$$\Omega_\Lambda = \frac{\Lambda}{3H^2} = 1$$

and since it is assumed that the inflationary phase lasts long enough for the vacuum energy to be completely dominant, when it ends we then have $\Omega_0 = 1$. Hence the universe is flat to an extremely good approximation.

12 Week 13 - Baryon Acoustic Oscillations and CMB Peaks