Lecture Notes: Physical Chemistry in Outer Space

Diethelm Johannsmann, Institute of Physical Chemistry, TUC

Last updated: 8/9/2024

1 Motivation

Because physical chemistry is cross-sectional, its principles apply in other fields as well, including astronomy. When looking up, we find the same laws as when looking right or left. Actually, we see a bit more when looking up. The same laws apply in outer space, but the conditions are different. Neutron stars are an example. Studying neutron stars changes our views on the more mundane parts of physical chemistry.[1]

The course focuses on topics that are particularly interesting from a physical-chemical point of view. Among the recurring questions are:

- To what extent must thermodynamics be adapted?
- What are the mechanisms of structure formation? Is there complexity in the outer space?
- What are the analogies between outer space and the laboratory?

Although this course is not intended as a lecture on astrophysics, an attempt is made to at least briefly touch on all the important topics (always from the perspective of physical chemistry). When the links to physical chemistry are not explored in depth, the book by Binney entitled "Astrophysics: A very short introduction" serves as a guide.[2] Binney is quoted so often that the reference to footnote 2 is omitted occasionally.

Disclaimer: These are lecture notes, this is not a textbook. Because lecture notes live and change, inaccuracies and errors occasionally creep in. If there is room for improvement, please send an e-mail to johannsmann@pc.tu-clausthal.de.

One more disclaimer: I spent time to diligently formulate the German version of this text. The translation was done with DeepL and I spent limited time to catch errors. Equations are labeled as “Glg.” (the abbreviation for German “Gleichung”). It would have been much work to change those labels into “Eq.” because these are targets for cross-references.

2 In heaven as on earth (?)

Following Binney, we start with a few philosophical remarks. It is not self-evident that the laws of nature should be the same "up there" and "down here". Deviating from Binney, we question this view. The doubt is motivated by the anthropic principle (section 2.1). In a second step, we accept the universality of the laws but question the universality of the fundamental constants, that is, the universality of the values of the parameters contained in the laws. In two cases, fundamental constants can be obtained from both astrophysics and laboratory experiments. These cases are a temporal drift of the fine structure constant ($\dot{\alpha} \neq 0$?, $\dot{\alpha}$ pronounced as "alpha-dot") and the neutrino mass.

The insight that the same laws apply in the sky and down here is usually attributed to Newton. He was sitting in a garden and saw an apple fall. According to Newton, the apple (in free fall) and the planets follow the same laws ("force = mass times acceleration" with the force in this case being gravity). Newtons view ("the same physical laws everywhere") is widely accepted today.[3] A spectacular success of this concept is the discovery of helium (Box 2.1). Binney in Ref. 2 starts with this observation and

---

[1] For this reason, the course occasionally digressed to certain topics from physical chemistry with analogies in astrophysics - yes - but without direct relevance to astrophysics. Astrophysics is then - for a while - only the teaser and we talk physical chemistry.

[2] Binney, J. *Astrophysics: A Very Short Introduction,* Oxford University Press 2016

[3] A historica side note: At the beginning of the 20th century, it was assumed that the atom had a similar structure to the planetary system. This "planetary model" was refined by Bohr. However, it lost its virtue for explaiing nature when the wave nature of electrons was recognized. It does not follow from "As in heaven, so on earth" that the 1s electron behaves in the same way as Mercury.

places it in the context of John 1:1: "In the beginning was the Word", which here is meant to say: the laws of nature are above the material world (says Binney).

Again: In the beginning was the Word. Or not? It is generally assumed that elementary logic applies everywhere. The equation 1 + 1 = 2 should also be true on Aldebaran, although it is not clear how this should be tested. Furthermore, there are no indications that the fundamental structure of the laws of nature would be different in outer space. The processes in space usually are modeled with partial differential equations, just like the weather or the mechanics of railway tracks.[4] It is quite possible, though, that certain natural constants have different values far away from here or had different values in the early universe. A somewhat separate question is the extent to which thermodynamics is valid in astrophysics. Although thermodynamics is not bound to specific physical laws, it has prerequisites, which are not always fulfilled. Of course, these prerequisites also exist on Earth, but in astrophysics the limits of thermodynamics are more impressive (sections 4.4.1 and 4.4.3).

Box 2.1: **Discovery of helium in the solar spectrum**
The discovery of helium is an example of how the universality of the laws of physics has astonished and continues to astonish people. Towards the end of the 19th century, our understanding of nature was perceived as more or less complete – with the exception of small clouds in the sky of natural science that were later given the name "quantum mechanics".
Line spectra were known (and understood later in the context of quantum mechanics) and the same lines were found as absorption lines in the spectrum of solar radiation (Fig. 2.1). The same lines? Not quite: There was a set of absorption lines in the solar spectrum that could not be assigned. A new element was postulated and named helium (from "helios", Greek for "sun"). A little later, this element was also found on Earth. Helium was discovered on the sun, but apart from that it has just as much to do with the sun as copper and lead. The periodic table is universal.

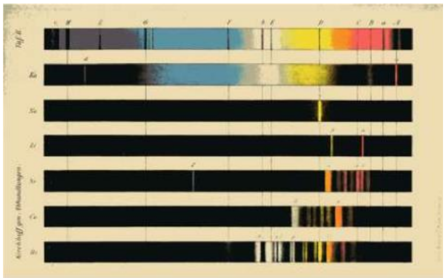


Fig. 2.1
Kirchhoff and Bunsen: "Chemische Analyse durch Spectralbeobachtungen" in: Annalen der Physik und Chemie, Vol. 110, No. 6, 1860, pp. 161-189.
Newly published by Gabriele Dörflinger, Heidelberg University Library. *www.haus-der-astronomie.de/3642696/04strahlung.pdf*

2.1 Fine tuning and the anthropic principle

The question of variable natural “constants” comes up in the context of the fine-tuning problem. It can be argued that life would not exist if the fine structure constant (governing the strength of electromagnetism, α) and the coupling constant of the strong interaction (g) had values slightly different from what we know. This fortunate circumstance may be viewed as a consequence of the anthropic principle. The anthropic principle applied to cosmology postulates that the universe is much larger than what we see.[5] We see to a horizon, which is 46 billion light years away.[6] The bubble visible to us is – according to the proponents of this variant of the anthropic principle – a single, small pinhead in a much larger overall universe. Only in this small area do α and g have the favorable values (favorable for us, the complex living entities). Elsewhere they are different. We live in a nice corner of this large universe. For the same reason, we live on Earth instead of Mars. That's nice but not surprising. We live here *because* it is nice here.

[4] The ordinary differential equations were developed by Newton *to* model celestial mechanics.

[5] Sometimes this larger universe is called the "multiverse", but the multiverse also has other meanings.

[6] 46 billion light years is the diameter of the visible universe *today*. The universe is only 13.7 billion years old. The fact that today's diameter is greater than 13.7 billion light years is due to cosmic expansion. The details are complicated.

In the following, we first introduce a mechanism that could bring about a very large universe (cosmic inflation, Section 2.1.3). Secondly, it is proposed that the natural constants should be temperature-dependent,[7] where high temperatures mean temperatures beyond 10^{15} K. Thirdly, it is proposed that the cooling that took place during the Big Bang may have led to different values for the low-temperature values of the natural constants in the different regions of the very large universe.

Box 2.2: We live where it‘s nice

The proponents of the fine-tuning problem claim that life is only possible in the narrow bar on the left in Fig. 2.2. If the nuclear forces were stronger (top left), there would be no hydrogen atoms, because two hydrogen nuclei would fuse to form a diproton. Without hydrogen, there would be no small molecules, because hydrogen atoms terminate the growth of (e.g.) carbon chains. Fluorine, chlorine, bromine and iodine are too rare to replace hydrogen in this respect. Without small molecules, there would be no liquids with low viscosity.

If nuclear power were much weaker, the stars would not shine because nuclear fusion would not be efficient enough (bottom left).

If the Coulomb repulsion between the protons in the nucleus were greater, carbon would decay radioactively just like uranium (bottom right). No carbon, no life ("carbon chauvinism").

If both α and g were large, the atoms would be so stable that they would no longer react chemically (top right).

This diagram has been criticized because it ignores alternative mechanisms. Other natural constants would not only close existing entries to the path of evolution, but also open new ones.

There are similar diagrams in which one of the axes is the strength of gravity.

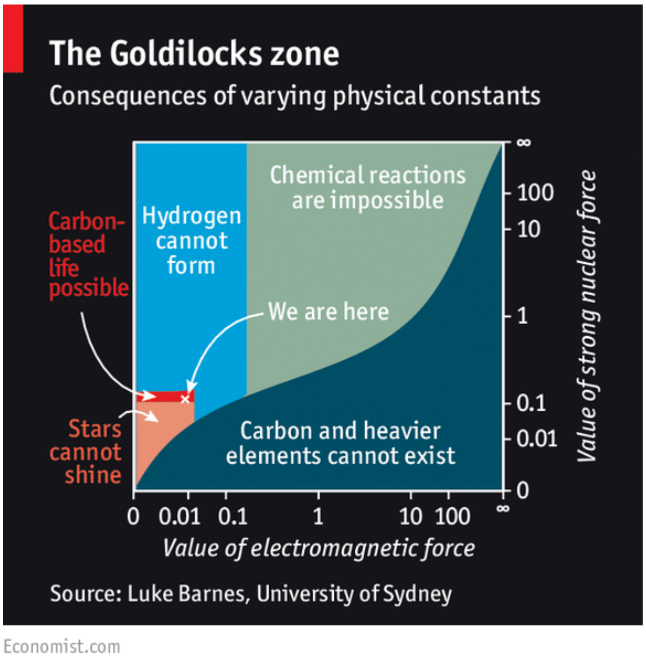


Fig. 2.2
A map of the problems that would arise if the fine structure constant α and the coupling constant of the strong interaction g had different values than they have. *https://www.niepelt.ch/wp-content/uploads/2015/08/20150815_BBC928_0.png*

2.1.1 General relativity in a nutshell, cosmic expansion

Soon after he had developed general relativity (GR), Einstein turned to cosmology and was disappointed. He had hoped to describe the cosmos as a three-dimensional space that curves into a fourth dimension (as a “hypersphere”). This idea charmingly avoids infinity. An observer flying on a straight line will eventually come back home, similar to a traveler on the equator.

However, as long as the diameter of this hypersphere was assumed as constant in time, this structure was incompatible with Einstein's field equations. The mass in the universe generates an attractive force, which results in an accelerated contraction of the entire universe. Einstein sighed and made the extensions to his theory which appeared to be necessary. Courage had left him. Instead of confidently asserting that the size of the universe cannot be constant in time, he modified his theory. He wrote

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \qquad \text{Glg. 2.1}$$

[7] Such "gliding coupling constants" are also contained in the standard model of particle physics.

The previous equation only had the first term on the left-hand side. The Einstein tensor $G_{\mu\nu}$ is the simplest divergence-free tensor that can be constructed from the curvature tensor.[8,9] The energy-momentum tensor $T_{\mu\nu}$ is also divergence-free. Einstein called Λ the "cosmological constant". Glg. 2.1 is the second simplest meaningful equation that links the curvature of space to the energy-momentum tensor.

When Hubble discovered the expansion of the universe on the basis of the redshift of distant galaxies, Glg. 2.1 was rehabilitated without a cosmological constant. Friedmann and Lemaitre had found a solution to Einstein's field equations without Λ, which has the form of the closed 3D hypersphere sought by Einstein, but requires a time dependence of the radius $a(t)$. The function $a(t)$ is linked by Friedmann and Lemaitre to the mass density of the universe. The larger the mass density, the more the rate of expansion slows down over time. Above a certain "critical" mass density, expansion eventually turns into contraction. The mass density of the universe is rather close to the critical value.

Later Einstein regretted that he had made doubts about the simple version of general relativity transparent. He described the cosmological constant as "the greatest folly of my life". That was the state of affairs until at the end of the 1970s. The cosmological constant returned a first time in the form of the inflationary universe (section 2.1.3). It returned a second time in the 1990s, when it was discovered that the cosmic expansion currently does not slow down (as predicted by Friedmann and Lemaitre), but on the contrary accelerates. Today, the cosmological constant is non-zero and is called "dark energy". The density of the dark energy is $6 \cdot 10^{-10}$ J/m^3.[10] Glg. 2.1 is the basis of the standard cosmological model. This is called the Λ–CDM model (also: “Lambda-CDM model”) "CDM" here stands for "cold dark matter".

In today's notation, the Friedmann equation reads

Glg. 2.2

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^3}{3}$$

$$2\frac{\ddot{a}}{a} = -\left(\frac{\dot{a}}{a}\right)^2 - \frac{kc^2}{a^2} + \Lambda c^2 = \frac{8\pi G}{c^2} p$$

ρ is the mass density (baryonic matter plus dark matter), p is a pressure and k is a constant that can take the values 1, 0 or –1. k is linked to the curvature. For this universe, $k = 0$. The Λ–CDM model is based on these equations and adjusts the constants ρ and Λ in such a way that there is agreement with the values of the expansion rate, which is derived from the redshift as a function of distance. In this case, Λ assumes a positive value. Furthermore, the mass density is larger than the density of the visible mass. There is also evidence for dark matter outside of cosmology (section 6). The expansion rate of the universe is the only evidence for dark energy.

2.1.2 The physics behind the dark energy

Λ is a term in the Einstein equations whose physical meaning is not obvious. (The meaning of the energy-momentum tensor $T_{\mu\nu}$ *is* known.) For reasons that can only be explained in the context of general

[8] In simple terms, the curvature tensor quantifies in what ways the sum of the angles in the triangle can differ from 180°. This difference is the central signature of curvature. The curvature becomes evident, even if the hypothetical fourth dimension, into which the three-dimensional space would curve, is invisible to the observer.

[9] Simplicity is considered a characteristic of general of relativity.

[10] New measurements of redshifts as a function of distance suggest that Λ may not be constant in time even today.

relativity, Λ is interpreted as the energy density of the vacuum.[11] Such a vacuum energy is known from vacuum fluctuations. A reminder: The quantum mechanical harmonic oscillator (potential: $V = ½\, \kappa x^2$ with κ the spring constants) has a zero-point energy. There is a similar zero-point energy for electric fields in vacuum. The energy is proportional to E^2 with E the electric field in the same way that the energy of the harmonic oscillator is proportional to x^2. The zero-point energy in electrodynamics is linked to vacuum fluctuations of the electric field. The vacuum is not empty. The vacuum fluctuations of the electric field are well known in physical chemistry because they cause spontaneous fluorescence. Spontaneous fluorescence is not really spontaneous. It is stimulated by the vacuum fluctuations.

However, the energy density of the vacuum fluctuations does not match the value of the constant Λ. It is too large by a factor of 10^{120} . This is called "the worst prediction in physics". Very well. For reasons explained in section 2.1.3, a strongly positive cosmological constant was suspected to have existed in the early universe long before the new measurements of accelerated cosmic expansion. However these are not caused by the vacuum fluctuations of the electromagnetic field of E.

2.1.3 The inflationary phase

The reason for formulating the inflation hypothesis was not the fine-tuning problem, but rather were two peculiarities of the Big Bang model.[12] First, the microwave background has almost the same temperature everywhere, although the different regions at the time were unable to communicate with each other due to the finite speed of light. The regions are mutually outside the respective "causal horizons". Heat did not flow between these regions. This is the "horizon problem". Second, the mass density of the universe is astonishingly close to the critical mass that marks the transition between a continuing expansion and a transition into the "big bounce". Researchers who are looking for a reason for this coincidence call this the "flatness problem".

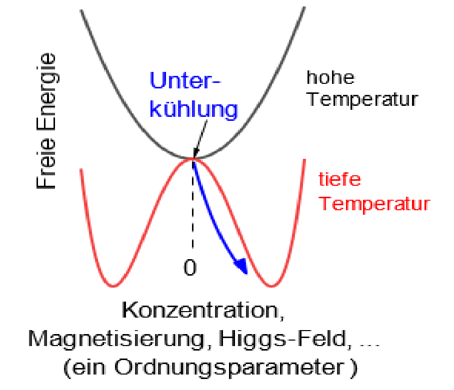


Fig. 2.3
Spontaneous symmetry break with hypothermia

Allan Guth explained these peculiarities in 1981 by claiming that the universe visible to us (the bubble enclosed by the microwave background) is a small part of a much larger universe. This would require the universe to have expanded much faster in an early phase than predicted by the conventional Big Bang model (Glg. 2.2). He proposed "cosmic inflation" as the reason. Since cosmic inflation, there are many

---

[11] In ref. 1, $g-\Lambda_{\mu\nu}$is understood as the stress tensor of the vacuum. The vacuum must look the same to all observers (with different velocities relative to each other) and therefore this tensor must be a multiple of the Minkowski metric (the metric tensor in flat space). The Minkowski metric is

$$g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

The simplest stress tensor has the form

$$\begin{bmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{bmatrix}$$

Comparison of the two tensors leads to the conclusion that the pressure, p, (assuming positive energy density) must be negative.

[12] According to some sources, the conventional Big Bang model has more open ends than mentioned here.

such bubbles and the bubbles are all locally flat in a certain sense ($k = 0$ in Glg. 2.2). The local flatness entails a mass density close to the critical mass density.

Inflation has an analogy in ferromagnetism, sketched in simplified form in Fig. 2.3. The importance of entropy increases at high temperatures. Maximum entropy exists at random orientation of the elementary magnets, where the macroscopic magnetization vanishes. At temperatures below the critical temperature (here: "Curie temperature"), the influence of enthalpy increases and the elementary magnets align themselves in parallel (all in the same direction). In between there is a 2nd-order phase transition, which causes a spontaneous symmetry breaking, because the magnets (together) might point in to all directions.

There is no such 2nd-order phase transition for electromagnetic fields in vacuum. The vacuum does contain quantum fluctuations of E, but E fluctuates around zero at all temperatures. $E = 0$ is the ground state. The ground state also corresponds to a field-free vacuum for the fields that cause the weak and the strong interaction. The respective particles (the vector bosons W^+, W^-, and Z_0, as well as the gluons) are created spontaneously for short times, but the wave functions still fluctuate around zero.

Only the Higgs particle is different (Fig. 2.4). The energy landscape above the Higgs field has its lowest energy at a finite value of the Higgs field. At temperatures above 10^{15} Kelvin, the Higgs field fluctuates around zero. This situation is the analog of a paramagnetic phase at a temperature above the Curie temperature.

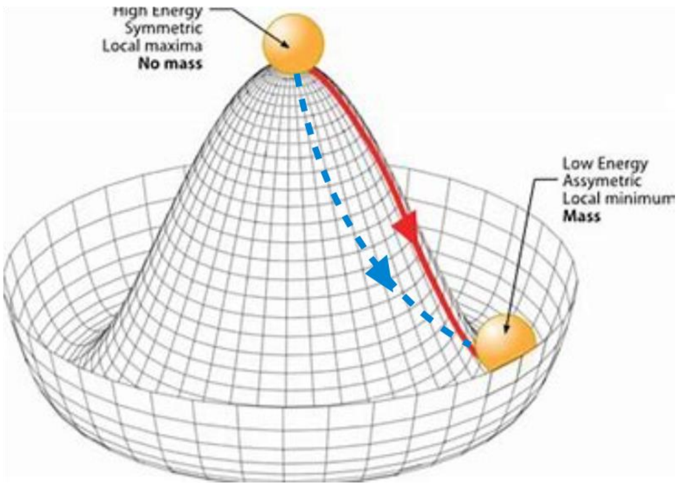


Fig. 2.4
The Higgs field does not have its energetic minimum at zero. This does not matter at high temperatures. Averaged over time, the magnitude of the Higgs field then is zero. At low temperatures (< 10^{15} K), the Higgs field moves towards the groove. This results in a spontaneous symmetry breaking. The blue dashed arrow indicates the possibility of supercooling. Adapted from *physics.stackexchange.com/questions/486003/how-can-the-higgs-field-sense-that-the-temperature-is-high-enough-to-restore-t*

Guth postulated two things. First, the Higgs field fluctuated around zero in the early universe. Second, the cooling was so incredibly fast that the Higgs field could not follow. This "supercooling"[13] is indicated by the dashed arrow in Fig. 2.4. The supercooling is associated with an energy in the same way that a supercooled melt contains a latent heat which is released when the crystallization is nucleated (as in heat cushions[14]). The analogy of the latent heat is the "dark energy" of the vacuum, which created a negative pressure and caused the universe to rapidly expand.

[13] This is a special form of supercooling in that the supercooled state is not metastable, but unstable.
[14] *en.wikipedia*.org/wiki/Latent *heat storage*

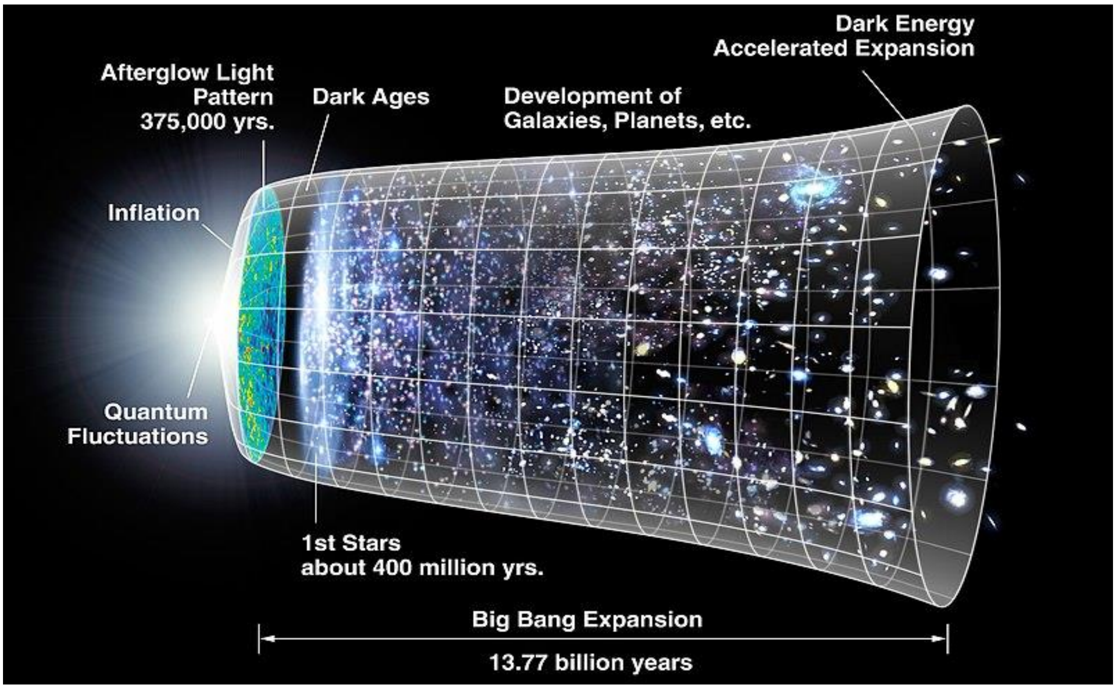


Fig. 2.5
The current cosmological model including inflation and accelerating expansion
en.wikipedia.org/wiki/Expansion_of_the_universe

Recent research has shown that inflation cannot be driven by the Higgs field. It is believed that there are other, currently unknown fields that have driven inflation.

Is there evidence that certain natural constants have had different values in the early universe? For the fine structure constant, the answer for the times when stars existed currently is no (section 2.2). However, there are some discrepancies between conditions shortly after the Big Bang and the conditions today:

- The amount of primordial lithium (lithium produced during the Big Bang) is a factor of three less than calculated. This is the "lithium problem".
- If the rate of expansion of the universe is calculated from the redshift of the microwave background, it is 7% smaller than if the redshift of the galaxies is extrapolated to this early point in time. The two values of the Hubble constants are 73 (km/s)/MPc for the galaxies and 67.7 (km/s)/MPc[15] for the microwave background. This difference is called "Hubble tension".
- The mass distribution in the universe today is slightly more homogeneous than what would be expected from calculation forward from the small inhomogeneities in the microwave background. The universe is currently less "clumpy" than expected (Box 4.2).
- If the inflation hypothesis is correct, the cosmological constant Λ had a different value shortly after the Big Bang than it does today.

There *are* indications that some things may be different in outer space. "Outer space" here means a time shortly after the Big Bang.

[15] MPc for mega parsec or mega parallax second. A star at a distance of 1 parsec changes its apparent position in the sky by one arc second over the course of half a year due to the movement of the earth around the sun. 1 parsec corresponds to 3.26 light years.

As a side remark, we include a well-known illustration, which summarizes the current standard model of cosmology (Fig. 2.5).

2.2 A time dependence of the fine structure constant?

The fine structure constant, α, can be measured with good precision and one can search for a temporal variation of α ($\dot{\alpha}$, pronounced alpha-dot). The name "fine structure constant" suggests that it would pertain to a detail from atomic spectroscopy. However, the fine structure constant may well be called the most important constant of nature. It is more important than the speed of light, *c*, Planck's quantum of action, *h*, and the vacuum permeability, ε_0, because *c*, *h* and ε_0 are not natural constants in the narrow sense. The values depend on the choice of the units meter, second and joule. In order to avoid this dependence of the natural constants on units, so-called "natural units" can be defined. This is done by writing $\hbar = c = 4\pi\varepsilon_0 = 1$. The units joule, meter and coulomb in this way disappear. For instance, the light-second is the unit of length. Only the second remains.[16] The fine structure constant – being dimensionless – does not depend on the definition of the units. α is defined as

Glg. 2.3

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 hc}$$

e is the elementary charge. The numerical value is $\alpha = 7.297352566 \cdot 10^{-3}$. The reciprocal value is also often noted $\alpha^{-1} \approx 137.035999$. In natural units, the fine structure constant is the square of the elementary charge. The value of the charge of the electron pervades all of physics. Implications of the fine structure constant for the physics of atoms and molecules are discussed in section 16.1.

Since the fine structure in the atomic spectra is caused by magnetic interactions (keyword: spin-orbit coupling), the fine structure constant can be determined spectroscopically from the absorption spectra of atoms. Atomic spectroscopy can be carried out with high precision both on earth and on light that reaches us from a large distance.

This raises the question of whether α has the same value on distant stars as down here. A distant star implies a distant past because of the light travel time. It could be that α changes on cosmic timescales and this is what the authors of Ref.17 have claimed. This is the one data point in Fig. 2.7, which is lower than the others and which is not compatible with the zero. Spectroscopy of stars and galaxies will always be less accurate than spectroscopy in the laboratory because the Doppler shift turns all lines into bands. This is also the case in Fig. 2.6. Two different regions from the spectrum of the quasar HE 0515–4414 are shown. The patterns in both spectra are similar because the Doppler shift (actually plural, the Doppler shifts from the different regions) has approximately the same effect on both lines. These are two iron lines in absorption. They are absorption lines because the iron in question is contained in the outer, colder regions of stars. Light with the relevant wavelength, which penetrates from the interior of the star through this photosphere to the outside, is absorbed in the colder outer layer.[18] (The same is true forthe Fraunhofer lines in Fig. 2.1.) Iron was chosen firstly because it is a relatively common element (section

---

[16] In the context of thermodynamics, $k_B = 1$ is also sometimes written. In this case, the temperature has the unit of energy. (The energy for this case is not reduced to the second by choosing $\hbar = 1$).

[17] J. K. Webb; M. T. Murphy; V. V. Flambaum; V. A. Dzuba; J. D. Barrow; C. W. Churchill; J. X. Prochaska; Wolfe, A. M., Further Evidence for Cosmological Evolution of the Fine Structure Constant. *Phys. Rev. Lett.* **2001,** 87, 091301

[18] Depending on the wavelength, you look into the star at different depths. At the wavelength of a strong line, the frequent absorptions and re-emissions mean that only the outer (colder and darker) areas of the photosphere are visible. Next to the line you see the hotter and brighter interior of the star.

11.1). Secondly, iron is a relatively rather element. Heavy elements have a strong spin-orbit coupling. The numerical value of α can be determined more easily and accurately using the iron spectrum than using the spectrum of sodium, for example (keyword: splitting of the sodium D-line).

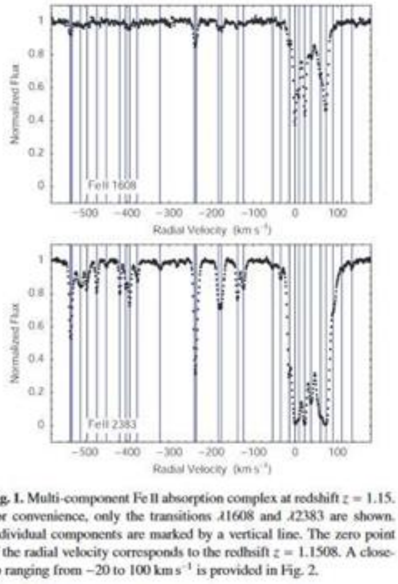


Fig. 2.6
Spectra of quasar HE 0515-4414.
The splitting results from the Doppler shifts in different regions of the source. The analysis focuses on the difference between the wavelengths of the two underlying iron lines (upper and lower graph).
Quast, R.; Reimers, D.; Levshakov, S. A., Probing the variability of the fine-structure constant with the VLT/UVES. *Astronomy & Astrophysics* **2004,** 415, (2), L7-L11.

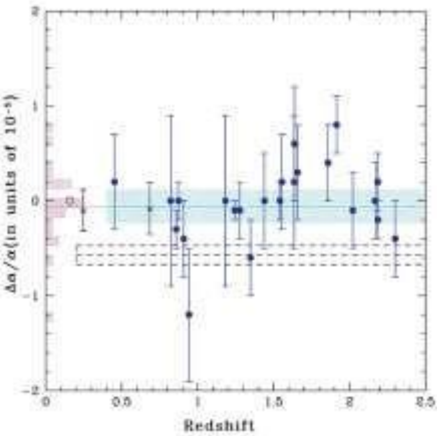


Fig. 2.7
The UVES measurement values of the relative change of the fine structure constant alpha from the sample of absorption systems, plotted as a function of the redshift and the corresponding look-back time. The open circle is the measurement from the Oklo natural reactor (near Gabon, West Africa). The horizontal long dashed lines show the area of the previous claim of variation of the fine structure constant. Clearly, the UVES data are inconsistent with this range.
www.eso.org/sci/publications/messenger/archive/no.116-jun04/messenger-no116-25-28.pdf

The authors of Ref. 17 remained alone with their results. Subsequent investigations have shown that even on a cosmic time scale the value of $\dot{\alpha}/\alpha$ is compatible with zero. Fig. 2.7 gives an overview. The 3σ–confidence interval for $\dot{\alpha}/\alpha$ from the astrophysical measurements is between $-2.5 \cdot 10^{-16}$ year^{-1} and $+1.2 \cdot 10^{-16}$ year^{-1}.

$\dot{\alpha}$ can also be determined in the laboratory.[19] For this purpose, the frequencies of two transitions of the ytterbium atom were measured with high precision and the measurement was repeated after some time. The result was $\dot{\alpha}/\alpha = (-0.7 \pm 2.1) \cdot 10^{-17}$ year^{-1}.[20] Since the uncertainty is larger than the mean the

[19] Godun, R. M.; Nisbet-Jones, P. B. R.; Jones, J. M.; King, S. A.; Johnson, L. A. M.; Margolis, H. S.; Szymaniec, K.; Lea, S. N.; Bongs, K.; Gill, P., Frequency Ratio of Two Optical Clock Transitions in Yb-171(+) and Constraints on the Time Variation of Fundamental Constants. *Physical Review Letters* **2014,** 113, (21).

[20] It is hoped that in the foreseeable future it will be possible to anchor the very precise optical clocks to an electromagnetic transition of an atomic nucleus (instead of to transitions in the electron shell). This will be the "thorium 229 clock". Because the processes in the nucleus are less disturbed by the environment than the processes in the shell, it is hoped that the thorium 229 clock will be a factor of 100 more precise than the current optical clocks. The problem is that the experiments on high-precision spectroscopy are already more accurate than the theory. This is unlikely to change because the radius of the proton must be taken into account. The protn radius results from the strong interaction and cannot be modeled with the required accuracy. So why the thorium 229 clock? Firstly, we will see. Secondly, the accuracy in the determination of $\dot{\alpha}/\alpha$ can be significantly increased once again.

result is compatible with zero. The bounds from the laboratory experiments are therefore closer to zero than the bounds from astrophysics.[21]

2.3 The neutrino mass

Although neutrinos are extremely numerous, they interact weakly with ordinary matter and are therefore largely invisible. But if you try hard enough, you can see them. Neutrinos are produced during nuclear reactions, especially in the sun.[22] In the famous Homestake experiment in the 1970s, the neutrinos that reach us from the sun were counted. A tank filled with tetrachloro-ethene was used for this purpose. Chlorine reacts with neutrinos to form ^{37}Ar, which decays to chlorine with a half-life of 35 days. To detect this decay, the tank was flushed with helium as a carrier gas every few weeks and the argon was collected in a cold trap. The collected argon was purified in several steps, fed into a specially designed proportional counter and analyzed for decay for 250 to 400 days. Due to the low count rate of around one event per week, special precautions had to be taken to suppress the ambient radiation.[23]

Firstly, the operators of the experiment believed that they had calibrated their detectors sufficiently. Secondly, they believed that they knew the number of neutrinos produced in the sun per day. With a calibrated detector, a known distance to the sun and a known production rate, they knew the expected count rate. However, the measured count rate was a factor of 3 lower, which constitutes the "solar neutrino problem". At the time, not everyone believed that the modeling of nuclear reactions in the sun was reliable down to a factor of 3 in the neutrino production rate. The calibration of the experiment was also always viewed as challenging. The nuclear physicists and those who ran the Homestake politely insisted and 30 years later it was confirmed that 2 out of 3 neutrinos do indeed turn into something, that the detector in the Homestake experiment cannot see, while they travel from the sun to the earth. These are the "neutrino oscillations". The neutrino oscillations prove that the neutrinos have a rest mass.

With neutrinos, a distinction must be made between “interaction states” and “energy eigenstates”. Fig. 2.8 puts this difference in the context of the pyramidal inversion of the ammonia molecule (which is often taught in physical chemistry as an example of a double-well potential). For more details to the pyramidal inversion, see Wikipedia. The chemical reactions that produce ammonia lead to tetrahedra, where the nitrogen is either located to the left of the triangle formed by the three hydrogen atoms (“L”) or to right (“R”). These are not energy eigenstates, though. The nitrogen may tunnel back and forth between L and R. The energy eigenstates are the symmetric (“symm”) and the antisymmetric (“anti”) superposition of L and R. The states “symm” or “anti” are also referred to as "mass states" (because $E = mc^2$, same as energy eigenstates). L and R are called "interaction states". The same happens when quarks or neutrinos are created via the weak interaction. Even then, the interaction states (with the names down, strange and bottom for the quarks and ν_e, ν_μ, and ν_τ for the neutrinos, as well as their antiparticles) are different from the mass states. In the case of neutrinos, this can be seen from the fact that the different forms transform into each other (in analogy to the pyramidal inversion).

---

[21] α could also be constant today and could have nevertheless increased or decreased in the early universe. Astronomical observations and laboratory experiments do not address the exact same question in this respect.

[22] They also occur in nuclear reactors and in elementary particle accelerators (and are also studied there). They are produced in large numbers during a supernova and have been observed in particular in coincidence with the supernova SN 1987A (a total of 12 counting events). High-energy neutrinos are mainly produced in the Milky Way and are the subject of "neutrino astronomy".

[23] Text adapted from *en.wikipedia.*org/wiki/Homestake-Experiment

From the neutrino oscillations, it is concluded that the neutrinos have a small mass. (This was different in an outdated version of the standard model.) If the different energy eigenstates have different frequencies (which lead to interference), they must also have different energies according to $E = h\nu$. This results in different masses because of $E = mc^2$. This in turn cannot be the case if all neutrinos have zero

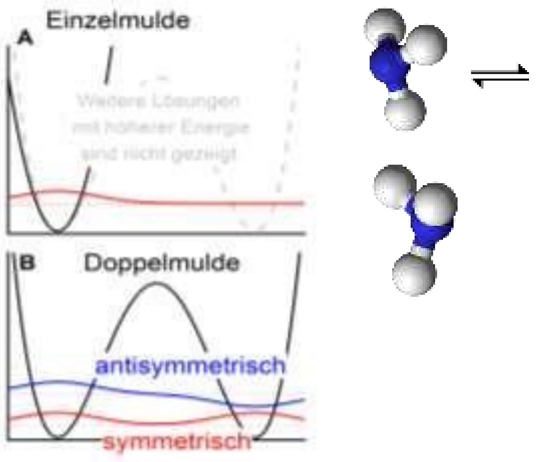


Fig. 2.8
A: The Gaussian function solves the time-independent Schrödinger equation for the parabolic potential (not proven). **B**: For two neighboring potential wells, the solution can be approximated by the sum and the difference of the solutions to the isolated parabolic potentials. The "pyramidal inversion" with calotte models is illustrated in the center.
commons.wikimedia.org/wiki/File:Nitrogeninversion-3D-balls.png

mass. A current attempt from Karlsruhe to determine the neutrino mass is described in the reference from footnote 24. According to this data, the neutrino mass is less than 0.8 eV[25] (as of 2022). The current best upper limit for the mass of neutrinos comes from astrophysics. The universe contains about 337 neutrinos per cm^3 [26]. If the sum of the three neutrino masses were larger than 0.12 eV, then the neutrinos would retard the expansion of the universe to an extent that is incompatible with the experimentally observed expansion.

Note: The neutrinos are part of the "hot dark matter", to be distinguished from the "cold dark matter". The latter takes part in the cluster formation of matter, which has led to the formation of galaxies. Because the neutrinos weigh so little, they fly through these clusters undisturbed. Cold dark matter exists, but neutrinos are not part of it.

2.4 Neutrino mixing could explain the imbalance between matter and antimatter

For the ammonia molecule, the conversion between {L,R} and {symm, anti} can be formulated as a matrix equation. In this case, symmetry enforces a fairly simple form of the matrix. The equation is

Glg. 2.4

$$\begin{pmatrix} \Psi_{symm} \\ \Psi_{anti} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{pmatrix} L \\ R \end{pmatrix}$$

In the case of ammonia, the matrix is not only simple, it is also purely real. What if it were complex? Then "anti" would have a phase relative to "symm" (details are irrelevant). Let the value of this phase be φ. The state "anti" would then lag behind the state "symm" in time (or run ahead if the phase is negative). A factor $\exp(-i\varphi)$ would be added to the factor $\exp(-i\omega t)$ with the different frequencies for "symm" and "anti". For example, the following could apply

---

[24] *en.wikipedia.org/wiki/KATRIN*
[25] Here, the mass was converted into energy using the relation $E = mc^2$.
[26] Photons from the CMB: 411 per cm^3 .

$$\Psi_{symm} = \frac{1}{\sqrt{2}}\left(L+R\right)\exp\left(\mathrm{i}\varphi+\mathrm{i}\omega t\right)$$
$$\Psi_{symm} = \frac{1}{\sqrt{2}}\left(L-R\right)\exp\left(-\mathrm{i}\varphi+\mathrm{i}\omega t\right)$$
Glg. 2.5

This is important, because such a running ahead or lagging behind would violate time-reversal invariance. If time were to run backwards, rushing ahead would become lagging behind and vice versa.

The relevant matrix for the neutrino mixture *is* complex. (A similar mixture with a complex matrix also exists for quarks.) This complex matrix violates the time-reversal invariance at the level of microscopic dynamical laws. This may be regarded as an "arrow of time", but it has nothing to do with the increase in entropy (which also is an arrow of time).

The time-reversal asymmetry is of secondary importance as far as time is concerned. It is of much importance because all gauge theories are "CPT-invariant".[27] "C", "P" and "T" stand for a reversal of charge, parity (handedness) and time. If a process violates T-invariance, it must also violate CP-invariance. "CP" transforms matter into antimatter. The violation of time-reversal invariance creates a difference between matter and antimatter. It can explain why the universe contains more matter than antimatter.

This argument can be made quantitatively. Specifically, the CP violation for quarks is too small to explain the observed imbalance between matter and antimatter. For neutrinos, CP violation is known in principle, but the value of the phase is not yet known precisely enough to relate it to the excess of matter in the universe. However, it is hoped that this will change in about 10 years. Large, planned and funded neutrino experiments will then provide the first data.

2.5 If it can happen it will (?)

Even if the laws of nature are the same everywhere, in the sky offers more to see than the earth. The argument goes back to Fritz Zwicky. He presumably referred to the fact that at some point the neutron stars that had been postulated in theory were actually found (section 7.4). Here is the quote from ref. 1: "In the 1930s the eccentric Swiss astronomer Fritz Zwicky restored the primacy of the heavens to some extent by asserting that 'if it can happen it will'. That is, anything that is permitted by the laws of physics will happen somewhere in the Universe, and with the right instruments and a bit of luck we can see it happening. Zwicky's Principle indicates that it is profitable to think hard about what weird objects and exotic events are in principle."

The following objects and phenomena were postulated before they were found:

- Black holes
- Neutron stars
- Neutrino flashes generated in type II supernovae
- The cosmic microwave background
- Gravitational waves

[27] The "gauge theories" form the basis of the standard model of physics.

The following phenomena were not on anyone's radar before they were found:
- Dark matter and dark energy
- Quasars
- Gamma ray bursts

The following phenomena or objects have not been seen, although the standard model is open to them:
- Magnetic monopoles
- Worm holes, i.e. solutions to Einstein's field equations that would allow a journey into the past.

Astrophysics is special in the following regards:
- In interstellar space, the collision rates between particles are lower than on Earth. The particles have up to 100 years before they collide for the first time.[28] If a molecule is in an excited state and if radiative deexcitation is forbidden or almost forbidden in some way,[29] collisional deexcitation will eventually occur on Earth, with the result that the emission line is not observed. In space, such lines become visible because the slow emission does eventually occur (section 12.1).
- It is much hotter inside stars than on Earth (center of the sun: $15 \cdot 10^6$ K). Nucleosynthesis takes place in the center of the sun, but not of heavy elements. Temperatures of 10^{10} K are reached in the shock fronts of supernovae. Heavy elements are produced there in the "r-process" (Section 11.1) .
- Somewhere in the universe (where and how exactly is not entirely clear) particles with an energy of 10^{20} eV are produced (Box 10.2). This is 10^8 times more than what the Large Hadron Collider (LHC) in Geneva achieves.
- The force of gravity on neutron stars is 10^9 times larger than on Earth. Inside a neutron star, the density is 10^{14} times larger than the density of rock on the Earth's crust.[30] The pressure in the center is larger than 10^{25} atmospheres.
- The magnetic field in the vicinity of a neutron star is between 10^4 and 10^{14} Tesla (section 7.4). For comparison: NMR magnets reach several Tesla.
- There are large-scale magnetic fields in interstellar space. Electric fields of this type do not exist because electric fields are usually screened by countercharges. This does not happen in the same way with magnetic fields because there are no magnetic monopoles.
- There are gravitationally bound, rotating disks (section 8) in several variants. They correspond in some ways to Saturn's rings, but the interactions between the particles are stronger. In the "accretion disks", friction heats up the material. The processes are not always extreme in terms of temperature and pressure, but you have to think your way into things because angular momentum plays a role not familiar to us.

Below, we will often talk about structure formation. Structure formation is a prerequisite for complexity. For this reason, much importance is attached to self-organization in physical chemistry (e.g. when micelles are formed). Structure formation processes in space include
- the Jeans instability (section 4.3)
- the baryonic acoustic oscillations (Box 4.2)

---

[28] A density of 1 particle per cm^3 and a temperature of 100 K were assumed,
www.astro.princeton.edu/~burrows/classes/204/ISM.pdf
The density of the intergalactic medium (a plasma) is usually given as 1 particle / m^3 .

[29] "Forbidden" here means that the electric dipole transition is forbidden. The other transitions (such as the magnetic dipole transition) are less efficient.

[30] The density in neutron stars is two to five times higher than in atomic nuclei on Earth. This makes modeling more difficult. Nuclear matter is basically known from nuclear physics, but the density there is lower.

– the spontaneous generation of magnetic fields (section 7.1)
– the focusing of magnetic fields into flux tubes (section 7.1)
– the streaming instability (section 8.1).

Some things are *not* to be found in outer spaces:

– There is nothing in the sky that is colder than 2.7 K, because this is the temperature of the microwave background.[31] There is no superconductivity of electrons (but see footnote 78).
– There are no high-purity materials anywhere (such as the high-purity silicon used in semiconductor production, defect density $\approx 10^{11}$ cm^{-3}).
– Nowhere is there as much complexity in the sense of complexity theory (keyword: "emergence") as on Earth.

These latter statements do not, of course, concern possible extraterrestrial civilizations.

2.6 It must hang together (?)

In a sub-chapter entitled "It must hang together", Binney marvels at the fact that the existing models not only describe astrophysics, our everyday lives and microscopic physics equally well, but that there are also relatively few such models that these – if formulated appropriately – come across as simple. Yes... Binney's fascination can be shared. Magnetism and electricity describe a large range of phenomena and they are parts of a common theory of electromagnetism, the basis of which can be condensed into four equations (Maxwell's equations). The search for a simple model also guided Einstein when he formulated general relativity. The standard model of physics can also be regarded as a general and – with all due caution – simple theory. In particular, there is a common formalism (the "gauge theories") for quantum electrodynamics, the weak interaction and the strong interaction. Perhaps "compactly formulated" is a more appropriate term than "simple".

However, the “Weltformel”, which Einstein had hoped for and dreamed of, does not yet exist (not even on the horizon). Also, the standard model is incomplete. There must be a theory of quantum gravity, but it is unclear what it should look like. The big picture exists ("It must hang together"), yes, but there are gaps, open ends and peculiarities whose simplicity and beauty are not obvious (an opinion). Furthermore, some constants of nature (perhaps all of them? perhaps even some laws of nature?) in our little corner of the universe have become the way they are purely by chance (section 2.1.3).

3 The three-body problem, randomness and ergodicity

The molecules in gases move randomly (it seems), although the underlying dynamic equations are deterministic. All equations of microscopic physics are deterministic. The "deterministic chaos" was discovered in celestial mechanics, just like Newtonian mechanics. The hero was Poincaré.

3.1 Deterministic chaos in the three-body problem

The solution to the two-body problem (Kepler, Newton) was a transformative success. If you want to solidify a success, you have to move on and apply the principles to other problems. This is how things are today. A natural next step in this case leads from the two-body problem to the three-body problem. This turned out to be a difficult matter, where "difficult" is a misleading term from today's perspective.

[31] This is almost correct: In the Boomerang nebula, a location has been found where, after adiabatic expansion and adiabatic cooling, the temperature according to the CO rotation spectrum (Box 12.1) is only 1 K. There are natural refrigerators in the universe. But nowhere in the universe is it much colder than 2.7 K.

The fact that the problem is so difficult is more significant than the way in which clever people have in some cases tamed the problem.

We encounter the three-body problem in astronomy, for example, in the movement of asteroids. These are attracted by other the bodies of the planetary system, but the sun and Jupiter play the largest role, hence the name "three-body problem". The asteroids are the good-natured cases in that there is a clear hierarchy (Sun >> Jupiter >> asteroid). Certain approximations can be made and these lead to good approximate solutions. Many non-experts today are familiar with the 2nd Lagrange point (L_2). The 2nd Lagrange point lies behind the earth on the axis formed by the sun and the earth. The distance from the Earth is just right for a body located at the 2nd Lagrange point to remain there. The James Webb space telescope lives and works at the 2nd Lagrange point of the Earth-Sun system.

Again, there are good-natured and less good-natured cases. The theory of the three-body problem stumbled along somewhat joylessly. In 1885, the editors of the journal *Acta Mathematica* wanted to tick the matter off and announced an award for its comprehensive and final treatment. It was foreseeable that the award would go to Poincaré, as Poincaré was the leading figure in the field. Poincaré sent something in and received the award. Before his contribution went to print, the editors had a few questions.. There were some gaps and as Poincaré tried to fill those, the hole got deeper and deeper. At this point Poincaré deserves credit. He didn't quietly turn away to more transparent problems but rather kept chewing on this one. Years later, it became clear that the three-body problem *cannot* always be tamed. Sometimes the three-body problem leads into chaos.[32] Chaos is more important for our view of the world than the 2nd Lagrange point. If we know the locations and momenta of three bodies at a certain point in time, we can always calculate the orbit for the next year – or even for the next 40 million years[33,34] – numerically.[35] The question is what happens in the long term. In the case of the three-body problem, this question sometimes cannot be answered.

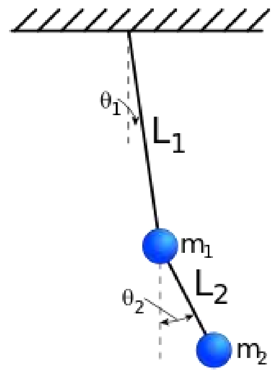


Fig. 3.1
The double pendulum
en.wikipedia.org/wiki/Double_pendulum

The fundamental dynamic laws of physics are deterministic. This applies to $F = ma$, it applies to Maxwell's equations (to electrodynamics), to the Schrödinger equation (to quantum mechanics) and to the standard model of particle physics. Anyone who – firstly – knows the locations and velocities of the sun, Jupiter and Ceres with infinite precision and who – secondly – can calculate with infinite precision can also calculate the orbit of Ceres for all future (and for all past as well). However, neither of the two prerequisites are open to humanoids. A small inaccuracy in our knowledge of the current locations and velocities of the sun, Jupiter and certain asteroids turns

32 The fact that the three-body problem sometimes leads chaos can also be understood from the fact that although there are many double stars, there are only a few triple stars. If three bodies with approximately the same mass orbit each other, sooner or later one of these bodies is ejected from the group, leaving behind a (stable) two-body system. The existing multiple-star systems are usually hierarchical in the sense that a single, particularly heavy star is at the center and dominates events.

33 According to Binney, the current measurement accuracy for the locations and momentum of the planets is sufficient to make a prediction for the next 40 million years. At this point you hit a kind of wall. If it is to be 60 million years, the measurement accuracy must be 10 times better. If it is to be 80 million years, the measurement accuracy must be 100 times better. The relationship between the prediction span and the required accuracy is exponential.

34 The time after which the predictions become qualitatively unreliable is also called Lyapunov time.

35 A side note from Binney: If you try to calculate the movement of our planetary system in the future, you find a danger that Mercury may enter into a resonance with Jupiter. Then its orbit will become elliptical and it will also destabilize the orbits of Venus and Earth. The probability of this happening is reduced if the theory of general relativity (GR) is included in this calculation. The effects of GR are tiny, but tiny things play a role in the three-body problem. Binney suspects a general connection here. GR makes the entrance doors to resonances somewhat narrower. Binney concludes: Without GR, we wouldn't exist.

into a complete ignorance of the orbit of the asteroid in question after a few tens of millions of years. (Not all asteroids behave chaotically in this sense.) The unknown orbits are the essence of chaotic behavior. In the words of Edward Lorenz: "[Chaos is realized when] the present determines the future, but the approximate present does not approximately determine the future." The concept of randomness does not mean that the deterministic dynamic laws would not hold. Randomness means that it is impossible for thinking beings to anticipate the events in question. Beings who create an internal representation of the world around them cannot avoid randomness.

Is there an analog to chaos in the planetary system in molecular physics? In principle, vibrationally excited molecules also exhibit chaotic behavior. If moderately large molecules are excited with a large amplitude, the energy wanders back and forth between the different vibrational modes. This results in mode coupling. The vibrational energy is distributed over many modes, it is "thermalized". However, this does not lead to thermodynamic equilibrium in the narrow sense, although the term "thermalized" suggests that. The movement is very complicated. Interestingly, there is a theory not based on molecular details, which models the influence, that thermalization of the vibrational energy after impact excitation has on the kinetics of chemical reactions in gases (keyword: Lindemann-Hinshelwood mechanism and its extensions). This theory is good enough to have a name. It is the RRKM theory.[36]

Fig. 3.2
The globular cluster 47 Tucunae is home to several million stars. In order to calculate the orbital motion of a single star within the framework of a mean-field theory, a potential would be calculated from the time-averaged positions of all the other stars.
en.wikipedia.org/wiki/ 47_Tucanae

The analogy has limits:

- Molecules in the gas phase are subject to quantum mechanics. Quantum chaos differs from classical chaos.
- Molecules behave less chaotically when the excitation energy is low. The potentials then are approximately parabolic. The restoring force is largely proportional to the excursion from the state of minimum energy. Such linear systems do not go into chaos. In gravity there is no such parabolic low-energy limit. The force of gravitation is never proportional to the distance.
  The transition to chaos at high excitation can be illustrated with the double pendulum (Fig. 3.1). The double pendulum has only two degrees of freedom (the two angles). At small angles, the restoring torques are almost proportional to the angle that $mg\ \Theta \approx mg\ \theta$. ($mg$ is the gravitational force). The double pendulum then oscillates regularly. There is a symmetric and an antisymmetric mode (same or opposite sign of both angles). Chaos only occurs at larger amplitudes.
- An analogy (with limits) to the "resonances" in planetary motion are the "Fermi resonances" in molecular physics. Sometimes a vibrational frequency shifts when the vibration couples to overtones or combination tones of the same molecule that have almost the same frequency.

[36] The letters stand for names. "M" stands for Marcus, who also devised the "Marcus theory". The Marcus theory is of much importance in electrochemistry. It describes the transfer of electrons and makes simplifications in an ingenious way, which at first glance appear to be presumptuous, but which in practice are valid and enable understanding.

Remember:
- For chaotic systems in the sense of chaos theory, the trajectories critically depend on the initial conditions and are therefore largely unpredictable in the long-term limit.

3.2 Side remarks: Transfer of momentum between Venus and Earth as a consequence of a correlation

In quantum mechanics (and elsewhere), one sometimes tries one's luck with a "mean-field theory". Fig. 3.2 outlines the procedure for globular clusters. Globular clusters contain several 100 000 stars, all of which interact gravitationally with each other. Since 100 000 coupled differential equations are difficult to handle, an averaged potential is calculated for each individual star, which results from the *time-averaged* positions of all other stars. In a first step, you simply assume some potential. You then calculate all the orbits one after the other, average them over time and obtain a new, better potential. This is repeated until the iterations no longer cause any changes. Self-consistency is then achieved.[37] However, the star in question exerts *instantaneous* forces on its neighboring stars, which result in a correlated motion. If the consequences of such correlated motions do not average out over time, they lead to an error in the mean-field calculation.

The procedure is similar in quantum mechanics: You choose an electron in the potential of the nucleus. At the beginning, you do not know the orbitals of the other electrons, so you cannot take them into account when calculating the average potential. An initial calculation of all orbitals is carried out, in which all interactions between the electrons are neglected, and a "0^{th}-order" approximation is obtained. A first-order potential is then calculated from these 0^{th}-order orbitals and the orbitals are recalculated using this potential. The procedure is repeated until the subsequent run no longer produces a change (until the result has "converged"). This so-called Hartree-Fock method does not take instantaneous correlations into account. (The Hartree-Fock method does take care of the antisymmetry of many-electron wave functions under permutation. It is special in this regard.)

We further explore correlations and their consequences for time-averaged dynamics using the orbits of Earth and Venus. Venus contributes to the gravitational potential in which the earth moves. One might distribute the mass of Venus evenly over its orbit and calculate an averaged, time-constant potential. This would correspond to the Hartree-Fock method. However, this approach falls short in that it does not take into account the transfer of angular momentum from Venus to Earth. When Venus passes Earth, it first slows down the movement of the Earth as long as it is behind the Earth. Later, when it has passed, Venus accelerates the Earth. These two effects do not cancel because the acceleration lasts slightly longer than the deceleration. After passing, Venus has transferred a small part of its own angular momentum to the Earth. This is a classic correlation effect that is only noticed if the *instantaneous* (rather than the time-averaged) positions of the objects are considered.

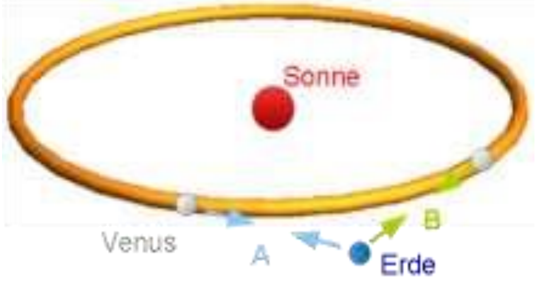


Fig. 3.3
An example of the effect of correlations from celestial mechanics. In order to capture the influence of Venus on the movement of the Earth, its mass could be distributed over its orbit. This representation does not take into account the fact that Venus accelerates the Earth when it passes the Earth. The net acceleration of the Earth occurs because the period of time during which Venus slows down the Earth (light blue) is slightly shorter than the period of time during which Venus accelerates the Earth (green).

[37] Instead of "mean field theory", we sometimes also say "self-consistent field theory".

Two cases must be distinguished. In globular clusters there are "encounters" between stars (such as between Earth and Venus), which change the trajectories of the stars. However, these are uncorrelated with each other. A "relaxation time" can be defined within which the stars have forgotten their original orbits. Statistical mechanics remains applicable. In the planetary system, however, Venus and Earth meet again and again. These "encounters" are correlated and statistical mechanics in the narrow sense does not apply.

3.3 Side remarks: Spin-orbit coupling

In molecular physics, one encounters the fact that only the total angular momentum of the electron is conserved. The orbital angular momentum and the spin are not conserved individually. If, for example, the spin has the value +1/2 at the beginning of an experiment, angular momentum is exchanged back and forth between spin and orbit over the course of time. The energy eigenstate has neither a defined orbital angular momentum nor a defined spin, but only a defined total angular momentum. This is an example where the "interaction states" and the energy eigenstates are different (see also section 2.3). An interaction state typically has a well-defined spin, because the process of formation dictates that.

The process of spin-orbit coupling can be illustrated using an example from astronomy. The moon has undergone such a coupling. It used to rotate around its own axis independently of its orbit around the Earth. Today, however, it performs a "bound rotation". Today, we always see the same side of the moon. In the case of the moon, the spin-orbit coupling was an irreversible process. It was brought about by the tidal forces.[38] Although the moon has no oceans, the rotation led to a small deformation, analogous to ebb and flow on Earth. On Earth, too, ebb and flow lead to spin-orbit coupling. Over time, these effects lengthen the days and shorten the years. On the time scale of a few years, other, stronger effects are superimposed on this effect (e.g. redistribution of mass between the Earth's core and mantle, which changes the moment of inertia). 400 million years ago, the day lasted only 21.5 hours.[39]

There are no tidal forces for electrons. This role is taken by magnetism. Place the center of the coordinate system at the location of the electron. (This is an accelerated coordinate system.) From the electron's point of view, the nucleus represents a circular current that generates a magnetic field at the electron's location. Whether the spin is parallel or antiparallel to this field makes a small difference to the energy. This is referred to as the "fine structure". Historically, this small difference gave rise to the definition of the "fine structure constant" (section 16.1).

3.4 Uncorrelated disturbances lead to statistical mechanics

The path from random behavior to statistical mechanics consists of two steps. In the first step, a distinction is made between the motion of a test body as calculated by mean-field theory and the disturbances of this motion. The undisturbed motion would result from the time-averaged gravitational potential, which is caused by all other bodies. The disturbances result from so-called "encounters".[40] In a second step, the correlations of the disturbances are examined. If the disturbances are uncorrelated with each other and if they are uncorrelated with the movement of the test body, they can be treated well analytically. The law of large numbers applies. The effects of the many disturbances can then be understood as

---

[38] Tidal forces deform a body that is on a circular path around another body because centrifugal force and centripetal force are only equal in the center of the body. The centrifugal force prevails on the outside, while the centripetal force prevails on the inside. The tidal forces got their name from the ebb and flow of the tide, but they also exist in many other places.

[39] The duration of a day can be deduced indirectly from the fact that the year had 405 days, as evidenced by the growth rings on trees, for example. Because the total angular momentum (Earth's spin and orbit around the sun) is constant, a longer year implies shorter days. The length of a day cannot be deduced directly from the fossils.

[40] "Encounters" are caused by gravitational interactions. A direct collision between stars is extremely unlikely.

sums of random numbers. These sums are themselves statistically distributed, but these distributions of the sums are sharp. Their maximum is easy to predict.

More specifically, the dynamics is then described by the Langevin equation or the Fokker-Planck equation. For the form of these equations, see Wikipedia. The Langevin equation describes many individual particles that are subject to uncorrelated disturbances. The Fokker-Planck equation deals with distributions instead of the individual particles, as if these were a continuum. A large number of uncorrelated disturbances are found more frequently in many-particle systems than in systems with only a few bodies. Both equations lead to "relaxations" in the course of which a previous state is forgotten.

The Langevin equation and the Fokker-Planck equation lead to a classical variant of statistical mechanics. In physical chemistry, quantum mechanics is usually taken for granted (meaning that there are discrete micro-states) and a partition sum (a Boltzmann-weighted number of microstates, the latter countable) is calculated. Furthermore, thermodynamic equilibrium is often assumed. The latter does not exist globally for self-gravitating systems. These differences acknowledged, globular clusters can be described by means of statistical mechanics (i.e. modelled with the Langevin equation or the Fokker-Planck equation).

The treatment of correlations can be rather difficult. This is why globular clusters are often modeled with "N-body simulations". Such simulations make no assumptions about the correlations.

Sometimes correlations can be taken into account easily, if they extend over short times or distances. An example is the orientation of the segments of polymer chains. Neighboring segments often have correlated orientations, but this can be taken into account by defining new, longer, "effective" segments in such a way that neighboring effective segments are uncorrelated with respect to their orientation. This procedure eliminates the problems arising from local chain stiffness, but it does not eliminate the consequences of self-avoidance. Two segments that are not adjacent on the chain, still are never allowed to occupy the same space. This self-avoidance implies a "long-range correlation". The consequences of self-avoidance are most efficiently investigated using simulations. These simulations lead to the Flory exponent (see Wikipedia).

There also is a short-range correlation in the movement of particles and liquids. This movement is not an ideal random walk, because collisions take place with particles that themselves collided with the test particle shortly before and have not forgotten this collision. This leads to the "long-time tails" in the velocity autocorrelation functions. These exist, but they are not very important.

<u>Remember:</u>

- The difference between chaos in the narrow sense (three-body problem, double pendulum, weather) and the subject of statistical mechanics (ergodic systems) is largely in the correlations.

4 The universe moves away from heat death

This section examines the special features of thermodynamics in astrophysics from various perspectives. The title is intended to sound optimistic. Clausius assumed that the sad final state of this universe would be a uniformly warm soup in thermodynamic equilibrium.[41] (It is said that this expectation contributed to the depressions, which plagued Boltzmann at the end of this life.) However, the universe *was* in such a state 380,000 years after the Big Bang. It consisted of a largely homogeneous plasma with a temperature of around 3000 K. Nevertheless, today there are complex entities that discuss the heat death. Clausius' argument is questioned below.

4.1 Cosmic expansion and black holes lead to global cooling

An important feature of the sky is its expansion. Because new space is constantly being created, entropy can constantly be dispersed into this new space. That structures form while the universe cools is not surprising (see the list at the end of section 2.5).

Less well known is the role of black holes. For a deeper understanding of black holes, we would have dive into general relativity.[42] We don't do that. The equations of general relativity have a so-called "Schwarzschild solution", which leads to objects from which there is no escape. For the outside observer, the black hole has a mass, an angular momentum, and possibly a charge. This list is complete.[43] For the outside observer, the black hole has no other properties, which would be independent from these.

One might think that if the black hole has so few properties, it would also have a low entropy. However, Bekenstein and Hawking were able to show that the black hole contains the numerous microstates that it has swallowed in the course of its formation.

In the case of gas, the "macro-state" (specified by e.g. pressure and temperature) implies a decision made by the experimenter (a "relevance concept"). There is a certain (limited) freedom in the choice of the relevance concept. You can – if you want to – be interested in the locations and velocities of the particles. This happens from time to time. This decision is not open to us in the case of the black holes. In quantum mechanics, it is emphasized that entropy quantifies "remediable ignorance". Entropy counts open questions that may be asked. (The question of whether Schrödinger's cat is alive or not is not among these permitted questions). This definition of entropy must be modified for the black hole. For the black hole, an entropy quantifies open questions that can only be pursued by those who embark on the journey into the interior of the black hole.

Bekenstein and Hawking showed that the entropy of the black hole is given as

Glg. 4.1

$$S_{BH} = \frac{k_B A}{4 l_P^4}$$

---

[41] As things stand today, the universe will not end in heat death, but instead in the "Big Freeze". For details, refer to Wikipedia. After 10^(10^{26}) to 10^(10^{76}) years (there is still a small uncertainty according to Wikipedia), iron stars become black holes by means of a tunnel effect. After that, nothing more happens.

[42] Do not confuse the black hole with the black body.

[43] This is the content of the "no-hair theorem". In this context, it was also proven that black holes are not a computational artifact that could be avoided if one were to dispense with perfect spherical symmetry in the Schwarzschild solution.

The index *BH* stands for black hole (or Bekenstein-Hawking). *A* is the surface area of the black hole[44] and $l_P = (G\hbar/c)^{31/2} = 1.6 \cdot 10^{-35}$ m is the Planck length (Box 4.1). *G* is the gravitational constant. The entropy of a black hole is extraordinarily large. Black holes swallow up large amounts of entropy, similar to the dark sky at night.

Now you can take the derivative of entropy with regard to energy and arrive at an (inverse) temperature. A lengthy calculation results in:

Glg. 4.2

$$T_H = \frac{\hbar c^3}{8\pi G m k_B}$$

The subscript *H* stands for Hawking.

In principle, this temperature can be measured based on a special form of thermal radiation. Due to non-local quantum correlations across the horizon of the black hole, particle-antiparticle pairs sometimes form at the edge of the black hole, where one of the two escapes. The escaped particles have a thermal energy distribution. It can be used to measure the temperature of the black hole.[45] Based on this type of "Hawking" radiation, holes evaporate over time. The evaporation time is

Glg. 4.3

$$t_{ev} = \frac{5120\pi G^2 M^3}{\hbar c^4} = 2.1 \cdot 10^{67} \text{ years} \left(\frac{M}{M_\odot} \right)^3$$

$M_\odot$ is the mass of the sun. Black holes are not necessarily an ultimate final state (or are they? See footnote 41).

A side note: Black holes have a negative heat capacity similar to self-gravitating gas clouds. If they swallow matter (and if their mass increases in the process), they become colder although energy is supplied to them. If they emit energy via Hawking radiation, they become lighter and hotter.

Black holes are subject to thermodynamics, but so far there is no simple way to assign an entropy to the curvature of space. If – as Einstein claimed – thermodynamics is universal, it should also be applicable to the curvature of space. There are a number of proposals and conjectures for "gravitational entropy", but there is no consensus.

<u>Remember</u>

- In principle, structure formation in the universe is not surprising, because the universe cools down. Cooling also leads to structure formation on Earth.
- A first reason for the cooling is the expansion of the universe.
- A second reason for the cooling are the black holes. These have an entropy and a temperature. The entropy is very high, the temperature is very low. As they swallow microstates, they cool their environment.

---

[44] More precisely: the "event horizon". Entropy is therefore not proportional to the volume of the black hole, although it is not entirely clear how this volume would be defined, because spacetime has a singularity at the center.

[45] Note the low temperature. A black hole with the mass of the sun has a Hawking temperature of 60 nanoKelvin.

Box 4.1: The Planck scale

The Planck scale of 10^{-35} m, on which the effects of quantum gravity become visible, arises if one postulates an elementary particle that forms a black hole on its own. According to everything we know, electrons are point particles. Why are they not black holes? A black hole is an object whose *entire* mass is located within the Schwarzschild radius associated with its mass (also: the "event horizon"), the latter given as

$$r_S = \frac{2Gm}{c^2}$$

G is the gravitational constant. To make the Schwarzschild radius plausible, multiply the equation above by mc^2 and divide by r_S :

$$mc^2 = \frac{2Gm^2}{r_S}$$

For black holes, the energy in gravity is larger than the energy in the rest mass.

In the context of quantum mechanics, even a point particle is not a black hole as long as the spatial uncertainty Δx is greater than the Schwarzschild radius. If the mass increases, the Schwarzschild radius becomes larger and the spatial uncertainty smaller (the latter due to $\Delta x \approx \hbar/(2\Delta p)$ with Δp proportional to the mass). Then, when Δx and r_S are equal, the Planck mass is reached.

The spatial uncertainty is written as $\Delta x = \hbar/(2\Delta p)$. We assume the relativistic energy-momentum relation ($E^2 = m c^{24} + p c^{22}$) and neglect the rest mass. For relativistic particles, $E \approx cp$ and $\Delta p \approx \Delta E/c$ apply. Now we assume $\Delta E \approx E$ and replace the mass m in the equation for the Schwarzschild radius with E/c^2 :

$$r_S = \frac{2GE}{c^4} \approx \frac{2G\hbar c}{c^4 \Delta x}$$

The factor 2 has disappeared behind the ≈. If one requires that $l_P = r_S = \Delta x$, it follows that

$$l_P = \sqrt{\frac{G\hbar}{c^3}}$$

Inserting the numerical values leads to $l_P \approx 10^{-35}$ m. Such a particle would have a mass of around 20 μg (the Planck mass), corresponding to 10^{19} proton masses.

– Due to quantum non-locality, black holes can decay – albeit very slowly.

4.2 The microwave background: An unstable equilibrium

The apparent (unstable) state of equilibrium at the beginning of the universe is accessible via the cosmic microwave background (CMB). The CMB is the thermal spectrum of a plasma. Charged particles and electromagnetic waves have a strong interaction. The temperature of the photons (the radiation temperature, to be determined from the spectrum of the black body) and the temperature of the particles (to be determined from the Maxwell distribution for the velocities) were the same. When the ions and electrons recombined to form atoms, the universe became transparent.[46] The photons of the CMB that we detect today were generated 13.7 billion years ago by the plasma that turned into a neutral gas at that time and have since traveled through space largely undisturbed. We measure the radiation temperature today as T = 2.725 K because the plasma is moving away from us with a redshift of $\Delta\lambda/\lambda = z \approx 1089$. Otherwise, the radiation temperature would be around 3000 K.

The CMB is among the phenomena in the sky that were predicted before they was discovered. The discoverers of the CMB, Penzias and Wilson, were unaware of this prediction. They reported a noise and are said to have only superficially understood the deeper significance of their discovery, even after win-

[46] Later there was an epoch of "reionization", but this did not make the universe opaque again due to the then lower density of the gas.

ning the Nobel Prize. Fig. 4.1 shows the spectrum of the CMB together with a fit to the Planck distribution. No deviations between fit and data can be seen in this plot. (The data were previously cleaned of some artifacts).

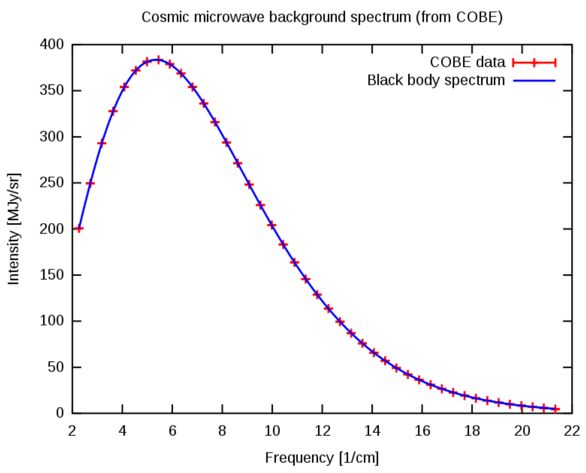


Fig. 4.1
Spectrum of cosmic microwave background measured by the FIRAS instrument on the COBE satellite. This is the most precisely measured black body spectrum in nature. The error bars are too small to be seen even in an enlarged image.
en.wikipedia.org/wiki/Cosmic_microwave_background

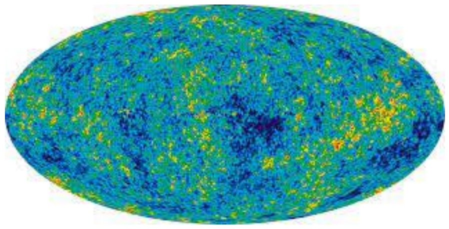


Fig. 4.2
Acquired by ESA's Planck space telescope, the most detailed map ever created of the cosmic microwave background - the relic radiation from the Big Bang - was released today revealing the existence of features that challenge the foundations of our current understanding of the Universe. The image is based on the initial 15.5 months of data from Planck and is the mission's first all-sky picture of the oldest light in our Universe, imprinted on the sky when it was just 380 000 years old.
www.esa.int/Science_Exploration/Space_Science/Planck/Planck_reveals_an_almost_perfect_Universe

Until the early 1990s, the CMB was considered isotropic. It took a lot of effort to measure the small angular variation, but this very small anisotropy today is a cornerstone of cosmology. The radiation temperature fluctuates around ± 18 µK, the latter being the standard deviation. Fig. 4.2 shows the famous map of temperature deviations from the mean value.[47] If you stare at this map for a while, you see that the yellow islands are not all the same size, but that the sizes are not completely randomly distributed, either. This impression is quantified by plotting the spectral power density against the multipole order (Fig.

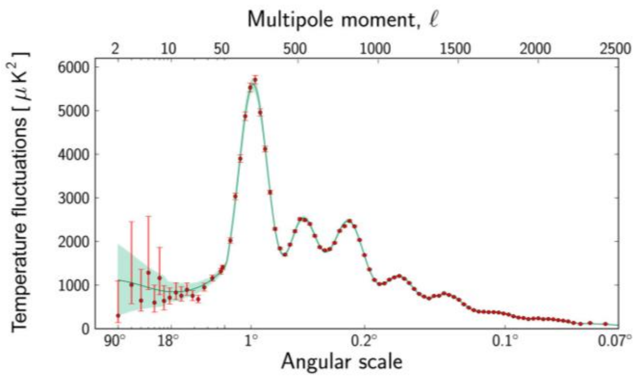


Fig. 4.3
This graph shows the temperature fluctuations in the Cosmic Microwave Background detected by Planck at different angular scales on the sky. This curve is known as the power spectrum. The largest angular scales, starting at angles of ninety degrees, are shown on the left side of the graph, whereas smaller and smaller scales are shown towards the right. The red dots correspond to measurements made with Planck; The green curve shown in the graph represents the best fit of the 'standard model of cosmology' - currently the most widely accepted scenario for the origin and evolution of the Universe - to the Planck data.
https://www.mpg.de/7044245/Planck_cmb_universe

[47] This map has been corrected for a "dipole component". This dipole component results from the fact that the Earth is moving relative to the CMB. The Doppler shift makes the CMB appear somewhat hotter in the direction of motion than in the opposite direction. If the dipole component is not subtracted, the standard deviation of the temperature is 57 µK.

4.3). A complication: Because this temperature distribution lives on a sphere and not on a plane, the temperature distribution is not Fourier decomposed (i.e. not represented as a sum of sine and cosine functions), but as a weighted sum of spherical harmonics $Y_{l,m}$ (θ,φ) (with *l* multipole order). The squares of the weights (the power spectral density) are shown. On the right in Fig. 4.3 are the small-scale structures (the high multipole components, the analog of the high Fourier components). On the left are the large-scale (more precisely: large-angle) structures.

The power spectrum of the CMB shows characteristic peaks corresponding to the preferred sizes of the red islands in Fig. 4.2. These are caused by the baryonic acoustic oscillations (BAOs) (Box 4.2).

Remember

- The CMB is the radiation from a plasma that was close to thermodynamic equilibrium. The radiation temperature was about 3000 K.
- There is a small dependence of the temperature on angle.

Box 4.2
Baryonic acoustic oscillations (BAOs) constitute the large-scale structure of the universe.
On the largest spatial scale, the universe consists of "voids" and "filaments" (Fig. 4.4). More galaxies and galaxy clusters can be found in the shells and filaments than in the largely spherical spaces between them.
In the early phase of the universe, the spherical interstices were areas of *increased* density. According to the virial theorem, these inhomogeneities were self-reinforcing (Sections 4.4.3 and 4.4.4). Then, however, a difference occurred between the ordinary ("baryonic") matter and the dark matter. In the course of the contraction, the temperature increased and the radiation pressure of the photons (which were emitted and absorbed by the ions) increased. However, this radiation pressure does not affect the dark matter. The photons produced an outward movement of the baryonic matter (only this matter). The interplay between gravity and radiation pressure led to large-scale waves that can be interpreted as sound waves. When the universe became transparent 380 000 years after the Big Bang, the movement of the baryonic matter was decoupled from the photon density. The baryonic "shells" moved further outwards and today form the "cosmic foam".

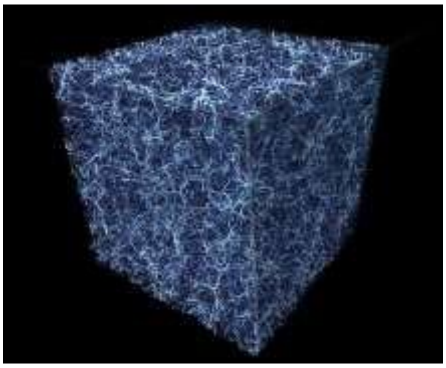


Fig. 4.4
On a large scale, the universe consists of "voids" and "filaments".
en.wikipedia.org/wiki/Filament_(cosmos)

4.3 Polarization of the microwave background, primordial gravitational waves

The fluctuations in the CMB originate from the certain vacuum fluctuations in the early universe. There are scalar fields (like the Higgs field), there are vector fields (like light) and there are gravitational waves (vacuum-gravitational fluctuations in the early universe, also called: "primordial gravitational waves"). The latter stand out because they may violate parity symmetry (the symmetry under a point reflection at the coordinate origin). As a consequence, a map of the polarization of the CMB may contain vortices. Vortices have a handedness (clockwise or counterclockwise).

How does a polarization of the CMB come about? It may come about by scattering of the light by the electrons and ions ("Thompson scattering"). If scattering occurs at 90°, the scattered light is polarized perpendicularly to the direction of propagation of the primary beam. This is the same with Rayleigh scattering, which lets the sky appear blue. The blue sky is polarized. If the intensity of the CMB fluctuates slightly, the CMB from areas next to the brighter areas should be slightly polarized. According to this explanation, the source of polarization is a scalar field (a fluctuation of density, pressure or temperature). Such scalar fields can never generate vortex fields (of whatever) for reasons of symmetry.

However, gravitational waves are described by a 2nd-rank tensor (by the metric tensor $g_{\mu\nu}$). In mathematics, a distinction is made between vectors (such as the electric field or the velocity) and "axial vectors" (such as the magnetic field or the vortex density of a velocity field). Axial vectors have three components and can be represented as arrows. However, unlike normal vectors, they *do not* change direction when space is reflected at the origin. The axial vectors result from a cross product. Write the magnetic field as the rotation of the vector potential ($\vec{B} = \vec{\nabla}\times\vec{A}$) or write the vortex density (the "vorticity",ξ) as $\vec{\xi} = \vec{\nabla}\times\vec{v}$. The x-component of $\vec{B}$ is $B_x = \partial_y A_z - \partial_z A_y$.[48] Under reflection at the coordinate origin, both $\vec{\nabla}$ and $\vec{A}$ change sign. $\vec{B}$ then does not. The vorticity does not either.

To avoid this strange behavior, consider the 2nd-rank tensor $\vec{\nabla}\otimes\vec{A}$ (with two indices, given as $(\vec{\nabla}\otimes\vec{A})_{ij} = \partial_i A_j$). This tensor has 9 entries. Such tensors can be split into a symmetric and an antisymmetric part:

$$(\vec{\nabla}\otimes\vec{A})_{ij} = (\vec{\nabla}\otimes\vec{A})_{sym,ij} + (\vec{\nabla}\otimes\vec{A})_{anti,ij} = \frac{(\vec{\nabla}\otimes\vec{A})_{ij} + (\vec{\nabla}\otimes\vec{A})_{ji}}{2} + \frac{(\vec{\nabla}\otimes\vec{A})_{ij} - (\vec{\nabla}\otimes\vec{A})_{ji}}{2}$$

The antisymmetric part contains zeros on the diagonals and 6 elements off the diagonals, but these are antisymmetric when indices are swapped ($(\vec{\nabla}\otimes\vec{A})_{anti,ij} = -(\vec{\nabla}\otimes\vec{A})_{anti,ji}$) . Three independent parameters remain, which can be written as a vector (i.e. as the cross product) if desired. The following relations apply: $B_x = A_{anti,yz}$, $B_y = A_{anti,zx}$ and $B_z = A_{anti,xy}$.

Because $g_{\mu\nu}$ is a tensor, it can have an antisymmetric part, and this part can create vortices on the CMB polarization map. Do these vortices exist? That is the question. These vortices also go by the name of "B-modes", to be distinguished from "E-modes" (Fig. 4.5). So far it has not been possible to nail down the -B-modes, but many people try.

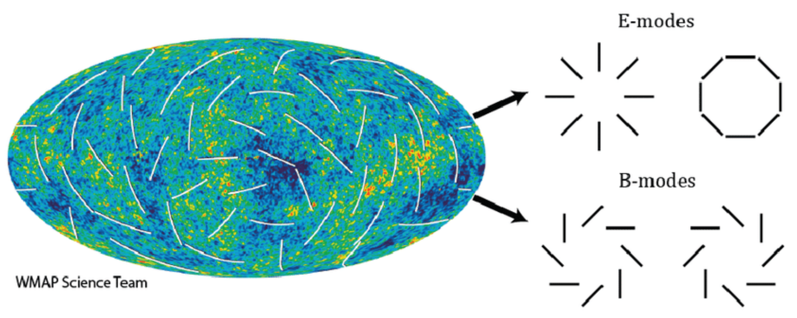


Fig. 4.5
The CMB anisotropy polarization map may be decomposed into curl-free even-parity E-modes and divergence- free odd-parity B-modes. Primordial B-modes are only created by tensor perturbations (inflationary gravitational waves).
www.researchgate.net/figure/The-CMB-anisotropy-polarization-map-may-be-decomposed-into-curl-free-even-parity-E-modes_fig1_263811882

4.4 Peculiarities in the thermodynamics of self-gravitating systems

4.4.1 In gravitationally bound systems, the internal energy is an extensive

In the case of self-gravitating systems, a distinction must be made between
- globular clusters
- clouds of plasma or gas.

In globular clusters, the gravitational interaction is the only interaction between the numerous stars. All problems discussed below become effective. There is no temperature and no thermodynamic equilibrium. With clouds of plasma or gas, the situation is more difficult because the molecules collide. If the energy in the interactions (the gravitational energy) is not counted as internal energy, then a temperature

[48] ∂_x is short for ∂/∂_x .

can be derived from the kinetic energy. In the case of the plasma, this temperature also is the radiation temperature of the photon field.[49]

The long-range nature of the gravitational interaction makes familiar ideas from thermodynamics inapplicable to globular clusters.[50] For the $1/r$-potential, the internal energy U is no longer an extensive quantity. Extensive quantities are additive under the combination of two systems. Extensive quantities in this sense are the amount of substance, the volume, and – often, but not always – the internal energy. (Intensive quantities are, for example, temperature or pressure.) The internal energy U is extensive if it is caused by the interactions of the molecules (more generally: the bodies) with only their neighbors. Then U is $\approx N\varepsilon z/2$ with N the number of molecules, ε the pair interaction energy and z the number of neighbors ($4 < z < 14$). The factor 1/2 arises because ε involves two partners. It is "≈" instead of "=" because a molecule on the surface has fewer neighbors than a molecule in the bulk and furthermore because this formula ignores three-body interactions. The compliations acknowledged: If $U \approx N\varepsilon z/2$ applies, the internal energy is proportional to the amount of substance and is therefore an extensive quantity.

If the interaction is proportional to r^{-n} with $n \leq 3$, U is no longer extensive. The energy of *each body individually* then depends on the size of the system. Fig. 4.6 illustrates the matter based on forces rather

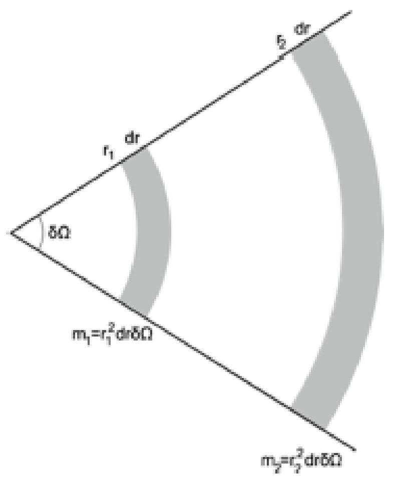


Fig. 4.6
Consider the star on the left-hand side of the diagram. Let the mean density of the other stars ρ be approximately constant. Then the stars on the shell on the far right exert a force of approximately the same magnitude on the star as the stars on the smaller shell further to the left. The force exerted by each individual star scales like $1/r^2$ but the number of stars on the shells scales like r^2 . There is no net distance dependence (except for a possible distance dependence resulting from a distance dependence of □).
https://doi.org/10.1007/s11229-018-02032-5.

than energy. For the math (below), we use energies. For the energy of the body we write

Glg. 4.4

$$\varepsilon_{\text{Körper}} = \int_{r\_\min}^{r\_\max} \frac{Gm}{r} \rho(r) 4\pi r^2 \mathrm{d}r$$

G is the gravitational constant, and $\rho(r)$ is the number density. r_{min} is the minimum distance between two bodies, r_{max} is approximately the system size. If Glg. 4.4 contained a term $1/r^n$ with $n > 3$ (instead of the $1/r$), the upper limit in this integral could be replaced with ∞ , because this integral would converge. However, this is not permitted for $n \leq 3$. With $V \propto 1/r$, the following applies to the internal energy

[49] Temperatures which only pertain to parts of a system often exist. These makes sense if the subsystems in question are not in equilibrium with the other parts of the system (or if this is questionable). This is sometimes referred to as "spin temperatures" or "rotation temperatures" (section 12.2).
[50] Mukamel, D., Notes on the Statistical Mechanics of Systems with Long-Range Interactions. *arxiv.org/pdf/0905.1457.pdf* **2009**.

Glg. 4.5

$$U \approx \sum_{\text{Körper}} \varepsilon_{\text{Körper}} \approx N \int_{r\_\min}^{r\_\max} Gm\rho 4\pi r \mathrm{d}r = NGm\rho 4\pi \left[r_{\max}^2 - r_{\min}^2 \right] \approx NGm\rho 4\pi r_{\max}^2$$

For simplicity, a homogeneous number density of the bodies, ρ, was assumed and geometrical factors were neglected. If eight systems are combined to one system with the same number density, r_{max} doubles (approximately). U is no longer proportional to N. In consequence, U according to Glg. 4.5 is no longer extensive. The temperature, defined by $1/T = \mathrm{d}S/\mathrm{d}U$, is no longer an intensive quantity. Usually, S and U are both extensive. The ratio is then intensive. However, this is not the case for self-gravitating systems.

The importance of short-range interactions is also emphasized in the discussion of the of surface en-

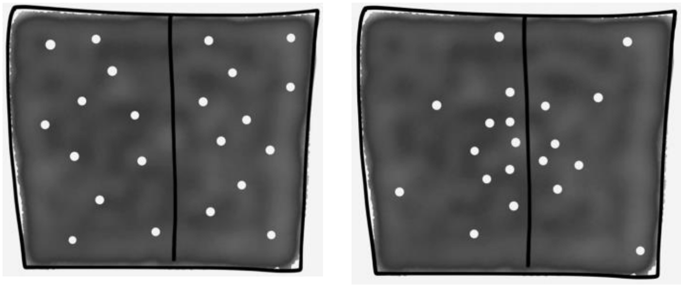


Fig. 4.7
If two systems with long-range interactions are combined, the particles are distributed as shown in the figure on the right (short-range: left). Then the internal energy and the entropy are not additive.
https://link.springer.com/article/10.1007/s11229-018-02032-5

ergy. The simplest model for the surface energy is the "model of missing neighbors". If only the interactions between neighbors are to be significant, the interaction must be short-ranged. In this context, however, short-ranged does not imply $V \propto r^{-n}$ with $n > 3$, but $V \propto r^{-n}$ with $n > 4.5$. The van der Waals interaction (proportional to r^{-6}) is therefore short-range in this sense. Exponentially decreasing interactions are also short-ranged. The H-bridges and the hydrophobic interaction are short-range. Systems with "weakly long-range interactions" ($3 < n < 4.5$) are a difficult topic.

A side note: The entropy is also not additive when systems with $1/r$ potentials are joined, because the particles are distributed differently in equilibrium after joining (Fig. 4.10).

Are there other long-range interactions other than gravity? Not many. Diverging integrals (as in Glg. 4.5 for the case $r_{\max} \to \infty$) always entail complications. If they were frequent, physical chemistry would be more tedious than it actually is.

<u>Remember:</u>

– For long-range interactions, the internal energy is not an extensive quantity. A temperature cannot be defined. Conventional thermodynamics cannot be applied.

4.4.2 The viral theorem

There are two variants of the virial theorem. One variant (discussed in the following) links the kinetic energy to the forces acting in a system (more precisely: to the virial, Glg. 4.6). A thermodynamic equilibrium or a temperature are not part of this argument. The virial theorem law has far-reaching consequences for $1/r$ potentials. A second variant of the virial theorem connects the virial with the temperature. This variant is described in section 16.2. It leads to the equipartition theorem and to the pressure of the real gas (keyword: virial expansion).

The non-thermodynamic version of the virial theorem states that

Glg. 4.6

$$\langle E_{kin} \rangle = -\frac{1}{2}\left\langle \sum_{ik} \vec{r}_k \cdot \vec{F}_k \right\rangle$$

The term on the right is the virial.

The following text is based on https://de.wikipedia.org/wiki/Virialsatz. The argument starts from

Glg. 4.7

$$2E_{kin} = 2\sum_i \frac{1}{2} m_i \vec{v}_i \cdot \vec{v}_i = \sum_i \vec{p}_i \cdot \vec{v}_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \vec{p}_i \cdot \vec{r}_i\right) - \sum_i \vec{r}_i \cdot \frac{\mathrm{d}}{\mathrm{d}t}\vec{p}_i = \frac{\mathrm{d}G}{\mathrm{d}t} - \sum_i \sum_j \vec{r}_i \cdot \vec{F}_{ij}$$

The last term on the right-hand side can be transformed as follows

Glg. 4.8

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \vec{r}_i \cdot \vec{F}_{ij} = \sum_{j=2}^{N}\sum_{i=1}^{j-1}\left(\vec{r}_i \cdot \vec{F}_{ij} + \vec{r}_j \cdot \vec{F}_{ji}\right) =$$
$$\sum_{j=2}^{N}\sum_{i=1}^{j-1}\left(\vec{r}_i \cdot \vec{F}_{ij} + \vec{r}_j \cdot \vec{F}_{ij}\right) = \sum_{j=2}^{N}\sum_{i=1}^{j-1}\left(\vec{r}_i - \vec{r}_j\right) \cdot \vec{F}_{ij} = \sum \vec{r}_{ij} \vec{F}_{ij}$$

$\vec{F}_{ij}$ is the force which particle *j* exerts on particle *i*. It was used that $\vec{F}_{ij} = -\vec{F}_{ji}$. It follows that

Glg. 4.9

$$2E_{kin} = \frac{\mathrm{d}G}{\mathrm{d}t} - \sum_{ij} \vec{r}_{ij} \cdot \vec{F}_{ij}$$

G is the sum of the scalar products of the momenta $\vec{p}_i$ and the locations $\vec{r}_i$ of all particles:

Glg. 4.10

$$G = \sum_{i=1}^{N} \vec{p}_i \cdot \vec{r}_i$$

The average of d*G/*d*t* over long time periods is zero*:*

Glg. 4.11

$$\left\langle \frac{\mathrm{d}G}{\mathrm{d}t} \right\rangle_{Zeit} = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^{\tau} \left(\frac{\mathrm{d}G}{\mathrm{d}t}\right)\mathrm{d}t = \lim_{\tau\to\infty}\frac{1}{\tau}\left(G(t) - G(0)\right) = 0$$

It was assumed that the velocities and locations of the particles are bounded (as is the case with periodic orbits, for example). Let the forces be equal to the gradients of a potential: $\vec{F}_{ij} = \mathrm{d}\vec{p}_i/\mathrm{d}t = -\mathrm{d}V/\mathrm{d}\vec{r}_{ij}$. If the potential *V* is proportional to $|\vec{r}_{ij}|^{-1}$, it follows that

Glg. 4.12

$$2\langle E_{kin} \rangle = \left\langle \sum_i \vec{r}_{ij} \cdot \frac{\mathrm{d}V}{\mathrm{d}\vec{r}_{ij}} \right\rangle = -\langle V \rangle$$

This is the virial theorem. It also applies to planetary systems,[51] to the hydrogen atom (even if treated quantum mechanically[52]), and to the kinetic energy of degenerate matter (section 5).

The process, by which the mean values predicted by the virial theorem are established, is also called "virialization". Virialization should not be confused with thermodynamic equilibrium. In globular clusters, virialization takes a time corresponding to approximately one orbit of a star around the center.

4.4.3 Self-gravitating systems are thermally unstable, the gravothermal catastrophe

In thermodynamics, a distinction is made between macro-states (sets of parameters that are known and robust) and micro-states (much larger sets of parameters that are not known and which, in their entirety, fully describe the system). If the microstates are all equally probable (Section 3.4), the most probable macro-state is the one to which the most corresponding micro-states correspond. This macro-state is the thermodynamic equilibrium.[53] The logarithm of the number of corresponding microstates is S/k_B with S the entropy. The entropy is maximum for closed systems at equilibrium.

Consider a system in an environment.[54] An environment is an infinite reservoir in the sense that its temperature does not change appeciably when heat is added to or removed from it. The system is in thermal equilibrium with the environment. Then the following applies

Glg. 4.13

$$0 = \frac{dS_{sys}}{dU_{sys}} + \frac{dS_{env}}{dU_{env}} = \frac{dS_{sys}}{dU_{sys}} - \frac{dS_{env}}{dU_{sys}}$$

$$\Rightarrow \frac{1}{T_{sys}} = \frac{1}{T_{env}} \Rightarrow T_{sys} = T_{env}$$

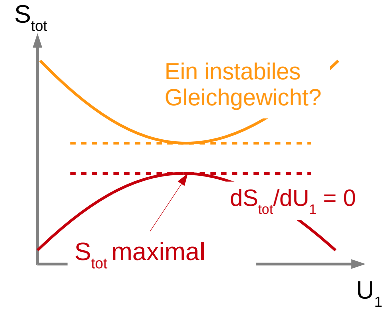


Fig. 4.8
For a stable thermal equilibrium, the total entropy is maximum. The tangent to the maximum is horizontal ($dS_{tot} /dU_1 = 0$).

dU is a small amount of heat, that is exchanged back and forth between the system and the environment. The subscripts *sys* and *env* denote the system and the environment. It was used that $1/T$ = dS/dU according to the definition of temperature.[55]

In thermal equilibrium, the temperatures of the system and its environment are the same. Over time, temperature differences will even out. At some point, the temperature will be the same everywhere. This is Clausius' heat death. The fact that the heat death does not occur under certain circumstances can be explained in purely thermodynamic terms. Thermal equilibrium is stable on Earth because the entropy in equilibrium is maximum ($d^2S/dU^2 < 0$):

---

[51] Mercury moves faster than Jupiter.

[52] The kinetic energy of the 1s-electron is longer than the kinetic energy of the 2-pelectron.

[53] More precisely, there is a family of closely related macrostates that together form the thermodynamic equilibrium. There are always small fluctuations in the macro parameters in equilibrium.

[54] The argument also applies more generally to any two systems in thermal contact.

[55] Strictly speaking, you have to write $\frac{1}{T} = \left(\frac{dS}{dU}\right)_V$ because $dU = -pdV + TdS$.

Glg. 4.14

$$\frac{\mathrm{d}^2 S_{tot}}{\mathrm{d}U_1^2} = \frac{\mathrm{d}^2 S_1}{\mathrm{d}U_1^2} + \frac{\mathrm{d}^2 S_2}{\mathrm{d}U_1^2} = \frac{\mathrm{d}^2 S_1}{\mathrm{d}U_1^2} + \frac{\mathrm{d}^2 S_2}{\mathrm{d}U_2^2}$$
$$= \frac{\mathrm{d}(1/T_1)}{\mathrm{d}U_1} + \frac{\mathrm{d}(1/T_2)}{\mathrm{d}U_1} = \left(\frac{\mathrm{d}(1/T_1)}{\mathrm{d}T_1}\right)\frac{\mathrm{d}T_1}{\mathrm{d}U_1} + \frac{\mathrm{d}(1/T_2)}{\mathrm{d}T_2}\frac{\mathrm{d}T_2}{\mathrm{d}U_1} = \left(-\frac{1}{T_1^2}\right)\frac{\mathrm{d}T_1}{\mathrm{d}U_1} + \left(-\frac{1}{T_2^2}\right)\frac{\mathrm{d}T_2}{\mathrm{d}U_2}$$
$$= -\frac{1}{T^2}\left(\frac{1}{C_{\mathrm{V},1}} + \frac{1}{C_{\mathrm{V},2}}\right)$$

C_V is the heat capacity at constant volume. C_V is positive for the terrestrial systems with which we are familiar. Gravitating gas clouds, however, have a negative heat capacity. The heat capacity is derived from the kinetic energy.[56] (The same applies to temperature.) From the virial theorem, Glg. 4.12 and $\langle E_\mathrm{kin}\rangle = 3/2\, k_\mathrm{B}T$ for particles without rotation and oscillation, it follows that

Glg. 4.15

$$\frac{3}{2}k_B T = -\frac{\langle V\rangle}{2}$$

The heat capacity C is negative:

Glg. 4.16

$$C := N\frac{\mathrm{d}\langle E_{kin}\rangle}{\mathrm{d}T} = -\frac{3}{2}Nk_B$$

Self-gravitating clouds of gas are therefore unstable in this sense. Contracted areas heat up, they radiate energy, they continue to contract and they continue to heat up (Fig. 4.9).[57] This process ends with the stars. In stars, the other interactions between the particles are no longer negligible.

Note: When it was said that the contracted areas radiate energy, it was assumed that they radiate more energy than they absorb from outside. This process requires a cold sky, caused by the expansion of the universe.

The gravothermal catastrophe is a special case of the thermal instability discussed above. Consider a self-gravitating cloud that has transported energy from the inside to the outside as part of a fluctuation. This is an energy transfer *within* the cloud. According to the virial theorem, the core then continues to heat up while the shell cools down and expands. Energy is then transferred from hot to cold (from the

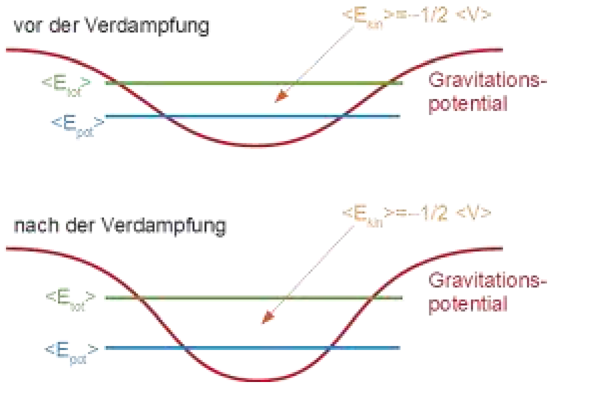


Fig. 4.9
For self-gravitating systems, the mean kinetic energy and thus also the temperature increases when particles evaporate or energy is radiated.

[56] Reservations against the definition of heat capacity via the kinetic energy lead to conceptual considerations of the most diverse kind. See for example: Wallace, David. "Gravity, Entropy, and Cosmology: In Search of Clarity". *The British Journal for the Philosophy of Science* 61 (July 3, 2009).

[57] This is the opposite of evaporative cooling. Heat is generated from the evaporation.

inside to the outside). The process is self-accelerating. Note: The envelope will allow material to evaporate into the environment.[58] This is an open system.

In clouds of gas, the process ends when a star has formed in the core. (Star formation is very complex.) In red giants the center condenses while the shell expands. In the case of open star clusters, the process ends in a binary star system. Double stars are largely stable. (Triple stars are not.) In the case of globular clusters (10^6 stars), the gravothermal catastrophe is underway, in principle, but it is slow.

An important consequence of the gravothermal catastrophe is that self-gravitating systems are never in a global thermodynamic equilibrium. However, local thermodynamic equilibria can exist (if there are frequent collisions).

<u>Remember:</u>

- Self-gravitating systems heat up spontaneously when they release energy.
- In self-gravitating systems, the core contracts and heats up, while the shell expands and cools. Examples are the red giants.
- Self-gravitating systems with no interactions other than gravity are never in a global thermodynamic equilibrium. However, there may be local thermodynamic equilibria.

4.4.4 The Jeans collapse

There also is hydrostatic instability, which, however, requires a certain minimum size. If the gain in gravitational energy during a contraction is larger than the increase in kinetic energy, the equilibrium becomes unstable. This happens as soon as the mass of the cloud is larger than the mass in Glg. 4.22 (larger than the "Jeans mass").

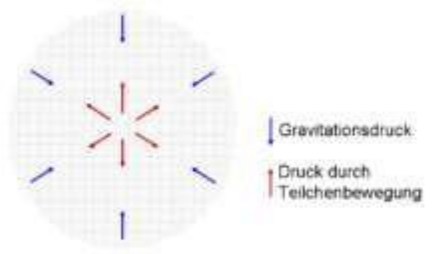


Fig. 4.10
In self-gravitating clouds, the gas pressure acts against the gravitational pressure.
en.wikipedia.org/wiki/jeans-criterion

The following text is adapted from https://en.wikipedia.org/wiki/Jeans_instability

Consider a homogenous spherical gas cloud with radius R. In order to compress this sphere to a radius $R - \mathrm{d}R$, work must be done against the gas pressure. During compression, gravitational energy is released. When this energy equals the amount of work to be done on the gas, the critical mass is reached. Let M be the mass of the cloud, T the (absolute) temperature, N/V the particle density, and p the gas pressure. The work to be done equals $p\mathrm{d}V$. Using the ideal gas law, according to which $p = N/V\,T$,[59] one arrives at the following expression for the work:

$$\mathrm{d}W \propto \rho_n T R^2\, \mathrm{d}R \qquad \text{Glg. 4.17}$$

[58] In the case of the red giants, the ejected material forms the "planetary nebulae", whereby the adjective "planetary" is historic. Planetary nebula have nothing to do with planets.

[59] The gas constant R has been absorbed here into a new definition of temperature.

The gravitational potential energy of a sphere with mass M and radius R is, apart from constants, given as:

Glg. 4.18

$$U \propto \frac{M^2}{R}$$

The amount of energy released when the sphere contracts from radius R to radius R – dR is obtained by differentiating this expression with respect to R:

Glg. 4.19

$$\mathrm{d}U \propto \frac{M^2}{R^2}\mathrm{d}R$$

The critical mass is reached when the released gravitational energy is equal to the work done on the gas:

Glg. 4.20

$$\frac{M^2}{R^2} \propto \rho_n T R^2$$

Next, the radius R must be expressed in terms of the particle density N/V and the mass M. This can be done using the relation

Glg. 4.21

$$M \propto \rho_n R^3$$

Elimination of R from the above two equations leads to the following expression for the critical mass:

Glg. 4.22

$$\text{I: } \frac{M^2}{R^2} \propto \rho_n T R^2, \qquad \text{II:} M \propto \rho_n R^3$$

$$\text{Ib: } \left(M^2\right)^{3/4} = \left(\rho_n T\right)^{3/4} R^3$$

$$\text{Ib/II: } \frac{\left(M^2\right)^{3/4}}{M} = \frac{\left(\rho_n T\right)^{3/4}}{\rho_n}$$

$$M^{1/2} = \frac{T^{3/4}}{\rho_n^{\ 1/4}}$$

$$M_\mathrm{J} = \left(\frac{T^3}{\rho_n}\right)^{\frac{1}{2}}$$

This last equation links the minimum mass for jeans collapse with temperature and density. The higher the temperature, the larger is the gas pressure that must be overcome by (scaling as $1/r$) the Jeans collapse.

$1/r$ potentials are rare on earth. Of course, the Coulomb potential is of long range, but the Coulomb potential is usually screened by countercharges. There is a playful experiment which emulates a "Jeans collapse" in the laboratory. Kavokine et al. found an $1/r$ potential in thermophoresis.[60] Colloidal particles sometimes move along a temperature gradient. The particles in question were absorbent and were heated with light. They then heated the surrounding fluid. Because these particular particles move towards the heat, there was a net attraction between the hot particles. A longer calculation (with some assumptions) led to an effective $1/r$ potential and the observed "implosion" indeed showed some peculiarities that are otherwise only known from astrophysics.

Remember:

– Self-gravitating systems collapse, if their size is above a certain minimum size.

4.4.5 Mass segregation

In self-gravitating systems, the heavy bodies migrate inwards over time. The lighter bodies move outwards. The following explanation is somewhat simplified.

The statistics of collision is complicated, but there is a tendency towards an even distribution of kinetic energy between the collision partners (related to the equipartition theorem).[61] Consider two bodies of different weight moving on orbits which are close to circular (but not strictly circular). Their velocity is close to the Kepler velocity, because the centripetal force and centrifugal force are about equal. Initially, the heavier body has the higher kinetic energy. After the collision, it has lost kinetic energy, meaning that it is now slower than would correspond to the Kepler velocity. Because the centrifugal force has decreased, it moves inwards. The reverse is true for the lighter body.

There is a connection between mass segregation and radial drift, which plays a major role in protoplanetary disks and accretion disks (Sections 8.1 and 8.4).

4.4.6 Local thermodynamic equilibria

The universe as a whole is not in thermodynamic equilibrium. Even self-gravitating systems as a whole are never in thermodynamic equilibrium. If the particles collide frequently, however, there are *local* thermodynamic equilibria. This is the case in the accretion disks (section 8.4) and sometimes also in protoplanetary disks (Section 8.1). Non-equilibrium thermodynamics then takes effect. In analogy to Rayleigh-Benard convection, there is a whole range of instabilities. The accretion disks show variabilities of all kinds.

Non-equilibrium thermodynamics was popular in the 1970s and 1980s. The claim was that non-equilibrium thermodynamics could explain dissipative structure formation. Dissipative structures only exist away from equilibrium. Examples are thermal convection, embryonic development and turbulence. However, there has been a shift in emphasis. Today, non-equilibrium thermodynamics primarily describes local equilibria and flows along the gradients of e.g. concentration or temperature (diffusion, heat conduction, etc.). As long as local equilibria exist, there is an extremum principle, as in equilibrium thermodynamics. In equilibrium thermodynamics, G tends to the minimum at fixed p and T. In non-equilibrium thermodynamics, all flows result from the requirement that the entropy production rate is maximum ($\dot{\sigma} \rightarrow \dot{\sigma}_{max}$). A few decades ago, it was a widespread opinion that dissipative structure formation should also follow from this principle. Rayleigh-Benard convection was cited as an example. If you heat

[60] Kavokine, N.; Zou, S. Y.; Liu, R. B.; Nigues, A.; Zou, B. S.; Bocquet, L., Ultrafast photomechanical transduction through thermophoretic implosion. *Nature Communications* **2020,** 11, (1).

[61] The local, ensemble-averaged kinetic energy is different in the different regions of the cloud, following the virial theorem.

a hotplate just a little, the heat is transported upwards via the usual heat conduction. There are local equilibria. The extremal principle $\dot{\sigma} \rightarrow max$ applies and the heat flow can be calculated from this. If you heat a little more, convection rolls form, which are also known from the weather. These "dissipative structures" disappear as soon as you stop heating and the system returns to equilibrium. (In contrast, the self-organized structures in soft matter, including micelles and vesicles, are stable in equilibrium).

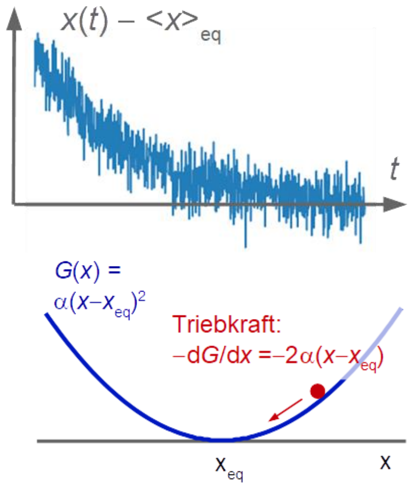


Fig. 4.11
If a system is close to thermodynamic equilibrium, the free enthalpy can be Taylor-evolved as $G(x) \approx$ □(x□x)$_{eq}^{2}$. The driving force for the development of the dynamic parameter x towards equilibrium is then □dG/dx = □2□(x x–$_{eq}$).
This results in the differential equation

$$\frac{d\left(x-x_{eq}\right)}{dt}=-2\alpha\left(x-x_{eq}\right)$$

The solution is the decreasing exponential function.
The nature of x is not specified here. It could be a concentration, a pressure difference, an electrical polarization or any other macroscopic parameter.
Attention: The movement towards equilibrium is superimposed on an errant flight.

Today, it is believed that there is no general extremum principle for these structure formation processes that could be used to model the multitude of processes in a common way. The extremum principle "$\dot{\sigma} \rightarrow \dot{\sigma}_{max}$" only applies to weak non-equilibria in which the fluxes are all proportional to the corresponding generalized forces. In astrophysics, too, there is a variety of structure formation processes, but these are most easily discussed in terms of the relevant dynamical laws (often non-linear partial differential equations). Entropy increases in all these processes, but the entropy production rate does not necessarily take its maximum value. An example: when soil dries, a pattern of cracks forms and entropy increases because elastic energy is dissipated. If you want to understand the cracks, you have to think about crack propagation. Crack propagation is part of mechanics. Of course, the second law thermodynamics applies, but it doesn't explain much. It is similar with the Jeans collapse, magnetic fields in plasmas and the streaming instability.

One could formulate it positively: If a situation is close enough to equilibrium, the approach to equilibrium takes place within the framework of a relaxation with a time dependence of the form exp(□t/□) where τ is the relaxation time (Fig. 4.11). Only those structures are formed that were already inscribed in the underlying potentials. The air in the mountains is thinner than at the sea, but this is not the result of spontaneous structure formation. This follows from the shape of the gravitational potential. All structures that are not inscribed in the potential *disintegrate* during relaxation. For example, differences in concentration often equalize if one waits for the relevant diffusion processes.

4.5 The biosphere needs the cold sky

In connection with the gravothermal catastrophe, it was emphasized that the radiation of energy heats up self-gravitating clouds. The radiation of energy requires a cold environment from which this energy does not return.

Similarly, the biosphere needs the cold sky as an entropy sink. One can argue with a Carnot process. The sun and the sky together do work on the atmosphere, do work on the oceans, and do work on the earth itself when they cause convection rolls and plate tectonics.[62]

A Carnot process is strictly required order in the biosphere. Photosynthesis in particular is different. Here, to create, the increase in entropy comes from the increasing number of photons. When the earth radiates heat, one solar photon ($\lambda \approx 500$ nm) turns into about 20 heat photons ($\lambda \approx 10\mu$m, Fig. 4.12). Each photon carries an entropy of about 3.6 k_B (section 10.6). The photon balance also causes the entropy to increase.

Kant is said to have said that the moral law within him and the starry sky above him filled his mind with awe. Arguably, the dark space between the stars is more important than the stars.

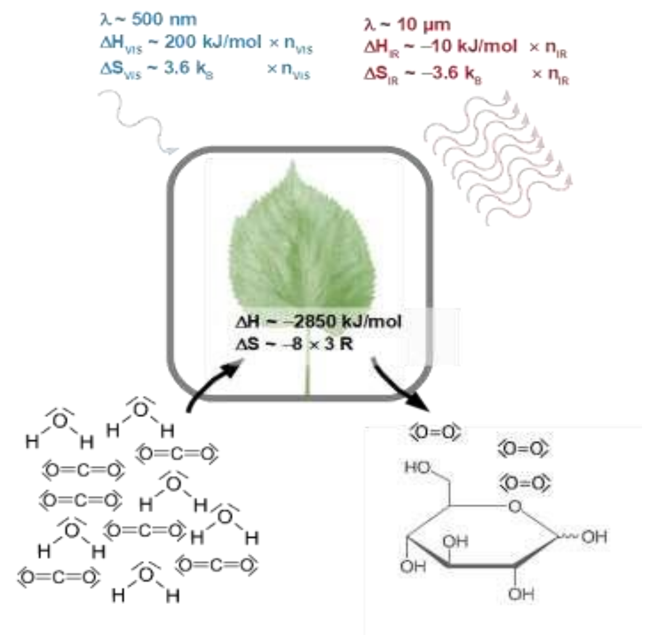


Fig. 4.12
Photosynthesis consumes entropy

5 Degenerate matter and the Chandrasekhar limit

"Degenerate matter" refers to states in which the diameter of the wave packets is larger than the distance between the particles. The particles have half integer spin (are "fermions") and therefore obey the Pauli principle. Two fermionic wave functions can never agree in all quantum numbers. This scenario is known from electrons in metals. The delocalized electrons form bands. The energy states are filled up from bottom to top. The boundary between the occupied and unoccupied states is the "Fermi energy". In space, white dwarfs and neutron stars consist of degenerate matter. The constituents of white dwarfs are ions and electrons (a plasma, but a degenerate plasma). The constituents of neutron stars are (mainly) neutrons. The nuclei of other stars also are sometimes degenerate.

In contrast to metals, the degenerate matter in space itself builds up the pressure and determines the pressure-density relation. (In metals, this is done by the ion lattice.) In degenerate matter, the pressure-density relation is different from that of gases (i.e. plasmas, in which kinetic energy makes the largest contribution to the total energy). The ideal gas law applies in these plasmas:

$$p = k_B T \frac{N}{V} \qquad \text{Glg 5.1}$$

In gases, the pressure depends on the temperature. If inceaser temperature, the plasma expands until a new hydrostatic equilibrium is found. The following applies to adiabatic compression

[62] The Earth's hot core also contributes to this work. Its energy comes partly from radioactivity and partly from a progressive contraction (i.e. from a progressive gravothermal catastrophe).

Glg 5.2

$$pV^{\gamma} = const \qquad p \propto \left(\frac{N}{V}\right)^{\gamma}$$

γ is the adiabatic exponent, given as

Glg 5.3

$$\gamma = \frac{c_p}{c_V} = \frac{c_V + R}{c_V} = \frac{f\frac{R}{2} + R}{f\frac{R}{2}} = \frac{f+2}{f}$$

The plasma under consideration here is monatomic. It has the 3 degrees of freedom of translation (no rotation, no vibration). Therefore, $\gamma = 5/3$. Again: we are concerned with adiabatic processes. These processes are too fast to dissipate any heat that may be generated. The Jeans collapse (section 4.4.4) adiabatic, because the temperature T in section 4.4.4 was assumed as constant.

The question now arises as to whether the gravitational pressure of self-gravitating systems increases more slowly than the pressure of the plasma, when the system contracts. If this is the case, the system is stable. We estimate the pressure as the derivative of energy with respect to volume:[63]

Glg 5.4

$$p = -\frac{\mathrm{d}E_{grav}}{\mathrm{d}V} \approx -\frac{\mathrm{d}}{\mathrm{d}V}\left(\frac{GM^2}{R}\right) \propto -\frac{\mathrm{d}}{\mathrm{d}V}\left(\frac{GM^2}{V^{1/3}}\right) \propto M^2 V^{-4/3}$$
$$pV^{4/3} = const$$

The gravitational pressure therefore increases more slowly than the plasma pressure (Fig. 5.1). [64]

The pressure of degenerate matter is also calculated as the derivative of energy with respect to volume. For each individual electron (for each individual particle) the following applies

Glg. 5.5

$$E = \frac{1}{2} m \mathrm{v}^2 = \frac{p_e^2}{2m}$$

The momentum is called p_e instead of p to avoid confusion with the pressure. Glg. 5.5 is the non-relativistic energy-momentum relation. To calculate the total energy, the energy of all particles must be added:

[63] Here we assume constant entropy, because $dU = -p\mathrm{d}V + T\mathrm{d}S$.

[64] Food for thought: Would a cloud of CO_2 collapse at room temperature? The following applies (found experimentally): $c_p \approx$ 37.12 J/(mol K) for CO_2 at room temperature.
The argument is of relevance in astrophysics. At some point during a type II supernova, the temperature rises to such an extent that the photons trigger nuclear fission ("photodisintegration"). This energy is then no longer available to counteract the gravitational pressure. A new degree of freedom is created.

Glg. 5.6

$$E_{tot} \propto \sum_{partikel} \frac{p_{e,i}^2}{2m}$$

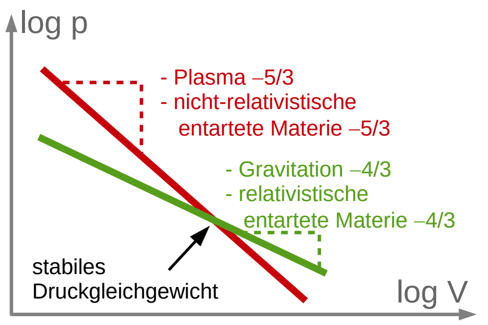


Fig. 5.1
Pressure-volume relation for the plasma, self-gravitating systems and the two forms of degenerate matter.

The particles successively fill up all possible momentum states (largely equivalent to the energy eigenstates) up to the Fermi edge. Call the "Fermi momentum" p_F. The momenta lie on spherical shells with the radius $4\pi p_e^2$. When integrating, it is important to ensure that the infinitesimal quantity dp^3_e is dimensionless (because we look for an energy). In quantum mechanics, the volume of a particle in phase space (approximately $\Delta x^3 \Delta p^3$) is given as $\hbar^3$ (numerical factors ignored). Therefore, dp^3_e must be replaced by $\Delta x^3/\hbar^3$. The result is

Glg. 5.7

$$E_{tot} = \int_0^{pF} \frac{p_e^2}{2m}\left(\frac{\Delta x^3}{\hbar^3}\right) d^3 p_e = \left(\frac{\Delta x^3}{2m\hbar^3}\right)\int_0^{pF} 4\pi p_e^2 dp_e = \frac{\Delta x^3 4\pi}{10m\hbar^3} p_F^5$$

The momenta are proportional to the wave numbers. These, in turn, are proportional to $1/\Delta x$:

Glg. 5.8

$$E_{tot} \propto \frac{\Delta x^3 4\pi}{10m\hbar^3}\frac{\hbar^5}{\Delta x^3} = \frac{4\pi}{10m}\frac{\hbar^2}{\Delta x^2} \propto \frac{4\pi}{10m}\frac{\hbar^2}{V^{2/3}}$$

The following results for the pressure

Glg. 5.9

$$p = -\frac{dE_{tot}}{dV} \propto \frac{2}{3}\frac{4\pi}{10m}\frac{\hbar^2}{V^{5/3}} \quad \Rightarrow pV^{5/3} = const$$

The adiabatic exponent is therefore 5/3, just like in the plasma. The white dwarfs are stable as long as the electrons are non-relativistic. Again: Degeneracy is a quantum phenomenon. Temperature plays no role as long as the diameter of the wave packets is larger than the distance between the particles.[65] The consequences are explained in sections 10.1 and 10.2. Furthermore, the pressure is lower than in ideal gases. Celestial bodies made of degenerate matter are therefore more compact than ordinary stars.

Remember: The temperature is not part of the algebra. The volume in the denominator is determined by the uncertainty relation (hence the ħ).

If the density continues to increase (because the gravitational pressure in the core of the red giant continues to rise), the volume per particle becomes smaller and smaller. The momentum becomes larger and larger due to the uncertainty relation. At some point, the particles become relativistic. For relativistically degenerate matter, the energy-momentum relation changes to

---

[65] The wave number increases with increasing kinetic energy. If the inverse wave number becomes smaller than the distance between two particles, the diameter of the wave packets can also become smaller than this distance.

$$E = cp_e$$ Glg. 5.10

The p_F^5 in Glg. 5.7 turns into p_F^4 . The adiabatic exponent turns into 4/3 (instead of 5/3). From then on, the celestial body is no longer stable. The maximum mass for white dwarfs is the Chandrasekhar limit (about 1.5 $M_\odot$). For neutron stars it is the Tolman-Oppenheimer-Volkoff limit (between 1.5 and 3 $M_\odot$).

Remember:

- In white dwarfs and neutron stars, the coherence length of the wave functions is larger than the distance between the particles. This form of matter is called degenerate.
- For degenerate matter, pressure is not a function of temperature.
- As long as the particles are non-relativistic, the adiabatic exponent is 5/3, just as for the ordinary atomic plasma. White dwarfs are therefore stable.
- At high pressure, the movement of the particles becomes relativistic. This changes the energy-momentum relation and ,in consequence the pressure-density relation. This leads to instability (an implosion).

6 Dark matter

- The velocities of the galaxies in the Conea cluster being incompatible with the virial theorem. The ex istence of dark matter is inferred[66,67]
- from the rotation curves of galaxies (Fig. 6.1)
- their effect as gravitational lenses
- the expansion rate of the universe (section 2.1.1)
- the velocities of stars in globular clusters (in conjunction with the virial theorem).

When "cold dark matter" is mentioned in the context of the Λ–XDM model, it is implied that thermodynamics also applies to this form of matter. "Cold" dark matter can be trapped in gravitational potential sinks. "Hot" dark matter has so much energy that it escapes these sinks. Cold dark matter is not only trapped in such sinks, it can even create these sinks itself by condensing - following the virial theorem - in certain regions.

At present, we only know what dark matter is probably not made of. It probably does not consist of neutrinos, because these would be hot and would not form clumps. It probably does not consist of a large number of compact, cold objects that cannot be seen by optical telescopes due to their low brightness. These "MACHOs" (massive compact halo objects) would be indirectly visible because they would act as gravitational micro-lenses. When gravitational micro-lenses enter the line of sight to a star, they produce a brief, transient brightening of the star (Fig. 6.3). This effect exists, but it is not very common. The rarity gravitational micro-lensing argues against the frequent occurrence of *all* forms of compact objects. It is conceivable that the dark material consists of many cold boulders and dark pieces of rock. However, these would cause micro-lensing.

Dark matter probably also does not consist of "weakly interacting massive particles" (WIMPS), i.e. particles that can observed based on the weak interaction (like neutrinos). WIMPs have been searched

[66] Remember: Accelerated motion of stars at this point does not contribute to evidence for dark matter. Velocities of stars (in addition to position) have been determined for many stars. For the acceleration, the 2nd time derivative of the position would be required. The current accuracy in determining star positions is rarely sufficient for this. (These positions were mainly determined with the Gaia satellite).

[67] Since dark matter is and remains so mysterious, one could assume that our current understanding of gravity is not quite correct instead.

for, but the researchers groups so far have come back empty-handed. At present, it looks as if dark matter is only evidencedvia the gravitational interaction.

There are "maps" of the distribution of dark matter. However, these maps are quite rough. They show, for example, that the dark matter in the spiral galaxies is less concentrated on the disk than the baryonic matter (stars, gas, dust).

Gravitational lensing is more efficient on large scales than on small scales. Gravitational lensing is best observed in clusters of galaxies. Whether or not one sees gravitational lensing depends on the ratio of the Schwarzschild radius of the object and the geometric radius of the object.[68] The Schwarzschild radius is larger than the radius (if the object is not a black hole). The "weak gravitational lensing" turns a circular galaxy into a galaxy with a slightly elliptical appearance. For any individual galaxy, a slight distortion due to gravitational lensing is indistinguishable from true ellipticity. However, the images of very many galaxies can be statistically analyzed (Fig. 6.2). This is the research goal of the "Euclid" space telescope.

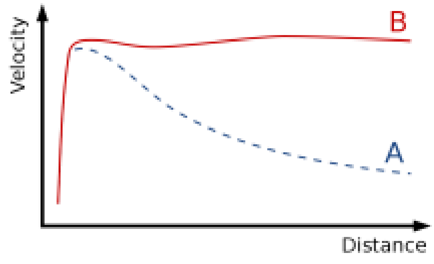


Fig. 6.1
Rotation curve of a typical spiral galaxy: A: Predicted from centripetal force exerted by visible matter:. **B**: Measured from the Doppler shift of the 21 cm line (Section 13.2).
The difference is attributed to dark matter.
en.wikipedia.org/wiki/Spiralgalaxy

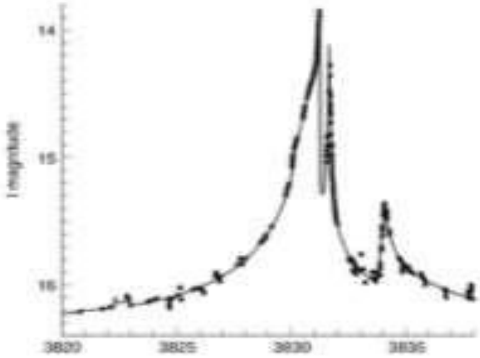
Remember:

– Dark matter is evident from the rotation curves of spiral galaxies, the speed of stars in globular clusters, the parameter ρ in the Friedmann equations for the expansion of the universe (Glg. 2.2), and gravitational lensing.
– There is more dark matter than baryonic matter.
– Dark matter is subject to thermodynamics. The dark matter in this universe is "cold" in the sense that it is trapped in minima of the gravitational potential.
– Dark matter probably consists neither of MACHOs nor of WIMPS.

[68] $R_{Schwarzschild/geometric}$ is generally the parameter that estimates how large the effects of general relativity are for the problem in question. The fact that this quotient is larger for galaxy clusters than for galaxies is not trivial.

32. Weak lensing in the galaxy cluster Abell 2218. The images of galaxies that lie behind the cluster are stretched by the cluster's gravitational field into arcs perpendicular to the field direction.

Fig. 6.2
If the galaxies in a certain regon are elliptical with a certain relative orientation to a center, one concludes to "weak gravitational lensing".
Source: Binney (footnote 1)



33. The micro-lensing event OGLE-2006-BLG-109. The brightness of the star measured at 12 observatories is plotted against time in units of a day. The gravitational fields of two planets generate extremely rapid brightness fluctuations. These data yield the mass of the star as 0.51 M? and the planets' masses as 231 M_e and 86 M_e similar to Jupiter and Saturn.

Fig. 6.3
If a star transiently becomes slightly brighter, this may be caused by a compact, invisible body in the line of sight. This is known as "microlensing".
Source: Binney (footnote 1)

7 Plasmas and magnetic fields in plasmas

7.1 Different types of plasmas

A plasma is a gas, but unlike ordinary gases, the particles are partially or fully ionized. There is no discrete transition temperature. The degree of ionization increases with increasing temperature.[69] Nevertheless, plasma is considered a separate state of matter because charged particles interact with each other differently from neutral particles.

There is not much chemistry in plasmas and there are few reactions that would lead to complex molecules. Most plasmas are chemically aggressive. However, there is plasma polymerization. In this process, a polymer – often a fluorinated polymer – is deposited on a surface. The polymer is usually highly cross-linked. The details of the reactions are difficult to control.

In plasmas, the ideal gas law applies at high pressure, but not the kinetic gas theory. (Box 16.1).

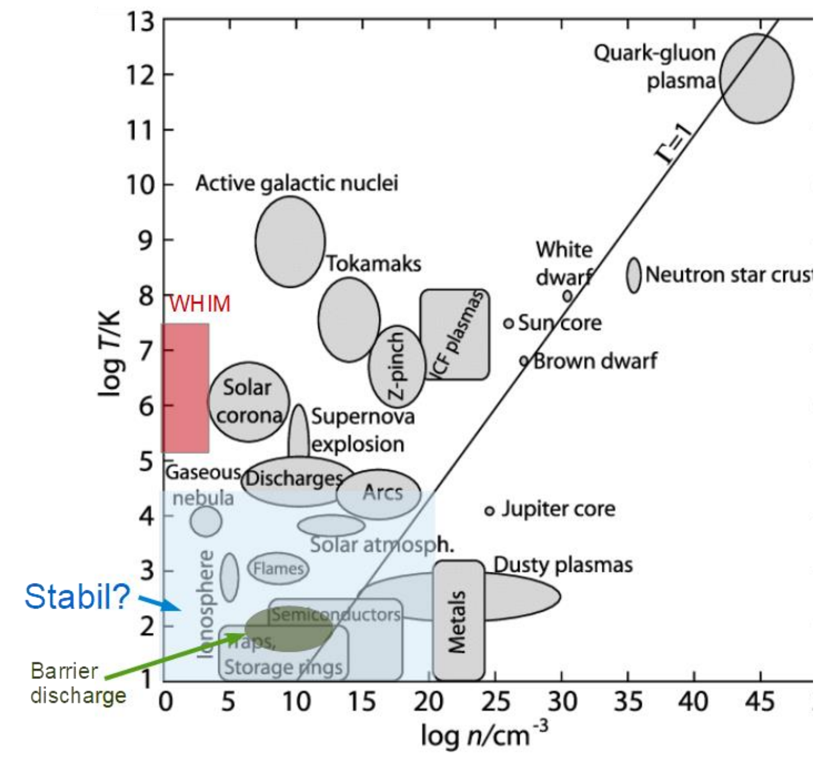


Fig. 7.1 Plasmas exist in a very wide range of pressures, densities and temperatures. Plasmas at room temperature and atmospheric pressure are unstable. They only exist as long as the source (e.g. a gas discharge in an arc or ionizing radiation) is active. *www.researchgate.net/figure/Different-types-of-plasmas-over-the-density-temperature-plane-Note-the-extremely-wide_fig1_2215721*

At a local level, there are strong electric fields. However, as in other places, these are largely shielded on a larger scale.[70] Electroneutrality also exerts a considerable force in plasmas. There is an analog to the Debye length. There is a length scale below which electroneutrality has no effect. This length depends on the square root of the charge density, just like the Debye length in the Debye-Hückel theory.

The following objects do not contain many ions:

– the molecular clouds (Sections, 10.5, 12.2, and 13.3)
– parts of the interstellar medium within a galaxy (keyword: 21-cm line, sections 13.1 and 13.2)
– the protoplanetary disks (Section 8.1)

[69] The transition is gradual in the same way that chemical equilibria depend gradually on temperature. Cooperative behavior would be necessary for a sharp transition. The degree of ionization is described by the Saha equation.

[70] On Earth, there are large-scale electric fields in the atmosphere. We know them from thunderstorms. Even when there is no thunderstorm, there is an electric field between the ground and the ionosphere. These fields should be regarded as the exception to the rule.

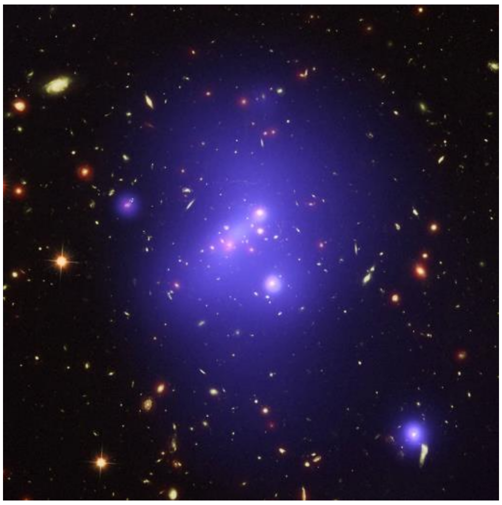


Fig. 7.2
Astronomers have made the most detailed study yet of an extremely massive young galaxy cluster using three space telescopes. This multi-wavelength image shows this galaxy cluster, called IDCS J1426.5+3508 (IDCS 1426 for short), in X-rays from the NASA Chandra X-ray Observatory in blue, visible light from the NASA/ESA Hubble Space Telescope in green, and infrared light from the NASA Spitzer Space Telescope in red. There is a region of bright X-ray emission (seen as blue-white) near the middle of the cluster, but not exactly at the center. ...
www.nasa.gov/mission_pages/chandra/images/galaxy-cluster-idcs-1426.html
The X-ray emission (blue) is not visible on black and white printouts.
The eROSITA satellite has created a map of galaxy clusters that shine in this way in the X-ray range.

7.2 The intergalactic medium

The space between galaxies is almost empty but only almost. It contains a plasma called WHIM, which stands for "warm and hot intergalactic medium". Warm and hot implies temperatures between 10^5 and 10^7 K. The high temperature is due to shock waves generated by material collapsing onto the WHIM from the voids. Outside the galaxies, the density of the medium is between 1 and 10 particles per m^3. The average collision time is a few hundred years. This means that a thermodynamic equilibrium is reached on astronomical time scales. The particles are in equilibrium with the other particles, but not with the CMB and also not with the electromagnetic radiation emitted by the surrounding stars.[71]

How do we know? The evidence is not overwhelming. The thermal radiation from this gas (in the X-ray range) is so weak that it is not picked up by current X-ray telescopes (especially Chandra[72]).[73] The gas can be detected by absorption lines in the X-ray spectra of bright X-ray sources behind the respective gas volumes. The absorption is caused by heavier elements that are not fully ionized. This is anything but trivial (remember the red shift), but according to Wikipedia, the WHIM has also been proven experimentally. The WHIM makes up 40-50% of the baryonic matter in the universe.

Remember:
- In the intergalactic medium (WHIM), the density is around 1 particle/m^3. The times between collisions are several hundred years.
- The particles reach a thermodynamic equilibrium despite the long collison times.
- The temperature is between 10^5 and 10^7 K. The heating is caused by shock waves, the latter generated by matter falling from the voids onto the WHIM.
- Thermodynamic equilibrium with the photon field is not achieved.
- The WHIM accouns for 40 - 50 % of the baryonic matter in the universe.

7.3 Consequences of magnetohydrodynamics

In plasmas (as elsewhere) magnetic effects are weaker than electric effects. This can be made plausible using the Bohr model (Section 16.1). However, because there are no magnetic monopoles, magnetic

---

[71] The collisions between particles and the CMB give rise to the GZK effect (after Greisen, Zatsepin and Kuzmin). It can also be regarded as an inverse Compton effect. In the Compton effect, a high-energy photon collides with a particle and transfers momentum. The photon loses energy and the particle gains energy. With the GZK effect, it is the other way around.
[72] *en.wikipedia*.org/wiki/Chandra_X-ray_Observatory
[73] Sometimes the plasma in galaxy clusters is also visible in emission, see Fig. 7.2.

fields are not screened in the same way as electric fields. There are interstellar magnetic fields and also intergalactic magnetic fields. Plasma in magnetic fields and magnetic fields in plasmas are described by magnetohydrodynamics (MHD). We report some results without going into the mathematics.

– If an ion flies at right angles to the magnetic field lines, it is forced onto a circular path by the Lorentz force and cannot be transported beyond this circular path at right angles to the magnetic field. However, transport along the field lines is possible. This coupling of the movement of the particles to the magnetic field also works in the other direction: if a plasma moves perpendicular to the magnetic field, it can pull the magnetic field along with it. The magnetic field and plasma then move together.
– Bodies with differential rotation (inside a little faster than outside) sometimes generate magnetic fields (assuming conductivity). This is referred to as a "dynamo" (earth dynamo, solar dynamo, dynamos in neutron stars). Leaving the details aside, the spontaneous amplification of magnetic fields can be made plausible as attempted in Fig. 7.3. There is a magnetic field in a conductive medium and there is

Box 7.1: Freezing of the field lines and magnetic pressure **F**
Freezing the field lines:
Combine Faraday's law with Ohm's law in Maxwell's equations, to arrive at the "law of induction" [1]

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \frac{1}{\sigma_0 \mu_0} \nabla^2 \mathbf{B} = \nabla \times (\mathbf{v} \times \mathbf{B}) + D_{mag} \nabla^2 \mathbf{B}$$

σ_0 is the conductivity. Without the first term, this equation would have the structure of Fick's 2nd law. D_{mag} is the "magnetic diffusivity". (Because the first term is non zero, diffusion takes place relative to the movement of the plasma). With perfect conductivity, the magnetic field line no longer diffuses. If $D_{\mathrm{mag}} \approx 0$, the movement of the magnetic field lines follows the movement of the plasma. Understanding this requires a little vector algebra. The field is then considered to be "frozen".

Magnetic pressure:
Start from an analog of the Navier-Stokes equation for magnetic forces

$$\rho\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right)\mathbf{v} = \mathbf{J} \times \mathbf{B} - \nabla p$$

This equation is also called the "equation of motion". On the left is the inertial force. The first term on the right is the Lorentz force. According to Ampere's law, the relation $\mu_0 \mathbf{J} = \nabla \times \mathbf{B}$ applies to the current density **J**. Furthermore, the following vector identity applies $\frac{1}{2}\nabla(\mathbf{B} \cdot \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{B} + \mathbf{B} \times (\nabla \times \mathbf{B})$

The relation $J \times \mathbf{B} = \frac{(\mathbf{B} \cdot \nabla)\mathbf{B}}{\mu_0} - \nabla \frac{\mathbf{B}^2}{2\mu_0}$ follows for the Lorentz force. The second term on the right has the form of a pressure gradient. It follows $\rho\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right)\mathbf{v} = \frac{(\mathbf{B} \cdot \nabla)\mathbf{B}}{\mu_0} - \nabla\left(\frac{\mathbf{B}^2}{2\mu_0} + p\right)$

Plasmas have a certain tendency to avoid zones with a high magnetic field. Conversely, the plasma sometimes concentrates the magnetic field in flux tubes. This leads to structure formation.
The first term in the equation above leads to a line tension ("magnetic tension force"). This term leads to a tendency to straighten curved field lines (as in the case of a stretched ribbon, shown above in Fig. 7.4).
[1] https://en.wikipedia.org/wiki/Magnetohydrodynamics

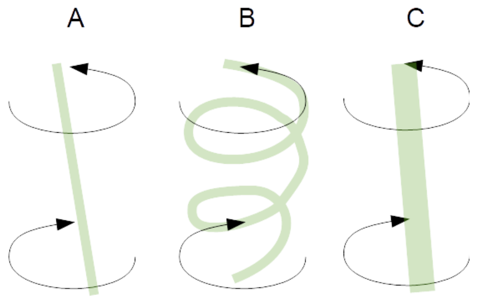


Fig. 7.3
Simplified sketch of the amplification of magnetic fields in plasmas with toroidal flow. A small initial magnetic field is dragged along by the flow and thereby elongated (B). The energy in the magnetic field increases. If it diffuses back and does not lose all its energy, the field has become stronger (C).
The effect is caused by an interplay of convective transport of magnetic fields (with the flow) and diffusive transport. The details are complicated

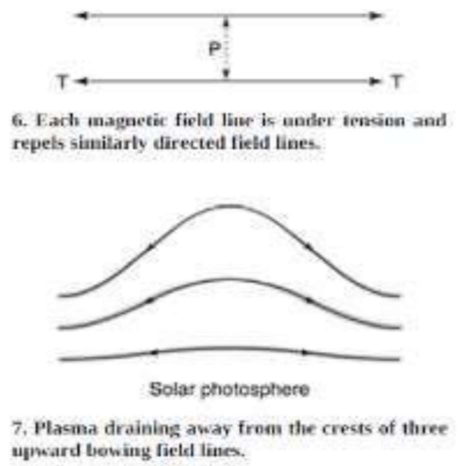


Fig. 7.4
Because the density of matter inside the tubes is lower than outside, a buoyant force is created near the surface of the sun and the magnetic field tubes move towards the surface of the sun. The material they carry with them rains back along the field lines in the direction of the gravitational field, making the tube even lighter and thus intensifying the process.
Source: Binney (footnote 1)

a circular flow in this medium.[74] If – as is often the case in plasmas – the magnetic field lines move together with the particles, straight magnetic field lines turn into helical field lines. This increases the energy contained in the magnetic field, because as the field lines become elongated. If there is a mechanism by which the field lines can diffuse back to the linear structure, then the interplay between elongation and back-diffusion results in positive feedback . The magnetic field is spontaneously enhanced.

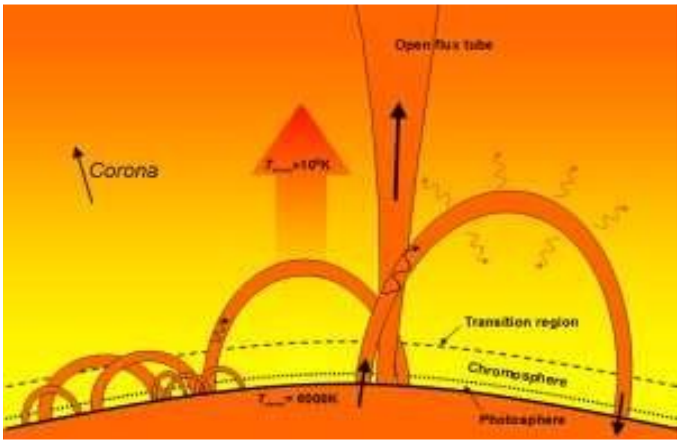


Fig. 7.5
Magnetic flux tubes on the sun
www.wikiwand.com/en/Flux_tube

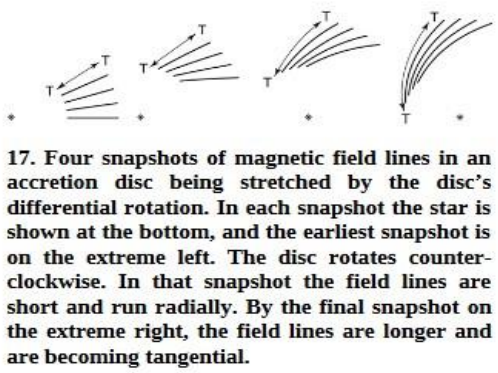


Fig. 7.6
The mechanism of the concentration of magnetic fields
Source: Binney (footnote 1

[74] There are further requirements for the flow field. There are a number of "no-dynamo theorems" that formulate these conditions.

– A high magnetic field generates pressure in the plasma (Box 7.1). The plasma is expelled from these regions. This is how magnetic confinement works in a fusion reactor. Conversely, a magnetic field is sometimes expelled from the plasma. It then is concentrates in flux tubes. This is a structure formation mechanism. The energy density in the tubes has a contribution from the magnetic pressure. Because the energy density is about the same as outside, the plasma in the tubes has a lower kinetic energy than outside. It is a little colder.
– Such tubes exist on the sun. There is an instability that causes these tubes to protrude from the surface of sun (Fig. 7.4). This flux tues give rise to the sunspots. They are somewhat darker than the rest of the sun's surface because the temperature there is lower. The magnetic field in the sunspots can be detected spectroscopically using the Zeeman effect.
– When magnetic field tubes merge, a lot of energy is released ("flux tube reconnection"). This energy contributes to the heating of the solar atmosphere (Fig. 7.5). Temperatures in the corona reach many millions of degrees.
– Fast, charged particles in strong magnetic fields emit synchrotron radiation (section 4.4), which can be recognized examining the spectrum. Synchrotron radiation is often emitted by jets. Fig. 7.7 shows the result of a simulation for the impact of a jet on interstellar gas.
– When a plasma becomes denser, the local magnetic field increases. This is what happens, for example, in the accretion disks (Fig. 7.6).

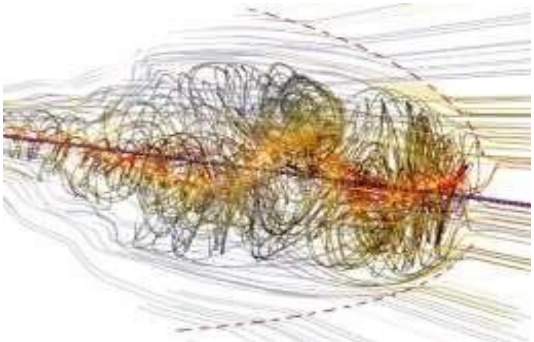


Fig. 7.7
Simulation of the magnetic fields at the tip of a jet. The equations of magnetohydrodynamics (MHD) apply. The lines are field lines of the magnetic field.
wwwmpa.mpa-garching.mpg.de/mpa/institute/news_archives/news1302_aaa/fig1.jpg

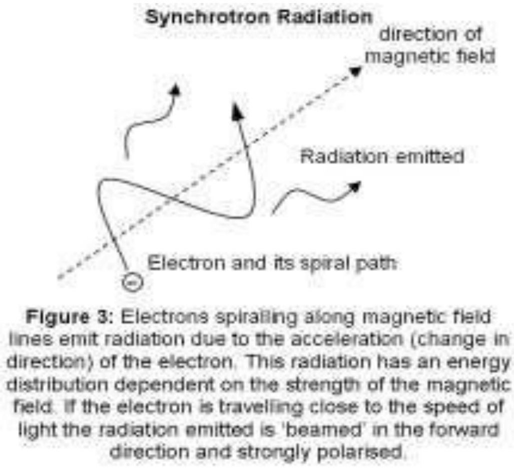


Fig. 7.8
On synchrotron radiation

Remember:
– There are often long-range magnetic fields in space.
– In plasmas, the movement of matter and the movement of the magnetic field are coupled.
– Spontaneous formation of magnetic fields can occur in plasmas with differential rotation (in dynamos).
– Magnetic fields are sometimes concentrated in flux tubes.
– The flux tubes are under tensile stress. The density and temperature of the plasma inside the tubes are lower than outside. The flux tubes therefore experience a buoyant force. Flux tube reconnection releases magnetic energy.

7.4 Side remarks: Neutron stars

Neutron stars were postulated soon after the discovery of the neutron. It was clear that these had to be compact objects with a high density. The neutron stars became accessible to observation as "pulsars", first discovered as radio sources. Understandably, these pulses (with a fixed frequency in the range of 1 Hz) were initially thought to be messages from other civilizations (this is the "LGM hypothesis", "LGM" for "little green men"). The assignment to neutron stars was based on a simple rule: fast processes must be caused by compact sources, because the finite light travel time does not allow cooperative processes over quantities that lie beyond $c\,t$ (c the speed of light and t the duration of the processes).[75] Pressure, temperature, magnetic fields and energy density are all extreme on the neutron stars and one wonders whether the familiar laws can be extrapolated to these extreme conditions. So far, the answer is: "By and large, yes".

The experimental basis for this statement consists of:

– Some neutron stars are pulsars.
– Some neutron stars emit jets.
– Some neutron stars are weak X-ray sources independent of their pulsation.

The consequences of the strong magnetic field are described in section 7.5. Other comments follow:

– The gravitational force on neutron stars is 10^9 times larger than on Earth. The pressure in the center is greater than 10^{25} atmospheres. The radius is about 10 km. For the internal structure see Fig. 7.9.
– Inside the neutron star, the density is 10^{14} times larger than the density of rock in the Earth's crust. The density in neutron stars is two to five times higher than in the atomic nuclei on Earth. Nuclear matter is known from nuclear physics, but the density there is lower.[76]
– There is a firm crust. Sometimes there are "star quakes". The star quakes can be inferred from occasional small changes in the speed of rotation.[77], because these slightly the moment of inertia.

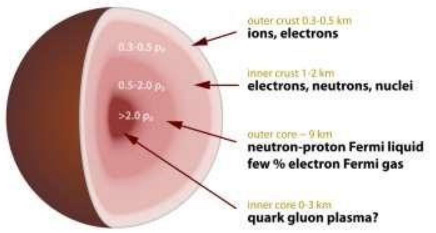


Fig. 7.9
Inner structure of neutron stars
en.wikipedia.org/wiki/Neutron_star

– Due to the conservation of angular momentum, the neutron stars rotate quickly (periods in the range of seconds).
– The millisecond-second pulsars have undergone further acceleration by an accretion disk, which formed from material falling in from a companion (Fig. 10.6).
– When an accretion disk forms around a neutron star (section 8.1), the energy in the disk is extremely high and there is a zoo of events with a lot of released energy. These are the "soft x-ray binaries".

---

[75] It also follows from this argument that the fast □-ray bursts (section 10.3) must originate from neutron stars or black holes of the same size (possibly from collisions of such objects). The other candidates - including supermassive black holes - are too large.

[76] A distinction is sometimes made between "nuclear forces" and "quantum chromodynamics" (QCD). On very small scales and at very high energies, the strong interaction is mediated by gluons. It then acts between the quarks. The interaction between the nucleons (protons and neutrons) is mediated by "mesons". These are particles consisting of a quark and an antiquark. Ultimately, the nuclear force also goes back to QCD, but there is an intermediate step. Sometimes this is portrayed as an to the difference between the electromagnetic interaction and the van der Waals interaction. In this picture, the vdW force is mediated by short-range virtual photons.
It is possible that QCD acts in the center of neutron stars instead of the nuclear forces.

[77] The length of the day also varies slightly on Earth because the distribution of mass changes over time. This is mainly due to plate tectonics and ocean currents. This variability on a short time scale is stronger than the slowing effect of the tides (Section 3.3)

– Many neutron stars have a high velocity relative to their neighboring stars. This leads to the conclusion that the previous type II supernova was not spherically symmetrical.
– Gravitational waves were indirectly detected for the first time being emitted the Hulse-Taylor pulsar. It is a pulsar with a companion. The orbital period decreases over time because gravitational waves are emitted. Although gravitational waves have a very small amplitude, they contain a lot of energy.

7.5 Extremely strong magnetic fields around neutron stars

An increase in the magnetic field during a compaction of a plasma takes on dramatic proportions in type II supernovae. Neutron stars (section 7.4) often have magnetic fields of the order of 10^{11} Tesla . Due to the spiral motion of the charges in magnetic fields, neutron stars only eject material at the magnetic poles. These are "jets". (Jets come in other variants.) When neutron stars rotate (they do rather quickly) and when the magnetic north pole is not on the axis of rotation, the neutron star acts like a lighthouse. The star appears bright at the moment the jet crosses our line of sight. These are the pulsars.

In special cases, a dynamo effect occurs and the magnetic field becomes a factor of 1000 stronger than the magnetic fields of ordinary neutron stars (up to 10^{14} Tesla). The superconducting protons play a role in this dynamo.[78] A magnetar is formed. Magnetars appear to the observer as "soft gamma repeaters" combined with "radio bursts". These bursts are generated by flux tube reconnection. Several 10 magnetars have been observed. This magnetic field decays in the course of about 10 000 years.

The energy density in this magnetic field is 10^4 times higher than the energy density of lead (calculated using $E = mc^2$). These magnetic fields are so strong that the vacuum becomes birefringent. Electromagnetic radiation can penetrate a plasma there because the magnetic field prevents the ions from following the alternating electromagnetic field.

The elementary particles with ultra-high energy (section 10.4) are accelerated by high magnetic fields in motion. This is a non-thermal mechanism (not entirely dissimilar to what happens in technical particle accelerators).

<u>Remember:</u>

– When a plasma collapses, the magnetic field collapses with it and then assumes extremely high values (10^{11} Tesla).
– The magnetic field of the neutron stars only allows matter (analogous to the solar wind) to escape at the magnetic poles. If the magnetic pole is not on the axis of rotation, a pulsar is formed.
– The emission at the magnetic poles leads to jets.
– In the case of magnetars, the magnetic fields are amplified by a factor of around 1000 as a result of a dynamo.

7.6 Side remarks: Synchrotron radiation

The following text is adapted from http://www.jeffstanger.net/Astronomy/emissionprocesses.html

The synchrotron emission mechanism is another example of an important astrophysical process (Figure 3 [Fig. 7.8]). Synchrotron radiation is observed in regions where relativistic electrons (those traveling close to the speed of light) spiral around magnetic field lines. This process results in strongly polarized

---

[78] In superconductivity, there is a coupling between two fermions mediated by the lattice vibrations in such a way that a quasiparticle with zero spin (a boson) is created. This is then subject to Bose-Einstein statistics with many consequences.
Such couplings exist in different variants. In particular, high-temperature superconductivity is still poorly understood. Coupling protons so strongly that they become superconducting requires high densities (and a penchant for visions). There are high densities on neutron stars.

radiation concentrated in the direction of the electrons motion (called "beaming"). Similar to bremsstrahlung, synchrotron has a characteristic shape of its spectra which is a power law spectrum. The shape of the spectrum produced is dependent on the energy distribution of the emitting electrons and is easily distinguishable from thermal blackbody radiation.

A knowledge of synchrotron radiation is essential in the study of a large group of astronomical objects called active galactic nuclei (AGN). A typical spectral profile for an AGN is shown in Figure 4 [Fig. 7.10 left].

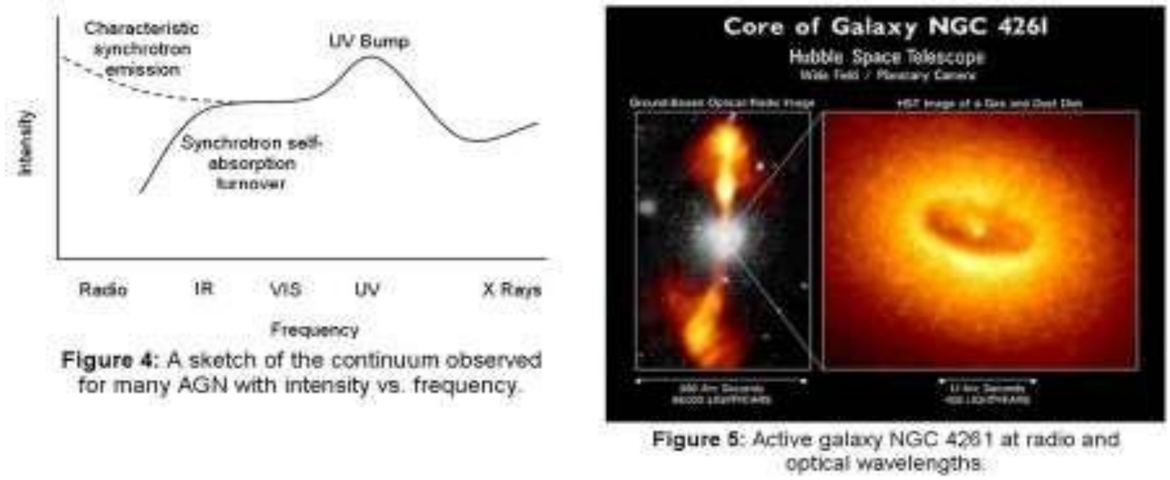


Fig. 7.10
Addressed in the text as "Figure 4" and "Figure 5".
http://www.jeffstanger.net/Astronomy/emissionprocesses.html

AGN are thought to contain massive black holes and the associated structure in the heart of the galaxy NGC 4261 can be seen in Figure 5. According to the unified theory for AGN the hot and dusty ion torus, shown to the right of Figure 5 [Fig. 7.10 right], emits unpolarized thermal radiation. This emission accounts for the 'UV bump' shown in Figure 4. This emission leads astronomers to believe that the thick ionized disk associated with this torus generates varying magnetic fields of up to 10^4 Gauss (very large) across its surface. This induces a large electric field and accelerates particles to relativistic speeds away from the disk (contributing to bipolar jets). These particles spiral along the magnetic field lines and produce synchrotron radiation which is up to 60% linearly polarized.

This synchrotron radiation is thought to be a contributing source to the power law component of the spectrum in Figure 4 [Fig. 7.10 left] (illustrated as the dashed line labeled as characteristic synchrotron emission). The spectrum shown in Figure 4 does not exhibit the synchrotron emission at low frequencies due to a common effect called synchrotron self-absorption. This is where the plasma that the synchrotron emitting electrons are part of becomes opaque to their synchrotron emission. This results in the turnover frequency that can be seen in Figure 4 (Fig. 7.10) at infrared (IR) wavelengths.

<u>Remember:</u>

- Fast-moving particles in magnetic fields emit synchrotron radiation.
- Synchrotron radiation has a characteristic spectrum and is polarized.

8 Accretion disks, protoplanetary disks and planetary systems

The gravitational collapse of gas and dust does not usually lead to a sphere, but rather to a disk, which is stabilized by the angular momentum and the resulting centrifugal force. The interaction between the bodies varies. There is always an interaction, even between Venus and Earth (section 3.2). If the interaction is weak, the bodies largely rotate at the "Kepler velocity", given as $(GM/r)^{1/2}$. The Kepler velocity results from the requirement that the centrifugal force is equal to the centripetal force. The orbits are almost circular, because if they were strongly elliptical, the interactions would be stronger and the system would be unstable. Examples of such systems with weak interactions between the orbiting bodies are our planetary system and Saturn's rings. Incidentally, our planetary system could well become unstable at some point, see footnote 35. Disks with stronger interactions usually consist of many rather small bodies or of gas. Examples are our planetary system in the early phase (Section 8.2) and the protoplanetary disks. Sometimes the interaction is so strong that the disk itself heats up. These are known as accretion disks. These exist in the vicinity of some neutron stars and in the vicinity of some supermassive black holes.

After an "encounter" has accelerated or decelerated a body on a Keplerian orbit, this body no longer has the Keplerian velocity and the orbit becomes slightly elliptical. The distance from the central body then "oscillates" around the value that the circular orbit would have and there are more encounters with other bodies. If these encounters are numerous, the oscillation is damped and the body returns to a circular orbit. This circular path has then shifted slightly inwards (keyword: "radial drift", section 8.1).

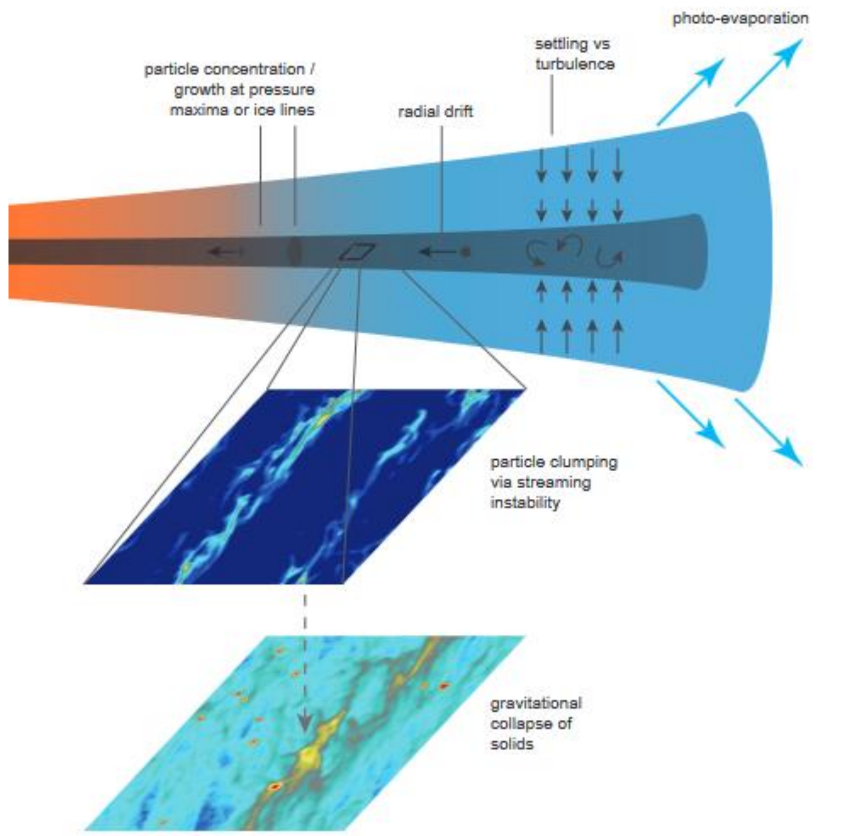


Fig. 8.1
The mechanism behind the streaming instability.
https://jila.colorado.edu/~pja/saas-fee/streaming.pdf

8.1 Radial drift and the streaming instability

When rocks larger than a few millimeters collide, they do not adhere permanently because the adhesion energy is not sufficient to dissipate the kinetic energy from the impact. The adhesion energy scales approximately like the surface area of the bodies, while the kinetic energy scales approximately like the volume. The surface-to-volume ratio decreases as the size of the bodies increases. Because of the lack of adhesion, collisions cannot explain the formation of planetesimals and planets.

This is where the streaming instability comes into play. The radial drift must first be explained. Basically, the orbital velocity for gas as well as for dust and stones results from the requirement "centripetal force = centrifugal force". For solid bodies, this means

$$\frac{mv^2}{r} = \frac{GmM}{r^2}$$ Glg. 8.1

m is the mass of the test body, *M* is the mass of the star. The gas pressure gradient also plays a role for gases. The gas is denser on the inside than on the outside. Gases obey the relation

$$\frac{mv^2}{r} = \frac{GmM}{r^2} - \nabla p$$ Glg. 8.2

The net centripetal force decreases and therefore the circulation speed of the gas decreases.

Fig. 8.2
This image of a protoplanetary disk (HL Tauri) from 2014 (from ALMA, see Box 12.1) has become famous. More recent images of other protoplanetary disks also show spiral arms.
https://technobyte.org/hl-tauri-alma-birth-solar-system-/

Fig. 8.3
The streaming instability is one of many instabilities that turn a homogeneous disk of gas and dust into a disk with structure.
jila.colorado.edu/~pja/saas-fee/saas-fee_turbulence pdf

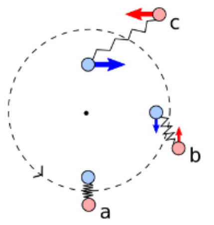
The dust particles also feel this pressure gradient, but it plays no role compared to the inertia and the gravitational force (because the dust particles have a larger mass). If there were no gas, the dust would orbit the star at Keplerian speed. However, the gas slows the dust down due to friction. The dust then flies somewhat slower than what centripetal force and centrifugal force would require. It drifts inwards. This is the "radial drift".

In the process of friction, the gas not only slows down the dust, at the same time the dust accelerates the gas. If the dust density is slightly higher somewhere, the gas flies a little faster there and the radial drift is smaller than in the neighboring zones. Now dust continues to drift into this region from the outside and further increases the dust density there. This further reduces the drift. The process experiences positive feedback. This is the "streaming instability". In this region, the dust density continues to increase. Filamentous structures form, which eventually collapse into spheres under the influence of gravity.

Today, protoplanetary disks can be observed better than a few years ago. A whole range of other instabilities are conceivable in protoplanetary disks (Fig. 8.3).

Section 8.2 traces the early history of our own planetary system. It is by no means trivial that a largely stable situation with many planets finally emerges.

There is a fairly solid understanding of the conditions in our solar system with regard to the objects on this side of Neptune's orbit. Beyond Neptune's orbit lie the trans-Neptunian objects. Some 10 000 of these objects with a radius > 10 km are suspected. Because these objects are cold, they are difficult to see. In particular, a suspected ninth planet beyond Neptune has not yet been identified. (A "planet" is a celestial body that has either swallowed up all other objects near its orbit or removed them from this orbit).



- Imagine there are two space wombats in nearby Keplerian orbits who are each holding one end of a spring
- The inner space wombat is moving faster than the outer one
- The inner space wombat gets pulled back while the outer one gets pulled forward
- Angular momentum gets transported outward

Fig. 8.4
The magnetorotational instability
https://lweb.cfa.harvard.edu/~namurphy/Lectures/Ay253_07_MHDinstabilities2.pdf

8.2 Side remarks: The young solar system

The following text is quoted from Binney.

It is thought that the young Saturn bumped into the young Jupiter in the way we have described and entered a 2:3 mean-motion resonance with Jupiter (three Jupiter years taking as long as two Saturn years) and then the two planets working in partnership pretty much stopped drifting inwards. Then the next planet out, we'll call it ice giant 1, drifting inwards encountered the almost stationary Jupiter-Saturn pair and entered a mean-motion resonance, probably again 3:2 with Saturn, and the system of three locked planets drifted only very slowly in radius. So along comes the next planet out, ice giant 2, and enters a mean-motion resonance with ice giant 1. This resonance may have been 3:4 . Now all four planets, working in partnership, remained in pretty much the same places while the young sun dispersed the gas in the disc.

Since the orbited time around a star increases with radius r as $r^{3/2}$, the time taken for dust to accumulate into asteroids and asteroids to gather into planets increases as we move outwards, and beyond ~ 20 AU this process was still incomplete when the Sun had dispersed the gas. So there was no ice giant 3 beyond ice giant 2, only a large number ~1 000 of Pluto-sized objects and zillions of asteroids. Once the gas was gone, three was nothing to damp the eccentricities of the asteroids that were excited by the gravitational fields of the Pluto-sized bodies. This ensemble of asteroids and Plutos surrounded the four locked planets but did not extend into the evacuated annulus of giant 2. The present-day Kuiper Belt of

Fig. 8.5
A sketch of the various objects in the planetary system. A distinction is made between
- Earth-like planets (mostly made of stone)
- the large gas planets (Jupiter, Saturn, Uranus and Neptune)
- the asteroids (made of stone, not spherical
- the dwarf planets (also known as "Plutos" or "Plutinos"), which are sperical but have not removed all the small bodies from their surroundings
- the objects in the Kuyper Belt (made of ice, i.e. water ice, ammonia ice or methane ice). They are sometimes in resonance with Neptune and sometimes not.
- the objects in the scattered disk. The orbits are eccentric and have a high inclination.
- the objects in the Oort cloud (made of ice, turn into comets from time to time)
https://www.pinterest.de/pin/309904018109936141/

asteroids is the descendent of this ensemble and we shall refer to the ensemble as the Primordial Kuiper Belt (PKB).

Once of the ice giants, probably ice giant 1, was on a slightly eccentric orbit and was able to exchange energy and angular momentum with the asteroids near the inner edge of the PKB at < 20 AU. If it wasn't locked in mean-motion resonances to the other planets, it would respond to this loss by moving to a fairly circular orbit of smaller radius. But it is locked, so it responds by moving to a more eccentric orbit. For a while its eccentricity steadily grows and then suddenly a secular resonance with the four-planet system causes the eccentricity of ice giant 1 to decrease, and the angular momenta of the planets to change in such a way the mean-motion resonant conditions are broken.

With the resonate conditions broken the planets can no longer exchange energy so any loss of angular momentum will lead to an increase in eccentricity (...). The eccentricities of the two ice giants quickly grow to large values so each of these planets crosses the other's orbit and possibly even Saturn's orbit. This is a time of great peril for the solar system, for a planet on a highly eccentric orbit is likely to induce other planets to move to eccentric orbits, and once Jupiter was on a highly eccentric orbit it would not be long before Jupiter would have driven every other planet either into the Sun or completely out of the solar system.

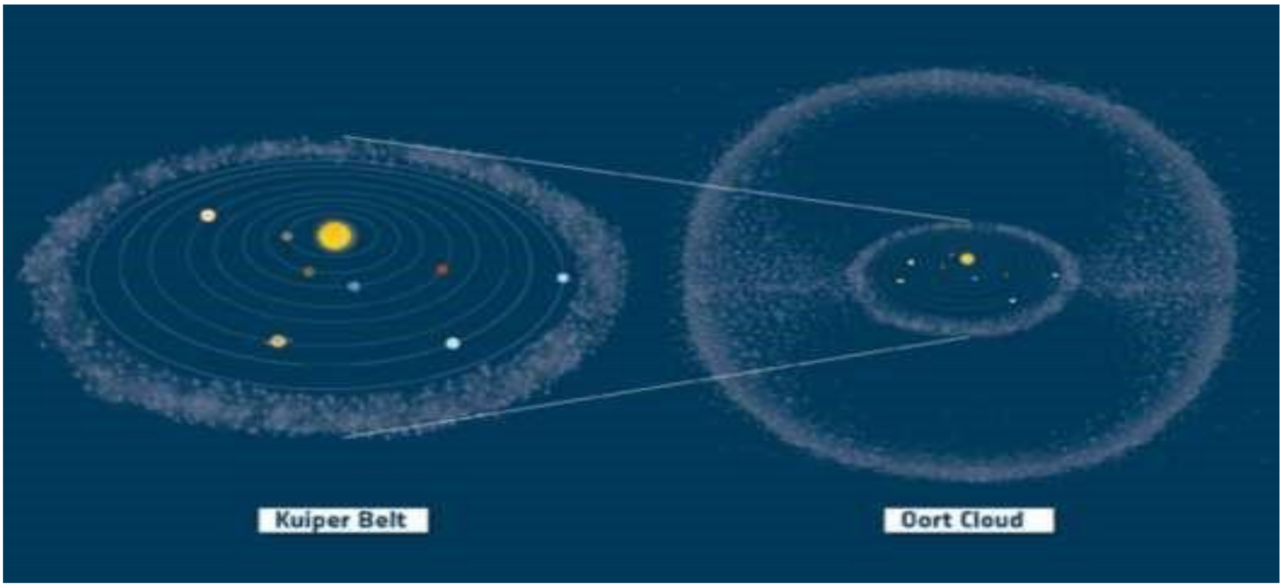


Fig. 8.6
Kuyper Belt and Oort Cloud
https://www.space.com/16144-kuiper-belt-objects.html

We shall see below that a catastrophe of this type has probably occurred in many planetary systems. We owe our existence to good fortune and the way the PKB acted as a fire bucket. As the eccentricities of the ice giants grow, they penetrated into the PKB and started to have close encounters with asteroids and Plutos. Scattering objects in the PKB damped the eccentricity of the ice giants, and the system settled to its present configuration. Neptune is now in a 1:2 mean-motion with Uranus and on an orbit of low eccentricity and semi-major axis α = 30.1 AU that places it far into the PKB.

Even the orbit of Uranus probably lies within the PKB (...). In some simulations of the evolution of the four-planet system after the resonance condition is broken, ice giant 1 ends up on a smaller orbit than ice giant two, and in other simulations it ends on the larger orbit. Thus we do not know which ice giant Neptune is.

The population of the PKB was decimated when the ice giants swept through it, For that reason the present-day Kuiper belt contains only ~ 0.07 M⊙ rather than the ~ 40 M⊙ from which it started. All but one of the ~ 1000 Plutos and many of the asteroids have been turfed out of the solar system [Fig. 8.7]. However, many of these objects at some stage appeared within the PKB and were scattered by the planets from Mercury to Saturn, pitting their surfaces and damping their eccentricities. Indeed, the rate at which asteroids hits the moon can be determined from the pattern of craters they made, and long before our cur-

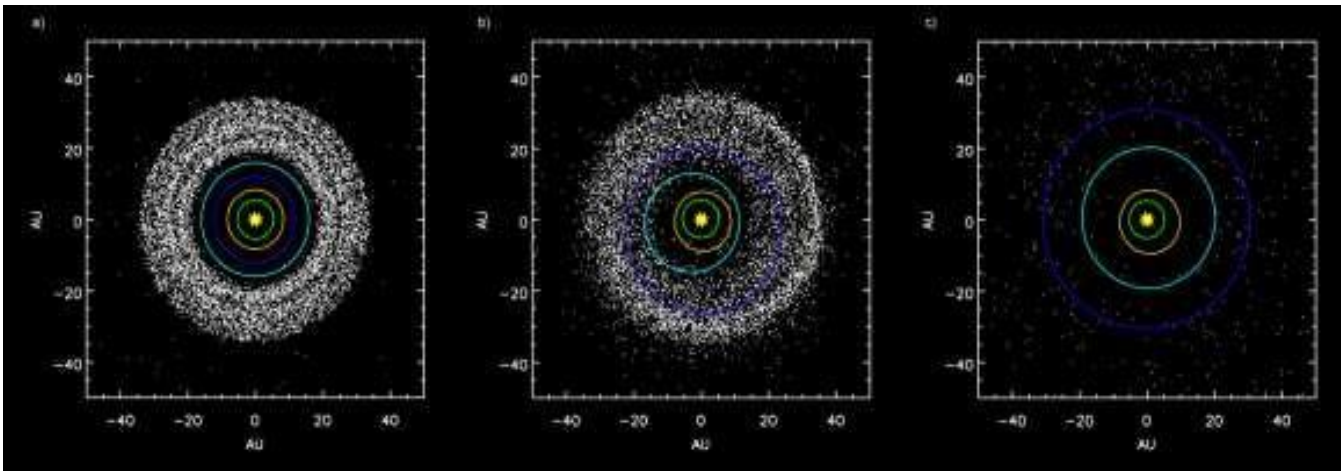


Fig. 8.7
Simulation showing outer planets and Kuiper belt: (a) before Jupiter/Saturn 1:2 resonance, (b) scattering of Kuiper belt objects into the Solar System after the orbital shift of Neptune, (c) after ejection of Kuiper belt bodies by Jupiter
en.wikipedia.org/wiki/Kuiper_belt

rent picture of the evolution of the solar system emerged it was known that there was a late heavy bombardment (LHB) of the Moon approximately 0.7 Gyr after the formation of the Sun 4.6 Gyr ago. Another likely legacy of this period of high asteroid density are the Trojan asteroids of Jupiter, which move on the same orbit as Jupiter but on the other side of the Sun. It is thought that Jupiter captured these asteroids at this time

8.3 Exoplanets

Planets outside the solar system receive an enormous interest. There are currently 5833 exoplanets with "confirmed" status (as of June 2023). The vast majority have been detected and observed indirectly. Among the indirect clues are a periodically changing Doppler shift in the spectrum of the star and a periodic small change in brightness following an occultation of the star by the planet. It is estimated that there are at least as many planets as stars in this galaxy.[79] It is estimated that about 20% of the sun-like stars have a planet in the "habitable zone". Liquid water could be present there. Spectra of planets can also be obtained indirectly if the planet passes behind the star and the spectrum of the entire system then changes very slightly.

The James-Web space telescope is equipped with a coronograph. This is a disk that blocks the light of the central star. An image of a planet was obtained with this device (so far only one, Fig. 8.8). Furthermore, several near-infrared spectra were obtained (in other words: obtained directly, in emission, not in the form of extinction of the light of the central star, Fig. 8.9).

---

[79] The phrase "as many planets as stars" expresses doubt that planetary systems like ours with many bodies are common.

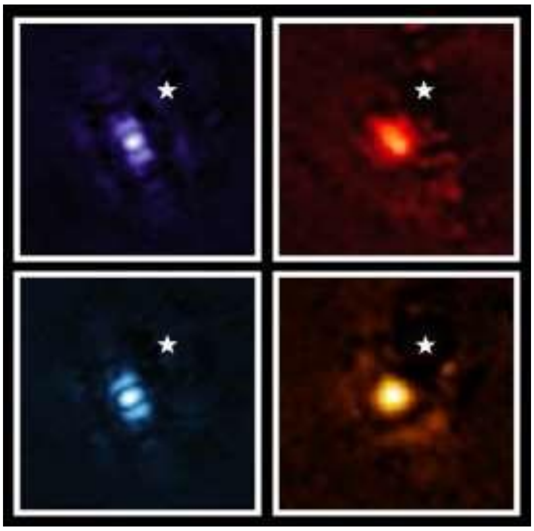


Fig. 8.8
Exoplanet HIP 65426 b shines in four different wavelengths in this image obtained with the James Webb Space Telescope. Purple represents 3 micrometers, blue is 4.44 micrometers, yellow is 11.4 micrometers and red is 15.5 micrometers. The shape of the planet is not a perfect circle because of the telescope's optics, in particular its hexagonal mirror.
The white star indicates the position of the central star covered by the coronograph.
theconversation.com/the-webb-telescope-has-released-its-very-first-exoplanet-image-heres-what-we-can-learn-from-it-189876
Text from
www.sciencenews.org/article/james-webb-spacetelescope-first-exoplanet-imag

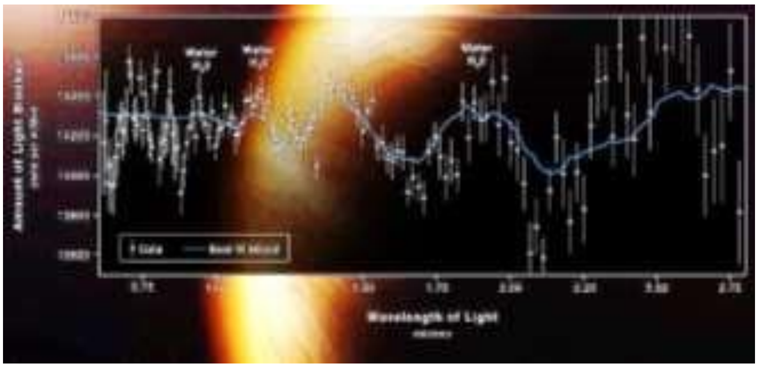


Fig. 8.9
NASA's James Webb Space Telescope has captured the signature of water, along with evidence for clouds and haze, in the atmosphere surrounding a hot gas giant planet orbiting a distant Sun-like star. WASP-96 b is one of more than 5,000 confirmed exoplanets in the Milky Way. Located roughly 1,150 light-years away in the southern-sky constellation Phoenix, it represents a type of gas giant that has no direct analog in our solar system. With a mass less than half that of Jupiter and a diameter 1.2 times greater, WASP-96 b is much puffier than any planet orbiting our Sun. And with a temperature greater than 1000°F, it is significantly hotter.

These are near-infrared spectra (NIR spectra). The bands correspond to overtones and combination tones.

www.nasa.gov/image-feature/goddard/2022/nasa-s-webb-reveals-steamy-atmosphere-of-distant-planet-in-detail

8.4 Accretion disks

Gas and dust are not aggregated into solid bodies in the accretion disks. The tidal forces would tear these bodies apart. Although the conditions in accretion disks are unusual, they are rather well understood. A distinction must be made between accretion disks around neutron stars or small black holes and accretion disks around supermassive black holes. The temperature curves and the luminosity curves are shown in Fig. 8.11. The accretion disks around the small objects form most of the steady X-ray sources in the sky.

There is a zoo of very energetic and brief events connected to accretion disks (e.g. the gamma ray bursts, Section 10.3). The accretion disks around the supermassive black holes are very bright, but not quite as hot and not the site of such fast, highly energetic events.

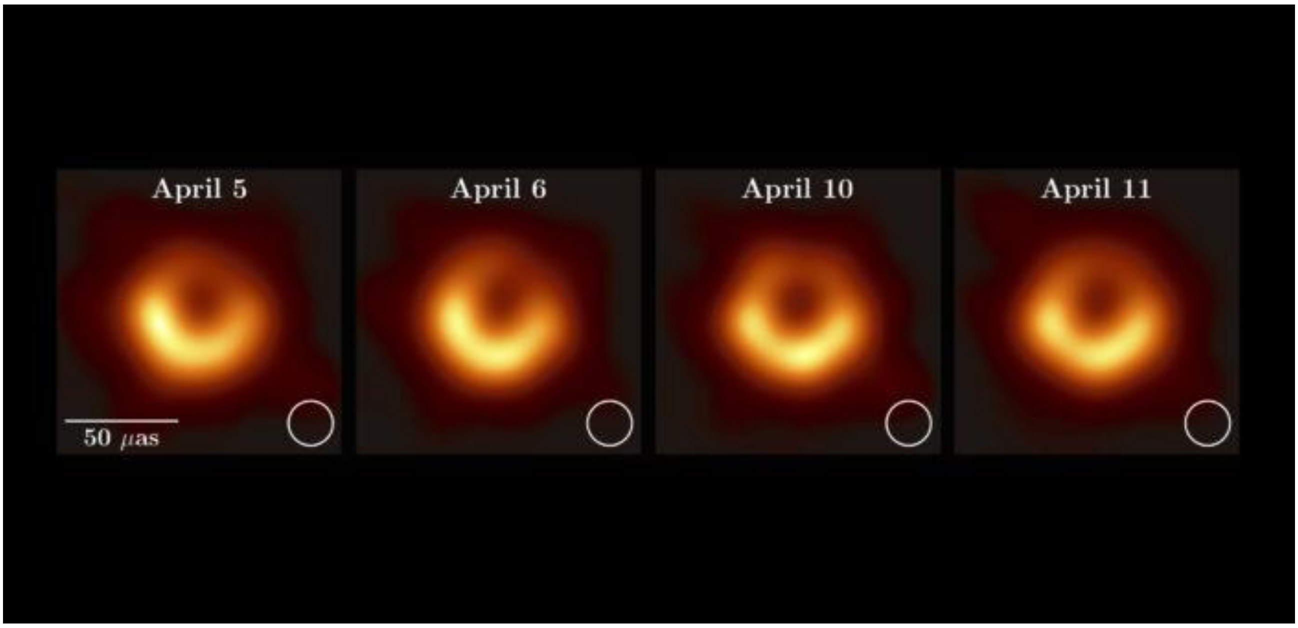


Fig. 8.10
In 2019, it was possible to spatialy resolve an accretion disk (around a black hole) for the first time. Until then, the models were largely based on spectroscopy, the variability of the objects and the jets that are sometimes emitted by the accretion disks. The variability can also be seen in this series of images.
aasnova.org/2019/04/10/first-images-of-a-black-hole-from-the-event-horizon-telescope/

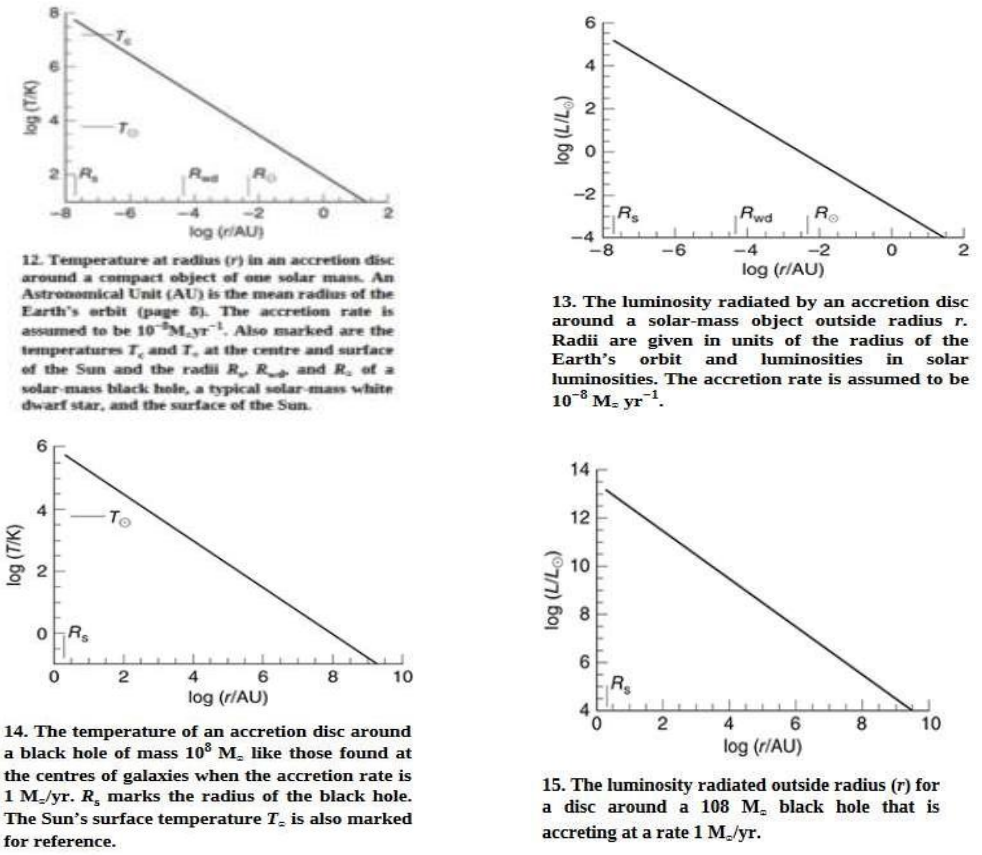


Fig. 8.11

Temperatures and luminosity in accretion disks:

Top: A black hole with a mass of about 1 $M_\odot$ Bottom: A supermassive black hole

Source: Binney (footnote 1)

9 Galaxies

9.1 Are galaxies discrete structural units?

Several forms of structural units are observed in the sky. These units can be regarded as "systems" in the sense of systems theory. Such systems have an "inside" and an "outside". The inner events are separate from the interaction with the outside, where not all inner events are significant for the interaction with the outside. In systems theory, we ask ourselves which properties of systems can lead to complexity.[80] Nowhere in astronomy is there such a pronounced complexity as in biology, but there are certainly more or less complicated and interesting systems. There are interactions between the systems. The planetary system in which we find ourselves can be considered interesting. The galaxies are also complicated and interesting.[81]

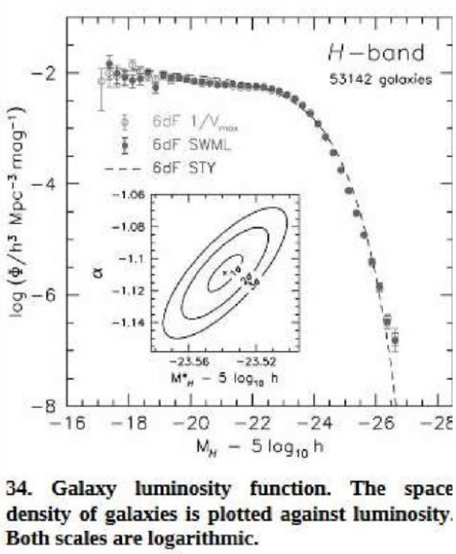


Fig. 9.1
There is a characteristic maximum size for galaxies (at "Schechter luminosity").
Source: Binney

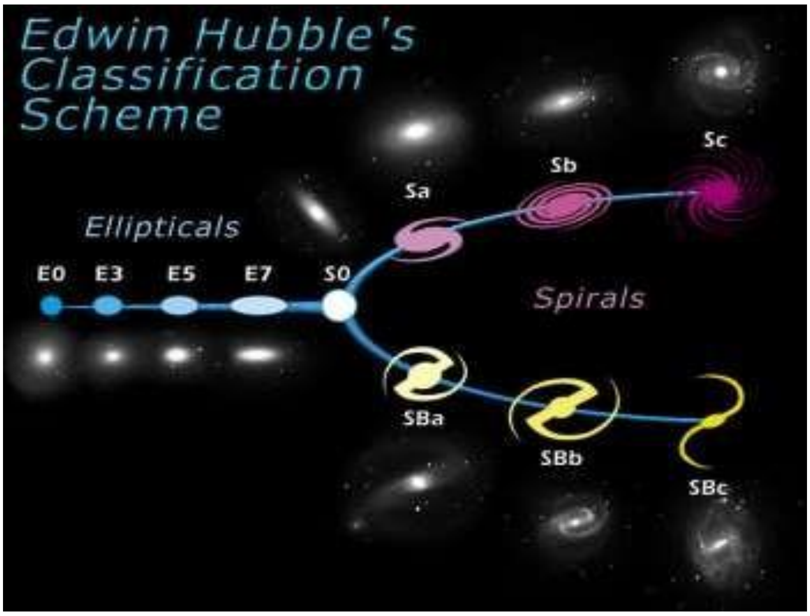


Fig. 9.2
Classification of galaxies
en.wikipedia.org/wiki/Galaxy_morphological_classification

---

[80] The term complexity is used in several ways. The "algorithmic complexity" of an object (an image, a text, a galaxy) denotes the minimum length of a code that describes the object. Boring texts can be summarized in a few words. The planetary system, the galaxies and the biological cell cannot be described in a few words. However, complexity also refers to the process by which such systems are created. This term is related to emergence. "Emergence" in the sense that other, more suitable concepts would lie above fundamental concepts is rather rare in astrophysics. You can be a good biologist without having understood quantum mechanics. Nowhere in astrophysics does physics take the back seat.

[81] There are about 10^{12} galaxies. The Milky Way contains around $2\cdot 10^{11}$ stars, but these are largely located in a specific area of the sky and are largely invisible. There is a difference here between the perception of the sky with the eye and the image that results with powerful telescopes. If these telescopes are not pointed at the Milky Way, you will see many more galaxies than stars. This is why galaxies can be examined quite well statistically. It is still possible to calculate average values if the ensemble of galaxies is subdivided according to age (i.e. redshift).

Binney has a big heart for the galaxies. The text is interesting, but the extent of his explanations exceeds the scope of this script. It is often about current research and sometimes - although Binney's didactic talent is truly remarkable - you cannot understand everything.

Are there other delimited units with an inner life in astrophysics? The molecular clouds that lead to star formation and end in an open star cluster should probably be mentioned here first and foremost (Section 10.5). Above the scale of galaxies there are "galaxy clusters" and "superclusters", but with these the outer edge and the inner life are in question. At the upper end of the scale, these various clusters merge into the "large scale structure of the universe" ("filaments", "voids", end of Section 2.1.1).

There is an upper limit for the galaxies at about 10^{14} solar masses (at the Schechter luminosity), Fig. 9.1). The value of this limit is not easily understood, but the fact that there is such a limit is related to the supermassive black holes at the centers of the galaxies. Once these have formed, the various neighboring galaxies live separately from each other (in much the same way that neighboring stars are separate). There are "mergers" (also of supermassive black holes), but there is no continuous coarsening of the distribution of the number density of stars (as is known from spinodal segregation). Once the supermassive black hole has formed, each galaxy has an identity and is different from the neighboring galaxies.

Binney also emphasizes the interactions between galaxies. Galaxies often merge. Virtually every galaxy has formed from several smaller galaxies. Furthermore, the halo of dark matter is extended in such a way that it can easily merge with the halo of a neighboring galaxy. According to Binney, galaxies are expressly not small worlds that stand alone. Binney would probably raise his eyebrows thoughtfully at the statement: "Galaxies are discrete structural units".

The classification of galaxies is less simple than that of stars. In the case of stars, the one central parameter its mass. The stars can be arranged on a line in the color-luminosity diagram, with the mass increasing from right to left. The classification of galaxies is drawn as a "tune fork-style" diagram according to Hubble (Fig. 9.2). A distinction is made between elliptical galaxies, spiral galaxies and barred spirals. Incidentally, the appearance of galaxies has changed as the universe has aged.[82]

The situation is more complex than with stars because there are more parameters that have an influence. Among these are

– the mass
– the angular momentum
– the content of dark matter
– the content of gas

This list is not complete and the classification of galaxies can be expanded (e.g. to include the "irregulars"). All confusion acknoledged: The classification scheme exists and one can explain when and why certain morphologies emerge.

Concerning of interactions within galaxies, a distinction must be made between

- the gravitational interaction between the individual star and the galaxy as a whole. This includes the interaction with the supermassive black hole at the center and the interaction with dark matter.
- the gravitational interaction between neighboring stars
- the gravitational interaction with the gas and the molecular clouds
- the forces exerted by supernovae
- forces that the magnetic field exerts on the gas.

[82] Details in *en.wikipedia*.org/wiki/Galaxy_morphological_classification

Remember:
- Many galaxies - especially spiral galaxies - have a complicated inner life.
- In the classification of galaxies, a distinction is made between elliptical galaxies, spiral galaxies and barred spirals.
- Parameters are the mass, the angular momentum, the content of dark matter and the content.

9.2 Modeling

The modeling of galaxies is particularly difficult but of course also particularly interesting. First of all, there is no thermodynamic equilibrium for stellar dynamics (Section 4.4.1). Dynamic models (models with a continued temporal evolution, which may be very slow) must be used. For stellar dynamics, a starting point is often an N-body simulation, i.e. the solution of N coupled ordinary differential equations for the N bodies ("acceleration = force / mass”. This cannot be implemented explicitly for galaxies (10^{11} stars), but smaller model systems can be studied. Nevertheless, N-body simulations quickly reach their limits when using large computers. Coarser models must then be used. One of these models is based on the Fokker-Planck equation (diffusion with potential gradients, diffusion here the diffusion of stars). In doing so, one must always try to take special features from the modeling at the lower level to the higher level. For example, if a binary star system flies past a third body, the third body will transfer energy to the binary star system more often than vice versa. The Fokker-Planck equation knows nothing of this fact.

Stellar dynamics does not cover the role of the gas (section 9.4). At this point, the micro/macro difference known from physical chemistry occurs twice. It occurs as a contrast between atoms and molecules on the one hand and macroscopic bodies on the other. It occurs a second time when the behavior of the stars is linked to the behavior of the galaxy.

Dark matter is a constant current source of worry. Dark matter far outweighs visible matter. Binney leaves the reader with the impresson that research can answer some of these puzzles.

9.3 The supermassive black hole at the center

How the supermassive black holes were formed is a subject of much research. In the early phase of the universe, these objects formed accretion disks that shone brightly. These are the "quasars" (for quasi-stellar object, see also section 8.1). Today, there are no more quasars close to us because these black holes have swallowed up everything there was to swallow over the course of time.

The conditions in the accretion disks in the vicinity of these supermassive black holes are remarkable in several respects. Firstly, the tidal forces tear apart the stars orbiting there.[83] Furthermore, up to a third of the rest mass of the particles is converted into radiation in these accretion disks. This means that the energy in the gravitational potential corresponds to about one third of the rest mass. (For comparison: in nuclear fusion, around 0.7 % of the mass is converted into energy). This energy turns into thermal energy and the disks glow brightly. They also allow brightly because the size of these objects far exceeds the size of ordinary stars. These objects are called "quasi-stellar" because they are point sources for tele-

---

[83] If a point like object orbits a central body, "centrifugal force = centripetal force" applies. For extended objects, however, the centripetal force is greater on the inside, while the centrifugal force is greater on the outside. On Earth, this leads to ebb and flow. On the moon, this leads to a very slightly prolate shape (a cigar-shaped deviation from the spherical shape). When compact bodies approach a black hole, they are torn apart by tidal forces and become part of the accretion disk. Entire stars can be "spaghettified" by supermassive black holes (*en.wikipedia.org/wiki/Spaghettification*). Such processes are not known to occur in planetary systems. There are tidal forces on Earth, but these leave the Earth intact.

scopes. The structure of these sources is not resolved because the distance is large and the angular diameter is correspondingly small. If the galaxy is spatially resolved by the telescope, it is referred to as a Seyfert galaxy.

Incidentally, it is not the case that the heaviest of these supermassive holes are also the brightest quasars. Heavy black holes have a large Schwarzschild radius. The larger this radius, the lower the negative gravitational potential at the Schwarzschild radius and the lower the tidal forces that tear the stars apart.

Sometimes active galaxies eject jets along the axis of rotation. When these collide with the surrounding medium, they shine brightly, especially in the radio range. These are known as BL Lacerta objects.

Remember:
- At the centers of all galaxies is a supermassive black hole.
- In the early phase of the universe, there were accretion disks in the vicinity of supermassive black holes. If these sources cannot be spatially resolved by telescopes, they are called quasars. In cases where the galaxy can also be spatially resolved, they are referred to as Seyfert galaxies.
- When the supermassive black hole ejects a jet, it is called a BL Lacerta object.

9.4 Elliptical galaxies and spiral galaxies, the special role of the gas

Elliptical galaxies are the simple case. They usually contain many old stars. There usually is little substructure (like the spiral arms of spiral galaxies). "Elliptical" usually means oblate, i.e. flattened, with the short axis parallel to the axis of rotation. As with all gravitating systems (section 4.4.1), mass moves towards the center and lowers the potential energy there. At the same time, outer areas expand and increase the entropy. However, this process is slow.

Spiral galaxies are the work of gas and dust. Galaxies that have lost all gas cannot become spiral galaxies. There are also few new stars there. Gas and dust produce a phenomenon that could be described by the term "friction". "Friction" causes stars, gas and, to a certain extent, dark matter to collect in the disk. In this way, the system gains potential energy without having to transport angular momentum outwards at the same time. Gas, dust and young stars are more concentrated than the older stars. The stars are more concentrated to the disk than the dark matter. The thin disk contains 65% of the visible mass of the galaxy, the thick disk only 5%. The central "bulge" did not participate in the formation of the disk. The bulge contains 3 % of the visible matter. For reasons that are not quickly explained, the central body sometimes has the shape of a bar. The galaxy then forms a barred spiral. The Milky Way is a barred spiral. The halo has not participated in the formation of disks. Therefore it mainly consists of old stars - often in the form of globular clusters. Dark matter also often remains in the halo.

The role of gas is quite complicated. First of all, it can certainly have a gravitational effect. If a galaxy (for example) contains as much gas in terms of mass as stars, then this gas will take an influence on all gravitational events. On the other hand, gas is also subject to hydrodynamics. Furthermore, the dynamics of the gas is influenced by magnetic fields. These are significantly weaker than the Earth's magnetic field (~1 nanoTesla, compared to 48 µTesla for the Earth's magnetic field in Europe), but nevertheless have an effect on a large scale. The gas does not act on the stars by directly generating a frictional force. It acts on the stars by forming molecular clouds in which supernova explosions occur a few million years later.[84] Supernovae have a substantial effect on the dynamics of all the matter around them. In particular, the shock waves generate new star-forming regions in the neighborhood.

Why exactly the spiral arms form is not quickly explained. It is a chain of instabilities.

Remember:
- The spirals of spiral galaxies are star-forming regions.
- There is a cascade of mechanisms that create new star-forming regions in the neighborhood of old star-forming regions.

9.5 The galactic halo, globular clusters

The galactic halo (with a roughly spherical shape) consists of scattered, old stars and globular clusters. The halo contributes only 1% of the visible matter. However, it contains 95% of the galaxy's total matter in the form of dark matter.[95] The stars in the globular clusters are all about the same age. The globular clusters must go back to large star-forming regions. Because the stars are more numerous than in most open clusters, the cluster has not disintegrated. Black holes of about 10^4 $M_\odot$ have recently been found in the centers of globular clusters. In terms of size, these black holes lie between ordinary black holes (a few $M_\odot$) and supermassive black holes (10^8 $M_\odot$).

[84] These star-forming regions contain HII regions. In these, the ionized hydrogen glows red during recombination. The HII regions make the spiral arms in spiral galaxies appear bright.[95] The text of this section is based on *de.wikipedia.org/wiki/Spiral-galaxy*.

10 Stars

10.1 Development over time, variable brightness

The larger part of a star's life is rather unspectacular. These stars are referred to as "main sequence stars". The main sequence consists of points in the color-luminosity diagram[85] (Fig. 10.1). In this diagram, a band runs from top left to bottom right. This is the main sequence, which is the consequence of a relationship between color and luminosity. The variable underlying both parameters is mass. The larger a star is, the hotter it is in the center, and the faster nuclear fusion takes place. Heavy stars are bright and blue (the latter because they are hot, Wien's law). Heavy stars live shorter than low-mass stars because the fuel is used up more quickly. Very heavy stars only live for a few million years. Low-mass stars can live longer than the universe is old now (mathematically: hundreds of billions of years).

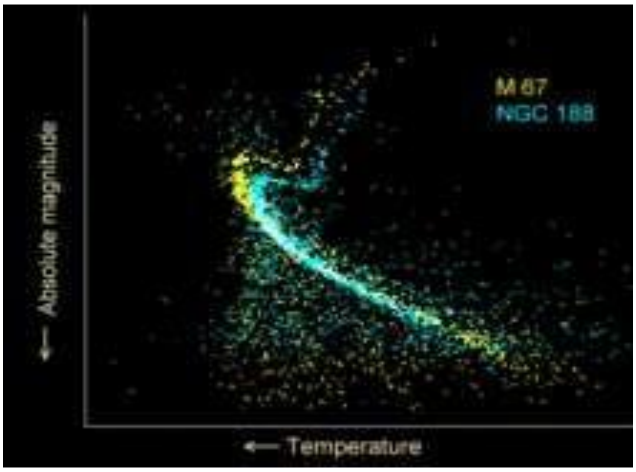


Fig. 10.1
A temperature-luminosity diagram for two open star clusters. NGC 188 (blue) is the older of the two and shows an earlier bend from the main sequence than M67 (yellow). Open star clusters (such as the Pleiades) are clusters of young stars. They are to be distinguished from globular clusters (made of old stars).
studylibde.com/doc/2494993/stars

Another initial parameter that influences the properties of stars is their "metallicity". In astronomy, all elements that are heavier than helium are considered metals. Metallicity was zero in the first generation of stars. Since then, it has increased because material is included in star formation that has already passed through one or more stellar lifetimes. The metallicity can be deduced from the spectra.[86] It has a small influence on the properties of the star (e.g. the mass-luminosity relationship).

The stellar radius can be determined optically for a few stars in the neighborhood of the Sun. A famous example is Betelgeuse (Box 10.1).

Nucleosynthesis - in other words, the various nuclear reactions that produce the heavier elements from hydrogen and deuterium and release energy in the process - is fairly well understood (section 11.1). Fig. 8.4 shows processes that consume hydrogen. At low temperatures (in the "brown dwarfs") hydrogen does not burn, but deuterium does (the latter is also always present because it was also produced in the Big Bang). At the end of a star's life, helium burns, simply because the hydrogen has been used up. Later, carbon, oxygen and silicon also burn. However, this only happens with sufficiently heavy and hot

[85] Somewhat sharper structures are obtained by plotting the so-called spectral class on the *x-axis* rather than the color (Fig. 10.5). The spectral class is derived from a rule that stars should be ordered according to the occurrence and strength of certain lines. This latter plot is called the Herzsprung-Russell diagram.

[86] Because the fraction of "metals" increases over time, the metallicity of a star can be used to infer the time of its formation. This analysis shows that some globular clusters consist of stars that were all formed at the same time (and are quite old). However, this is by no means always the case. Many globular clusters contain two or more populations, so they have undergone some kind of merging process.

stars. In the case of iron, the chain ends because iron has the highest binding energy per nucleon of all nuclei

Stars are spherical or almost spherical. Deviations occur in binary star systems due to tidal forces. Furthermore, convection can lead to irregularities in shape and brightness.

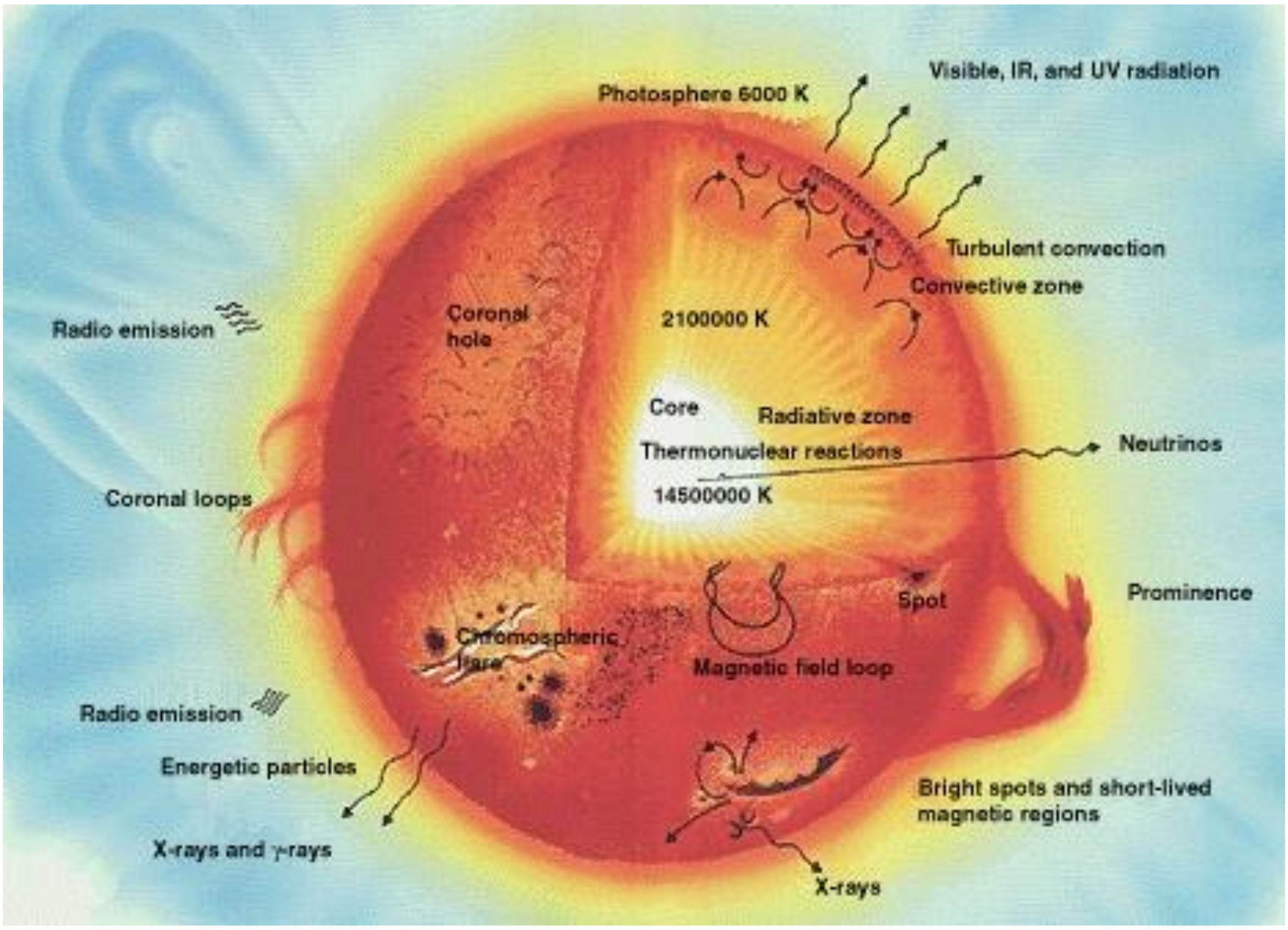


Fig. 10.2
Sketch of a sun-like main sequence star, *en.wikipedia.org/wiki/star-structure*

Fig. 10.2 outlines the internal structure of stars. There is a temperature distribution $T(r)$ and a pressure distribution $p(r)$. There are zones with and without convection. Different nuclear reactions take place in the different zones. Most stars have an atmosphere. The magnetic field tubes play a special role here. Where they emerge, the surface of the star is somewhat darker. On the sun, this is referred to as sunspots (see also Fig. 4.7).

The majority of stars have a largely constant luminosity over time. But there are also "variable" stars.[87] Causes for variability can be
- An occultation by planets (small effects)
- a rotation, combined with a small ellipticity, the latter caused by the tidal forces in binary star systems

[87] Nowadays, changes in luminosity in the order of magnitude of 10^{-4} mag can be detected.
The "mag" is a logarithmic measure of the apparent luminosity. A difference in apparent brightness of 5 mag corresponds to a difference in luminosity of a factor of 100, whereby bright stars have a small magnitude. The magnitude can also be negative. The details and calibration are quite complicated. Polaris has a magnitude of 2.0 mag.

- Effects of stellar atmospheres and flares in particular (as on the sun), produced by the fusion of magnetic flux tubes. Flares are only visible in less luminous stars. Otherwise their brightness is insignificant in comparison to the total luminosity.
- the "helium flash"

  This mechanism underlies the "strip of instability" in the Hertzsprung-Russell diagram (Fig. 10.5). If the fuel is present as degenerate matter (as it is in this case), the pressure is governed by the density (according to the uncertainty relation, section 5) and is no longer a function of temperature. This eliminates an otherwise stabilizing negative feedback according to "reaction rate ↗ $\Rightarrow T$ ↗ $\Rightarrow p$ ↗ $\Rightarrow$ expansion ↗ $\Rightarrow \rho$ ↘ $\Rightarrow$ reaction rate (. The reaction rate increases until the expansion of the wave packets ($\Lambda = h/(2\pi m k_B T)^{1/2}$) becomes smaller than the distance between the particles at a very high kinetic energy. Then the degeneracy is lost and the flash follows.
- Instabilities resulting from a complicated temperature dependence of the opacity in the outer areas of the star.
- different mechanisms for eruptions, which are often associated with accretion disks (Section 8.1). The accretion disks often occur in binary systems where one partner is a red giant and material passes from the red giant to the other star.

Some variables have a fixed relationship between period and luminosity. The distance can then be calculated from the period and the apparent brightness. This works particularly well with the δ cepheids The δ cepheids were important standard candles when measuring distances in astronomy was more difficult than it is today. [88]

On the surface of stars there is a corona with extremely high temperatures (sun: several million K). The mechanisms of heating are sound and the fusion of magnetic flux tubes (Section 7.1). The temperatures are so high because the density becomes lower and lower towards the top. The sound waves travel upwards, but no longer find any material into which they could dissipate their energy. This can be thought of as an analogy to the crack of a whip.[89]

Because the corona is very hot, the ions in the solar wind have a high kinetic energy. Lines were also found in the emission spectra from the solar corona that could not be assigned. An element with the name “coronium” was postulated. It turned out that the radiation in question was emitted by highly ionized atoms. Such highly ionized atoms do not exist on Earth.

---

[88] The Gaia satellite has much improved our understanding of the astronomic distance scale. For stars close to the earth, the distance can be calculated from the parallax, i.e. from the change in apparent position in the sky when the earth orbits the sun. The distance measure commonly used in astronomy is the parallax second, parsec. 1 parsec corresponds to $30 \cdot 10^{12}$ km and 3.26 light years. With earthbound telescopes, distances of up to around 100 parsecs can be determined. Gaia has increased this distance to around 8000 parsecs and thus anchored the "astronomic distance ladder" more firmly at the lower end.

[89] Goethe also commented on this question in the prologue to Dr. Faustus. He has the archangel Raphael say:

The sun resounds in the old way,
In brotherly spheres, in song of competition,
And its prescribed journey,
it completes with thunder.

In this particular case, J.W. was right: it is indeed very loud on the sun. I owe this argument to E. Pitz, Heidelberg.

At the end of a star's life, there are again several scenarios. When the fuel is used up, ash accumulates in the center. Depending on weight and temperature, the ash can consist of different nuclei. In the light stars, carbon is part of the ash because the temperature is not sufficient to ignite the carbon. In heavy stars, the ash mainly contains heavier elements (but not elements heavier than iron, because the series of fusion processes ends with iron).

Box 10.1: A special case: Betelgeuse
The star Betelgeuse is particularly interesting in that it is one of the few stars the shape of which can be spatially resolved. The images sometimes show a particularly bright spot (a hot bubble). Some of the images show dust. Betelgeuse is a supergiant whose brightness fluctuates. The strong fluctuations in recent years have suggested that a supernova could be imminent. More recent results indicate that this is not the case.

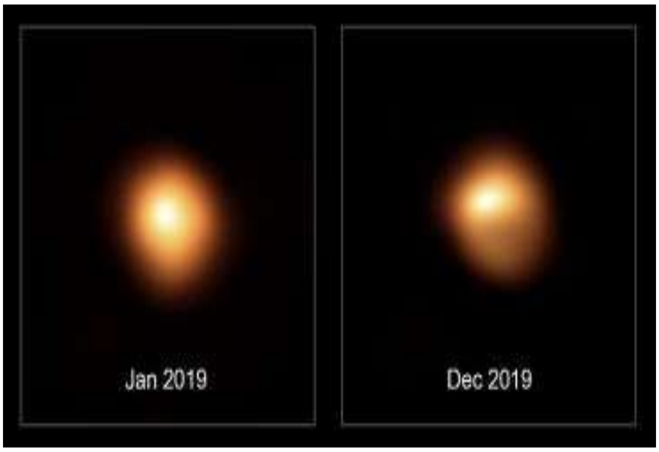


Fig. 10.3
Betelgeuse in January and December 2019, taken with the Very Large Telescope of the European Southern Observatory. Diameter of the disks: 0.05 arc seconds.
en.wikipedia.org/wiki/Beteigeuze

Above the core of ash, the star may continue to burn. This is known as shell burning. Because there is less heavy material above the burning areas than with the main sequence stars, the star expands. It is quite as hot as before. It becomes a red giant. Many of the stars seen in the sky with the naked eye are red giants. Because the gravity on the surface of the red giant is relatively small (due to the large radius), more material escapes into space than with ordinary stars. The sun also emits a solar wind (a particle radiation). This wind is stronger for red giants. In particular, it contains heavy elements that condense into dust. The dust could be described as soot, but the star dust mostly contains silicates rather than carbon.

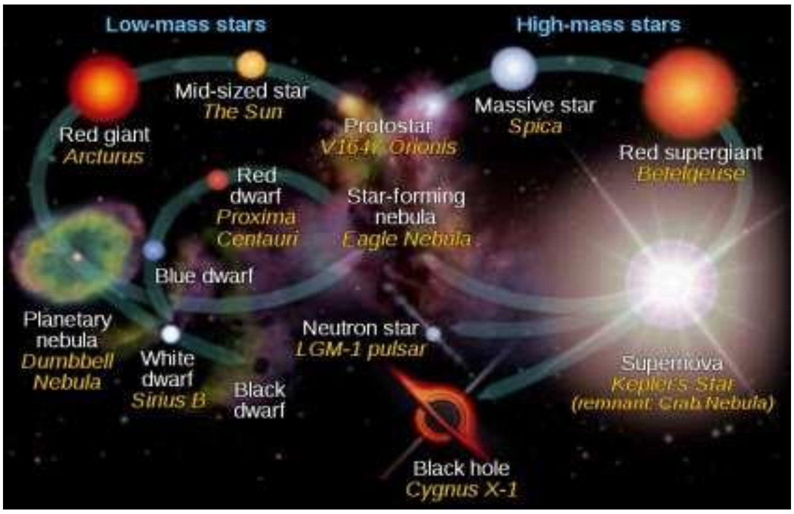


Fig. 10.4
An overview of stellar evolution
en.wikipedia.org/wiki/Stellar_evolution

At the very end of a star's life there sometimes are violent explosions. These are discussed in section 10.2.

Seismology can also be carried out on the sun. Binney says that the models agree well with the results of seismology-basically, but that there are significant deviations. It could be that the course of the nuclear reactions is not modeled quite correctly.[90] It could be that the opacity of the solar matter is not correctly estimated. And it is possible that dark matter is involved in an energy transport from the inside to the outside.

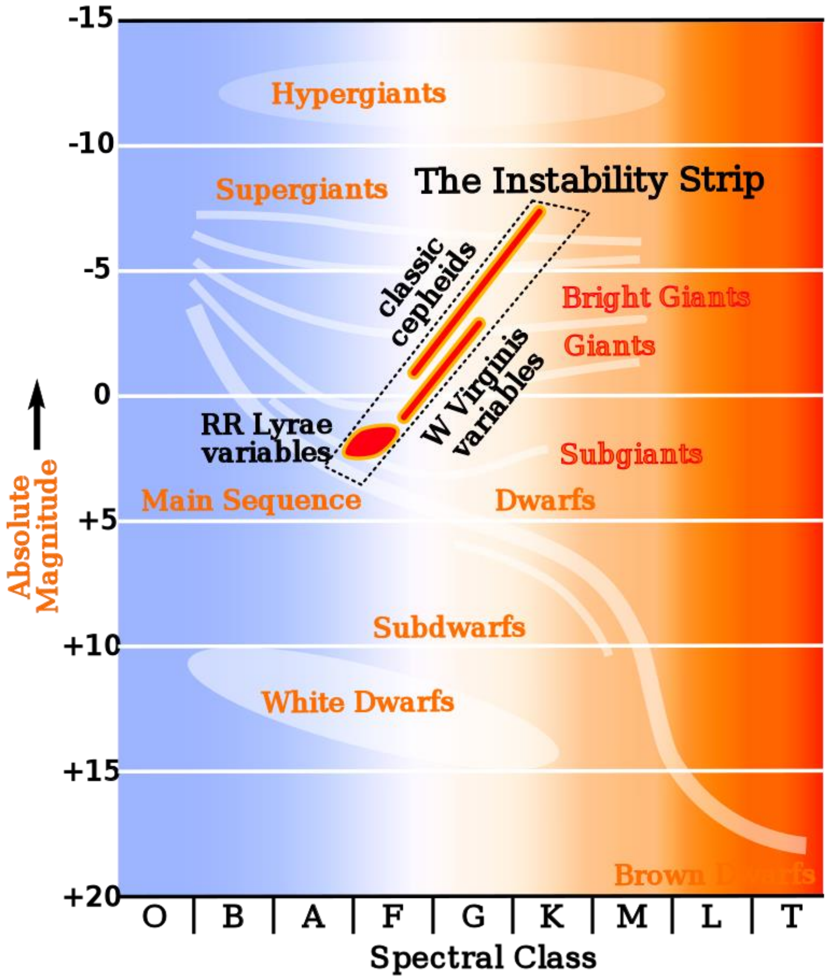


Fig. 10.5
A schematic Hertzsprung-Russell diagram with the instability strip.
The Hertzsprung-Russell diagram is similar to the temperature-luminosity diagram, whereby the colors are replaced by the "spectral classes". Here the stars are arranged according to spectra.
en.wikipedia.org/wiki/Instability strips

Remember:

- The most important parameter of an ordinary star is its mass. Heavy stars are bright, shine in blue, and only live for a short time (a few million years).
- Another parameter (of minor importance) is metallicity, i.e. the fraction of heavy elements. Very old stars have low metallicity.
- The ordinary stars form the main series in the color-luminosity diagram.
- The nuclear reactions in stars are well understood. The internal structure of stars (with or without convection zones) are understood, as well.
- There are various causes of variable brightness, including helium flashes and accretion disks of matter flowing in from a second star.
- Stars have an extremely hot corona.
- Towards the end of their lives, shell burning turns the stars into red giants.

10.2 Novae and supernovae

A nova often is at the end of a giant or supergiant. In ordinary stars, the reaction region is limited on the outside by the overlying matter. When the shell burning reaches the outer area of the star, the reactor no longer has a wall, so to speak, and the reaction mixture expands. The temperature rises enormously because the reaction heat can no longer be dissipated. The reactions then take place away from equilibrium. In particular, elements heavier than iron are formed. The s-process takes place, where "s" stands

[90] The rates of the reactions are known from laboratory experiments. They are also referred to as "reaction cross-sections". The rates are converted into reaction cross-sections by postulating hypothetical particles with a certain size of shadow cast, which always react strictly according to the relevant reaction when a collision occurs.

for "slow" (section 11.1). A novas end in "planetary nebulae", i.e. extended celestial objects that evidence the previous explosion (left in Fig. 10.4). The name is historical; the planetary nebulae have nothing to do with planets.

Supernovae are more violent than novae. They come in two forms. In type I supernovae, a neutron star has a red giant as a companion. This is not unusual insofar, as binary stars are common and every heavy star eventually becomes a red giant. The red giant transfers material to the neutron star. This material first orbits the neutron star in the form of an accretion disk (Fig. 10.6), but ultimately collapses onto the neutron star. Due to the high pressure on the neutron star, this material is degenerate. Further is is also combustible. Because of the degeneracy, a negative feedback loop which lets the fire burn steadily in normal stars is eliminated. This mechanism has already been described in the context of helium flashes. Type I supernovae are caused by thermonuclear reactions. They are also referred to as "deflagration supernovae". This instability occurs at a fairly well-defined threshold. This is why type I supernovae all have the same brightness. The apparent brightness on Earth can then be used to deduce the distance. Type I supernovae are among the standard candles used to anchor the cosmic distance scale.

Type II supernovae follow from a different mechanism, which is related to degenerate of matter on the central star (Section 5). When the pressure in a white dwarf increases to the extent that the particles become relativistic, an implosion occurs. What happens next depends on the mass. If the mass is sufficiently large (> 0.75 $M_{\odot}$), a black hole is formed.[91] With a slightly lower mass, however, the neutrons produced in the reaction $p^{+} + e^{-} \rightarrow n + \nu_e$ again form a degenerate matter. The implosion comes to a standstill. The temperature rises enormously and produces an expansion, which ejects large quantities of material and leaves a neutron star behind.

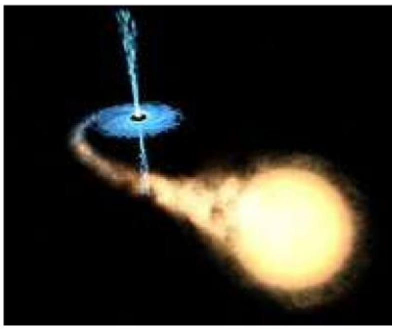


Fig. 10.6
On the accretion of matter ejected by a red giant onto a neighboring binary star
simple.wikipedia.org/wiki/Accretion_disk

The neutrinos from the fusion of electrons and protons are so numerous that they carry a significant fraction of the energyreleased by the supernova. They are so numerous that the supernova 1987A was detected on Earth with neutrino detectors (11 events in 13 seconds in the Kamiokande detector). Finally, they are so numerous that they exert an influence on the expanding shell despite the extremely rare collisions. The neutrinos carry 99% of the energy in the supernova. The details are complicated and the models have been revised several times in recent years in the light of simulations with high-performance computers. In particular, spherical symmetry is not always given. This can be seen from the fact that neutron stars often move quite quickly. The momentum is the recoil from the somewhat unbalanced explosion (Section 7.4).

During a supernova, an energy of 10^{44} J = 10^{51} erg is released over the course of a few seconds. This energy is also called "foe" (for "fiftyone erg"). The sun generates approximately this amount of energy over its entire lifetime. For a short time, a supernova shines as brightly as an entire galaxy.

<u>Remember:</u>
- Novae are explosions at the end of a star's life in which the shell burning ends in an expansion.
- Type I supernovae are thermonuclear. A companion transfers material to a neutron star until it ignites. It ignites violently because temperature and pressure are decoupled due to the degeneracy.

[91] This maximum mass of the neutron star is the analog of the Chandrasekhar limit for white dwarfs. It is called the "Tolman-Oppenheimer-Volkoff limit". The numerical value is somewhat uncertain because the interior of the neutron star is poorly understood (quark-gluon plasma? strange quarks?).

- Type II supernovae follow an implosion of a white dwarf when a neutron star is formed (instead of a black hole).
- Elements heavier than iron are mainly produced in novae and supernovae.

10.3 Side remarks:gamma ray bursts

From time to time, extreme amounts of energy are released in space over short periods of time. This refers to both the absolute energy and the energy of individual particles. Binney writes:

In 1963 the UK, the USA, and the Soviet Union signed a treaty banning tests in the atmosphere of nuclear devices. Neither side trusted the other the USA and the USSR launched top-secret satellites that would detect gamma rays emitted by elicit tests. To everyone's surprise many bursts of gamma rays were detected.

The bursts lasted from seconds to a minute, and they occurred too often to be plausibly generated by nuclear devices. After the military experts on both sides had puzzled over the data in secret, each side learnt that the other saw these events, and it became clear that the sources were astronomical. In 1973 the data were made public and it was the turn of the astronomers to be puzzled. The events seemed to be uniformly distributed over the sky, which indicated that their 188 sources were not associated with stars in our Galaxy as most X-ray sources are. The sources had to be either within ~ 0.1 kpc of the Sun or spread through a volume much bigger than our Galaxy. But the timescales of the sources were much too short for them to be associated with active galaxies, and nobody could come up with a credible source close to the Sun. In 1986 Bohdan Paczynski had the courage to posit that, despite their small timescales, they are at cosmological distances, and probably associated with some kind of exploding star. In 1997 this conjecture was proved correct when the William Herschel telescope took photographs of the region around a burst that had just been detected, and the rapidly fading optical after-glow of the event was seen in a distant galaxy. Since then optical afterglows have been routinely detected, and we have optical spectra of the underlying objects. These data establish that many gamma ray bursts are indeed associated with exploding stars. It has also emerged that there is more than one kind of source of gamma ray bursts, and our understanding of these objects is incomplete.

Two more comments on gamma ray bursts, not covered by:

– The integral energies of the event become somewhat smaller if one assumes that the gamma rays are not emitted isotropically but directionally, i.e. if one postulates a connection with jets. Even then, the integral energies released in these events remain extremely high.
– Gamma ray bursts would certainly be dangerous if they originated from a star close to Earth. In particular, there is a hypothesis that the "late Ordovician mass extinction mass event" 443 million years ago was triggered by a gamma ray burst. The gamma radiation would have destroyed the ozone layer, with various consequences not only in the form of increased UV pollution on the ground, but also in the form of climate change. Wikipedia describes this scenario as the subject of a "minority hypothesis".

10.4 Side remarks: Cosmic radiation with extremely high energy

High-energy particles are detected with the air-shower experiments (Fig. 10.7). The record holder had an energy of $2 \cdot 10^{20}$ eV.[92] Propabilitydistributions as a function of energy in Fig. 10.8 are shown. A possible mechanism for generation mechanism is shown in Fig. 10.9.

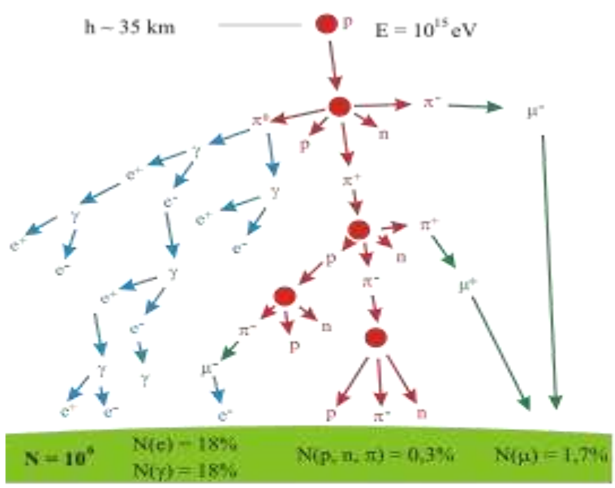


Fig. 10.7
In the context of "air-showers", high-energy particles from cosmic rays become pancake-like formations of secondary particles, which emit Cherenkov radiation and thus become optically visible. Cherenkov radiation is produced when a particle is faster than light waves in the medium in question. The medium must have a refractive index greater than one. (This is the case in air.)
en.wikipedia.org/wiki/Air_shower_(physics)

Box 10.2: Particles with ultra-high energy

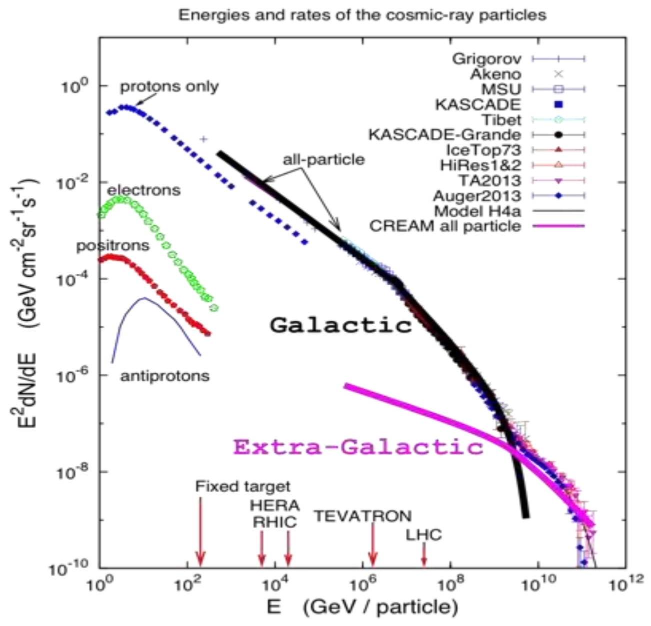


Fig. 10.8
Probability distribution of UHECRs (ultra-high-energy cosmic rays**)**

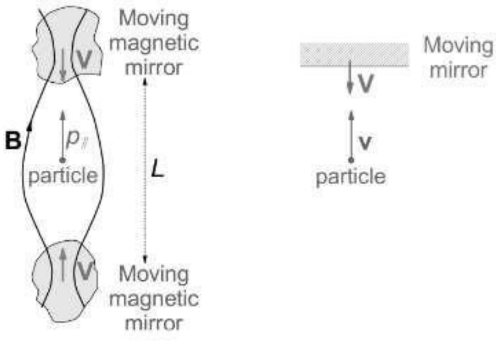


Fig. 10.9
Fermi acceleration of a particle trapped between to magnetic mirrors (left) and the equivalent acceleration of a particle upon head-on reflection by a moving mirror.

The processi is a non-thermal process. Electromagnetic fields are central (just like in particle accelerators).
*lesia.obspm.fr/perso/nicolemeyer/ Master/AccelerationParticulesp419- 450.***pdf**

[92] For comparison: the Large Hadron Collider (LHC) in Geneva generates energies of up to 10^{14} eV.

10.5 Today, stars are mostly formed inside molecular clouds

Molecular clouds are complex structures that are not fully described by the adjective "molecular". (Molecules and their observation using rotational spectra are discussed in section 13.3.) Molecular clouds also contain dust and they are the cradle of young stars ("stellar nurseries").

At the beginning of a molecular cloud there is some kind of condensation of matter. The condensation can be caused by gravitation. Condensation also occurs when two galaxies collide and the clouds of gas they contain collide, as well. Instabilities related to the fact that the galaxy rotates faster on the inside than on the outside also contributes to the formation of molecular clouds.

Star formation in molecular clouds is depicted in Fig. 10.10 in more detail. The large, hot and short-lived stars emit UV light, which ionizes the surrounding hydrogen leading to so-called H II regions.[93] A few million years later, the large stars eject heavy elements. These winds, the supernovae and the radiation pressure finally disperse the molecular cloud. However, because the material (the dust, in particular) has to go somewhere, new molecular clouds form in the neighborhood. The spiral arms glow due to such cascades.

The time period from the formation of the first star to the end of star formation in the respective region is a few tens of millions of years, which is a rather short period on a cosmic time scale. The process ends with an "open" cluster of young stars. The Pleiades are such an open cluster.

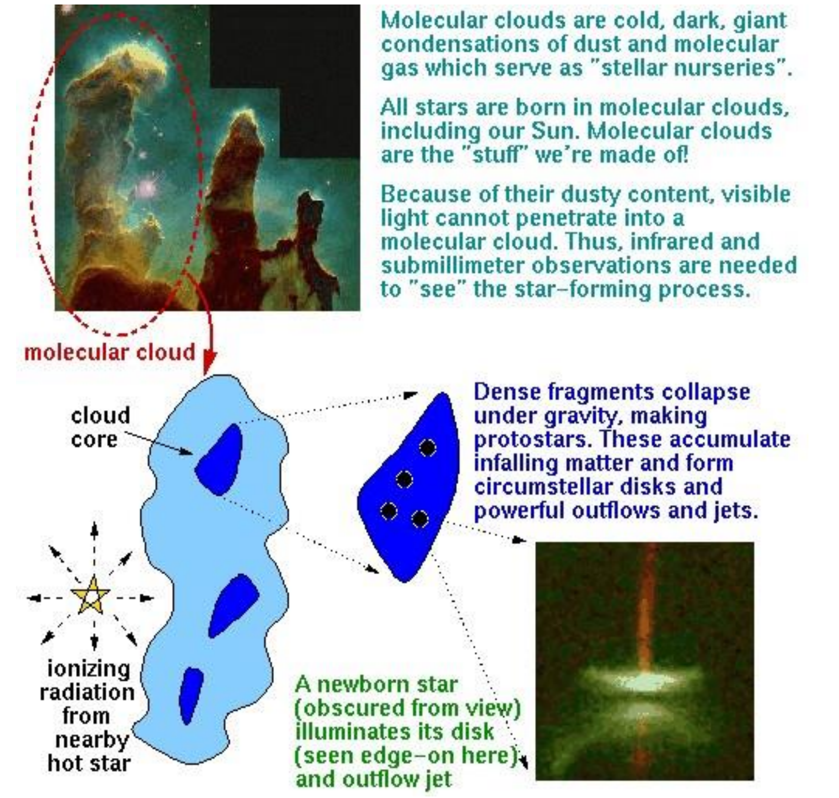


Fig. 10.10
On star formation in molecular clouds
loke.as.arizona.edu/~ckulesa/research/overview.html

Dust plays an important role. Its temperature is around 50 K. Dust heats up by absorption, where even a single absorbed photon causes the temperature to rise substantially due to the low thermal mass. Following to the Stefan-Boltzmann law, the effectiveness with which a warm body radiates its heat decreases sharply as the temperature falls. A temperature of 50 K creates a largely stationary situation, but not an equilibrium. The dust absorbs light and thus lets of the cloud appear as dark.

The gravitational collapse of gas, dust and stars sometimes leads to a disk rather than a sphere,[94] which is stabilized by the angular momentum and the resulting centrifugal force. A disk of this type can be seen at the bottom right in Fig. 10.10. The disks can be seen in emission in the infrared. Luminous

93 "HII" stands for singly ionized hydrogen. ("III" would stand for doubly ionized, which is of course only possible for elements with $Z \geq 2$). HII regions glow in a characteristic red, caused by an emission line at λ = 656.3 nm, the so-called H-α line from the Balmer series.

94 The exceptions are elliptical galaxies and globular clusters. In these, the friction is not large enough.

dust clouds were among the most impressive sources found by the first space-based infrared telescopes. Protostars are also characterized by jets that are ejected perpendicular to protoplanetary disks (also shown at the bottom right of Fig. 10.10).

Remember:
- Molecular clouds contain gas, dust and young stars.
- Molecular clouds are star-forming regions.
- Molecular clouds only exist for a few tens of millions of years.
- The spiral arms of spiral galaxies glow due to the star-forming regions they contain. The HII regions glow.

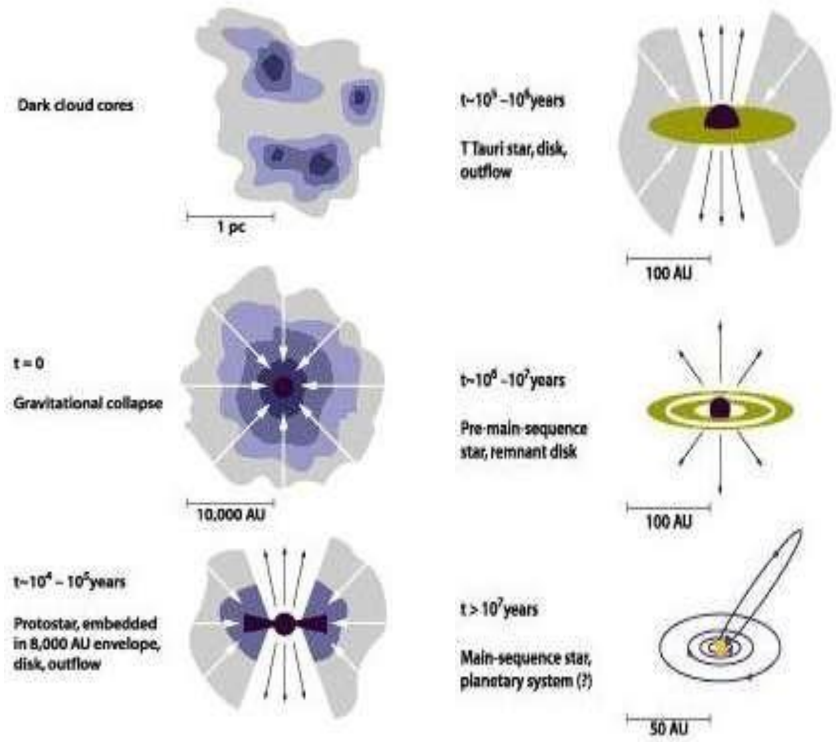


Fig. 10.11
The process of star formation as a large cloud of cold molecular gas and dust is transformed into a much smaller planetary system over millions of years. Water exists in the gas phase in the dark cloud cores and during gravitational collapse, though water and other molecules begin to freeze out onto dust grains in the dense center of the collapsing core. After 10^5 – 10^6 years, the protostar emerges, and water frozen onto dust is released along with the molecules that were formed by reactions on dust grains during the collapse phase. These molecules react in the hot gas at this stage to form prebiotic molecules such as amino acids. By the T Tauri phase, water is present either in the inner disk as vapor or as ice far from the star in the remnant envelope or the outer disk.
(Courtesy NASA/JPL-Caltech)
www.ifa.hawaii.edu/UHNAI/article5.htm

10.6 Side remarks: Photon gases

The following section is based on the Ref. 94 . The term "photon gas" exists, but some differences to particle gases should be kept in mind:

– Photons can be generated and destroyed. The number of photons is a function of temperature and volume.
– Photons do not scatter in a vacuum.[95] A thermal equilibrium can only be reached if it is mediated by sufficiently opaque matter (from which the do photons scatter). This matter may consist of hot walls. It can also consist of a plasma. Neutral gases usually are too transparent.

[95] An exception is conceivable at very high energies. Photons can polarize the vacuum. If an intense laser beam generates such a polarization, a second beam can be scattered off this induced polarization. The effect has not yet been observed experimentally.

Photons do not have a rest mass, but they do have an energy and a momentum. The relation $E = h\nu = \hbar\omega$, $p = h/\lambda = \hbar k$, and $E = cp$ apply. Photons can exert pressure on a wall, as is known from light mills (Fig. 10.12).[96]

The relationship between energy and momentum ($E = cp$ instead of $E = p^2/(2m)$, the latter being the relationship for the monatomic gas) is the first peculiarity. The second essential feature is that the number of particles is not constant. Photons can be created and destroyed. The number of particles is a function of temperature and volume, which cannot be set independently. Therefore, the definition of a chemical potential ($\mu := (\mathrm{d}G/\mathrm{d}n)_{p,T}$) does not make sense for the photon gas.

At this point, some relations need to be accepted without proof:

Glg. 10.1

$$U(V,T) = bVT^4 \qquad a)$$

$$p(T) = \frac{b}{3}T^4 \qquad b)$$

$$b = \frac{8\pi^5 k_B^4}{15h^3c^3} = 7.56\times10^{-16}\ \frac{\mathrm{J}}{\mathrm{K}^4\mathrm{m}^3}$$

Since the constant b (the "Stefan-Boltzmann constant") contains Planck's quantum of action, h, quantum mechanics must be involved in its derivation. We justify the T^4-dependence of the energy based on blackbody radiation. (The formula for the pressure in b) is justified in a similar way).

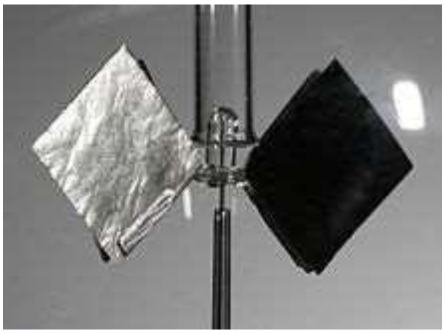


Fig. 10.12
The impeller of a light mill. On the left the unblackened side of a mica plate, on the right the blackened side.
en.wikipedia.org/wiki/Lightmill

The spectrum of the black body is universal, i.e. independent of material properties. The spectrum of the CMB is well described by the Planck curve, see Fig. 4.1. The energy density of the photon field as a function of frequency is given by

Glg. 10.2

$$U(\nu,T)\mathrm{d}\nu = 2\,4\pi k^2 h\nu \frac{1}{\exp\left(\dfrac{h\nu}{k_BT}\right)-1}\mathrm{d}\nu = 8\pi\frac{h\nu^3}{c^3}\frac{1}{\exp\left(\dfrac{h\nu}{k_BT}\right)-1}\mathrm{d}\nu$$

We do not prove Glg. 10.2, but make the factors plausible. The factor 2 on the left results from the two polarizations. The factor $4\pi k^2$ is the surface area of a sphere in k-space. $h\nu$ is the energy of each photon. The factor $1/(\exp(h\nu/(k_BT))-1)$ results from the Bose-Einstein distribution. Note: This factor is different from the Boltzmann factor. The latter would be $\exp(-h\nu/(k_BT)) = 1/\exp(h\nu/(k_BT))$.

The Planck distribution applies not only to photons (light waves) but also to phonons (sound waves), as long as the phonon wavelength is large compared to the distance between atoms. (There is a complication because longitudinal and transverse sound have different velocities).

96 Note: A vacuum must be created so that the photon pressure drives the light mill. Otherwise, a different mechanism is stronger than the pressure exerted by the photons. The air in front of the black surfaces can also heat up, expand and thus exert a force on this surface.

The total energy is proportional to T^4 . This is the content of the Stefan-Boltzmann law. The T^4 dependence can be understood from Glg. 10.2 by replacing the variable ν with the variable $\nu' = h\nu/k_B$ T. With $\nu = k_B T\nu'/h$ one obtains

Glg. 10.3

$$E_{tot} = \int_0^\infty 8\pi \frac{h\nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right)-1} \mathrm{d}\nu = \int_0^\infty 8\pi \frac{h}{c^3} \left(\frac{k_B T \nu'}{h}\right)^3 \frac{1}{\exp(\nu')-1} \mathrm{d}\left(\frac{k_B T \nu'}{h}\right)$$

$$= 8\pi \frac{h}{c^3} \left(\frac{k_B T}{h}\right)^4 \int_0^\infty \nu'^3 \frac{1}{\exp(\nu')-1} \mathrm{d}\nu'$$

Since the integral does not depend on the temperature, the total energy scales as T^4 (see Glg. 10.1a). The value of the integral is $\pi^4/15$. If the internal energy of a solid is mainly contained in the thermally excited sound waves, the derivative of the internal energy with respect to temperature (the specific heat capacity) is proportional to T^3. This is the Debye law. C_V tends towards zero with decreasing temperature as T^3.

The Planck distribution itself is not needed in the following. The Planck distribution is important insofar, as various integrals, whose intergrands contain the Planck distribution as a weight function, lead to the energy, the number of particles, and other properties of the photon gas (see e.g. Glg. 10.3). We continue with some rather elementary thermodynamic considerations.

Based on the relation for the pressure (Glg. 10.1b), the expansion in ΔV for the work

Glg. 10.4

$$W = -\frac{b}{3} T^4 \Delta V$$

The internal energy changes by

Glg. 10.5

$$\Delta U = bT^4 \Delta V$$

Because $\Delta U = Q + W$, Q obeys the relation

Glg. 10.6

$$Q = \frac{4}{3} bT^4 V$$

With $\Delta S = Q/T$ we get

Glg. 10.7

$$\Delta S = \frac{4}{3} bT^3 \Delta V$$

If zero volume is assigned a vanishing entropy, the following applies to entropy as an absolute quantity

Glg. 10.8

$$S = \frac{4}{3} b T^3 \Delta V$$

If the entropy remains constant (as in adiabatic processes), it follows that

Glg. 10.9

$$T^3 V = const$$
$$pV^{4/3} = const$$

The photon gas therefore has an adiabatic exponent of 4/3, just like relativistically degenerate matter (section 5). Note that c_p does not exist for the photon gas because the temperature cannot be varied at constant pressure. An equation of the form $\gamma = c_p/c_V$ would not make sense for the photon gas.

In Fig. 10.13 the differences between the photon gas and the ideal gas are tabulated. If you insert numbers, you get an energy per photon of 2.7 $k_B T$ (ideal monatomic gas: 1.5 $k_B T$), a pressure of 0.9 $k_B TN/V$ (ideal gas: $k_B TN/V$) and an entropy per photon of 3.6 k_B (no comparable simple relation for the particle gas[97]).

Table III. Numerical comparison of classical ideal and photon gas functions. Here the ideal gas is 1.00 mol of monatomic argon at $P = 1.01 \times 10^5$ Pa, $V = 2.47 \times 10^{-2}$ m^3, and $T = 300$ K.

Function	Classical ideal gas	Photon gas
N	6.02×10^{23} atoms	1.35×10^{13} photons
U	3.74×10^3 J	1.51×10^{-7} J
P	1.01×10^5 Pa	2.04×10^{-6} Pa
S	155 J/K	6.71×10^{-10} J/K

Fig. 10.13
Some numerical values for comparing photon gas and ideal gas. The pressure of the photon gas at room temperature is negligible compared to the pressure of the molecular gas. This is the case even in the center of the sun ($T \approx 15 \cdot 10^6$ K). Only in the center of large and bright stars does the photon gas make a significant contribution to the pressure. The temperature there is around 10^9 K.
Source: Footnote 52

Examples of the importance of the photon gas are the star-forming regions and the baryonic acoustic oscillations, BAOs, the latter associated with the galactic voids (Box 4.2). Star formation phases are processes in the molecular clouds is limited in time. As soon as the stars have formed, the photon gas disperses the ordinary gas and the dust. It "blows" this matter away, so to speak. The intergalactic bubbles (voids) are large, roughly spherical spaces in the universe that are almost empty. These voids were created in the early phase of the universe. The material in the clouds heated up (following the virial theorem, section 4.4.2) and a photon gas developed. Because photons do not feel gravity, the photon gas expands. Because the matter is ionized, the photons scatter off the ions and drag the plasma with them. The plasma forms an expanding shell and initially leaves the dark matter behind. At the moment, when the ions and electrons recombine to form atoms, the photon gas decouples from the ordinary matter. The ordinary matter ("baryonic matter") continues to expand until neighboring expanding shells collide and form large-scale flat structures or filaments. These are the regions in which many galaxies can be found today.

Because the pressure of the photon gas is proportional to T^4, it usually plays no role at room temperature compared to the pressure of the particle gas. Otherwise, the ideal gas law would have to be modified accordingly. See the table in Fig. 10.13. Even in the center of the sun ($T \approx 15 \cdot 10^6$ K), the photon pressure is rather insignificant. It only plays a role in very hot stars.

---

[97] The Sackur-Tetrode equation applies to the ideal gas. For solids, the entropy is again roughly proportional to the number of atoms. This is a consequence of Dulong and Petit's rule.

Remember:

– Photons have an energy and a momentum, but no mass.
– Because photons do not scatter from other photons, a thermal equilibrium (resulting in a blackbody spectrum) only forms in contact with matter that has a similar temperature.
– In contrast to particle gases, the number of photons in photon gases is a function of volume and temperature.
– The pressure and energy of the photon gas scale as T^4. In stars, the pressure of the photon gas is only significant at high temperatures (only in the cores of heavy stars).
– The photon gas and the stellar winds blow molecular clouds apart.
– Photon pressure created the voids in the early universe (Box 4.2).
– The photon gas has an adiabatic exponent of 4/3.
– The photon gas has an entropy per photon of 3.6 k_B.

11 Nucleosynthesis

11.1 *Origin of the elements*

Heavier elements account for about 2 % of the baryonic matter. Among them, elements with an even number of nucleons (C, O, Si, ...) are more common than those with an odd number of nucleons (N, P, Cl, ...). Elements heavier than iron are rare because their binding energy is lower than that of iron. They are mainly produced in the explosions at the end of a star's life. Iron is more common than the elements next to it in the periodic table because it has the highest binding energy of all nuclei. However, iron is also rarer than C, O and Si. Fig. 11.1 shows the periodic table with color coding for the synthesis mechanism. Fig. 11.2 shows the relative abundance of elements.

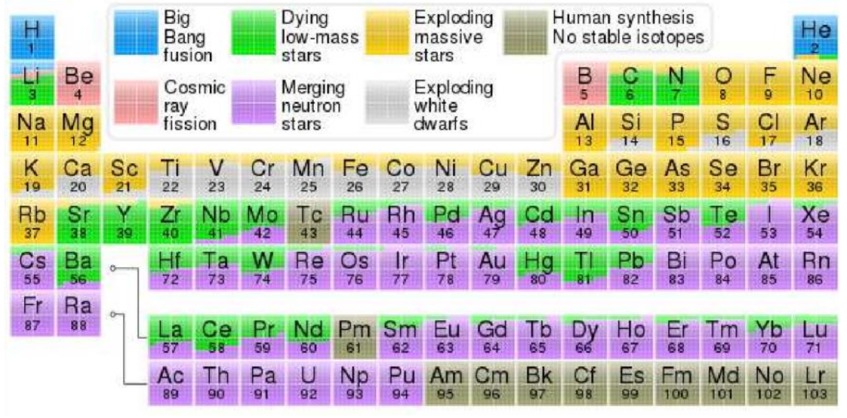


Fig. 11.1
List of elements with the nucleosynthesis pathways
en.wikipedia.org/wiki/Nucleosynthesis

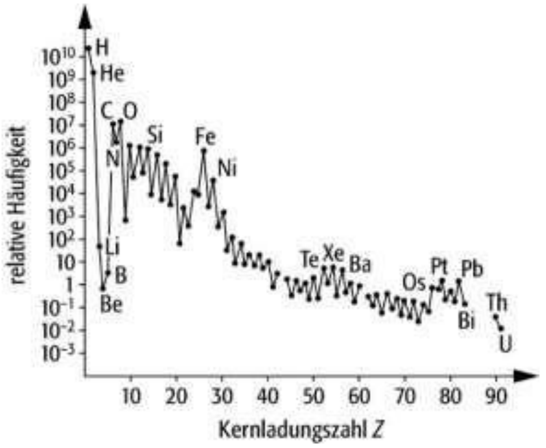


Fig. 11.2
The mean cosmic element abundance.
www.spektrum.de/lexikon/physik/elementhaeufigkeit/4241

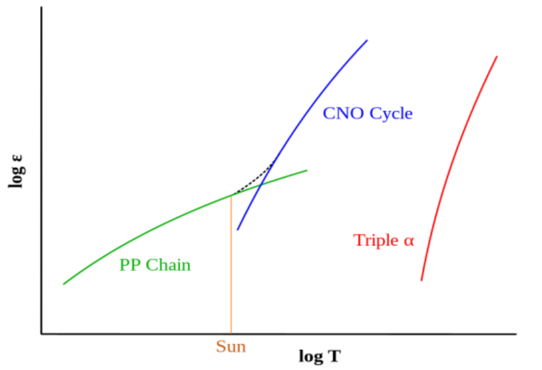


Fig. 11.3
Logarithm of the relative energy output (ε) of proton-proton (PP), CNO and Triple-α fusion processes at different temperatures. The dashed line shows the combined energy generation of the PP and CNO processes within a star. At the Sun's core temperature, the PP process is more efficient.
en.wikipedia.org/wiki/Stellar_nucleosynthesis

Wikipedia lists the following mechanisms of nuclear synthesis:[98]

- Big bang nucleosynthesis
- Stellar nucleosynthesis
- Explosive nucleosynthesis
- Neutron star collisions

[98] *en.wikipedia.org/wiki/Nucleosynthesis*

– Black hole accretion disk nucleosynthesis
– Cosmic ray spallation

We discuss the first three entries, only.

Primordial nucleosynthesis is the formation of atomic nuclei with more than one nucleon shortly after the Big Bang. Deuterium, helium and traces of lithium are formed. The heavier elements that can be observed today originate from fusion and other nuclear reactions in stars and thus from at later times. The relative abundances of deuterium and^3 He are very well explained by the theory. For lithium, however, there is a deviation between the measured value and the theoretically calculated value that. The latter is larger by a factor of 3. This is known as the primordial lithium problem.[99]

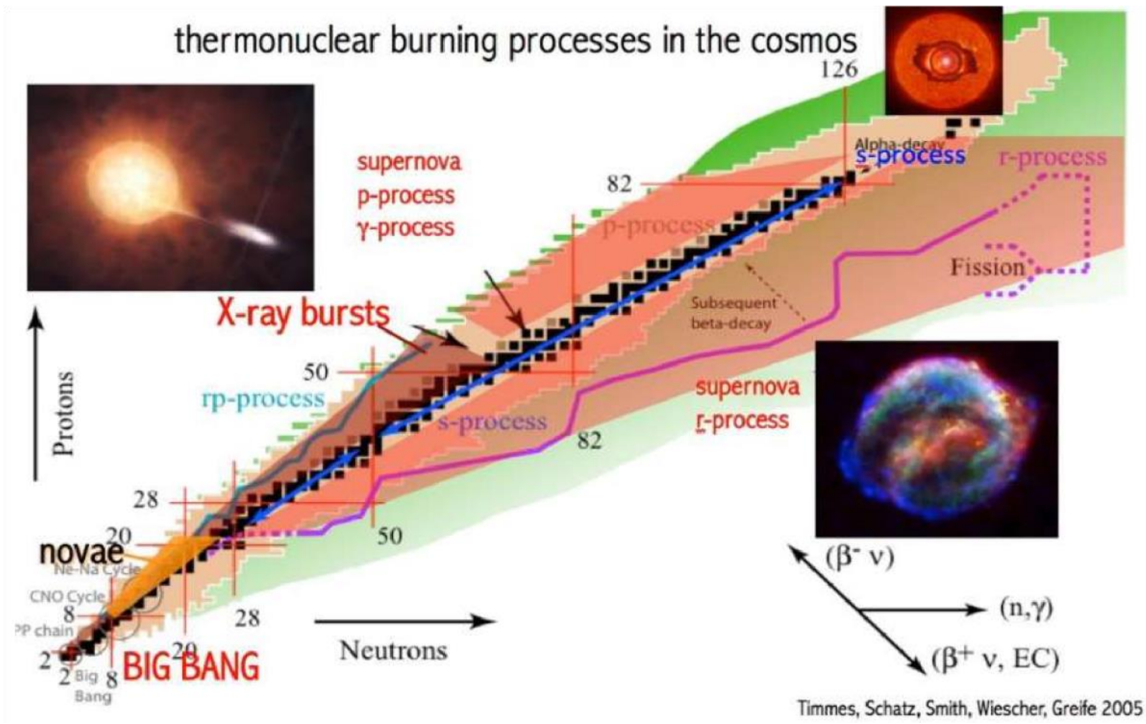


Fig. 11.4
For r-process and s-process
www.asc.ohio-state.edu/physics/ntg/6805/slides/rprocess.pdf

Fig. 11.3 shows the efficiency of three stellar nucleosynthesis processes. Note that the 3-α- process requires a 3-particle-collision of helium nuclei. Such events are inefficient. The 3-particle collision is necessary because ^{8}Be (in principle conceivable as the product of a 2-particle-collision of He nuclei) immediately decays into two helium nuclei. The fact that the 3-α process is so efficient has to do with details of nuclear physics. There is an excited state of ^{12}C whose energy is just right (a "resonance"). Without this resonance, there would be much less carbon in the universe and probably no life, either.

[99] This section is based on *en.wikipedia*.org/wiki/Primordial_Nucleosynthesis

At extremely high temperatures, elements heavier than iron are also formed. If $k_B T$ is much larger than the binding energies, a large number of nuclear reactions take place (more or less energetically favorable). The entropy favors the formation of many different elements. Many unstable nuclei are also forme, as well. After cooling, the metastable nuclei remain. This includes lead and uranium.

A distinction is made between the s-process (at the end of the star's life, novae, lasting up to a few thousand years) and the r-process (in type II supernovae). The difference between the two is illustrated by Fig. 11.4. On the nuclide map, the nuclei form an energy landscape. An elongated valley also runs through this map. For the light elements, $A \approx 2Z$ (where A is the number of nucleons and Z is the charge). The nuclei contain approximately the same number of neutrons and protons. This is a consequence of the strong interaction. Heavy nuclei contain more neutrons than protons because the protons repel each other electrically, whereas the neutrons do not. The lowest point in the valley is iron.

The s-process runs along the trough. Fusion processes repeatedly lead to nuclei outside the trough, but then β decays or inverse β decays (i.e. conversions between neutrons and protons) lead these objects back to the trough. β decays are mediated by the weak interaction and are therefore relatively slow. The s-process ("s" for slow) allows time for these β decays.

The r-process ("r" for rapid) is faster. It is favored when there is a high neutron excess. Neutrons are present in excess in supernovae of type II and in collisions of neutron stars. Nuclei are then formed, that lie far below the valley of stability on the nuclide map. Only after the many neutron captures are the neutrons converted into protons in a chain of β δεχαψσ, resulting in stable nuclei.

12 Special features of spectroscopy in astrophysics

12.1 Forbidden lines

Text adapted from en.wikipedia.*org/wiki/Nebulium*

Sometimes, lines are seen in the sky that are not known form the laboratory because collional dexcitation does not compete with radiative deexcitation in these highly rarefied media. Because these lines are "forbidden", deexcitation is slow (but not completely suppressed).

The best-known example is “nebulium”. In 1864, William Huggins discovered bright lines with wavelengths of 3726, 3729, 4959 and 5007 Å in the planetary nebula NGC 6543 (Cat's Eye Nebula, constellation Dragon). These produce the characteristic green appearance of this nebula. He suggested an element not yet found on Earth, the "nebulium", as a possible explanation. It was not until 1927 that Ira S. Bowen identified the supposed nebulium lines as forbidden lines of twice ionized oxygen. These forbidden lines can only occur at the very low density of gas nebulae.

12.2 CO rotation temperature

The rotational temperature of CO is discussed in Box 12.1.

12.3 Cosmic masers

Most remarkable about cosmic masers is their existence as such. As a reminder, the word "maser" stands for "microwave amplification by stimulated emission of radiation." Stimulated emission requires a population inversion. Population inversion requires a non-equilibrium. In equilibrium, the populations of the states are Boltzmann-distributed and the lower states are more populated than the higher states. Absorption is then more likely than stimulated emission.

Are there cosmic **l**asers in the universe?[100] No, these do not exist. Stimulated emission is always in competition with spontaneous emission. If an initial level for a possible stimulated emission has been populated by some kind of pumping process, the molecule is int the excited state for a limited time. If a stimulating photon passes by during this time: good. Otherwise: just as well. The molecule spontaneously thenfalls back to the lower state. For reasons that are not quickly explained, the rate of spontaneous emission has a pre-factor of ν^3 with ν the frequency. If one uses the values 10^{15} Hz and 10^9 Hz for the frequencies ν (optical frequencies and microwave frequencies), one sees that microwave-excited states live about 10^{18} times longer than optically excited states. Maser activity is less demanding than laser activity.[101] On earth, both masers and lasers have mirrors for feedback. A certain photon is repeatedly sent through the active medium. There are no such mirrors in cosmic masers. The size of the active region is so enormous that the amplification and coherence occur even without feedback.

How is maser recognized? Which molecules are active? A maser is recognized by a single strong (and coherent) line. The first of these lines (at 1.665 GHz) was discovered in 1965. The surprise was expressed by postulating a new form of interstellar matter with the name “Mysterium”. However, it was the

---

[100] "l" for "light", i.e. optical wavelengths

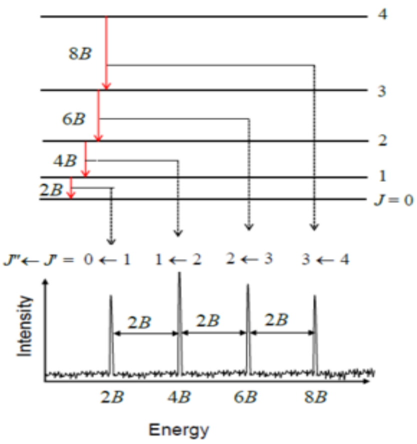
[101] The laser was also developed later on Earth than the maser. Invented in the sense that someone thought up the possibility of such coherent radiation, the maser and laser were invented at the same time. Pumping is more complex for the laser than for the maser. Pumping is particularly complex for UV lasers (required in semiconductor lithography) and for X-ray lasers. The best X-ray lasers at present are free electron lasers (FELs), e.g. in Hamburg. Extreme-UV lithography (EUV lithography) at λ = 13 nm makes use of non-coherent radiation.

OH molecule. The transition is somewhat exotic. One speaks of Λ-doubling. The nuclear spin is involved, just as with the 21-cm line. Since then, masers have also been found at many other wavelengths. Even in the atmosphere of Jupiter there is stimulated coherent emission.

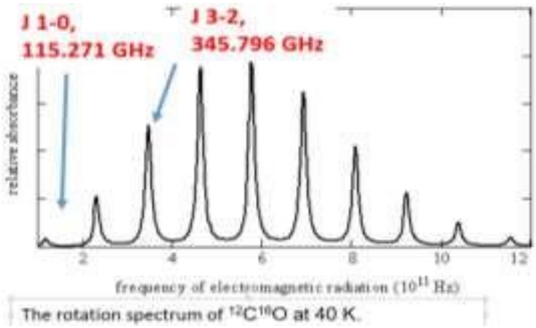
13 Gases composed of neutral atoms and molecules

13.1 The gas in the galaxies

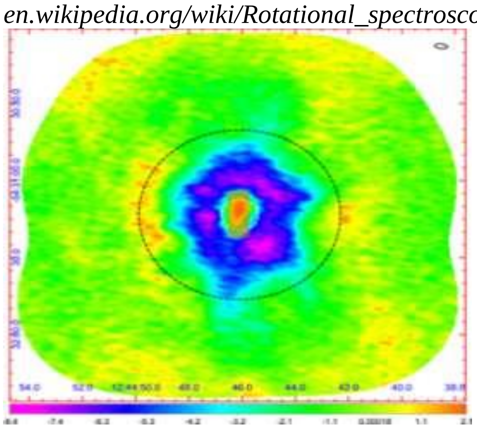
Box 12.1: Rotation temperature of CO in the boomerang nebula



en.wikipedia.org/wiki/Rotational_spectroscopy/



https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Book%3A_Quantum_States_of_Atoms_and_Molecules_(Zielinksi_et_al)/07%



Map on the J1-0 line.
The rectangle in the center approximately corresponds to the image on the right.
core.ac.uk/download/pdf/84870536.pdf

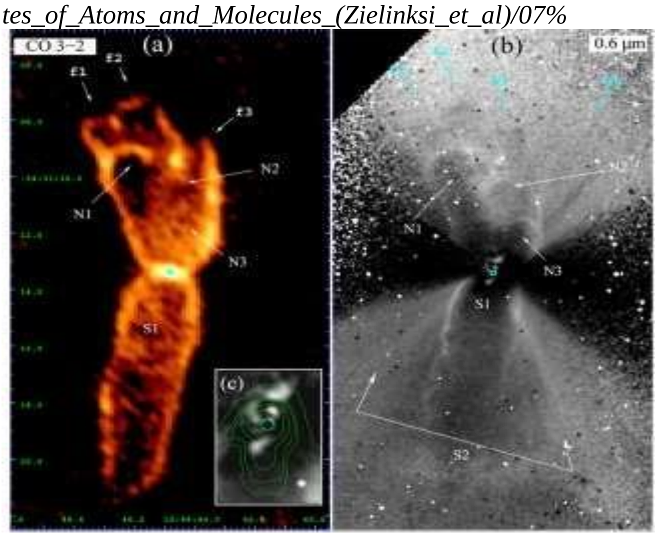
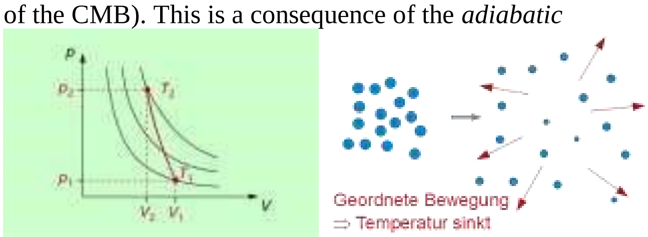


Image of the Hubble telescope overlaid with an image of the ALMA observatory map on the J 3-2 line
core.ac.uk/download/pdf/84870536.pdf

en.wikipedia.org/wiki/Atacama_Large_Millimeter/submillimeter_Array

The rotation temperature in the outer areas of the boomerang nebula is only about 1 K (i.e. lower than the temperature of the CMB). This is a consequence of the *adiabatic*



www.lernhelfer.de/schuelerlexikon/physik-abitur/article/adiabatic-state-changes

The role of gas in galaxies is quite complex. First of all, the density there is so high that the shock waves that heat up the intergalactic gas do not penetrateto the inside of the galaxy. The temperature is therefore usually so low that the gas is not ionized. For hydrogen, this can be inferred from the 21 cm line emitted by atomic hydrogen (Section 13.2). On the other hand, the gas is not so cold that molecular hydrogen (H_2) would form. H_2 (and other molecules) is only found in the molecular clouds. There, the dust determines the local temperature (~ 50 K). At the same time, the dust catalyzes the reactions to form molecules. These dust clouds are only rarely found in very old galaxies, because the elements heavier than helium are rare there. Molecular clouds, dust, and star formation always come together.

Only in the vicinity of young stars is the hydrogen ionized due to the UV-radiation from the stars. It

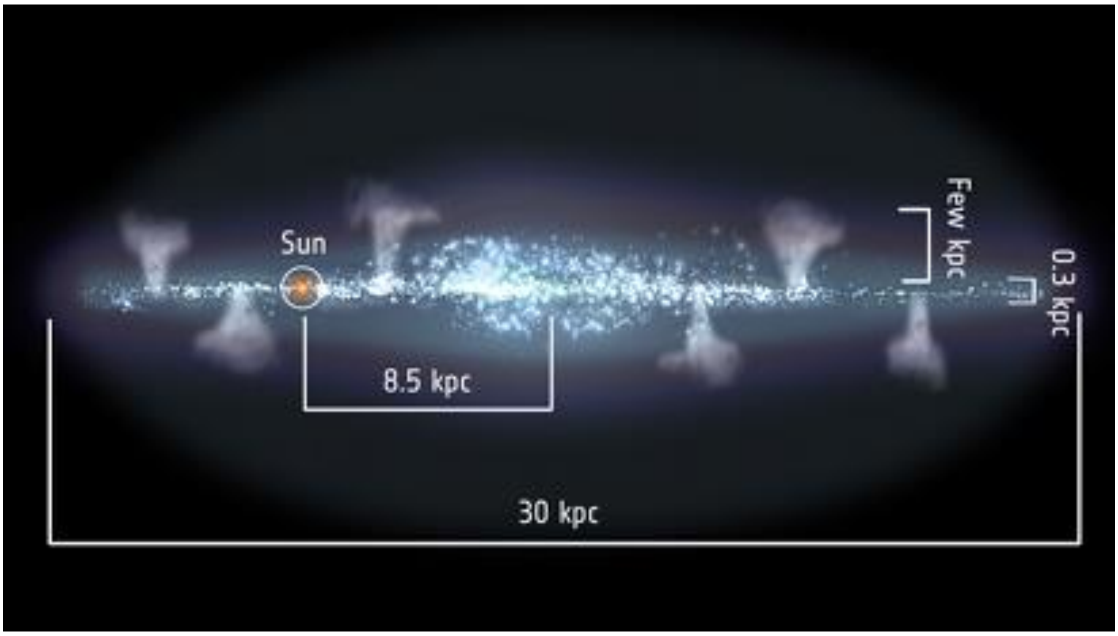


Fig. 13.1
This illustration of the Milky Way shows the galactic fountain scenario: supernova explosions in the galactic disc heat the interstellar medium and can drive hot gas out of the disc, creating so-called galactic fountains that contribute to the formation of a halo of hot gas around the Milky Way. As the gas rises to above and below the disc, reaching heights of a few kiloparsecs, it emits radiation and thus becomes cooler, condensing into clouds which then fall back into the disc, in a way that resembles a fountain.
sci.esa.int/web/xmm-newton/-/47988-illustration-of-galactic-fountains

then glows in a characteristic red color. These are referred to as the HII regions ("II" for singly ionized). HII regions are indicators of star formation, because only young stars are hot enough to produce HII (section 10.5).

The mass of the gas relative to the mass of the stars can vary between 10^{-2} and 10. Gas is required for star formation. In spiral galaxies, the gas is enriched in the spiral arms. There are shock waves (partly caused by gravitation by the central bar of the galaxy). The shock waves then trigger star formation. Because of star formation, the spiral arms shine brightly.

Sometimes hot gas is ejected up and down from the disk in star-forming regions (from novae, supernovae and hot stars). This gas is ionized. “Galactic fountains" are formed (Fig. 13.1). The gas then cools down and "rains" back onto the disk.

13.2 Side remarks: The 21-cm line

In contrast to the warm and hot intergalactic medium (WHIM), the somewhat colder, non-ionized gas of atomic hydrogen in the galaxies is very easy to observe. It can be studied using the famous 21-cm line, which is seen with radio telescopes. Atomic (i.e. non-ionized) hydrogen is mainly found in galaxies. The Doppler shift of the 21- cm line can be used to measure the rotation curves of galaxies (the angular velocity as a function of the distance from the center). The 21- cm line is worth a digression insofar as the lifetime of the excited state is 11 million years, which can be understood quite well. As a reminder, typical fluorescence lifetimes are a few nanoseconds. A lifetime of one nanosecond is considered long insofar as other electronic processes take place on the femtosecond scale, i.e. 10^6 times faster. Nuclear movements take place on the picosecond scale, so they are still much faster than fluorescence. The excited state of the 21-cm line is 10^{26} times more long-lived than the excited states of fluorescence.

Radio radiation with a wavelength of 21 cm (corresponding to 1.4 GHz) is emitted when the nuclear spin of atomic hydrogen changes from parallel to antiparallel orientation relative to the spin of the electron. Such an influence of the nuclear spin on the energy levels of hydrogen exists. It is part of the "hyperfine structure" and is not worth more than a few sentences in the conventional teaching of molecular physics. The hyperfine structure can be seen in the optical spectra using high-resolution laser spectroscopy. Back to the 21-cm line: The first question is why there is no collisional deexcitation over a period of 11 million years. Collisional deexcitation always competes with ordinary fluorescence. Collisions are rare in the interstellar medium, but the collision times of a few years are considerably shorter than the lifetime of the spin-parallel H atom. However, collisions generally have no influence on the nuclear spin. This is known from nuclear spin-polarized ^{3} Ho. This material can be stored for many hours. If you like, you can take it anywhere by car. All that is needed are containers with non-magnetic walls.

The long lifetime results from two separate facts, namely the low frequency and a selection rule. Firstly, the frequency: We quote Wikipedia for the rate of spontaneous fluorescence: [102]

The rate of spontaneous emission (i.e., the radiative rate) can be described by Fermi's golden rule. The rate of emission depends on two factors: an 'atomic part', which describes the internal structure of the light source and a 'field part', which describes the density of electromagnetic modes of the environment. The atomic part describes the strength of a transition between two states in terms of transition moments. In a homogeneous medium, such as free space, the rate of spontaneous emission in the dipole approximation is given by:

Glg. 13.1

$$\Gamma_{rad}(\omega) = \frac{\omega^3 n |\mu_{12}|^2}{3\pi\varepsilon_0 \hbar c^3} = \frac{4\alpha\omega^3 n |\mu_{12}|^2}{3c^2}, \qquad \frac{1}{\pi\varepsilon_0 \hbar c} = 4\alpha$$

where ω is the emission frequency, n is the index of refraction, μ_{12} is the transition dipole moment, ε_0 is the vacuum permittivity, $\hbar$ is the reduced Planck constant, c is the vacuum speed of light, and α is the fine structure constant. [...] The above equation clearly shows that the rate of spontaneous emission in free space increases proportionally to ω^3 .

Important to Glg. 13.1 are the frequency (to the third power) and the square of the (magnetic) transition dipole moment. For this calculation, let the fluorescence wavelength with which we are comparing be 200 nm, corresponding to $1.4 \cdot 10^{15}$ Hz (1.4 petahertz). Based on the pre-factor ω^3, we conclude that the lifetime of the spin-parallel atomic hydrogen is at least $10^{3\times 6} = 10^{18}$ times longer than the lifetime of the excited fluorescent dye. This would result in a lifetime of 10^{18} nanoseconds (i.e. 10^9 sec). However, 11 million years correspond to 2.5 10^{14} seconds: a factor of 10^5 is still missing. This factor results from

[102] *en.wikipedia*.org/wiki/Spontaneous_emission

the fact that the spin flip is a magnetic dipole transition (instead of an electric dipole transition). The rule of thumb that magnetic effects are smaller than electric effects applies here (section 16.1).

Remember

- The 21-cm line results from a flip of the nuclear spin in atomic hydrogen.
- The lifetime of the excited state is around 11 million years. The long lifetime results from the term ω^3 in the rate of spontaneous emission and also from the fact that it is a magnetic transition.
- The Doppler shift of the 21-cm line can be used to measure the rotation curves of galaxies.

13.3 Molecules in space

Molecules in space are primarily identified using microwave spectra (i.e. rotational spectra).[103] A preliminary remark is necessary here. In the course on molecules and spectroscopy, the spectrum of the rigid linear rotator is treated thoroughly. Example molecules are CO_2 , NO, CO, H_2 , N_2 and O_2. For these, the rotation around the molecular axis is an electronic excitation which is frozen at room temperature. There are only two quantum numbers (J and J_z). The energy eigenvalues are given as $hcB\,J$(J+1) with B the rotational constant. The rotational spectra consist of evenly spaced peaks. (This is so because of the selection rule $\Delta J = \pm 1$). If the molecule is not a linear rotator but is still a symmetric top (examples are methane and ammonia[104]), the situation is more complicated but still manageable even for non-experts. There then is one more quantum number K and there are still explicit formulas for the energy levels. If the molecule is no longer a symmetrical top (example: H_2 O), the energy levels must be calculated numerically. The spectrum extremely complicated.[105] Also: the nuclear spins play their part to keep things interesting. So far, microwave spectra (an example is shown in Fig. 13.2) have been used to identify around 200 molecules.[106] Due to UV-induced dissociation, it is assumed that there are not many more molecules in space. On the other hand, the assignment of certain lines or groups of lines to molecules sometimes leaves questions unanswered.

Talking of the "Number of known molecules": Polycyclic aromatic hydrocarbons (PAHs) are not counted here. *There is* a certain diversity among these molecules. The different PAHs are difficult to distinguish from each other and they tend to form clusters.

It is sometimes emphasized that amino acids can also be found in space. This does not have much to do with life. Amino acids were also found in experiments by Urey and Miller (which also – as it later turned out – had little to do with prebiotic chemistry[107]).

The chemistry in space is less rich than on Earth. This firstly, is due to the fact that there often is UV radiation in many places in space and that photodissociation therefore competes with the formation of molecules. It is also related to the fact that the formation of molecules takes place either in the gas phase or heterogeneously catalyzed on the surface of dust grains. There is no liquid-phase synthesis. More than half of the known molecules contain carbon. The metastability of the four-bond carbon plays about the same role as on Earth.

---

[103] The James Webb telescope recently detected $CH_3^{\cdot}$ with an IR spectrometer. In this study, H_2 was also observed, visible from the quadrupole transition at $\lambda = 2.12$ µm.

[104] For the symmetrical top, two of the three main moments of inertia are the same.

[105] Because the rotational spectrum of H_2O is so well populated, humid air is not transparent. Clusters of water walso play a role to IR radiation

[106] If you want to make yourself interesting at a party: There is more alcohol in the Orion Nebula than in all of Scotland. In fact, there is even more alcohol there than water on Earth.

[107] According to current opinion, life originated in hydrothermal vents in the deep sea. Lightning - as used in the experiments by Urey and Miller - played no role there.

Dust also plays a role for the molecular clouds insofar, as it brings about a largely uniform temperature of around 50 K. Due to collisions, this is also approximately the temperature of the gas.

One particular molecule deserves a, which is CO. Although CO is much rarer than H_2, it is still common enough to serve as a thermometer for gas clouds. As a reminder: H_2 has no dipole moment and therefore no rotational spectrum. CO has such a rotation spectrum and the temperature can be deduced from the relative strengths of the lines. The occupation ratios follow the Boltzmann distribution

Glg. 13.2

$$\frac{n_{J'}}{n_{J''}} = \frac{2J'+1}{2J''+1}\exp\left(-\frac{E(J') - E(J'')}{k_B T}\right)$$

$2J + 1$ is the degeneracy. The lower energy state has two primes. When converting the line widths to the temperature, there are a few subtleties to consider, which we will ignore. In Box 12.1 an example of the application of Glg. 2.4 is worked through. In the Boomerang nebula, the rotational temperature of CO is surprisingly low.

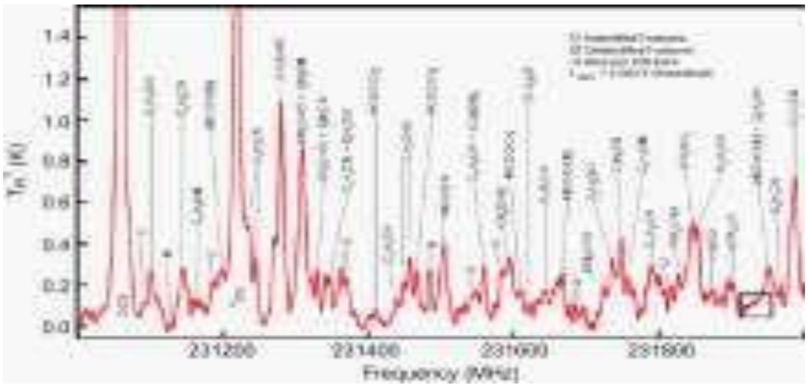


Fig. 13.2
A microwave spectrum from a molecular cloud
This λ = 1.3 mm spectrum of the molecular cloud SgrB2(N) near the Galactic center is completely dominated by molecular lines from known and unknown (U) species (Ziurys et al. 2006, NRAO Newsletter, 109, 11). More than 140 different molecules containing up to 13 atoms have been identified in space.
www.cv.nrao.edu/course/astr534/MolecularSpectra.html

Meteorites provide a further access of chemistry in space. The carbonaceous chondrites, i.e. meteorites with a high carbon content, stand out. Many PAHs are found. "Cosmochemistry" in this context often refers to a detailed isotope analysis (Section 14.1). There are "fingerprints". In particular, it is believed that some meteorites have found their way to us from Mars (or even from the moon). One detail: the amino acids in the Murchison meteorite are preferentially left-handed (for whatever reason[108]). There are considerably more than 200 different molecules (thousands) in meteorites.

The interstellar medium in the molecular clouds is not optically dense. There is no equilibrium between the temperature of the particles (approximately Maxwell distributed) and a temperature of the photons, whereby the latter is not well defined because there are other high-energy photons in addition to the CMB. The low density means that line radiation plays a greater role in the energy balance than on Earth. The Earth heats up in the thermal radiation field of the sun (≈ 5800 K); it cools down by emitting (by and large) thermal radiation into space (≈ 300 K). In molecular clouds, a major mechanism of heating is the absorption of Lyman-α photons from the neighboring HII regions.[109] A molecule absorbs the photon and,

[108] One can imagine that only molecules of one handedness are incorporated during crystallization. If the other form then evaporates more easily and there are only a few crystals, an asymmetry can arise by chance. The preferred handedness in the Murchison meteorite does not prove that life comes from outer space

[109] The Lyman-α transition of atomic hydrogen leads from 2p to 1s. The wavelength is 121.-567 nm. HII regions surround young hot stars. See also footnote 79.

for example, releases the energy to the gas as kinetic energy after dissociation. Among other things, cooling takes place via an emission line of CII at 157.74 μm ($^2P_{3/2} \rightarrow {}^2P_{1/2}$). After a collision, a (relatively low-lying) electronic excitation occurs. The excitation is optical and removes kinetic energy from the gas.

Remember

- About 200 molecules have currently been identified on the basis of microwave spectra. Many of them contain carbon.
- These 200 molecules do not include polyaromatic hydrocarbons (PAHs). These exist in a large variety. Characterization in detail is difficult. PAHs occasionally condense into soot.
- Molecules are also found in the meteorites.
- The formation of some of these molecules is heterogeneously catalyzed (on dust grains).
- The microwave spectrum of CO allows to derive the local temperature.
- UV-induced photodissociation prevents the formation of many more molecules.

14 Condensed matter

Compared to gases, the universe contains rather little condensed matter. One of the reasons for this is that there are relatively few atoms with a mass number of $A > 4$ (helium). Condensed matter requires attractive interactions. The attractive interactions are weak for hydrogen and helium because they are not polar and because they can hardly be polarized due to the small number of electrons

Liquids are only known from planets, because liquids are usually in a material equilibrium with a vapor phase.[110] In a logarithmic pT diagram, the liquids occupy a relatively small area (Fig. 14.1). Only sufficiently large planets can bind the vapors gravitationally. In free space, droplets evaporate immediately. The main molecules that form liquids are

- Water
- Ammonia
- Hydrocarbons

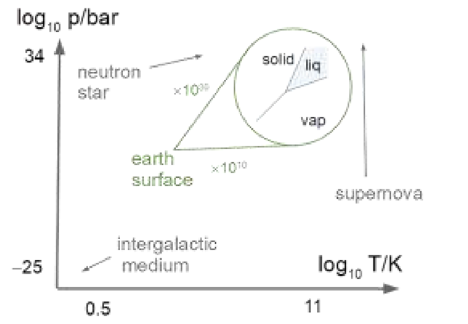


Fig. 14.1
Molecular liquids are not very common in space because they are only stable in very narrow ranges of pressure and temperature. These conditions exist on Earth and on some other planets, but only there.

Liquid hydrocarbons (including methane) exist on Saturn's moon Titan. Liquid ammonia is not known from the solar system.[111] Silicon rarely forms small molecules (which could be liquid as condensed matter). In particular, SiH_4 is a very unstable molecule. This is because the H atoms cannot shield the central atom as well as methane and the other hydrocarbons due to their lack of size. The famous metastability of carbon compounds does not occur with silicon compounds. Silicates are virtually non-existent in the gas phase. The binding energy of silicate crystals (in various morphologies) is so high that silicates immediately form dust. In more technical terms, the Kelvin radius for silicates is only slightly larger than the SiO_2 molecule. There is virtually no nucleation barrier.

In free space, the atoms H, C, N and O initially form small molecules, which become visible via their rotational spectra (section 13.3). The formation of the molecules can be catalyzed heterogeneously on the surfaces of dust grains, but does not have to be. In a cold environment, these molecules can condense into solids. Comets consist to a large extent of ice. They are "dirty snowballs".

Carbon can condense into soot particles in the form of polycyclic aromatic hydrocarbons (PAHs). However, the dust in the interstellar dust clouds is mainly silicate. Soot is sometimes found in meteorites (the "carbonaceous chondrites").

There is experimental evidence for the high silicate content in the dust in the form of a broad emission at 10 μm (Fig. 14.2 and Fig. 14.3). Such broad absorption can also be seen in model dusts. The constituents (Si, C, Al, other elements) are first evaporated and then the vapor is deposited on a cold plate ($T \approx 70$°K) in the presence of argon. The argon forms a solid matrix which traps small particles that have formed on the way to the plate. One then spectroscopes these films and hopes for a match with the emission spectra from the dust clouds. This agreement occurs mainly for silicate particles.

[110] Exceptions are ionic liquids and silicone oils

[111] Ammonia is more polar in liquid than the hydrocarbons. The chemistry in liquid ammonia could be more interesting than the chemistry in the lakes and seas on Titan (the latter consisting mainly of methane).

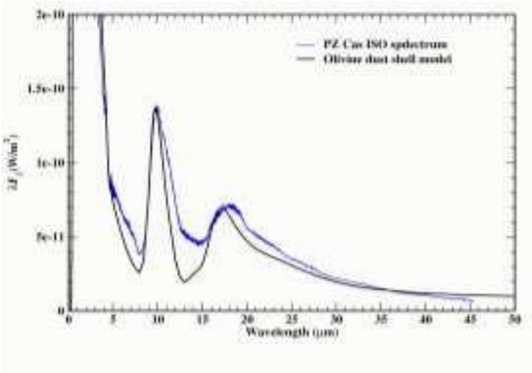


Fig. 14.2
The figure above shows why we think these silicates are present in space. The ISO spectrum of a typical evolved star, PZ Cas (which has a spectral type of M3Iab, and which is a semi-regular variable of period 900 days, V magnitude range between 9.8 and 12.7), is shown along with a model spectrum for a dust shell of Olivine dust around a cool star. The model spectrum is calculated using the properties of Olivine dust measured in the laboratory. Although the match is not at all perfect, the two big features seen for amorphous Olivine grains appear to match the ISO spectrum features reasonably well. The features are due to oscillation modes in Si-O bonds. The longer wavelength feature is at 17 microns in the Olivine dust model spectrum and at 18 microns in the PZ Cas spectrum, indicating that the dust in this star is slightly different from terrestrial Olivine.
*www.*stsci.edu/~volk/features1a.html

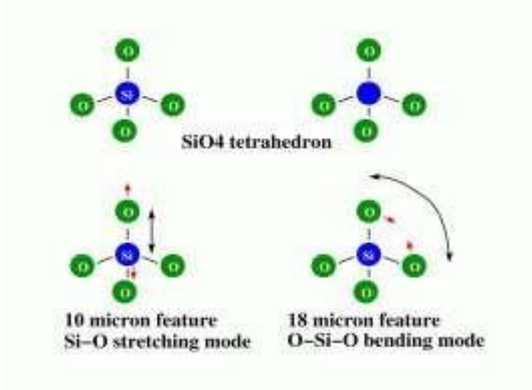


Fig. 14.3
The above figure shows a schematic of the two modes that are responsible for the two features. The red vectors show the direction of motion of atoms, and the black arrows show the general oscillation modes. Exactly what the frequencies of the oscillations are depends on the details of the mineral structure, and so it can vary from one type of silicate to another.
*ww.stsci.edu/~volk/features1a.*html

There is a historical side note to this. At some point, four bands were seen in these spectra from the laboratory that were not known from the sky or the earth. The scientist in question kept the matter in mind and at some point met a colleague who was familiar with the theoretically calculated IR spectra of C_{60} . The soot in question contained C . The soot in question had contained C_{60} . In the ^{13}C NMR, only one line was seen. After that, the C_{60} wave took its course. Since then, C_{60} has existed not only as a hypothesis and as a peak from mass spectroscopy, but also as a material.

Planets and other clumps sometimes form from the dust (section 8.1). If these hit the Earth as meteorites, an isotope analysis can be carried out. This can also be done with the rocks brought back from the moon by the Apollo missions. Studies of lunar rocks suggest that the Moon was ejected from the Earth by a violent collision with a third body at an early stage in the Earth's history. Meteorites from Mars can be recognized as such by their isotope stamp.[112] It is estimated that there are between 95 and 148 "parent bodies" for the meteorites that hit the Earth. Isotope analysis also plays a role here. The paths that lead to certain isotope stamps are described in section 14.1 goes into more detail.

Remember:
- The condensed matter in the universe mainly consists of dust.

[112] For evidence that meteorites with this isotopic stamp are made of Martian rock, see *en.wikipedia.*org/wiki/Martian_meteorite

- Dust consists mainly of silicates, but also partly of carbon. Ice - and of course other atoms and molecules - also occur.

14.1 Side remarks: Isotope fractionation

The influence of the nuclear mass (which has different values for the different isotopes) is rather small in chemistry. By far the most important nuclear property for chemistry is the electrical charge (not the weight). Because chemical and physico-chemical processes are weakly dependent on the weight of the nucleus, the isotopic fingerprints can be time capsules.

Sometimes an accumulation of certain isotopes is caused by cosmic radiation. In other cases, radioactive decay plays a role. In the following, the main focus is on why isotopes sometimes accumulate or deplete in certain places.

Many isotope effects can be traced back to the vibration term scheme. A vibration can be both an intramolecular vibration and a vibration of atoms and molecules in the crystal lattice. If the mass changes, the energy eigenvalues change due to $E_v \approx \hbar\omega(v+1/2)$ and $\omega = (\kappa/\mu)^{1/2}$. μ is the reduced mass. The isotope effects are particularly large if you replace a CH bond in a molecule with a CD bond (-when you deuterate, analogous for OH or NH). In this case, the reduced mass almost doubles. This is evident from the fact that heavy water (D_2O) is moderately toxic.[113] In the other cases (C/ C/^{121314} C, O/^{1618} O, many others) the effects are mainly interesting for analytical purposes. Mars meteorites can be identified by their isotope ratios. The O/^{1816} O ratio at a certain depth of a drill core from Antarctic ice says something about the climate at the time of formation. Natural products can be distinguished from synthetic products. The origin of ores (from the Congo? such ores are subject to an embargo) can be determined using the isotope stamp. Such isotope stamps can also be obtained from small sample quantities using mass spectroscopy. There is not always a single, conclusive physical interpretation for an isotope stamp. For example, it can be seen that the abundance of^{15} N (relative to^{14} N) increases along the food chain. Plants contain less^{15} N than animals. However, this finding cannot be modeled ab initio.

The phase equilibria between crystals and vapor are particularly important for isotope fractionation in the bodies of the solar system. An increase in mass decreases the distance between the terms in the vibrational term scheme and consequently increases the relevant state sums. Because $A = -k_B T \ln Q$ (with A the free energy and Q the sum of states[114]), the free energy then decreases. Because the intermolecular interactions also have an influence on the sum of states, this results in synergism. Phase coexistence results in different free energy differences between the two phases for the various isotopes. The partition coefficients are therefore different for the various isotopes.[115] The effect is small, but widespread. Even if the details are not understood, ores from different sources can be distinguished empirically on the basis of their isotopic composition. Something similar can be observed in condensation: Water molecules -containing the heavier^{18}- O atoms tend to be the first to condense and rain out. Snow falling in Canada contains more $H_2{}^{18}O$ than rain falling in Florida.

---

[113] Furthermore, isotope separation is much easier for hydrogen and deuterium than for the various uranium isotopes, for example. Heavy water is not as expensive as you might think.

[114] $Q = \Sigma_i \exp($□□E_i □$E_0)/(k_B T))$

[115] If a substance is distributed (as a minority component) over two phases α and □, $a_\alpha / a_\beta = \exp($□□$\mu/(RT))$. *a* is the activity. You still have to convert to concentrations. The quotient of the concentrations is the partition coefficient. It is important in 2-phase extraction.

15 New instruments

Astronomy and astrophysics have benefited enormously from new devices in recent years. This development is continuing. A whole series of new large-scale devices are planned - both as satellites and earth-bound. Here are some facts in bullet points.

– With so-called "adaptive optics", mirrors are enabled to compensate for image distortions caused by the atmosphere ("seeing"). The resolution of the telescope then no longer corresponds to the "seeing disk" (diameter is ~ 1 arcsecond) but to the diffraction image of the mirror.
– Large new optical telescopes rely on much more light. These can be used to study faint galaxies at large distance.
– Another class of telescopes – also optical – relies on a large field of view. Firstly, this allows to study weak gravitational lensing to study (section 6, the Euclid satellite). Weak lensing requires good statistics. Furthermore, these instrument search for rare events. This includes microlensing (also section 6).
– IR telescopes look through dust clouds, they see warm objects (star-forming regions, exoplanets) and they see galaxies at large red shifts. The flagship is the James Webb space telescope.
– In the radio range, digital interferometry achieves a resolution that corresponds to a pupil as large as the diameter of the earth.[116] The achievable resolution is in the range of a few 10 micro-arcseconds. Here, the square kilometer array (SKA) should bring about a leap in the quantity and quality of data by a factor of around 1000.
– Incremental progress has been made with the X-ray satellites and the gamma-ray satellites.
– Neutrino astronomy is primarily possible with high-energy neutrinos. These do not originate from the sun or from terrestrial nuclear reactors. It is possible to obtain directional information. Unfortunately, cosmic radiation creates a background. This is why the instrument (the "Ice Cube" in Antarctica) is particularly suitable for studying the northern sky.[117] Recently, however, a map of the Milky Way (largely in the southern sky) was also presented.
– The detection of gravitational waves was of course spectacular. So far, the event in question has only been identified optically in one case. This is the famous event GW170817. The current instruments only see the merging of objects with masses of the order of the mass of the sun. The merging of supermassive black holes (often present in the centers of galaxies) is too slow. Evidence for gravitational waves with this origin has been obtained with pulsar timing arrays. The millisecond pulsars are such good clocks that gravitational waves can be inferred - at least in principle - from slight irregularities in the apparatus frequency. There are indications of gravitational waves from the merging of supermassive black holes, but these indications take the form of low-frequency noise. A satellite experiment is also being planned (LISA), which should have better sensitivity at low frequencies. This instrument may even be able to detect primordial gravitational waves (Section 4.3).

[116] Digital interferometry is possible because the radio antennas are coherent detectors. The difference between coherent and incoherent detectors (with and without phase information) is discussed in the lecture Molecular Structure and Molecular Spectroscopy in connection with the quantum mechanical measurement process. As with incoherent detectors, the phase information is lost in the quantum mechanical measurement process.
[117] Neutrinos pass through the earth. The cosmic radiation generated in the northern sky does not.

16 Appendices

16.1 The role of fine structure constants in the physics of atoms and molecules

The role of the fine structure constant is explained below using three examples:
A) The speed of the electrons on their orbits around the nuclei is significantly less than the speed of light.
B) The diameter of an atom is significantly smaller than the wavelength of light.
C) Magnetic interactions in chemistry are significantly weaker than electrical effects.

To A): Comparison of the speeds of the electrons in the molecules with the speed of light

This is based on Bohr's atomic model. According to this, the angular momentum is quantized in units of ħ:

Glg. 16.1

$$\mu \mathrm{v} r = n\hbar$$

μ is the reduced mass, v is the velocity, r is the radius of the circular path and n is an integer. It follows

Glg. 16.2

$$\mathrm{v} = \frac{n\hbar}{\mu r}$$

Now we apply Newton's second axiom (force = mass × acceleration). For circular motion, the acceleration a is given as

Glg. 16.3

$$a = \frac{\mathrm{v}^2}{r}$$

If we use the electrostatic attraction to the nucleus for the force, we get

Glg. 16.4

$$\frac{Ze^2}{4\pi\varepsilon_0 r^2} = \frac{\mu \mathrm{v}^2}{r}$$

Z is the charge of the nucleus, e is the elementary charge and ε_0 is the dielectric permittivity of the vacuum. All the quantities that we derive below follow from Glg. 16.1 and Glg. 16.4. When we insert values, we always consider the ground state of the hydrogen atom (n = 1, Z = 1). This state is also known as the "H1s state".

To calculate the speed of the H1s electron, we set *one* of the two v's in Glg. 16.4 the Glg. 16.2 (with n = 1):

Glg. 16.5

$$\frac{Ze^2}{4\pi\varepsilon_0 r^2} = \frac{\mu \mathrm{v}^2}{r} = \frac{\mu}{r}\mathrm{v}\frac{\hbar}{\mu \mathrm{r}}$$

We obtain (with $Z = 1$ for the H atom)

Glg. 16.6

$$\mathrm{v} = \frac{e^2}{4\pi\varepsilon_0\hbar}$$

We divide by the speed of light and find

Glg. 16.7

$$\frac{\mathrm{v}}{c} = \frac{e^2}{4\pi\varepsilon_0\hbar c}$$

The right side in Glg. 16.7 is the *fine structure constant* α:

Glg. 16.8

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c}$$

To B): Comparison of the size of atoms with the wavelength of light

The circulation time T is

Glg. 16.9

$$T = \frac{2\pi}{\omega} = 2\pi\frac{r}{\mathrm{v}} = 2\pi\frac{a_0}{c\alpha}$$

a_0 is the Bohr radius (0.5 Å). Its value also follows from the Bohr model, but is not derived here. We form Glg. 16.9 once again to

Glg. 16.10

$$Tc = \lambda = \frac{2\pi}{\alpha}a_0$$

Tc is the wavelength of light that has the frequency $1/T$. For T in the femtosecond range, the corresponding wavelengths are in the UV-Vis range. ($c \cdot (1\ \mathrm{fs}) \approx 300$ nm.) Atoms and molecules are therefore smaller than the wavelength of the light that causes electronic excitation by a factor of around $\alpha/(2\pi)$.

Re C): Relative strength of magnetic and electrical effects in chemistry

We limit ourselves to dia- and paramagnetism. For demonstration purposes, we choose the interaction energy of two dipoles (electric or magnetic) at a distance r. The following relations apply

Glg. 16.11

$$E_{dip,el} = \frac{1}{4\pi\varepsilon_0}\frac{\mu_{el}^2}{r^3} f\left(\theta_1,\theta_2,\phi\right)$$
$$E_{dip,mag} = \frac{\mu_0}{4\pi}\frac{\mu_{mag}^2}{r^3} f\left(\theta_1,\theta_2,\phi\right)$$

f is a function of the angles, which need not be of interest here. It will be truncated further down in the ratio formation. μ_0 is the magnetic permeability of the vacuum.

A typical value for an electric dipole is ea_0 with e the elementary charge and a_0 the Bohr radius. Magnetic dipoles are found for circular currents. The magnetic dipole moment is the product of the area of the circle and the current. A typical value is $\pi a_0^2\, e/T = \pi a_0^2\, e\mathrm{v}/(2\pi a_0) = e\mathrm{v}a_0/2$ with v the velocity of the electron according to the Bohr model.

For the ratio of the interactions one finds

Glg. 16.12

$$\frac{E_{dip,mag}}{E_{dip,el}} = \mu_0 \varepsilon_0 \frac{\mathrm{v}^2}{4}$$

The product $\mu_0\, \varepsilon_0$ is identical to $1/c^2$ (c is the speed of light). The result is

Glg. 16.13

$$\frac{E_{dip,mag}}{E_{dip,el}} \approx \frac{1}{4}\frac{\mathrm{v}^2}{c^2}$$

Due to $\mathrm{v}/c \approx \alpha \ll 1$, this ratio is much smaller than one.

16.2 The thermodynamic variant of the virial theorem

In this section, the theorem $\langle E_{\mathrm{kin}} \rangle = -1/2\, \langle V \rangle$ is placed in a thermodynamic context for systems that interact via a $1/r$-potential. However, it also applies to systems that are not in thermodynamic equilibrium (section 4.4.2).

A classical dynamic system consisting of N particles is described by N spatial coordinates, q_i , and N momenta, p_i . These $2N$ coordinates span the "phase space". The infinitesimal phase space volume is called d□. There is an energy function - the "Hamilton function" - with the name $H(\{q_i ,p_i \})$. According to Boltzmann, the probability of finding the system at the coordinates $\{q_i ,p_i \}$ is given by $\exp(-H(\{q_i ,p_i \}))$. The following relation must be proven:

Glg. 16.14

$$\left\langle q_i \frac{\mathrm{d}H}{\mathrm{d}q_j} \right\rangle = k_B T \delta_{ij}$$

$\langle ... \rangle_{th}$ denotes the Boltzmann-weighted mean value. δ_{ij} is the Kronecker delta with the properties $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. A very analogous relation (which is also used below) applies to the pulses, p_i .

Ref. 118 calls Glg. 16.14 “the virial theorem". Wikipedia calls Glg. 4.6 the virial theorem and proves it differrently from what ref. 118 does. It does not prove it from the temperature and the Boltzmann distribution. In the proof according to ref. 118 the equipartition theorem and the inner virial are included as secondary results. On the other hand, it requires thermal equilibrium.

For the proof of Glg. 16.14 we first write out the mean value:

Glg. 16.15

$$\left\langle q_i \frac{\mathrm{d}H}{\mathrm{d}q_j} \right\rangle = \frac{\int q_i \frac{\mathrm{d}H}{\mathrm{d}q_j} \exp\left(-\frac{H}{k_B T}\right) \mathrm{d}\Gamma}{\int \exp\left(-\frac{H}{k_B T}\right) \mathrm{d}\Gamma} = \frac{1}{Z} \int q_i \frac{\mathrm{d}H}{\mathrm{d}q_j} \exp\left(-\frac{H}{k_B T}\right) \mathrm{d}\Gamma$$

---

[118] Schwabl, F., *Statistical Physics*. Springer **2006**, Chapter 2.6.4.1

In the second step, the denominator was renamed Z. The denominator is also called the "partition integral". We use the relation

Glg. 16.16

$$\frac{\mathrm{d}}{\mathrm{d}q_j}\exp\left(-\frac{H}{k_B T}\right)=\exp\left(-\frac{H}{k_B T}\right)\left(-\frac{1}{k_B T}\right)\frac{\mathrm{d}H}{\mathrm{d}q_j}$$

It follows

Glg. 16.17

$$\frac{\mathrm{d}H}{\mathrm{d}q_j}\exp\left(-\frac{H}{k_B T}\right)=-k_B T\frac{\mathrm{d}}{\mathrm{d}q_j}\exp\left(-\frac{H}{k_B T}\right)$$

Insert in Glg. 16.15 results in

Glg. 16.18

$$\left\langle q_i\frac{\mathrm{d}H}{\mathrm{d}q_j}\right\rangle=\frac{1}{Z}\int q_i\frac{\mathrm{d}}{\mathrm{d}q_j}\exp\left(-\frac{H}{k_B T}\right)\left(-k_B T\right)\mathrm{d}\Gamma$$

We assume that the function exp(-$H(\{q_i, p_i\})$) tends to zero at infinity. This is the case for parabolic potentials, for example. There should be no walls. The case with a wall is discussed (briefly) below in connection with the inner virial. If exp(-$H(\{q_i, p_i\})$) becomes zero at infinity, we can use Glg. 16.18 partially integrate according to

Glg. 16.19

$$\left\langle q_i\frac{\mathrm{d}H}{\mathrm{d}q_j}\right\rangle=\left[\frac{1}{Z}\int x_i\exp\left(-\frac{H}{k_B T}\right)\left(-k_B T\right)\mathrm{d}\Gamma\right]_{-\infty}^{\infty}-\frac{1}{Z}\int\frac{\mathrm{d}q_i}{\mathrm{d}q_j}\exp\left(-\frac{H}{k_B T}\right)\left(-k_B T\right)\mathrm{d}\Gamma$$

The square brackets refer to the evaluation on the edge of the integration volume (i.e. at infinity). This term is omitted according to the assumption. The following applies: $\mathrm{d}q_i/\mathrm{d}q_j = \delta_{ij}$ with δ_{ij} the Kronecker-δ. We subtract the constant terms from the integral and obtain

Glg. 16.20

$$\left\langle q_i\frac{\mathrm{d}H}{\mathrm{d}q_j}\right\rangle=k_B T\delta_{ij}\frac{1}{Z}\int\exp\left(-\frac{H}{k_B T}\right)\mathrm{d}\Gamma=k_B T\delta_{ij}$$

In the last step, we used the fact that the integral on the right is just the state integral. This proves Glg. 16.14.

Now consider a harmonic oscillator with the potential

Glg. 16.21

$$V=\frac{1}{2}\kappa x^2$$

κ is the spring constant. We assume that the oscillator is coupled to a heat bath. It then bumps frequently and the probabilities for the occurrence of a certain value of x are subject to the Boltzmann distribution. We replace in Glg. 16.14 H by V. We can do this because the mathematics is independent of whether we consider the total energy or only the potential energy. We then obtain

Glg. 16.22

$$\left\langle x \frac{\mathrm{d}H}{\mathrm{d}x} \right\rangle = \left\langle \kappa x^2 \right\rangle = 2\langle V \rangle = k_B T$$

The mean potential energy of the oscillator is therefore 1/2 $k_B T$.

We consider the kinetic energy of a particle in one dimension, given as

Glg. 16.23

$$E_{kin,1D} = \frac{1}{2m} p_x^2$$

We replace in Glg. 16.14 H with E_{kin} and arrive at

Glg. 16.24

$$\left\langle p_x \frac{\mathrm{d}E_{kin,1D}}{\mathrm{d}p_x} \right\rangle = \left\langle \frac{1}{m} p_x^2 \right\rangle = 2\left\langle E_{kin,1D} \right\rangle = k_B T$$

and

Glg. 16.25

$$\left\langle E_{kin,1D} \right\rangle = \frac{1}{2} k_B T$$

In three dimensions

Glg. 16.26

$$\left\langle E_{kin} \right\rangle = \frac{3}{2} k_B T$$

These arguments can be repeated for all quadratic contributions to the Hamilton function. (In particular, they can be repeated for the rotational energy, given as 1/2 $I\omega^2$ with I the moment of inertia and ω the angular velocity). The mean energy associated with such quadratic contributions is 1/2 $k_B T$ in all cases. (It is 3/2 $k_B T$ for the kinetic energy because of the three dimensions of space). If the Hamiltonian function has only quadratic contributions, the number of these quadratic contributions is called f and the contributions are referred to as "degrees of freedom". The CO_2 molecule has 13 such degrees of freedom: 3 for the translation, 2 for the rotation, 4 for the potential energy in the 4 vibration modes and 4 for the kinetic energy in the 4 vibration modes. There are only 2 degrees of freedom for rotation because rotation around the molecular axis is an electronic excitation that is frozen at room temperature. (The vibrational degrees of freedom are also partially frozen).

If quantization is ignored, the mean thermal energy of a system with f degrees of freedom is given by f /2 $k_B T$. The following relations apply

Glg. 16.27

$$\left\langle \grave{E} \right\rangle = \frac{f}{2} k_B T, \quad \bar{U} = \frac{f}{2} RT, \quad c_V = \frac{f}{2} R$$

It should be emphasized again that a degree of freedom is a quadratic contribution to the energy function. If the potential is not a parabolic potential, the equipartition theorem does not apply. For example, the equipartition theorem does not apply to the potential energy of charges in the diffuse double layer above charged surfaces.

The potential is often only approximately harmonic. An example is the 6-12 potential is close to its minimum. Caution is then required. Such potentials often lead to condensed phases. Then the distances are close to the minimum-energy distance and the potential can be Taylor-expanded around the minimum (into a parabola). In this case, the equipartition theorem applies and the specific heat capacity is positive. However, it might also be a real gas. In this case, the attractive part ($\propto r^{-6}$) is more significant because the particles rarely come approach enough to feel the repulsion. In this case, the temperature dependence of the internal energy is not trivial. (The specific heat capacity of noble gases deviates very slightly from the equipartition theorem).

On earth, the gas is always confined in a container. The forces exerted by the wall must then be taken into account. A somewhat length calculation results in the relation[118]

Glg. 16.28

$$pV = \frac{2}{3}\langle E_{kin}\rangle - \frac{1}{6}\left\langle \sum_{m,n} |x_{mn}| \frac{\mathrm{d}V(|x_{mn}|)}{\mathrm{d}|x_{mn}|} \right\rangle$$

For gases, the first term on the right-hand side is identical to $Nk_{\mathrm{B}}T$. The second term on the right-hand side is called the "inner virial". (There are different conventions regarding the prefactor.) For non-interacting systems (such as ideal gases) the inner virial is zero. Then reproduces Glg. 16.28 reproduces the ideal gas law. For real gases, the virial is not equal to zero. Evaluating the virial in detail is anything but trivial. As part of the "virial development", the virial is developed according to the number density N/V and thus arrives at the "virial coefficients". The second virial coefficient is of particular importance.

The consequences of the virial theorem in self-gravitating systems are discussed in Section 4.4 is discussed.

<u>Remember:</u>

- For systems in thermodynamic equilibrium, the relations $\langle q_i\, \mathrm{d}H/\mathrm{d}q_j \rangle_{th} = k_{\mathrm{B}} T\, \delta_{ij}$ and $\langle p_i \mathrm{d}H/\mathrm{d}p_j \rangle_{th} = k_B T \delta_{ij}$ hold. and p_{i} are the locations and momenta (or angular momenta).
- The equipartition theorem follows from the relation above for the quadratic contributions to the energy function.
- The relation above clarifies the role of the "inner virial" for the real gas laws.

Box 16.1 Does the ideal gas law apply in the center of the sun?

Atkins' textbook reports that the ideal gas law applies approximately in the center of the sun, although the density of $\rho = 150$ g/cm^3 is greater than the density of condensed matter on earth. This is justified by the fact that the atoms are completely ionized and that the nuclear radius must therefore be used instead of the atomic radius for the particle radius. Because the nuclear radius is much smaller than the mean distance between the particles, the kinetic theory of gases can be applied. The kinetic theory of gases states that the particles usually fly uniformly in a straight line through space and only rarely experience collisions. The distance between two collisions is the mean free path . λ is the central parameter of the kinetic theory of gases.

This argument is problematic insofar as the particles primarily interact electrostatically with each other. The nuclear forces do play a role (because nuclear fusion occurs), but they do not dominate the pressure-volume relationship. As far as the interactions are concerned, the particle radius must not be equated with the nuclear radius. In plasmas, electrons almost never fly in a straight line because the Coulomb interaction ($\propto 1/r$) prevents this due to its long range. Plasmas are not described by the kinetic theory of gases.

Why does the ideal gas law still apply so well? This has to do with a variant of the virial theorem, reported in Glg. 16.28. This equation only assumes thermodynamic equilibrium, not the validity of the kinetic theory of gases. This equation contains the "inner virial" as the second term on the right-hand side. The inner virial causes the difference between ideal and real gases. If the mean energy in the interactions is smaller than the mean kinetic energy, then the correction to the ideal gas law according to Glg. 16.28 is small. This is the case for real gases and for the center of the sun. The ideal gas law applies because the kinetic energy has the relatively larger share of the total energy.