Appendix A
With a little bit of algebra

Someone told me that
each equation I included in the book would halve the sales.
Stephen William Hawking[1]

As long as algebra is taught in school,
there will be prayer in school.
Cokie Roberts[2]

This text is part of the volume *Pioneer of Galactic astronomy: A biography of Jacobus C. Kapteyn* by Pieter C. van der Kruit, published in the series Springer Biographies.In contrast to appendix A in the published book, this version does not shy away from mathematical details and equations.

This appendix aims at providing some additional background and understanding, but reading is not necessary for the reader not interested in such details. Fort further details and more background see [2].

A.1 Vibrating flat membrances

To give an idea of what Kapteyn's study of vibrating flat membrances in his PhD thesis entailed, I give here a short summary.

The questions addressed are related to the manner in which a membrane can vibrate, depending on shape and support. In those days it was not really necessary to add much original work to the thesis and Kapteyn's is for a large part indeed a survey

[1] From: *A Brief History of Time* [1].

[2] Cokie Roberts, née Mary Martha Corinne Morrison Claiborne Boggs (1943–2019), American journalist and author.

of the literature. But he does use observations that had been obtained by others to compare theory to observations. He starts out by writing down the equations that govern the vibrations.

Take an infinitessimally small element of the membrane with size ∂x by ∂y. Then

$$Fdy\frac{\partial^2 w}{\partial x^2}dx + Fdx\frac{\partial^2 w}{\partial y^2}dy = \rho dxdy\frac{\partial^2 w}{\partial t^2}. \tag{A.1}$$

On the left-hand side we have the total force, on the right mass times acceleration, both in the vertical direction. So it really is Newton's law. First look at the part on the right. There ρ is the mass-surface-density (in for example grams per square centimeter) and that is then multiplied by the surface of the element; this yields its mass. Furthermore, $\partial^2 w/\partial t^2$ is the acceleration in the (vertical) w direction. After all, if w is the position, then the first derivative $\partial w/\partial t$ is the change in w per unit of time t and thus the vertical speed; the second derivative $\partial^2 w/\partial t^2$ is then the change in the speed per unit of time and thus the acceleration.

On the left hand side there are two very similar terms. It is the total force *in the vertical direction*, where one term comes from the x direction and the other y. Each term includes the force acting on the membrane exerted by stretching it in the corresponding direction. F is the stretching force per unit length in both the x- and the y-directions. These are then multiplied by the size of the element in the relevant direction; the force in the x direction on the element is then Fdy and that in turn is multiplied by a factor that takes into account the projection of that force onto the w direction (in fact, the difference between the sines of the projection angles on the w-direction at both ends of the element). Together the left side then is the force in the w direction as a result of the stretching of the membrane.

Add to this the boundary conditions, namely that the displacement of the membrane at the edges has to be zero and that all w as well as dw/dt are equal to zero at the starting time $t = 0$. Kapteyn then reqrote the equations as

$$\frac{\partial^2 w}{\partial t^2} = c^2\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) \quad \text{with} \quad c^2 = \frac{F}{\rho}. \tag{A.2}$$

In this form it was first derived by Poisson.

The solution can be written in terms of waves, which are sine and cosine functions with time as a variable. For a rectangular membrane with sides a and b, for example, Kapteyn found a sum of sines and cosines as a solutions to the equation, which together comprise all the ways in which the membrane can possibly vibrate:

$$w = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(A_{ij}\cos\gamma t + B_{ij}\sin\gamma t)\sin i\pi\frac{x}{a}\sin j\pi\frac{y}{b}. \tag{A.3}$$

The constants A_{ij} and B_{ij} can be determined by applying the boundary conditions. Each term ij then represents an oscillation with a frequency

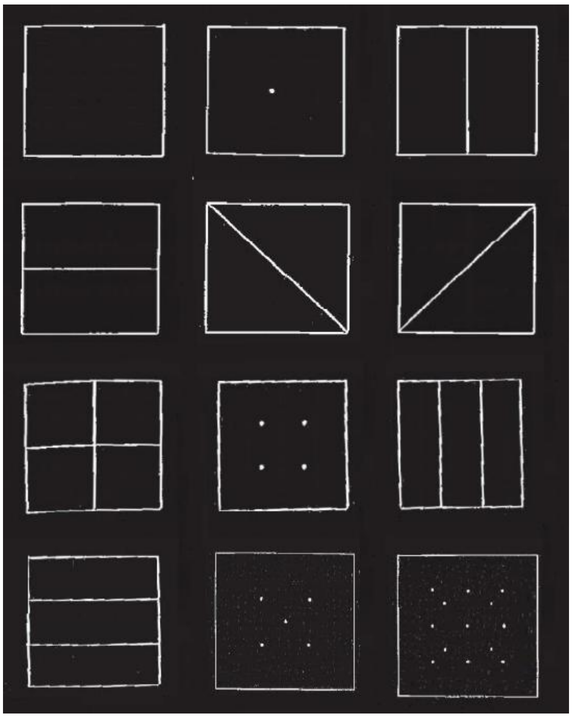

Fig. A.1 Linear nodes and pointlike nodes for a vibrating square membrane according to the PhD thesis of Kapteyn. Top-left: wavelength twice the lengths of sides; then there is no node (except the edges of course). Next we see some examples of 'overtones', where the wavelengths in each direction are equal to the length of the sides, two-thirds of this or half. Kapteyn Astronomical Institute, University of Groningen.

$$N(i,j) = \gamma/2\pi = \frac{c}{2}\sqrt{\frac{i^2}{a^2} + \frac{j^2}{b^2}}.$$

Fig. A.1 shows some of Kapteyn's results in terms of nodes. These are either pointlike or straight lines.

Kapteyn compared his results with observations. These were performed by others and consisted of covering a membrane with powder and then making it vibrate by striking it with a violin bow. The powder then collects in the nodes. Kapteyn concluded that the pointlike nodes must exist (apparently a new result), but they could not be seen in observations. He proposed that this was probably due to small imperfections in the thickness and elasticity that membranes in reality have.

A.2 Distances and luminosities

The distance of a star not too far from the Sun can be measured directly using the so-called annual parallax. In fig. A.2 we see how the annual motion of the Earth around the Sun is reflected in an elliptical orbit of the star on the sky. The semi-major axis of that ellips (which is equal to the angle p at the top of the triangle) then is a measure for the distance of the star. The radius of the Earth's orbit (the

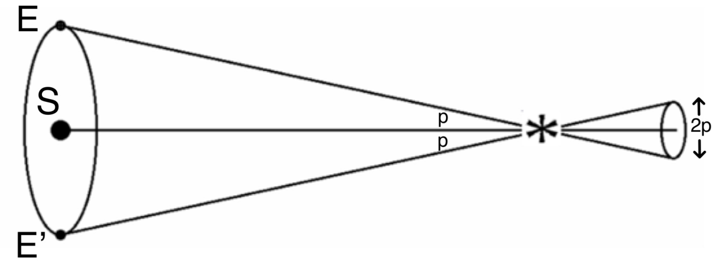

Fig. A.2 The annual circular motion of the Earth *E* around the Sun *S* on the sky in a projected elliptical motion by a star, of which the size of the ellipse ($2p$) depends on the distance to the star. The angles p are the parallax. Figure by the author.

Astronomical Unit) is 1.4960×10^{11} meters. When that angle p, the parallax, is 1 arcsecond, the distance of the star is 3.0857×10^{16} meters. This is one *parsec*. It is equal to 3.26 lightyears (one lightyear is the distance traveled by light in vacuum during one year).

The apparent magnitude of a star is a measure of its brightness in the sky; the concept originates from Antiquity, when the brightest stars were assigned magnitude zero and the faintest that the human eye could see, magnitude six. This system was already used in the star catalog of Hipparchus of Nicaea in the second century BC. The British astronomer Norman Robert Pogson redefined it in 1856 by proposing a scale where 5 magnitudes were exactly a factor of 100, so that one magnitude corresponds to a factor of $\sqrt[5]{100}$ =2.512. The difference in magnitude of two stars, which have a flux of F_1 and F_2 in energy per surface area per solid angle (for example, Joules per m^2 per steradian), is then

$$m_1 - m_2 = -2.5 \log_{10} \left(\frac{F_1}{F_2} \right). \tag{A.4}$$

The absolute magnitude (designated with capital M) is defined as the magnitude a star would have at a distance op 10 parsec. If the distance is r or the parallax p, then

$$M = m + 5(1 + \log_{10} p) = m + 5(1 - \log_{10} r). \tag{A.5}$$

This definition goes back to Kapteyn.

A.3 Sines functions of higher orders

It is well known that the sine and cosine of an angle x (in radians) can be expanded as an infinite series:

$$\sin x = \sum_{n-0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.$$

$$\cos x = \sum_{n-0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots.$$

The brothers Jacobus and Willem Kapteyn made a study of a set of generalized series of these that are designated by 'higher-order-sines':

$$\varphi_\mu(x) = \sum_{n=0}^{\infty} \frac{x^{kn+\mu}}{(kn+\mu)!} = \frac{x^\mu}{\mu!} + \frac{x^{\mu+k}}{(\mu+k)!} + \ldots. \tag{A.6}$$

$$\psi_\mu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{kn+\mu}}{(kn+\mu)!} = \frac{x^\mu}{\mu!} - \frac{x^{\mu+k}}{(\mu+k)!} + \ldots. \tag{A.7}$$

They published an extensive paper on this in 1886, that had been preceded by another one on the special case $k = 4$.

A.4 Kepler's equation

Planets, asteroids and comets move around the Sun in elliptical orbits according to the laws of Kepler. These laws are a direct result of the solutions to the equations Isaac Newton defined in his theory of gravity in the so-called *Two-body Problem*. I will treat this schematically. For more details see the notes on the two-body problem accompanying my lecture course on *Dynamics of Galaxies* at www.astro.rug.nl/∼vdkruit/jea3/homepage/two-body.pdf.[3]

Take two masses m_1 and m_2 with position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Then there are two fundamental equations:

$$m_1\ddot{\mathbf{r}}_1 = -Gm_1m_2\frac{\mathbf{r}_1 - \mathbf{r}_2}{r^3}, \tag{A.8}$$

$$m_2\ddot{\mathbf{r}}_2 = -Gm_1m_2\frac{\mathbf{r}_2 - \mathbf{r}_1}{r^3}. \tag{A.9}$$

Here $\ddot{\mathbf{r}}$ denotes the second derivative with respect to time $d^2\mathbf{r}/dt^2$. Now go to a co-moving coordinate system, in which the center of gravity is the origin:

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = 0.$$

Then

$$\ddot{\mathbf{r}}_1 = -G\frac{m_2\mathbf{r}_1 - m_2\mathbf{r}_2}{r^3} = -GM\frac{\mathbf{r}_1}{r^3}.$$

[3] The two-body problem can be solved analytically. This is not true for the three-body problem, where numerical integration is necessary in the general case. There is a 'restricted three-body problem' for circular orbits of two components around each other, and a third body of negligible mass. See www.astro.rug.nl/∼vdkruit/jea3/homepage/three-body.pdf.

$$\ddot{\mathbf{r}}_2 = -G\frac{m_1\mathbf{r}_2 - m_1\mathbf{r}_1}{r^3} = -GM\frac{\mathbf{r}_2}{r^3}.$$

So for $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ (the vector between the two bodies)

$$\ddot{\mathbf{r}} = -GM\frac{\mathbf{r}}{r^3}$$

We have three differential equations of the second order, so we will have in principle **6** constants of integration. Multiply with $\mathbf{r}$. Then

$$\mathbf{r} \times \ddot{\mathbf{r}} = -GM\frac{\mathbf{r}\times\mathbf{r}}{r^3} = 0$$

Integrate this equation with respect to time. This gives

$$\mathbf{r} \times \dot{\mathbf{r}} = \text{constant} = \mathbf{h}. \tag{A.10}$$

So angular momentum is conserved, which is *Kepler's second law*.

Some more vector manipulations give

$$\frac{h^2}{GMr} = 1 + \frac{\mathbf{P}}{GM}\cdot\frac{\mathbf{r}}{r} = 1 + \frac{\mathbf{P}}{GM}\cdot\hat{r}$$

Designate the angle between $\hat{r}$ and $\mathbf{P}$ by ν (the so-called *true anomaly*), so that $\mathbf{P}\cdot\hat{r} = P\cos\nu$. Then

$$\frac{h^2}{GMr} = 1 + \frac{P}{GM}\cos\nu \tag{A.11}$$

This is called the *equation of motion*. Note that this is the general equation for a *conic section* in polar coordinates for the case that one of the foci is located at the origin, which is

$$r = \frac{q}{1+e\cos\nu}.$$

From this we see that the excentricity e is

$$e = \frac{P}{GM}$$

For the case of an ellipse we have

$$\frac{h^2}{GM} = q = a(1-e^2)$$

So the motion of one of the bodies is in an ellipse with the other body in one of the foci. This is *Kepler's first law*. The geometry is illustrated in Fig. A.3. The position in an orbit follows from the value of the true anomaly ν from

$$r = \frac{a(1-e^2)}{1+e\cos\nu} \tag{A.12}$$

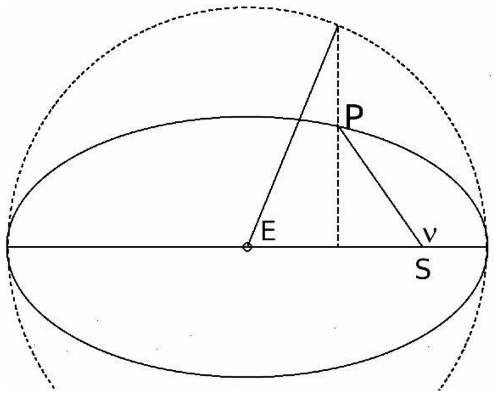

Fig. A.3 The orbit of a planet, asteroid or comet (P) is an ellipse with the Sun (S) in one of the foci. The true anomaly is ν and the construction shows the definition of the excentric anomaly E. For more explanation see the text. Figure by the author.

Now define the *excentric anomaly E* as

$$r = a(1 - e\cos E)$$

Differentiate

$$\dot{r} = ae\frac{dE}{dt}\sin E$$

and substitute this

$$dt = \sqrt{\frac{a^3}{GM}}(1 - e\cos E)dE.$$

Integrate over the full orbit

$$\int_0^T dt = \int_0^{2\pi}\sqrt{\frac{a^3}{GM}}(1 - e\cos E)dE.$$

So

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

This is *Kepler's third or harmonic law*, which can also be written as

$$\frac{T^2}{a^3} = \frac{4\pi^2}{G(m_1 + m_2)}. \quad \text{(A.13)}$$

Suppose you want to find the position of a planet in the orbit at a certain time. Then

you need to know the angle ν, the 'true anomaly'. The object moves faster in its orbit when it is closer to the Sun, so ν does not change uniformly with time. Kepler suggested the method to solve for ν as shown in Fig. A.3, where the angle E has been defined. This angle is called 'eccentric anomaly'. If you know E, ν is easy to find (and vice versa!) with:

$$\sqrt{\frac{1-e}{1+e}}\tan\left(\frac{\nu}{2}\right)=\tan\left(\frac{E}{2}\right),$$

where e is the eccentricity of the orbit (0 for a circle, 1 for a line). But E also does not increase uniformly with time. Therefore, Kepler defined an imaginary *mean anomaly*, M, which does run uniformly with time, but cannot be constructed in the figure. If the period is P and $T_\circ$ the time of the last perihelion passage (passing the point closest to the Sun), then M is by definition:

$$M=\frac{2\pi}{P}(t-T_\circ).$$

If you know the time t, M is easy to calculate. But then you have to find E and for that you have to use Kepler's equation, which connects M and E:

$$M=E-e\sin E. \tag{A.14}$$

If you know E, it's easy to calculate M by substitution, but that is not what you want. In practice you want to do the opposite and that is much more difficult. The difference between M and E, both of which go through 360° in the period P, is small when the eccentricity is small, but with asteroids, e is sometimes quite different from 0, and with a comet, the orbit is usually very much elongated and e is sometimes close to 1. E and M can sometimes be very different in parts of the orbit.

There are two general ways of tackling this problem. The first is by *iteration*, i.e. repeatedly calculating a better approximation until the result is sufficiently accurate. An example: write the equation as $E=M+e\sin E$ and take $E=M$ as a first approximation. Substitute that in the equation and then a better approximation is found as $E'=M+e\sin M$. Substitute that again and then $E''=M+e\sin E'$ is a still better approximation, and so on, until there is no significant improvement. If e is significantly different from 1, this may take a long time.

The second method is by *series expansion*. An example of this is the solution proposed by Joseph-Louis Lagrange:

$$E=M+\sum_{n=0}^{\infty}a_n e^n,\ \text{with}$$

$$a_n=\frac{1}{2^{n-1}n!}\sum_{k=0}^{n/2}(-1)^k\binom{n}{k}(n-2k)^{n-1}\sin[(n-2k)M]\ \text{ and}$$

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}=\frac{n(n-1)(n-2)....(n-k+1)}{k!}.$$

Calculate the terms until the improvement in the sum is no longer significant. For an asteroid in a fairly eccentric orbit, that may take some time.

The solution proposed by Kapteyn falls into the second category.

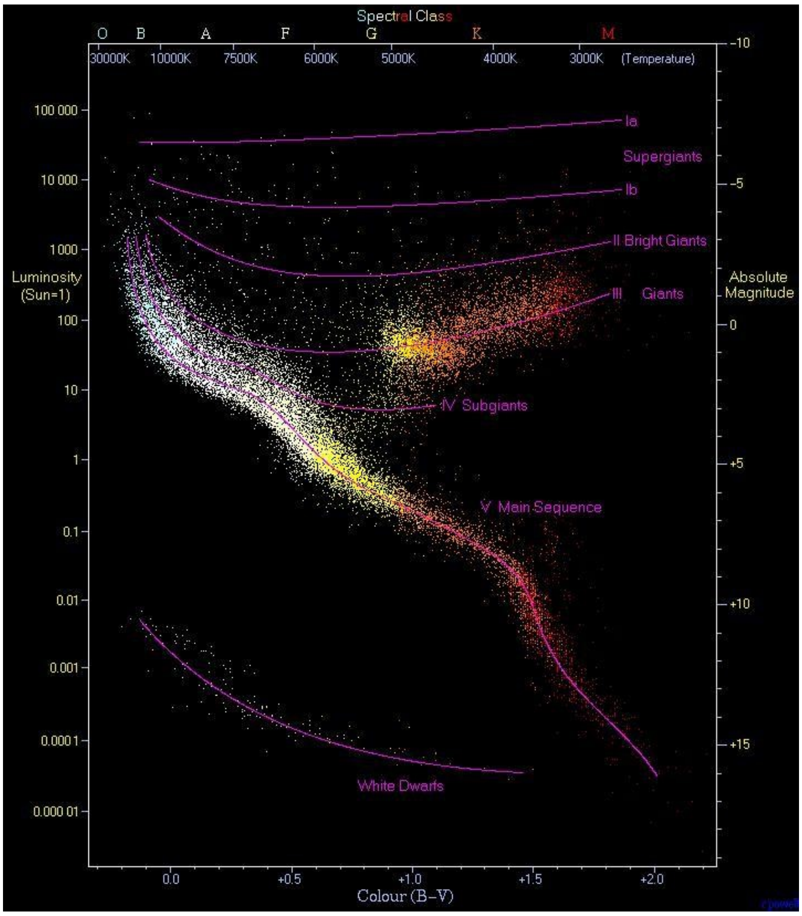

Fig. A.4 Hertzsprung-Russell diagram of stars. On the horizontal axis the color index $(B-V)$ (bottom) and at the top both spectral type and surface temperature in Kelvin. The vertical axis has on the right the absolute magnitude and on the left the luminosity expressed in that of the Sun. The lines and Roman numerals indicate luminosity classes. See also the text. From *An Atlas of the Universe* [3], with permission.

A.5 Stellar evolution

The fundamental diagram in astrophysics is named after the Danish astronomer Ejnar Hertzsprung and American Henri Norris Russell (see Fig. A.4). It is a diagram that relates the temperature at the surface of stars to the luminosity. The vertical axis is the luminosity, but for that also the absolute magnitude can be used. On the horizontal axis we have the temperature at the surface of the star, but one can also use for this the color index, for example the difference between the magnitudes of the star in a blue (B) and a visual (V) band. If the star is relatively bright in blue, then the star is relatively hot on the surface. But the spectral type can also be used for this, because the absorption lines in the spectrum of the star are created by atoms or ions in the outer parts, which absorb light at specific wavelengths; which atoms or ions are present and which lines are prominent, is strongly influenced by the temperature and also other physical conditions. Spectral types are indicated by the letters $OBAFGKM$ and a decimal subdivision within that. By definition, the color index $(B-V)$ is zero for an $A0$ star. O-stars are relatively blue; along the sequence from top-left to bottom-right, the color becomes increasingly redder, while the temperature at the surface of the star decreases. In O-stars, the lines of ionized helium are strongest, in B-stars those of neutral helium, in A-stars those of hydrogen, while in F-stars, G-stars and K-star lines of ionized calcium and metals gradually become stronger. In M-stars, lines of molecules (e.g., titanium oxide) are strong.

The radiation of a star is so-called black-body radiation; every body of a temperature T emits an amount of radiation, which depends on the temperature T. According to Stefan-Boltzmann's law, the energy integrated over all wavelengths per surface unit per second is equal to σT^4 (with the constant σ also named after Josef Stefan and Ludwig Eduard Boltzmann). Now, since the area of a sphere with radius R is equal to $4\pi R^2$, it follows that for the luminosity L of a star we have

$$L = 4\pi\sigma R^2 T^4. \tag{A.15}$$

Most stars lie along the Main Sequence from the top left to the bottom right. After being formed from a cloud of gas cloud, contraction and release of potential energy will make the inside of a star hotter, and as the density increases more energy is added than can be radiated away. At ten million degrees, nuclear reactions start, converting hydrogen into helium. The stars at the top are bright, heavy, and hot and for them this period is shortest, the M-stars are faint, light and cool and they live longer than the present age of the Universe. If L is the luminosity, T the surface temperature, M the mass, and τ the time on the Main Sequence, then the following applies in reasonable approximation

$$L \propto T^6 \quad L \propto M^3 \quad \tau \propto M^{-2}.$$

The derivation in very general terms is as follows (see also my Lecture Course *Introductory Astronomy* [2].

Take a column with a cross section ds in a star at distance r from the center and take the mass in the star within r equal to $M(r)$. A small element dr along the column than has a volume $dr\, ds$. On this element the gravitational force downward is (with ρ the density en G the gravitational constant)

$$P_- = \rho \frac{GM(r)}{r^2}\, ds\, dr.$$

This will have to be compensated by the difference on pressure over the element

$$P_+ = P\, ds - (P + dP)\, ds = -\frac{dP}{dr}\, dr\, ds.$$

So for *hydrostatic equilibrium* we have

$$-\frac{1}{\rho}\frac{dP}{dr} = \frac{GM(r)}{r^2}. \tag{A.16}$$

Now look at radiation and assume a radiation flux H. The energy absorbed in that element then is

$$E_- = \kappa H\rho \; ds\, dr,$$

with κ the absorption coefficient per unit mass. And since each photon has a momentum $h\nu/c$ (h is the Planck constant, ν the frequency and c speed of light), so the absorbed momentum is

$$P_- = \frac{\kappa}{c}H\rho\, ds\, dr.$$

For equilibrium this should be equal to the difference in radiation pressure $P_{\rm rad}$ over the element:

$$-dP_{\rm rad}\, ds = -dr\,\frac{dP_{\rm rad}}{dr}ds = \frac{\kappa}{c}H\rho\; dr\; ds.$$

With the radiation pressure $P_{\rm rad} = \frac{1}{3}aT^4$ with $a = 4\sigma/c$ this becomes

$$H = -\frac{c}{\kappa\rho}\frac{dP_{\rm rad}}{dr} = -\frac{4}{3}\frac{ac}{\kappa\rho}T^3\frac{dT}{dr},$$

and the energy that flows through the sphere with radius r becomes

$$L_{\rm r} = 4\pi r^2 H = -4\pi r^2\;\frac{4ac}{3}\;\frac{T^3}{\kappa\rho}\frac{dT}{dr}. \tag{A.17}$$

This is the *radiative transfer*.

Finally we have a *equation of state* for which we can take the ideal gas law

$$P = nkT = \frac{k}{\mu}\rho T. \tag{A.18}$$

Hydrostatic equilibrium, radiative transfer and equation of state together give the fundamental equations for stellar structure. The properties of matter (composition, ionization, etc.) are all included in the κ. For a full analysis also the energy production needs to be considered.

In a dimenensional analysis we can derive the basic proprotionalities in the relations of stars. This means that we write, e.g. for the temperature which is $T(0)$ in the center and $T(R)$ at the surface with $T(R) << T(0)$:

$$\bar{T} \approx \frac{T(0)+T(R)}{2} \quad ; \quad \frac{d\bar{T}}{dr} \approx \frac{T(R)-T(0)}{R} = \frac{-2\bar{T}+2T(R)}{R} \approx -\frac{2\bar{T}}{R},$$

and

$$\bar{\rho} = \frac{3}{4\pi}\frac{M}{R^3}.$$

So hydrodynamic equilibrium becomes

$$\frac{2\bar{P}}{\bar{\rho}R} \approx \frac{GM(\frac{1}{2}R)}{(\frac{1}{2}R)^2} = \frac{GM}{2R}.$$

We then can find

$$\bar{P} \approx \frac{GM}{4R}\bar{\rho} \quad ; \quad \bar{P} \approx \frac{k}{\mu}\bar{\rho}\bar{T} \quad ; \quad \bar{T} = \frac{\mu}{k}\frac{GM}{4R}$$

This is not bad, since for the Sun we then get $\bar{T} = 6 \times 10^6$K. Now ignoring proportionality constants we get

$$L \propto R^2 \frac{\bar{T}^3}{\bar{\rho}} \frac{\bar{T}}{R} \propto \frac{R\bar{T}^4}{\bar{\rho}} \propto R\left(\frac{M}{R}\right)^4 \frac{R^3}{M} \propto M^3. \tag{A.19}$$

This is the *mass-luminosity relation*, which has been found empirically to be about the same as this.

If we assume $T_{\text{surface}} \propto \bar{T}$, we find

$$T_{\text{surface}} \propto \bar{T} \propto \frac{M}{R} \propto \frac{L^{1/3}}{R},$$

and with the relation just found this gives

$$L \propto T^6_{\text{surface}}, \tag{A.20}$$

which is the *Main Sequence*.

Assume that the total energy a star can produce is proportional to its mass (say each star uses over its lifetime a similar fraction, about 10%) as fuel. Then the *lifetime* on the Main Sequence

$$\tau \propto \frac{M}{L} \propto M^{-2}. \tag{A.21}$$

The derivation of these formulae comes from my Lecture Course *Introductory Astronomy* [2].

Eventually all hydrogen in the central parts of a star will be used up, so that the core, which now consists entirely of helium, will be extinguished. Initially 'hydrogen burning' continues for a while in a shell around the core and the star becomes brighter. A star like the Sun then moves up along the sub- and the giant branches (along numbers IV and III in Fig. A.4). The temperature T at the surface drops and the star becomes redder, but according to the formula above, the radius R will then increase. Because energy is no longer produced there, the core contracts, but then gets hotter. When it becomes hot enough, helium burning will start there, converting it into carbon and oxygen. A star like the Sun is then on the 'clump' halfway along the line marked with III. Then the star expels its outer layers and forms a so-called planetary nebula. The star then cools down to a white dwarf on the line at the bottom of the figure. The matter becomes very compact, but at a certain moment the contraction stops because Fermi's exclusion principle and Heisenberg's uncertainty relationship together forbid that the electrons come even closer together. The uncertainty relationship says that the position and velocity of an electron (or other elementary particle) cannot be measured infinitely accurately at the same time. Then two different electrons can in principle not get so close together and have the same speed so accurately that they would be indistinguishable. The Fermi principle prohibits this and so matter cannot become more compact than when all particles can just be distinguished in position and speed. This is called degeneration pressure. This is explained in my Lecture Course *Introductory Astronomy* [2].

In a more massive star than the Sun, the pressure in the core becomes so great that the center continues to contract; the star then becomes so hot that even heavier chemical elements are formed. But in the long run this process also stops and the central parts contract even further. The gravitational contraction force then is so great that the degeneration pressure can no longer compensate for it; then the electrons are, as it were, merged into the protons and form neutrons. This process is so fast that an enormous pressure wave propagates through the star, blowing itself up like a supernova. More chemical elements are formed and this material is thrown out into space. The remaining central parts are then a neutron star, which is in equilibrium under the influence of the degeneration pressure of the neutrons, or, if the intial star mass is even large and the contraction forces too strong, a black hole.

A.6 The 'star ratio'

This concept was used by Edward Pickering in 1903 to obtain information about the distribution of stars in space. Assume that all stars have the same intrinsic luminosity and are uniformly distributed in space. Take a certain distance from the Sun, then all stars at that distance have the same apparent brightness or magnitude, say m. Stars of a magnitude weaker, i.e. $m+1$, in the sky are according to the definition of the magnitude scale a factor $\sqrt[5]{100} = 2.512$ fainter and have to be therefore a factor $\sqrt{2.512}$ further away. The stars between apparent magnitude m and $m+1$ fill a shell. Now take stars that are another magnitude fainterer, so $m+2$. They are a factor of 2.512 fainter than the stars of magnitude $m+1$ and are a factor of $\sqrt{2.512}$ further away. The shell between $m+1$ and $m+2$ is a factor 2.512 larger than the one between m and $m+1$ and the thickness is a factor $\sqrt{2.512}$ larger. So the volume of the shell is a factor $2.512 \times \sqrt{2.512} = 2.512^{3/2}$ larger and so is the number of stars in it.

This means that the number of stars between magnitude $m+1$ and $m+2$ should be the same factor $2.512^{3/2} = 3.981$ times larger than the number between m and $m+1$. And this holds for every value of m. This ratio, the factor of 3.981, was called the 'star ratio' by Pickering and served as a reference value. In practice, he did not use the factor 3.981 itself, but actually its logarithm, which is exactly 0.6.

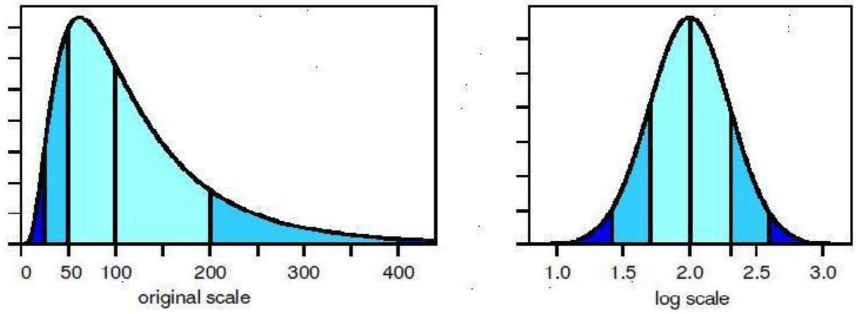

Fig. A.5 On the left a log-normal probability distributions. On the right the same distribution but plotted on a logarithmic horizontal scale. After [4].

In reality on the sky, this ratio is smaller than this theoretical value; this may be due to changes in the density of the stars with changing distances or to the distribution of their intrinsic luminosity, but also to absorption of starlight by dust in interstellar space.

A.7 Skew probability distributions

An example of a skew probability distribution that Kapteyn studied, is the so-called 'log-normal' distribution. The well-known Gaussian probability distribution is

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[x-M]^2\right\}. \tag{A.22}$$

The property M is the central value and σ is called the dispersion. The log-normal probability distributions then is

$$P'(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[\ln(x)-M]^2\right\}. \tag{A.23}$$

An example of a log-normal distribution has been drawn on the left in Fig. A.5. This is for the values $M = 100$ and $\sigma = 2$. To the right the same distribution on a logarithmic (base 10) scale on the x axis, so that it becomes the normal Gaussian distribution (then the parameters become $M = 4.61$ and $\sigma = 0.693$). The color changes with integer values of σ (34.1% of the total from the median up to 1σ, 13.6% from 1σ up to 2σ, etc.).

In normal distribution, the mode (the maximum or most common value), the median (half larger and half smaller) and the regular average have the same value (in the figure 10^2 or 100). At the log-normal, the average is $e^M + \frac{1}{2}\sigma^2 = 127.8$, but the mode is at $e^{M-\sigma^2} = 62.2$ and the median at $e^M = 100$.

Kapteyn went on to study a general family of skewed distributions characterized by

$$P_{\text{JCK}}(x) = \frac{1}{F'(x)\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left[F(x)-M\right]^2\right\}, \qquad \text{(A.24)}$$

where $F'(x)$ is the first derivative $dF(x)/dx$ of a function $F(x)$. Kapteyn chose the function $F'(x) = (x+\kappa)^q$. The special case $\kappa = 0$, $q = -1$, which he used extensively, is the log-normal distribution, which can be seen when you note that $F = \ln x$, $F' = 1/x$.

An excellent general discussion of the theory of log-normal dsitributions has been given by Eckhard Limpert and collaborators [4]. An extensive, technical discussion of Kapteyn's work in the field of statistics has been published an article by Ida H. Stamhuis and Eugene Seneta [5].

The degree of correlation (linear regression) between two parameters among n objects (e.g. mass and luminosity of n stars), x_i and y_i ($i = 1$ to n), is often described with the 'Pearson product moment correlation-coefficient', which makes use of the second moments of the distribution. These are

$$M_{xx} = \sum_{i=1}^{n}(x_i-\bar{x})^2 \;\;;\;\; M_{yy} = \sum_{i=1}^{n}(y_i-\bar{y})^2 \;\;;\;\; M_{xy} = \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}).$$

Here $\bar{x}$ and $\bar{y}$ are the mean values of x and y. The definition of the correlation-coefficient then is

$$r = \frac{M_{xy}}{\sqrt{M_{xx}M_{yy}}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_x}\right)\left(\frac{x_i-\bar{y}}{\sigma_y}\right), \qquad \text{(A.25)}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \;\;;\;\; \sigma_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

and

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \;\;;\;\; \sigma_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2}.$$

The absolute value $|r|$ then varies from 0 in the case of absence of any correlation to 1 in case of complete correlation.

A.8 The fundamental laws of statistical astronomy

Here I use the notation used in the thesis of Kapteyn's student Willem Schouten, written in 1918 [6].

$D(r)$, density (number of stars per unit volume) of stars at distance r.
h_m, the apparent birghtness (in energy per unit of time per unit of area) of a star of magnitude m; $h_\mathrm{m} = 1$ for $m = 0$.
i, luminosity, corresponding to absolute magnitude M.
$N(h_\mathrm{m})dh_\mathrm{m}$, observed number of stars of magnitude m, that is between h_m and $h_\mathrm{m} + dh_\mathrm{m}$.
$\Pi(h_\mathrm{m})$, the average parallax of stars of magnitude m.
$\varphi(i)di$, percentage of stars between i and $i = di$.

Using $i = h_\mathrm{m}r^2$ and $di = r^2 dh_\mathrm{m}$, the functions $D(r)$ and $\varphi(i)di$ then follow from a non-trivial, simultaneous inversion of two integrals:

$$N(h_\mathrm{m}) = 4\pi \int_0^\infty D(r)\varphi(h_\mathrm{m}r^2)r^4 dr, \tag{A.26}$$

$$N(h_\mathrm{m})\Pi(h_\mathrm{m}) = 4\pi \int_0^\infty D(r)\varphi(h_\mathrm{m}r^2)r^3 dr. \tag{A.27}$$

References

1. *A brief history of time: From the Big Bang to black holes*, by Stephen W. Hawking, Bantam Dell, ISBN 978-0-553-10953-5 (1988).
2. See my lectures on *Introductory astronomy* on www.astro.rug.nl/∼vdkruit/Inleiding.html.
3. *The Hertzsprung Russell Diagram* by Richard Powell, www.atlasoftheuniverse.com/hr.html en en.wikipedia.org/wiki/File:HRDiagram.png. ‘This file is licensed under the Creative Commons Attribution-Share Alike 2.5 Generic license.’
4. E. Limpert, W.A. Stahel & M. Abbt, *Lognormal distributions across the sciences: Keys and clues*BioScience 51, 341-352 (2001) (with permission).
5. I.H. Stamhuis and E. Seneta, Pearson’s statistics in the Netherlands and the Astronomer Kapteyn, *International Statistical Review* 77, 96-117 (2009).
6. Willem Johannes Adriaan Schouten (1893–1971): *On the determination of the principal laws of statistical astronomy*, PhD Thesis, University of Groningen, 1918.