

GENERAL QUALIFYING EXAM SOLUTIONS: PHYSICS

Jessica Campbell, Dunlap Institute for Astronomy & Astrophysics (UofT)

Contents

1 Physics	2
1.1 Question 1	2
1.2 Question 2	9
1.3 Question 3	14
1.4 Question 4	15
1.5 Question 5	18
1.6 Question 6	22
1.7 Question 7	27
1.8 Question 8	32
1.9 Question 9	39
1.10 Question 10	43
1.11 Question 11	44
1.12 Question 12	53
1.13 Question 13	54
1.14 Question 14	56
1.15 Question 15	57
1.16 Question 16	67
1.17 Question 17	68
1.18 Question 18	69
1.19 Question 19	70
1.20 Question 20	79
1.21 Question 21	80
1.22 Resources	81

1 Physics

1.1 Question 1

Draw the geometry of gravitational microlensing of one star by another, and estimate the angular displacement of the background star's image.

1.1.1 Short answer

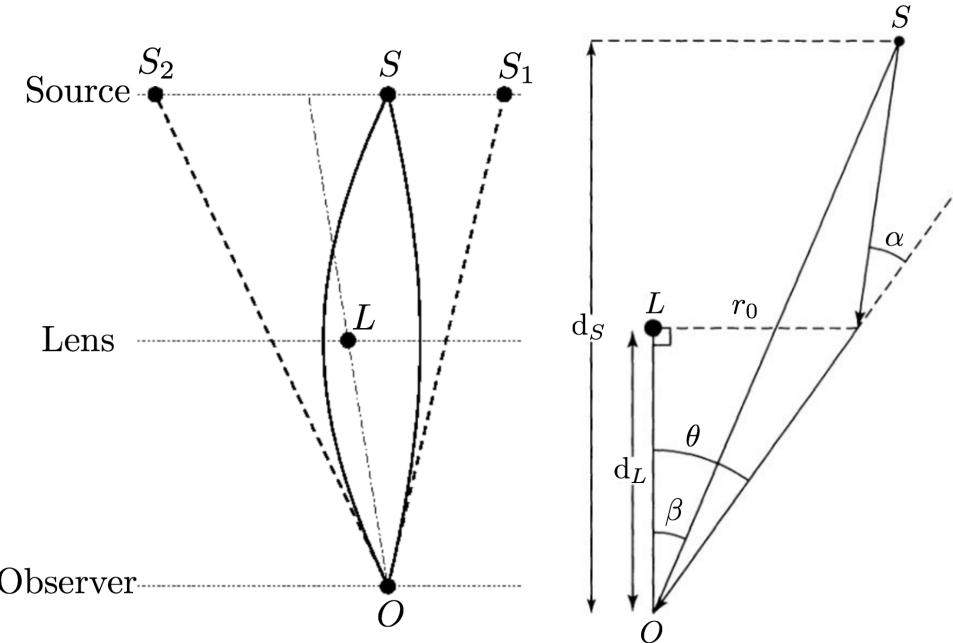


Figure 1: (**Left**): Setup of a gravitational lens. Figure taken from Abdo. (**Right**): The geometry for a gravitational lens. The path of light is taken from a source at point S , as it is deflected through an angle α , by the gravitational lens due to a point mass M at point L . Figures adapted from Carroll & Ostlie (2007).

The angular displacement of the background star's image from a mass M is

$$\alpha = \frac{4GM}{r_0 c^2} \text{ [rad]},$$

which includes the factor of 2 as a correction from the “Newtonian” value.

1.1.2 Additional context

The first written account of the deflection of light by gravity appeared in the article “On The Deflection Of Light Ray From Its Straight Motion Due To The Attraction Of A World Body Which It Passes Closely,” by Johann Soldner in 1804. In his article Soldner predicted that a light ray passing close to the solar limb would be deflected by an angle $\alpha = 0.84$ arcsec. More than a century later, in 1919, Albert Einstein directly addressed the influence of gravity on light. At this time Einstein’s General Theory of Relativity was not fully developed. This is the reason why Einstein obtained (unaware of the earlier result) the same value for the deflection angle as Soldner had calculated with Newtonian physics. In this paper Einstein found $\alpha = 2GM_\odot/c^2R_\odot = 0.83$ arcsec for the deflection angle of a light ray grazing the Sun (here M_\odot and R_\odot are the mass and radius of the Sun, c and G are the speed of light and the gravitational constant respectively).

Einstein emphasized his wish that astronomers investigate this equation. In 1913 Einstein contacted the director of Mount Wilson observatory, George Hale, and asked if it will be possible to measure the positions of stars near the Sun during the day in order to establish the angular deflection effect of the Sun. The first observational attempt to test Einstein’s prediction for the deflection angle was in 1914. This attempt was never accomplished which was fortunate for Einstein whose predicted value for the angular deflection was actually wrong. With the completion of the General Theory of Relativity in 1916, Einstein was the first to derive the correct formula for the deflection angle α of a light passing at a distance r from an object with mass M as

$$\alpha = \frac{4GM}{c^2} \frac{1}{r} [\text{arcsec}].$$

The additional factor of two (compared to the “Newtonian” value) is due to the spatial curvature which is missed if photons are just treated as particles.

For the Sun, Einstein obtained

$$\alpha_{\odot} = \frac{4GM_{\odot}}{c^2} \frac{1}{R_{\odot}} = 1.74 [\text{arcsec}].$$

This value was first confirmed to within 20% by Arthur Eddington and his group during a Solar total eclipse in 1919. Recently this value was confirmed to within 0.02%.

In 1924 Chwolson mentioned the idea of a “factious double star.” He also mentioned the symmetric case of one star exactly behind another star, resulting in a circular image. In 1936 Einstein also reported about the appearance of a “luminous” circle of perfect alignment between the source and the lens, such a configuration is called **Einstein Ring**. In 1937 Fritz Zwicky pointed out that galaxies are more likely to be gravitationally lensed than a star and that one can use the gravitational lens as natural telescope. **The geometry of gravitational lensing:** Gravitational lensing results when light follows the straightest possible worldline (ie., a **geodesic**) as it travels though the curved spacetime around a massive object. It is analogous to the normal refraction of light by a glass lens that occurs as the light crosses the lens surface, passing from one index of refraction n to another, where $n \equiv c/v$ is just the ratio of the speed of light in a vacuum to its speed v in the medium.

The apparent speed of light, the rate at which the spatial coordinates of a photon change, is called the **coordinate speed of light**. Starting with the Schwarzschild metric with $ds = 0$,

$$0 = \left(cdt \sqrt{1 - \frac{2GM}{rc^2}} \right)^2 - \left(\frac{dr}{\sqrt{1 - \frac{2GM}{rc^2}}} \right)^2 - (rd\theta)^2 - (2\sin\theta d\alpha)^2,$$

we can calculate the coordinate speed of a vertically travelling photon. Inserting $d\theta = d\alpha = 0$ shows that, in general, the coordinate speed of light in the radial direction is

$$\frac{dr}{dt} = c \left(1 - \frac{2GM}{rc^2} \right) = c \left(1 - \frac{r_S}{r} \right) [\text{m s}^{-1}],$$

When $r \gg r_S$, $dr/dt \simeq c$, as expected in flat spacetime. However, at $r = r_S$, $dr/dt = 0$.

Thus, the effective “index of refraction” is

$$n = \frac{c}{dr/dt} = \left(1 - \frac{2GM}{rc^2} \right)^{-1} \simeq 1 + \frac{2GM}{rc^2} [\text{dimensionless}]$$

for radially travelling light, assuming that $r_S/r \ll 1$. At a distance of 10^4 pc from a galaxy with a mass of $10^{11} M_{\odot}$, the effective index of refraction is $n = 1 + 9.6 \times 10^{-7}$. (Of course, the light passing by the point mass will never be travelling exactly radially. This was merely used to estimate the magnitude of the effect of gravity in a gravitational lens.) Obviously, the deviation of the light from a straight line will be extremely small.

Figure 1 shows the setup of a gravitational lens (left) as well as its detailed geometry (right). The path of light is taken from a source at point S , as it is deflected through an angle α , by the gravitational lens due to a point mass M at point L .

The angular deviation of a photon passing a distance r_0 (very nearly the distance of closest approach) from a mass M is

$$\alpha = \frac{4GM}{r_0 c^2} [\text{rad}],$$

which includes the factor of 2 as a correction from the “Newtonian” value. The distance to the source is $d_S/\cos\beta \simeq d_S$ where $\beta \ll 1$, and d_L is the distance to the lensing mass. It is then a matter of simple trigonometry to show that the angle θ between between the lensing mass and the image of the source must satisfy the equation

$$\theta^2 - \beta\theta - \frac{4GM}{c^2} \left(\frac{d_S - d_L}{d_S d_L} \right) = 0$$

where θ and β are measured in radians.

This quadratic equation indicates that for the geometry shown in Figure 1 (left) there will be two solutions for θ , and so two images will be formed by the gravitational lens. Designating these solutions as θ_1 and

θ_2 , these angles can be measured observationally and then used to find the values of β and M . The results are

$$\beta = \theta_1 + \theta_2 \text{ [rad]}$$

and

$$M = -\frac{\theta_1 \theta_2 c^2}{4G} \left(\frac{d_S - d_L}{d_S d_L} \right) [\text{M}_\odot].$$

Referring back to Figure 1 (left) note that $\beta = \theta_1 + \theta_2$ implies θ_1 and θ_2 have opposite signs. As a result, two images are formed on opposite sides of the gravitational lens, so M will be positive.

Einstein rings and crosses: If a quasar or other bright source lies exactly along the line of sight to the lensing mass, then it will be imaged as an **Einstein ring** encircling the lens (this phenomenon was described by Einstein in 1936). In this case, $\beta = 0$ in Figure 1 (left) and so the quadratic equation for θ can be solved immediately for the angular radius of the Einstein ring:

$$\theta_E = \sqrt{\frac{4GM}{c^2} \left(\frac{d_S - d_L}{d_S d_L} \right)} \text{ [rad].}$$

Of course, for a point source, the change of an exact alignment with the lensing mass is essentially zero. For an extended source, the requirements for an Einstein ring are that $\beta < \theta_E$ and that the line of sight through the lensing mass must pierce the extended source.

The value of θ_E can be calculated for any gravitational lens, regardless of the alignment of the lens and the source. Although the image may not be a ring, θ_E does provide a useful parameter for describing the properties of any gravitational lens. If $\beta < \theta_E$, as shown in Figure 1 (left), there will be two images formed by the point mass. If $\beta \gg \theta_E$, the position and brightness of the source are only slightly altered, but a secondary image appears close to the lensing mass that is reduced in angular size by a factor of $(\theta_E/\beta)^4$.

A point mass is clearly a crude representation of an actual galaxy. A better model of the lensing galaxy is provided by an isothermal sphere around a central core, similar to the model used for the central bulge of the MW. Another improvement is to depart from spherical symmetry and use an isothermal ellipsoid, which can produce either three or five images (an extended distribution of mass will produce an odd number of images.)

Luminous arcs in galaxy clusters: Another striking example of gravitational lensing is the formation of arcs by light passing through a cluster of galaxies. One such arc is in the cluster Abell 370. Up to 60 additional **arclets** and several distorted distant background galaxies have also been observed in that cluster. The source of the large arc must be a resolved object such as a galaxy rather than the star-like nucleus of a QSO. According to one model of Abell 370, the lensing mass (visible galaxies and DM) needed to produce the images in Abell 370 is at most $5 \times 10^{14} \text{ M}_\odot$. Taken with the combined luminosity of a few $\times 10^{11} \text{ L}_\odot$ for the lensing galaxies, this implies a mass-to-light ratio of at least $1000 \text{ M}_\odot/\text{L}_\odot$, indicating the presence of large amounts of DM.

Time variability of multiple images: An interesting effect occurs when the source for a pair of images increases its luminosity. Because the light from the source takes different paths on its way to the observer, there will be a time delay between the brightening of the lensed images. A time delay of about $1.4 - 1.5 \text{ yr}$ has been measured for the original double quasar Q0957+561. It turns out that the time delay is also inversely proportional to the Hubble constant. This offers a way of determining the value of H_0 that is independent of any other distance measurement.

Microlensing: Microlensing refers to the special case of gravitational lensing where the multiple images produced are too close together on the sky to be observed as separate images. However, the lensing can still be detected because these multiple images appear as a single object of increased apparent brightness. Although this is not detectable in a one-off observation (since we do not know the ‘normal’ brightness of the source), with the passage of time the lens moves across the Earth-source line and the amount of brightening changes. Typically the source will appear to brighten, reach a maximum and then fade symmetrically back to normal over the course of a few weeks or months; this is called a **microlensing event**. The major application of microlensing, suggested by Paczynski in 1986, is in the search for DM which is strongly believed to exist from rotation curves of spiral galaxies etc. Since the lensing effect depends only on lens mass, it can be used to search for very faint or invisible objects such as brown dwarfs, neutron stars, old white dwarfs or black holes, which might make up DM. These are collectively known as **massive compact halo objects** or MACHOs, in contrast to the hypothetical **weakly interacting massive particles** or WIMPs.

To understand the basics of microlensing, consider a small massive object (the lens) situated exactly on the line of sight from Earth to a background star and consider a number of light rays radiating from the

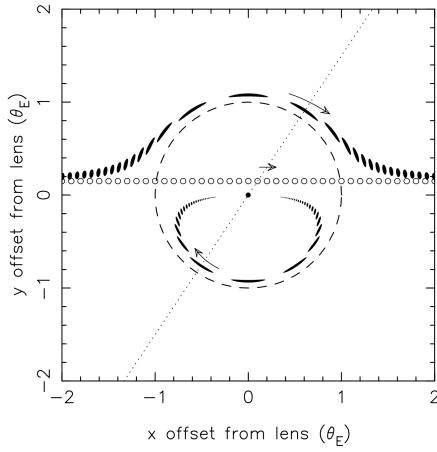


Figure 2: A microlensing event seen at ‘perfect’ resolution. The axes show angular offsets on the sky from the lens (central dot) in units of the Einstein angle; the dashed circle is the Einstein ring. The series of small open circles shows the ‘true’ source position at successive timesteps. For each source position, there are two images (solid blobs) collinear with the lens and source, as indicated by the dotted line; the arrows illustrate their motion. Figure taken from Encyclopedia of Astronomy and Astrophysics (2001).

star passing the lens at different distances and being bent towards the lens. Since *the bending angle for a light ray increases with decreasing distance from the lens*, it is clear that there is a unique ‘miss distance’ such that the ray will be deflected just enough to hit the Earth; this distance is called the **Einstein radius**. By rotational symmetry about the Earth-star axis, an observer on Earth with perfect resolution would see the star lensed into an annulus centered on its ‘true’ position, called an **Einstein ring**. As the lens is moved slightly off the line of sight (e.g., by 0.1 Einstein ring radii), the Einstein ring splits into two banana-shaped arcs, one on the same side of the lens as the source, one on the opposite side. As the lens moves further off (more than 1 Einstein radius), the arcs become more circular, the ‘opposite-side’ arc fades very rapidly and the ‘same-side’ arc turns into a slightly deflected and nearly circular image of the star. Figure 2 illustrates a sequence of such images for a typical microlensing event.

Although the perfect alignment giving the Einstein ring will rarely occur in practice, it is still a very important concept because the size of the hypothetical Einstein ring sets the length scale over which substantial brightening will occur. As we will see, for a typical lens in our Galaxy the radius of the Einstein ring r_E is roughly $8\sqrt{M/M_\odot}$ AU, where M is the lens mass. Knowing this scale allows us to understand most of the general characteristics of microlensing: it is extremely small compared with the typical distance to a lens, so the angular separation of the two images will be too small to resolve, hence the ‘micro’lensing. However, it is considerably larger than either the size of a star or the size of a MACHO, so we can usually approximate the lens and source as point-like, which leads to a simple prediction for the light curve shape. Also, r_E is very small compared with the typical separation of objects in the Galaxy, which implies that microlensing will be a very rare phenomenon. Another notable feature is that r_E is proportional to the square root of the lens mass. This means that the area of sky ‘covered’ by a lens (at fixed distance) is proportional to its mass, so the total fraction of sky covered depends only on the total mass density in lenses, not the individual lens masses. This fraction is called the **optical depth** τ , and is $\sim 10^{-6}$ for Galactic microlensing. The duration for a microlensing event is given by the time for the lens to move by $2r_E$ relative to the Earth-star line; for typical Galactic speeds of 200 km s^{-1} , this is $\sim 130 \text{ days} \times \sqrt{M/M_\odot}$.

For perfect alignment, simple geometry gives the (small) deflection angle of the light ray meeting Earth as $\alpha = r_E/d_L + r_E/(d_S - d_L)$, where d_L is the observer-lens distance and d_S is the source distance. Requiring this to equal the general relativity deflection, $\alpha = 4GM/c^2$, we obtain

$$r_E = \sqrt{\frac{4GMd_L}{c^2} \left(\frac{d_S - d_L}{d_S} \right)} [\text{AU}].$$

The angular Einstein radius is just $\theta_E \equiv r_E/d_L$. If we now introduce a small offset of the lens by a distance b from the Earth-source line (i.e., an angle $\delta \equiv b/d_L$, a simple generalization gives the two image angular positions (relative to the lens) as

$$\theta_{\pm} = 0.5 \left[\delta \pm \sqrt{\delta^2 + 4\theta_E^2} \right] [\text{rad}].$$

Since lensing preserves surface brightness, the magnification A_i of each image is given by the ratio of image to source areas, which for a ‘small’ source and any axisymmetric lens is just

$$A_i = \left| \frac{\theta_i \, d\theta_i}{\delta \, d\delta} \right| [\text{dimensionless}].$$

For a point lens, this leads to a total observed magnification as the sum of the two image magnifications,

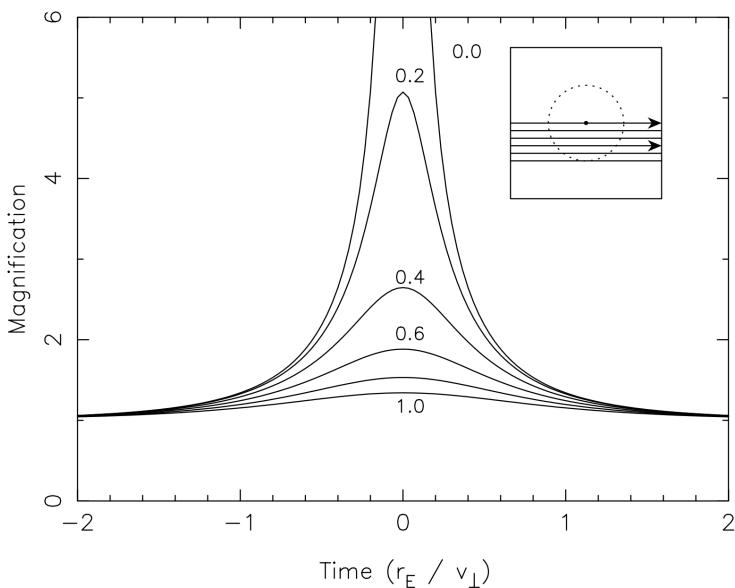


Figure 3: Microlensing event light curves (magnification versus time) for six values of the impact parameter $(\delta/\theta_E)_{\min} = 0.0, 0.2, \dots, 1.0$ as labelled. Time is in units of the Einstein radius crossing time r_E/v_{\perp} . The inset illustrates the Einstein ring (dotted circle) and source paths relative to the lens (dot) for the six curves. Figure taken from Encyclopedia of Astronomy and Astrophysics (2001).

$$A = A_+ + A_- = \frac{(\delta/\theta_E)^2 + 2}{(\delta/\theta_E)\sqrt{(\delta/\theta_E)^2 + 4}} \text{ [dimensionless].}$$

This behaves as $A \approx (\delta/\theta_E)^{-1}$ for $(\delta/\theta_E) \lesssim 0.5$, so the magnification may be large, but as $A \approx 1 + 2(\delta/\theta_E)^{-4}$ for $(\delta/\theta_E) \gtrsim 2$, so the magnification rapidly becomes negligible at large (δ/θ_E) . For uniform motions, we will have $(\delta/\theta_E)(t) = \sqrt{(\delta/\theta_E)_{\min}^2 + [v_{\perp}(t - t_0)/r_E]^2}$ where v_{\perp} is the transverse velocity of the lens relative to the line of sight and $(\delta/\theta_E)_{\min}$ is the value of (δ/θ_E) at closest approach, which occurs at time t_0 . Substituting this into the equation for the magnitude A gives the light curve $A(t)$; some examples are shown in Figure 3.

Observations of microlensing: To search for microlensing by MACHOs, since it is so rare we need millions of source stars and they should be far enough away to give a good path length for lensing, but not so far away that the individual stars are faint. The obvious targets for dark-matter lenses are the LMC and SMC, the largest of our MW's many satellite galaxies, since the sight line to them passes mainly through the Galactic halo. The Galactic bulge is also an interesting target, although here because of the high stellar density the microlensing should be mainly due to intervening faint stars rather than dark matter. Around 1991, several teams of astronomers called MACHO, EROS and OGLE began large projects using wide-field 1 m class telescopes to search for microlensing events. The novel feature of these experiments is their massive data volume, requiring large amounts of disk and tape storage and special high-speed photometry programs. The expected number of microlensing events is much smaller than the fraction of intrinsic variable stars, but fortunately microlensing has many strong ‘signatures’ that are distinct from all previously known types of variable star: the most important is the ‘uniqueness’, (i.e., any given star should be microlensed at most once in a human lifetime), while almost all variable stars are periodic or quasi-periodic. Also, most events should have a symmetrical shape as in Figure 3, and they should be achromatic (the source’s colour should not change during the event), since lensing is independent of wavelength.

The MACHO, EROS and OGLE teams announced their first candidate microlensing events near simultaneously in late 1993 (Figure 4). Since then, they have accumulated much more extensive datasets, and completed detailed calculations of their microlensing detection efficiency as a function of event duration; thus a number of general results have become clear:

1. Hundreds of events have been observed towards the Galactic bulge. They are distributed ‘randomly’ across the color-magnitude diagram, and their distribution of peak magnifications matches the theoretical prediction, which provides a very convincing check that microlensing works as predicted and that the experiments can detect it. The event rate towards the bulge is greater than initially predicted by Galactic models, favoring a barred structure for the Galaxy. The distribution of timescales is consistent with most of the lenses being low-mass stars as expected and can provide constraints on the stellar mass function at low masses.
2. Towards the Magellanic Clouds, no ‘short’ events (timescales from a few hours up to 20 days) have been seen by any group. This places strong limits on ‘Jupiters’ in the dark halo: specifically, compact objects in the mass range $10^{-6} - 0.05 M_{\odot}$ contribute less than 10% of the dark matter

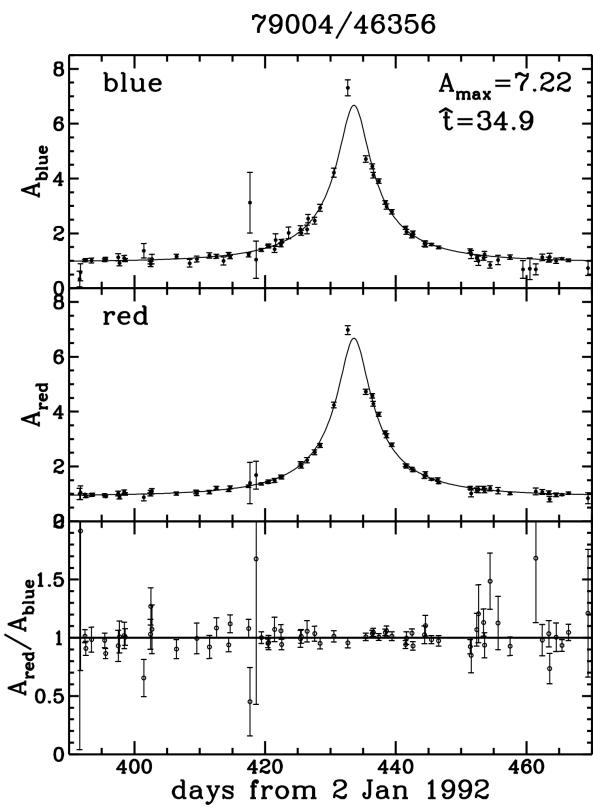


Figure 4:
The first LMC microlensing candidate from the MACHO project. (Expanded view: 6 yr of constant data are outside the plot). Upper and middle panels show brightness versus time in blue and red passbands respectively, in units of the baseline value. Points with error bars are observations, the curve is the best microlensing fit. The lower panel shows the ratio of red/blue flux, illustrating the lack of color change. Figure taken from Encyclopedia of Astronomy and Astrophysics (2001).

around our Galaxy. This is a very important result, as these objects were previously thought to be the most plausible form of baryonic DM, and (for masses below $0.01 M_\odot$) they would have been virtually impossible to detect directly.

3. Approximately 18 events with durations 30 – 200 days have been observed towards the Magellanic Clouds. The implied optical depth is well above expectation from ‘known’ stars, and is roughly 1/3 of the value expected ($\tau \approx 5 \times 10^{-7}$) for an all-MACHO dark halo. The relatively long durations imply lens masses of roughly $0.3 – 0.7 M_\odot$, which is a puzzle.
4. Many events are now detected in real time and announced on the world-wide web. This enables more frequent monitoring of ongoing events, giving more precise tests of the shape. This has attracted considerable interest since, where the lens is a low-mass star (probably the case for most of the events towards the Galactic bulge), a planet around the lens star may give rise to a short-duration ‘blip’ superimposed on the smooth microlensing light curve: this method is potentially sensitive to low-mass planets down to a few Earth masses, below the range accessible to radial-velocity planet searches. Two groups called PLANET and MPS are now searching for this effect.

Current questions in microlensing: The nature of the objects giving rise to the lensing towards the Magellanic Clouds is a significant puzzle at present; the inferred lens masses are well above $0.1 M_\odot$, which means that they cannot be ordinary hydrogen-burning stars in a spherical halo, as these would be visible in large numbers in deep images (e.g., the Hubble Deep Field). There are two classes of solution: either the lenses are in the halo but are much fainter (e.g., old white dwarfs or possibly primordial black holes), or they are low-mass stars but in a non-halo population (i.e., preferentially concentrated in the foreground of the LMC). The white-dwarf solution has significant problems associated with early metal production but has received some support recently from a possible detection of proper motions of faint objects in the Hubble Deep Field. Various classes of non-halo population have been proposed (e.g., a ‘thick’ LMC disk, a small dwarf galaxy projected in front of the LMC, a warped Milky Way disk etc.), but several searches for such populations have been negative.

There are several ways of testing these possibilities: one is to measure distances to individual lenses. In a ‘standard’ microlensing event as in Figure 3 it is impossible to measure the lens distance since the one physical parameter (event duration) depends on three unknowns, the lens mass, distance, and transverse velocity. However, as proposed by Gould and others, a small fraction of observed events should deviate from the standard shape for one of several reasons: the non-uniform motion of the Earth, the finite size of the source, a binary lens etc. In these cases we can obtain an additional observable which constrains the location of the lens. At present there are only one or two such cases for Magellanic Cloud lensing.

Future satellite observations may help to measure lens distances, either by measuring the ‘parallax’ effect or the small centroid shift during the event.

Another active area is the search for microlensing towards the Andromeda galaxy, M31. This is considerably more challenging since the greater distance means that there are multiple unresolved stars per resolution element, and sophisticated image differencing techniques are necessary. However, it has the advantage that microlensing from M31’s own halo will produce substantially more events towards the far side of M31’s disk than towards the near side, owing to the nearly edge-on inclination. Two groups called AGAPE and MEGA have done pilot studies and are currently undertaking large-scale searches towards M31.

Cosmological microlensing: Although microlensing within the Local Group is the main application at present, microlensing of ‘point’ sources at cosmological distances (e.g., QSOs, SNe, or gamma-ray bursts) may also be observable. In the case of a QSO lensed into multiple images by a foreground galaxy, the individual stars in the galaxy may act as microlenses. Here the microlensing appears as brightness changes unique to one image of the multiplet (after correcting for the time delay); in this case the optical depth is typically close to unity (because the surface density must be of order the ‘critical’ surface density to produce multiple imaging), so the interpretation of the light curve is complex. Other searches are in progress for microlensing of quasars behind the Virgo Cluster by dark objects in the cluster, and the Next Generation Space Telescope should be easily able to detect microlensing of stars in the Virgo galaxy M87.

1.1.3 Follow-up Questions

- Can you derive α from a Newtonian approach?
- Why/how does gravitational lensing magnify a star?

1.2 Question 2

A two-element interferometer consists of two telescopes whose light is combined and interfered. Sketch the response of such an interferometer to a nearby red giant star, as a function of the (projected) separation between the two telescopes. The red giant subtends one-fiftieth of an arc second on the sky, and the telescope operates at a wavelength of 2 microns.

1.2.1 Short answer

In a similar way that a single slit experiment results in an interference pattern depending on the slit size, an analogous interference pattern will emerge from a two-element interferometer and will depend on the projected baseline. Since the red giant star is resolved to 1/50 arcseconds, let's determine the projected baseline this corresponds to. Before we can do this, we must first convert the angular size from arcseconds into radians:

$$\theta = \left(\frac{1}{50} \text{ arcsec} \right) \left(\frac{1 \text{ deg}}{3600 \text{ arcsec}} \right) \left(\frac{1 \text{ rad}}{57 \text{ deg}} \right) = 10^{-7} \text{ rad.}$$

Now to determine the projected baseline:

$$\begin{aligned} \theta &\sim 2.2 \left(\frac{\lambda}{D} \right) [\text{rad}] \\ D &\sim 2.2 \left(\frac{\lambda}{\theta} \right) [\text{m}] \\ &= 2.2 \left(\frac{2 \times 10^{-6} \text{ m}}{10^{-7} \text{ rad}} \right) = 44 \text{ [m].} \end{aligned}$$

Thus, the resulting response function will be a Bessel function whose minima are at integer multiples of the 44 m projected baseline:

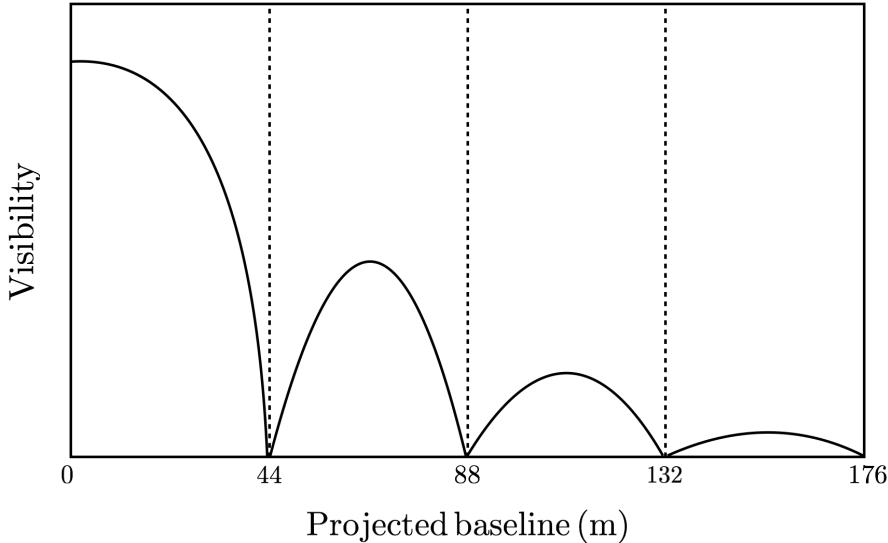


Figure 5: Response function of a two-element interferometer as a function of the projected baseline observing a resolved red giant star with an angular resolution of 1/50 arcseconds.

1.2.2 Additional context

The two-element quasi-monochromatic interferometer: The simplest radio interferometer is a pair of radio telescopes whose voltage outputs are correlated (multiplied and averaged), and even the most elaborate interferometers with $N \gg 2$ antennas, often called **elements**, can be treated as $N(N - 1)/2$ independent two-element interferometers. Figure 6 shows two identical dishes separated by the **baseline vector** \vec{b} of length b that points from antenna 1 to antenna 2. Both dishes point in the same direction specified by the unit vector \hat{s} , and θ is the angle between \vec{b} and \hat{s} . Plane waves from a distant point source in this direction must travel an extra distance $\vec{b} \cdot \hat{s} = b \cos \theta$ to reach antenna 1, so the output of antenna 1 is the same as that of antenna 2, but it lags in time by the geometric delay

$$\tau_g = \frac{\vec{b} \cdot \hat{s}}{c} [\text{s}].$$

For simplicity, we first consider a quasi-monochromatic interferometer, one that responds only to radiation in a very narrow band $\Delta\nu \ll 2\pi/\tau_g$ centered on frequency $\nu = \omega/(2\pi)$. Then the output voltages of antennas 1 and 2 at time t can be written as

$$V_1 = V \cos[\omega(t - \tau_g)] [\text{V}]$$

$$V_2 = V \cos(\omega t) [\text{V}].$$

These output voltages are amplified versions of the antenna input voltages; they have not passed through square-law detectors. Instead, a **correlator** multiplies these two voltages to yield the product

$$V_1 V_2 = V^2 \cos[\omega(t - \tau_g)] \cos(\omega t) = \left(\frac{V^2}{2} \right) [\cos(2\omega t - \omega\tau_g) + \cos(\omega\tau_g)] [\text{V}]$$

which follows directly from the trigonometric identity $\cos x \cos y = [\cos(x + y) + \cos(x - y)]/2$. The correlator also takes a time average long enough ($\Delta t \gg (2\omega)^{-1}$) to remove the high-frequency term $\cos(2\omega t - \omega\tau_g)$ from the **correlator response** (output voltage) R and keep only the slowly varying term

$$R = \langle V_1 V_2 \rangle = \left(\frac{V^2}{2} \right) \cos(\omega\tau_g) [\text{V}^2].$$

The voltages V_1 and V_2 are proportional to the electric field produced by the source multiplied by the voltage gains of the two antennas and receivers. Thus the correlator output amplitude $V^2/2$ is proportional to the flux density S of the point source multiplied by $\sqrt{A_1 A_2}$, where A_1 and A_2 are the effective collecting areas of the two antennas.

Notice that the time-averaged response R of a multiplying interferometer is zero. There is no DC output, so fluctuations in receiver gain do not act on the whole system temperature T_s as for a total-power observation with a single dish. Uncorrelated noise power from very extended radio sources such as the CMB and the atmosphere over the telescopes, also averages to zero in the correlator response. Short interference pulses with duration $t \ll |b|/c$ are also suppressed because each pulse does not reach both telescopes simultaneously. Likewise, a multiplying radio interferometer differs from a classical **adding interferometer**, such as the optical Michelson interferometer, that adds the uncorrelated noise power contributions.

The correlator output voltage $R = (V^2/2) \cos(\omega\tau_g)$ varies sinusoidally as the Earth's rotation changes the source direction relative to the baseline vector. These sinusoids are called **fringes**, and the **fringe phase**

$$\phi = \omega\tau_g = \frac{\omega}{c} b \cos \theta [\text{rad}]$$

depends on θ as follows:

$$\frac{d\phi}{d\theta} = \frac{\omega}{c} b \sin \theta = 2\pi \left(\frac{b \sin \theta}{\lambda} \right) [\text{dimensionless}].$$

The **fringe period** $\phi = 2\pi$ corresponds to an angular shift $\theta = \lambda/(b \sin \theta)$. The fringe phase is an exquisitely sensitive measure of source position if the **projected baseline** $b \sin \theta$ is many wavelengths long. Note that fringe phase and hence measured source position is not affected by small tracking errors of the individual telescopes. It depends on time, and times can be measured by clocks with much higher accuracy than angles (ratios of lengths of moving telescope parts) can be measured by rulers. Also, an interferometer whose baseline is horizontal is not affected by the plane-parallel component of atmospheric refraction, which delays the signals reaching both telescopes equally. Consequently, interferometers can determine the positions of compact radio sources with unmatched accuracy. Absolute positions with errors as small as $\sigma_\theta \approx 10^{-3}$ arcsec and differential positions with errors down to $\sigma_\theta \approx 10^{-5}$ arcsec $< 10^{-10}$ rad have frequently been measured.

If the individual antennas comprising an interferometer were isotropic, the interferometer point-source response would be a sinusoid spanning the sky. Such an interferometer is sensitive to only one Fourier component of the sky brightness distribution: the component with angular period $\lambda/(b \sin \theta)$. The response R of a two-element interferometer with directive antennas is that sinusoid multiplied by the product of the voltage patterns of the individual antennas. Normally the two antennas are identical, so this product is the power pattern of the individual antennas and is called the **primary beam** of the interferometer. The primary beam is usually a Gaussian much wider than a fringe period, as indicated in Figure 6. The convolution theorem states that the Fourier transform of the product of two functions is the convolution of their Fourier transforms, so the interferometer with directive antennas responds to

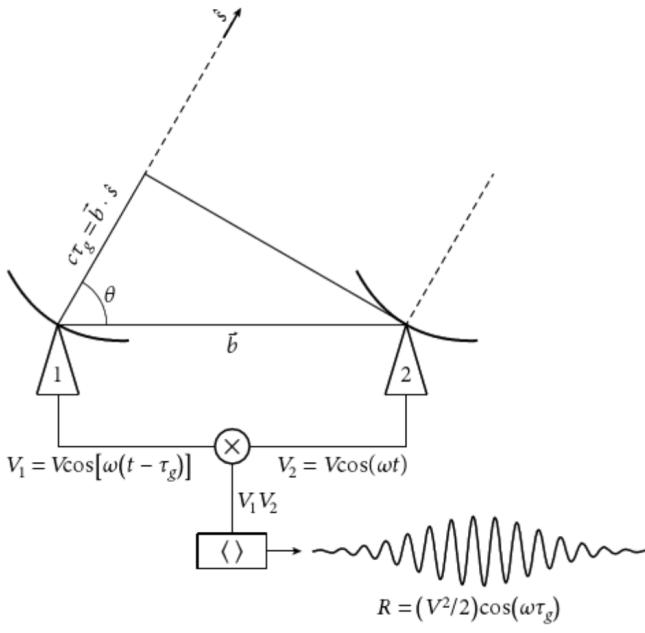


Figure 6: This block diagram shows the components of a two-element quasi-monochromatic **multiplying interferometer** observing in a very narrow radio frequency range centered on $\nu = \omega/(2\pi)$. \hat{s} is the unit vector in the direction of a distant point source and \vec{b} is the baseline vector pointing from antenna 1 to antenna 2. The output voltage V_1 of antenna 1 is the same as the output voltage V_2 of antenna 2, but it is retarded by the geometric delay $\tau_g = \vec{b} \cdot \hat{s}/c$ representing the additional light-travel delay to antenna 1 for a plane wavefront from a source at angle θ from the **baseline vector**. These voltages are amplified, multiplied (\times), and time averaged ($\langle \rangle$) by the **correlator** to yield an output response whose amplitude R is proportional to the flux density of the point source and whose phase $(\omega\tau_g)$ depends on the delay and the frequency. The quasi-sinusoidal output **fringe** shown occurs if the source direction in the interferometer frame is changing at a constant rate $d\theta/dt$. The broad Gaussian envelope of the fringe shows the primary-beam attenuation as the source passes through the beam of the dishes. Figure taken from Condon & Random (2016).

a finite range of angular frequencies centered on $b \sin \theta / \lambda$. Because the antenna diameters D must be smaller than the baseline b (else the antennas would overlap), the angular frequency response cannot extend to zero and the interferometer cannot detect an isotropic source – the bulk of the 3 K CMB for example. The missing short spacings ($b < D$) can be provided by a single-dish telescope with diameter $D > b$. Thus the $D = 100$ m GBT can fill in the missing baselines $b < 25$ m that the $D = 25$ m VLA dishes cannot obtain.

Improving the instantaneous point-source response pattern of an interferometer requires more Fourier components; that is, more baselines. An interferometer with N antennas contains $N(N - 1)/2$ pairs of antennas, each of which is a two-element interferometer, so the instantaneous **synthesized beam** (the point-source response obtained by averaging the outputs of all of the two-element interferometers) rapidly approaches a Gaussian as N increases. The instantaneous point-source responses of a two-element interferometer with projected baseline length b , a three-element interferometer with three baselines (projected lengths $b/3, 2b/3$, and b), and a four-element interferometer with six baselines (projected lengths $b/6, 2b/6, 3b/6, 4b/6, 5b/6$, and b) are shown in Figure 7.

Most radio sources are stationary; that is, their brightness distributions do not change significantly on the timescales of astronomical observations. For stationary sources, a two-element interferometer with movable antennas could make $N(N - 1)/2$ observations to duplicate one observation with an N -element interferometer.

Slightly extended sources and the complex correlator: The response $R = (V^2/2) \cos(\omega\tau_g)$ of the quasi-monochromatic two-element interferometer with a “cosine” correlator (Figure 6) to a spatially incoherent slightly extended (much smaller than the primary beamwidth) source with sky brightness distribution $I_\nu(\hat{s})$ near frequency $\nu = \omega/(2\pi)$ is obtained by treating the extended source as the sum of independent point sources:

$$R_c = \int I(\hat{s}) \cos(2\pi\nu\vec{b} \cdot \hat{s}/c) d\Omega = \int I(\hat{s}) \cos(2\pi\vec{b} \cdot \hat{s}/\lambda) d\Omega [V^2].$$

Notice that the even cosine function in this response is sensitive only to the even (inversion-symmetric) part I_E of an arbitrary source brightness distribution, which can be written as the sum of even and odd (anti-symmetric) parts: $I = I_E + I_O$. To detect the odd part I_O we need a “sine” correlator whose output is odd, $R = (V^2/2) \sin(\omega\tau_g)$. This can be implemented by a second correlator that follows a $\pi/2$ rad = 90° phase delay inserted into the output of one antenna because $\sin(\omega\tau_g) = \cos(\omega\tau_g - \pi/2)$. Then

$$R_s = \int I(\hat{s}) \sin(2\pi\vec{b} \cdot \hat{s}/\lambda) d\Omega [V^2]$$

The combination of cosine and sine correlators is called a **complex correlator** because it is mathematically convenient to treat the cosines and sines as complex exponentials using Euler’s formula

$$e^{i\phi} = \cos \phi + i \sin \phi [\text{rad}].$$

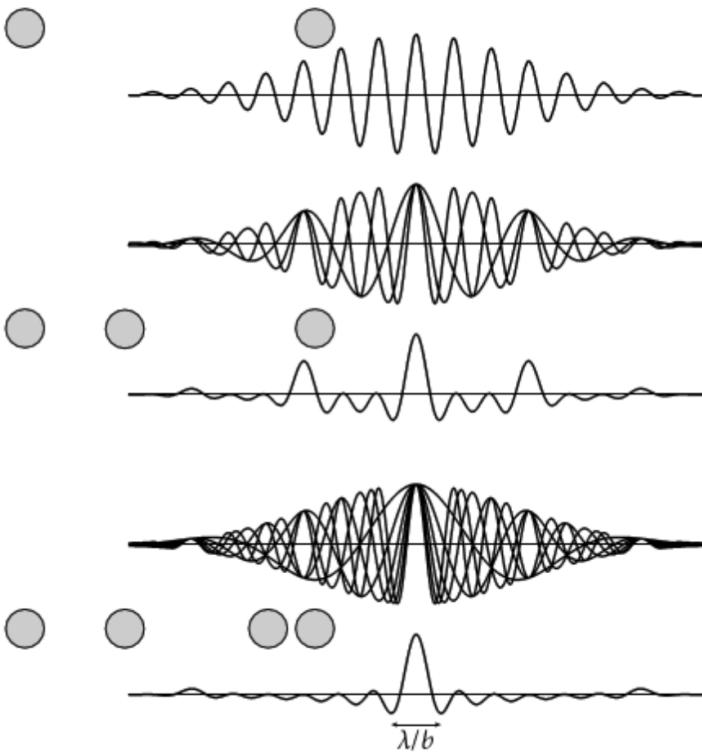


Figure 7: The instantaneous point-source responses of interferometers with overall projected length b and two, three, or four antennas distributed as shown are indicated by the thick curves. The synthesized main beam of the four-element interferometer is nearly Gaussian with angular resolution $\theta \approx \lambda/b$, but the side-lobes are still significant and there is a broad negative “bowl” caused by the lack of spacings shorter than the diameter of an individual antenna. Thus the **synthesized beam** is sometimes called the **dirty beam**. The instantaneous dirty beam of the multi-element interferometer is the arithmetic mean of the individual responses of its component two-element interferometers. The individual responses of the three two-element interferometers comprising the three-element interferometer and of the six two-element interferometers comprising the four-element interferometer are plotted as thin curves. Figure taken from Condon & Random (2016).

The **complex visibility** is defined by

$$\mathcal{V} \equiv R_c - iR_s \text{ [V}^2\text{]}$$

which can be written in the form

$$\mathcal{V} = Ae^{-i\phi} \text{ [rad]},$$

where

$$A = \sqrt{R_c^2 + R_s^2} \text{ [dimensionless]}$$

is the **visibility amplitude** and

$$\phi = \arctan \left(\frac{R_s}{c} \right) \text{ [rad]}$$

is the **visibility phase**. The response to an extended source with brightness distribution $I(\hat{s})$ of the two-element quasi-monochromatic interferometer with a complex correlator is the complex visibility

$$\mathcal{V} = \int I(\hat{s}) \exp(-i2\pi\vec{b} \cdot \hat{s}/\lambda) d\Omega \text{ [V}^2\text{].}$$

Fringe patterns and Fourier transforms: Interferometry begins with the Youngs slits fringe pattern (Fig. 1). With a single point source emitting coherent radiation, interference fringes are observed, with constructive and destructive interference observed as the relative delay of the two interfering rays changes; the separation of the fringes is $/d$, the wavelength of the light divided by the slit separation.

If the source is made wider (Figure 8b), we can think of it as a sequence of point sources each of which emit radiation which is uncorrelated with the emission from the others. It follows that the total interference intensity pattern is the sum of the individual patterns. Since an angular displacement in the source produces an equal angular displacement in the fringe pattern, as the source size approaches λ/d the fringe patterns will add to give a constant illumination (Figure 8c). In this case, the fringe visibility (defined as the difference between maximum and minimum intensity, normalized by the sum of maximum and minimum intensity) drops to zero. Conversely, when the angular size of the source is λ/d , the fringe visibility is 1; this corresponds to a situation in which the source size is smaller than the angular resolution of the interferometer, and only an upper limit of order λ/d can be obtained on it.

Now suppose that the slit spacing d is decreased. For the same size of source, this produces less “washing-out” of the fringes, because the same displacement of the source now produces much less displacement of the fringe patterns as a fraction of the fringe separation λ/d (Figure 8d). The smaller the slit separation,

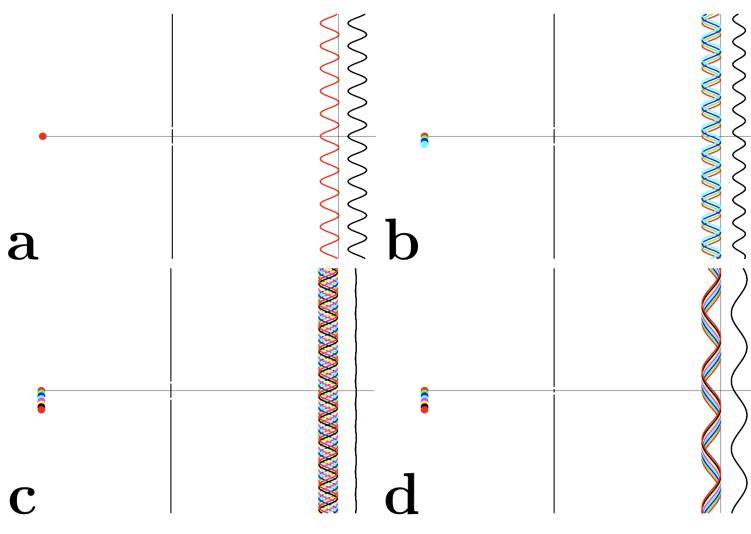
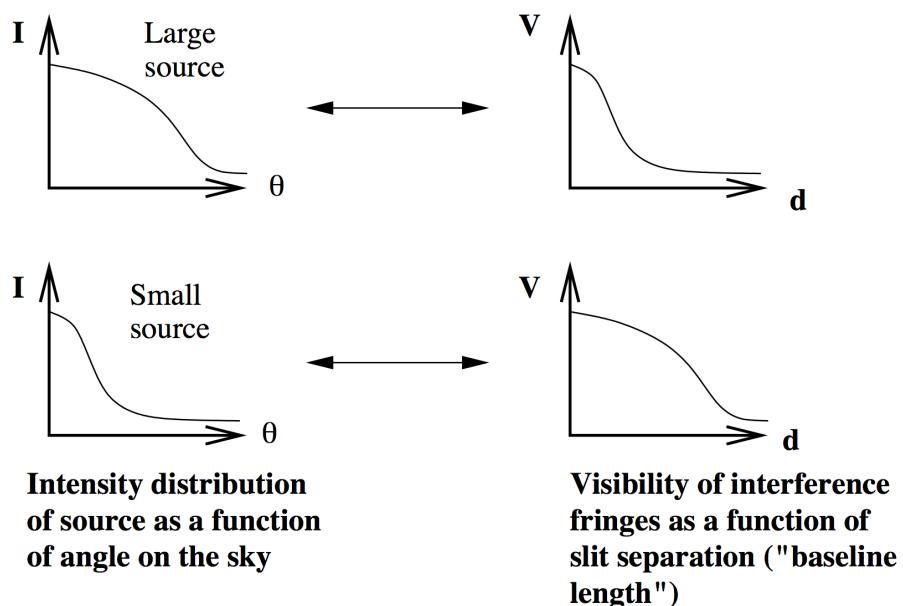


Figure 8: Young's slits in various situations. In each panel the source is shown on the left, and on the right of the slit are shown the fringe patterns separately for each part of the source and then the added fringe pattern. (a): The basic two-slit pattern, showing fringes an angular distance λ/d apart. (b): The effect of increasing the source size. An angular shift of the source position by θ shifts the fringe patterns by θ the other way. Since the patterns come from mutually incoherent sources, the intensity patterns add to give a pattern of reduced visibility. (c): When the size of the source reaches λ/d , the fringes add to give zero visibility. (d): If the slit spacing is then reduced, the fringe spacing increases, and the same size of source is still able to give visible fringes: the source would need to be increased in size to λ/d_{new} in order to wash out the fringes. Figure taken from Jackson.

Figure 9:
Relation between source brightness as a function of angular distance and visibility of interference fringes as a function of slit separation (baseline length). Figure taken from Jackson.



the larger the source size that can be probed using interferometry.

The situation is summarized in Figure 9. If we plot, for a given source distribution, the way in which visibility varies with slit separation, it can be seen that for small sources the visibility remains high out to large slit separation (in the limit of point sources, to infinite slit separation), whereas large sources produce visibility patterns which fall off quickly as the slit separation increases.

The relation between $I(\theta)$ and $V(d)$ represented here is one which maps a large Gaussian into a small Gaussian, and vice versa, and it is fairly obvious that it is a Fourier transform.¹

1.2.3 Follow-up Questions

- What do the minima in the response function tell you?

¹This is known as the Van Cittert-Zernicke theorem.

1.3 Question 3

What's the minimum mass of a black hole you could survive a fall through the event horizon without being ripped to shreds? Why would you be ripped to shreds for smaller black holes? How does this relate to the BH mass range for which we expect tidal disruption flares caused by shredding main-sequence stars?

1.3.1 Short answer

We can find the minimum mass using the equation for the tidal force one would experience:

$$F_T = \left(\frac{dg}{dr} \right) \ell = \frac{d}{dr} \left(-\frac{GM}{r^2} \right) \ell = -GM\ell \frac{d}{dr} r^{-2} = -GM\ell(-2)r^{-3} = \frac{2GM\ell}{r^3} [\text{N}].$$

Setting the radius equal to the Schwarzschild radius $r_s \equiv 2GM/c^2$,

$$F_T = \frac{2GM\ell}{r_S^3} = 2GM\ell \left(\frac{c^2}{2GM} \right)^3 = \frac{\ell c^6}{4G^2 M^2} [\text{N}],$$

we can now solve for the mass of the black hole M :

$$M = \sqrt{\frac{\ell c^6}{4G^2 F_T}} [\text{kg}].$$

Let's set $\ell \sim 1 \text{ m}$ for an order-of-magnitude length of the human body, and $F_T \sim 100 \text{ g} = 10^4 \text{ N}$ for the maximum force that the body can withstand. We can now use this to find the minimum mass of a black hole that you could survive a fall through the event horizon without being ripped to shreds:

$$M = \sqrt{\frac{(1 \text{ m})(2.998 \times 10^8 \text{ m s}^{-1})^6}{4(6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2})^2 (10^4 \text{ N})}} = 10^{34} \text{ kg} \sim 10^3 M_\odot.$$

Since we can see that the tidal force $F_T \sim M^{-2}$, you would be ripped to shreds for smaller black holes because the resulting tidal force on the body would be greater the smaller the black hole mass is.

1.3.2 Follow-up Questions

- How would you estimate the maximum tidal acceleration a star can withstand?
- Why is it enough to know if the surface of the star will be disrupted?

1.4 Question 4

How is synchrotron radiation generated, and how was it used to demonstrate the energy required to power radio galaxies?

1.4.1 Short answer

Synchrotron radiation is generated by relativistic cosmic rays being accelerated by local magnetic fields via the Lorentz force. The non-relativistic limit is cyclotron radiation, but at the relativistic speeds of cosmic rays, the Lorentz factor γ must be included for the relativistic case of synchrotron radiation. This type of radiation is non-thermal and linearly polarized. As cosmic rays are accelerated by local magnetic fields, they emit dipole radiation which is beamed in the direction of its velocity vector with an opening angle $\sim 1/\gamma$.

1.4.2 Additional context

Synchrotron radiation from the Galaxy was the first emission detected in radio astronomy, although it was not until the 1950s when detailed radio maps of the Galaxy were produced, that the connection was made between radio emission and the synchrotron mechanism. Synchrotron radiation, also known as **magnetobremsstrahlung radiation**, is a type of **non-thermal radiation** emitted by relativistic electrons being deflected by the magnetic field of the ISM. It occurs in the diffuse ISM and in SN remnants, due in part to electron acceleration associated with the supernova blastwave and in part to increased magnetic field strengths in the shocked gas. This radiation dominates the sky brightness at frequencies $\nu \lesssim 1$ GHz. Figure 10 shows the all-sky radio map at 408 MHz.

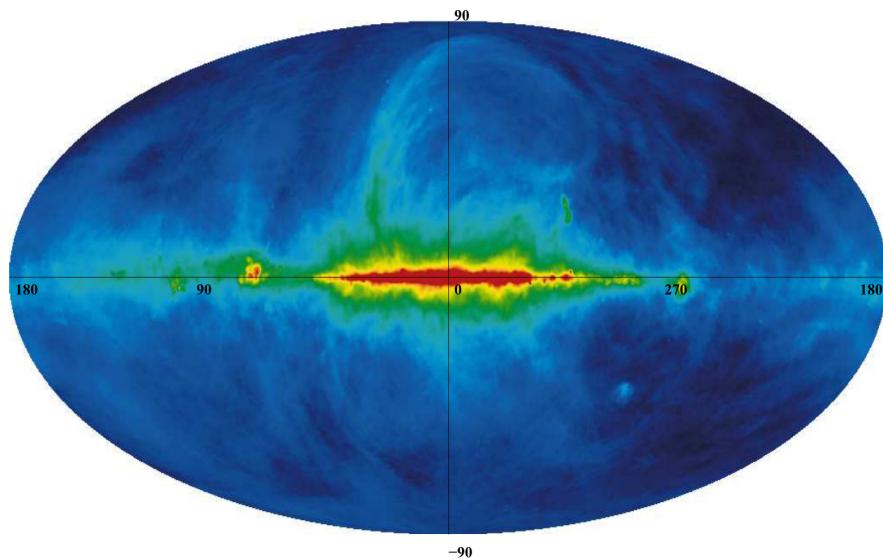


Figure 10: The synchrotron emission at 408 MHz across the entire sky in Galactic coordinates. As expected the emission is concentrated along the Galactic plane. The noticeable feature known as Loop I is clearly arching up from $\ell = 55^\circ$ towards the North Galactic Pole. This figure is adapted from Haslam et al. (1981). Figure taken from Newton-McGee (2009).

Cosmic rays (i.e., high energy nuclei and electrons) are an important component of the ISM along with magnetic fields and interstellar gas. While the origin of cosmic rays is still not comprehensively understood, it is believed that the electrons with an energy approaching 100 – 1000 TeV have been accelerated in SNRs, suggesting this is one mechanism for the primary formation of cosmic rays.

Magnetic fields accelerate these charged particles via the **Lorentz force**. The acceleration is given by

$$\frac{d}{d\tau}(m\gamma v) = q \left(\frac{\vec{v}}{c} \times \vec{B} \right) [N]$$

where m is the mass of the particle with charge q , τ is the retarded time, the Lorentz factor $\gamma = (1v^2/c^2)^{-1/2}$, and \vec{B} is the magnetic field. The trajectory of the particle is *helical*, with a **gyration frequency** ω_B and **pitch angle** θ (where $\theta = 90^\circ$ for a circular orbit, and $\theta = 0^\circ$ for a particle trajectory parallel to \vec{B}).

A charged particle deflected by magnetic fields while moving at relativistic speeds emits synchrotron radiation during this acceleration. The main properties of the synchrotron radiation are the following:

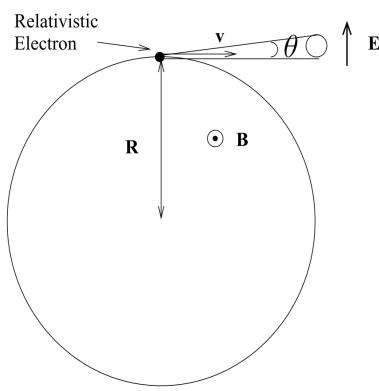


Figure 11:
The trajectory of a relativistic electron spiraling in a magnetic field B with a pitch angle θ . The electric field E is polarized parallel to the orbital plane. The synchrotron emission is emitted in a highly beamed cone. This figure is adapted from Kraus (1986). Figure taken from Newton-McGee (2009).

- high intensity;
- very broad and continuous spectral range from infrared up to the hard x-ray region;
- natural narrow angular collimation;
- high degree of polarization;
- pulsed time structure.

The graphical representation of synchrotron radiation, depicted in Figure 11, shows that the velocity vector of the particle generates the surface of a cone, with the synchrotron emission beamed within the surface of the cone, at an angle of:

$$\theta = \pm \frac{mc^2}{E} \text{ [rad].}$$

On the surface of the cone the radiation is 100% linearly polarized with its electric field in the direction $-\vec{v} \times \vec{B}$, where v is the direction of propagation of the electron. The degree of linear polarization, called the polarization fraction p , is independent of frequency and depends on the spectral index of the cosmic ray electrons:

$$p(\alpha) = \frac{3 - 3\alpha}{5 - 3\alpha} \text{ [dimensionless].}$$

In order to understand the angular and spectral distribution of the emitted radiation of synchrotron radiation, let us remind ourselves of the emission from a classical electron moving at a speed v much lower than the speed of light ($v \ll c$). In this case the emitted pattern is similar to that of an oscillating dipole with its maximum of intensity in the direction perpendicular to the acceleration and does not depend on the electron speed. For a relativistic effect, when the speed of the emitting electrons increases to relativistic values ($v \approx c$) the radiation pattern is compressed into a narrow cone in the direction of motion, resulting into an emission tangential to the particle orbit. The vertical half-opening angle θ is given by

$$\theta \approx \frac{mc^2}{E} \approx \frac{1}{\gamma} \text{ [rad].}$$

See Figure 12.

The radiated power from the accelerating charge is given (in cgs units) by

$$P = \frac{2}{3} \frac{q^4}{m^2 c^5} (\gamma \nu)^2 (B \sin \theta)^2 = 2.37 \times 10^{-15} \left(\frac{E}{\text{erg}} \right)^2 \left(\frac{B \sin \theta}{\mu\text{G}} \right)^2 \text{ [erg s}^{-1}\text{]}$$

where the final equality is given for electrons in particular, since the dependence on the mass-to-charge ratio means that electrons are by several orders of magnitude dominant over protons or ions in generating synchrotron radiation. The synchrotron power peaks at the **characteristic frequency** (in Hz) of

$$\nu_c = \frac{3}{4\pi} \frac{q}{mc} \gamma^2 (B \sin \theta) = 6.26 \times 10^3 \left(\frac{E}{\text{erg}} \right)^2 \left(\frac{B \sin \theta}{\mu\text{G}} \right) \text{ [Hz].}$$

This equation tells us that for radio observations (frequencies at or above tens of MHz) and with typical magnetic field strengths of order μG , the observed population of electrons have $\gamma \gg 1$. The radiation is

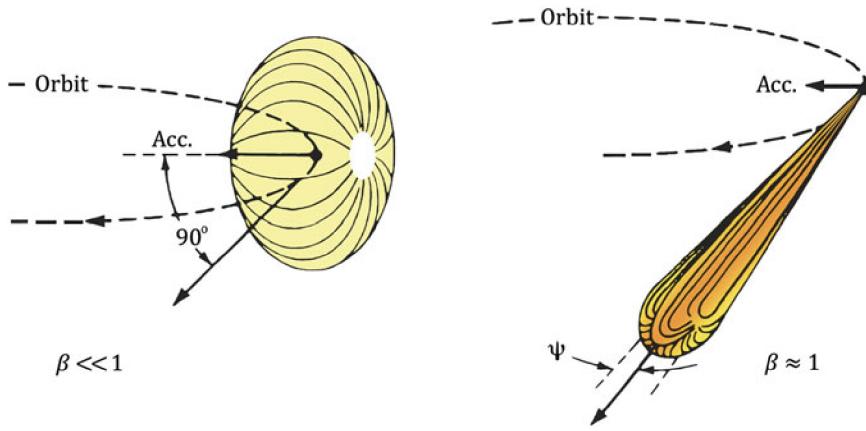


Figure 12: Qualitative radiation patterns related to charged particles moving in a circular orbit. The dipole pattern is achieved for slow particles when $v/c \ll 1$ (left); this is distorted into a narrow cone when $v/c \approx 1$ (right). Figure taken from Mobilio et al. (2015).

thus highly beamed (width $\sim \gamma^{-1}$) in the direction of the velocity vector. The population of electrons has a power-law energy distribution given by

$$n(\gamma)d\gamma = n_0\gamma^{-p}d\gamma \text{ [eV}^{-1}\text{]}$$

where a typical value for the power law index is $p \approx 2.5$. The resulting synchrotron luminosity is thus also a power law,

$$L_\nu \propto \nu^{-(p-1)/2} \equiv \nu^{-\alpha} \text{ [erg s}^{-1}\text{]}$$

with a typical spectral index $\alpha = 0.75$ for $p = 2.5$. The particles lose energy due to the power emitted with a characteristic timescale of

$$\tau_{\text{syn}} = \frac{\gamma}{d\gamma/d\tau} = 4\pi \frac{mc}{\sigma_T} \gamma^{-1} (B \sin \theta)^{-2} \text{ [yr]}$$

where the **Thomson cross-section** $\sigma_T = 8\pi q^4/3m^2c^4$. Since τ_{syn} is shorter for more energetic electrons (higher γ) the power law in the number density $n(\gamma)d\gamma$ steepens with time (often corresponding to distance from the site of cosmic ray acceleration), therefore increasing the value of α . This is referred to as **synchrotron aging**. Substituting for the characteristic synchrotron frequency ν_c , we can see that

$$\left(\frac{\tau_{\text{syn}}}{\text{yr}} \right) = 1.06 \times 10^9 \left(\frac{B \sin \theta}{\mu\text{G}} \right)^{-1.5} \left(\frac{\nu_c}{\text{GHz}} \right)^{-0.5}.$$

The spectrum of synchrotron radiation must be related to the detailed variation of the electric field as seen by an observer. Because of beaming effects the emitted radiation fields appear to be concentrated in a narrow set of directions about the particle's velocity. The observer will see a pulse of radiation confined to a time interval much smaller than the gyration period. The spectrum will thus be spread over a much broader region than one of order $\omega_B/2\pi$. This is an essential feature of synchrotron radiation.

Non-thermal cosmic radio emission was first thought to originate in stellar atmospheres (the **radio star hypothesis**). This hypothesis seemed quite reasonable at first glance in view of the existence of quite intense sporadic radio emission from the Sun. It is easy to see, however, that to explain the observed data, these hypothetical radio stars would need to possess extremely unusual properties. The discovery of a quasi-spherical component of the general galactic radio emission imposed severe demands on the radio star hypothesis. For instance, it became clear in this connection that sources of non-thermal galactic radio emission were distributed principally in the galactic halo, the discovery of which had been made only a short time before. Nevertheless, the radio star hypothesis was still not abandoned. If, on the other hand, we equate the general galactic radio emission with synchrotron radiation we then obtain completely probable and useful estimates of the intensity of interstellar fields and the concentration of relativistic electrons. Such estimates are also encouraging in the case of discrete sources. The radio star hypothesis was thus abandoned as early as the beginning of 1953 and the synchrotron character of the major part of non-thermal cosmic radio emission was accepted.

1.5 Question 5

What are “forbidden lines” of atomic spectra? In what conditions are they observationally important? In what conditions do they control the temperature of interstellar material?

1.5.1 Short answer

First, “allowed” transitions are those that obey the following selection rules for electric dipole transitions:

1. parity ($\sum_i \ell_i$) must change;
2. $\Delta L = 0, \pm 1$;
3. $\Delta J = 0, \pm 1$, but $\Delta J = 0 \rightarrow 0$ is forbidden;
4. only 1 electron wave function $n\ell$ changes with $\Delta\ell = \pm 1$;
5. spin does not change: $\Delta S = 0$.

“Spin-forbidden” transitions (i.e., semi-forbidden, inter-combination, inter-system transitions) are those that satisfy the first four selections rules but not the fifth (i.e., $\Delta S \neq 0$). These transitions are considerably weaker than the allowed transitions and are denoted with a single square bracket.

“Forbidden” transitions are those that fail to satisfy at least one of the first four selection rules. Such transition probabilities vary wildly and are denoted with two square brackets.

1.5.2 Additional context

Brief overview of atomic structure: Atoms consist of three subatomic particles: protons, neutrons, and electrons. A proton has a positive charge and a neutron has no charge. Both protons and neutrons are found in the densely packed, positively charged nucleus. The nucleus contains essentially all of the mass of the atom. Electrons carry a negative charge and are found in electron shells surrounding the nucleus. The mass of the electron is considered to be negligible.

All atoms have an atomic number, Z , and a mass number, A . The atomic number, Z , represents its number of protons and its mass number, A , represents its number of protons plus its number of neutrons. All atoms of a particular element will have the same atomic number, however they may differ in their number of neutrons and electrons. Two atoms with the same atomic number but different mass numbers (different numbers of neutrons) are referred to as **isotopes**. The average atomic mass found on the periodic table represents a weighted average of the naturally occurring isotopes for a particular element. Atoms with unequal numbers of protons and electrons produce charged atoms or **ions**.

Electrons within an atom are found in particular **orbitals**. An atomic orbital is a mathematical function that describes the wave-like behavior of either one electron or a pair of electrons in an atom. This function can be used to calculate the probability of finding any electron of an atom in any specific region around the atom’s nucleus. The term atomic orbital may also refer to the physical region or space where the electron can be calculated to be present, as defined by the particular mathematical form of the orbital. Electrons within an atom can be assessed according to the shell, subshell, and orbital to which they are assigned. These assessments are based on the quantum mechanical model. **Shells** are numbered as $n = 1, 2, 3, 4$, etc. and increase in size and energy as they get further away from the nucleus. Shells can be subdivided into subshells.

The maximum number of **subshells** is equivalent to the shell number. For example, when $n = 1$ (first shell), only one subshell is possible and when $n = 2$ (second shell), two subshells are possible. There are four different types of subshells. These various types of subshells are denoted by the letters s , p , d , and f . Each subshell has a maximum number of electrons which it can hold: s : 2 electrons, p : 6 electrons, d : 10 electrons, and f : 14 electrons. The s subshell is the lowest energy subshell and the f subshell is the highest energy subshell. As was mentioned previously, the shell number is equal to the possible number of subshells. Thus, when $n = 1$, the only subshell possible is the $1s$ subshell. When $n = 2$, two subshells are possible: the $2s$ and $2p$. When $n = 3$, three subshells are possible: the $3s$, $3p$, and $3d$. This means that in the first shell only two electrons are possible and they would be found in the $1s$ (2 electrons) subshell. In the second shell, 8 electrons are possible and would be found in the $2s$ (2 electrons) and the $2p$ (6 electrons) subshells.

Each subshell is further divided into **orbitals**. An orbital is defined as a region of space in which an electron can be found. *Only two electrons are possible per orbital.* Thus, the s subshell may contain only one orbital and the p subshell may contain three orbitals. Each orbital has its own distinct shape. An s orbital found in an s subshell is spherical, p orbitals found in p subshells are two-lobed, and d orbitals found in d subshells are four-lobed. Since there are three possible orbitals per p subshell, each orbital

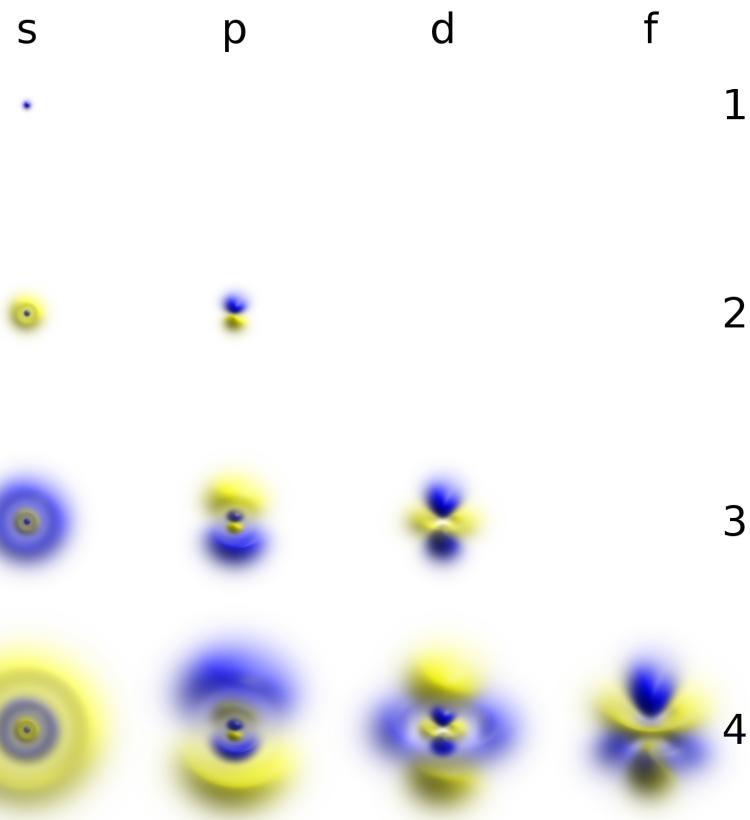


Figure 13: 3D views of some hydrogen-like atomic orbitals. Figure taken from Wikipedia.

adopts its own orientation. The p_x orbital lies along the x axis, the p_y orbital lies along the y axis, and the p_z orbital lies along the z axis.

When writing electron configurations for atoms, the shorthand symbol for a subshell followed by a superscripted number which represents the number of electrons in that subshell is used. For example, a Carbon atom with 6 electrons would have the electron configuration: $1s^2 2s^2 2p^2$. A possible arrangement within the $2p$ subshell would be to find one of the electrons in the $2p_x$ orbital and the second electron in the $2p_y$ orbital.

Figure 13 shows hydrogen-like atomic structure for s , p , d , and f orbitals.

Selection rules for radiative transitions: Some energy levels are connected by strong radiative transitions; in other cases, radiative transitions between the levels may be extremely slow. The strong transitions always satisfy what are referred to as the **selection rules for electric dipole transitions**. Here, we summarize the selection rules for the strong electric dipole transitions, and we also give the selection rules for **inter-system** and **forbidden transitions** that do not satisfy the electric dipole selection rules but nevertheless are strong enough to be astrophysically important. We will use the ion NII as an example; the first nine energy levels of NII are shown in Figure 14.

Allowed: Electric dipole transitions: The strongest transitions are electric dipole transitions. These are transitions satisfying the following selection rules:

1. Parity must change.
2. $\Delta L = 0, \pm 1$.
3. $\Delta J = 0, \pm 1$, but $\Delta J = 0 \rightarrow 0$ is forbidden.
4. Only one single-electron wave function $n\ell$ changes, with $\Delta\ell = \pm 1$.
5. $\Delta S = 0$: Spin does not change.

An allowed transition is denoted *without* square brackets, for example,

$$\text{NII } 1084.0 \text{\AA} \ ^3P_0 - ^3D_1^o.$$

This is a transition between the $\ell = 1s^2 2s^2 2p^2 \ ^3P_0$ and $u = 1s^2 2s^2 2p3s \ ^3D_1^o$ levels of NII, with a wavelength $\lambda_{u\ell} = 1084.0 \text{\AA}$. The transition has $A_{u\ell} = 2.18 \times 10^8 \text{ s}^{-1}$. This decay is very fast – the lifetime of the $^3D_1^o$ level against this decay is only $1/A_{u\ell} = 4.6 \text{ ns}$!

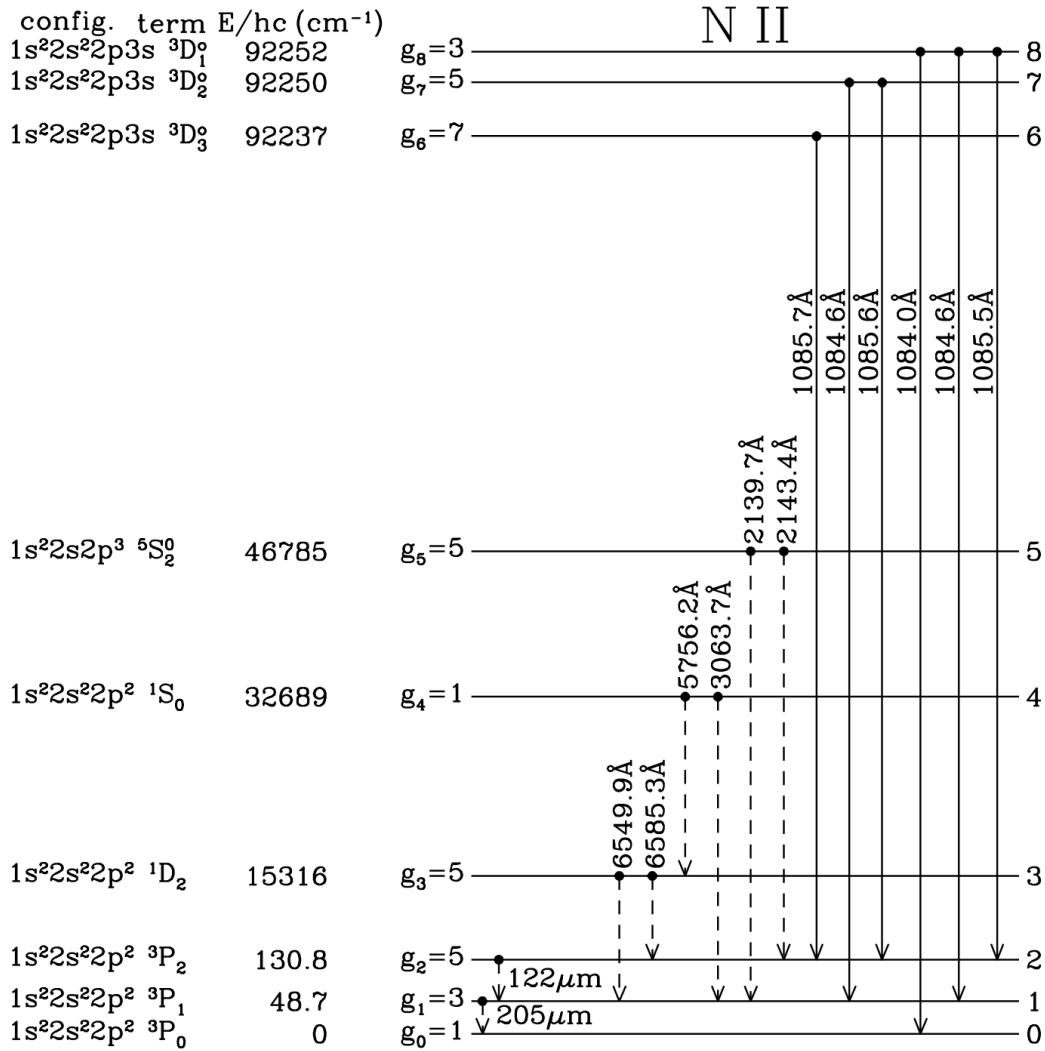


Figure 14: First nine energy levels of NII. Forbidden transitions are indicated by broken lines, and allowed transitions by solid lines; forbidden decays are not shown from levels that have permitted decay channels. Fine-structure splitting is not to scale. Hyperfine splitting is not shown. Figure taken from Draine (2015).

Spin-Forbidden or inter-system transitions: These are transitions that fulfill the electric dipole selection rules 1 to 4 but have $\Delta S \neq 0$. These transitions are considerably weaker than allowed transitions. Such transitions are sometimes referred to as **semi-forbidden**, or **inter-combination**, or **inter-system transitions**; the latter is the terminology that we will use. An inter-system transition is denoted with a single right bracket – for example,

$$\text{[NII]} 2143.4\text{ Å} {}^3P_2 - {}^5S_2^o.$$

This is a transition between the $\ell = 1s^2 2s^2 2p^2 3P_2$ and $u = 1s^2 2s 2p^3 {}^5S_2^o$ levels of NII, with a wavelength $\lambda_{ul} = 2143.4\text{ Å}$ and $A_{ul} = 1.27 \times 10^2 \text{ s}^{-1}$.

Forbidden transitions: Forbidden transitions are those that fail to fulfill at least one of the selection rules 1 to 4. The transition probabilities vary widely, depending on the values of the electric quadrupole or magnetic dipole matrix elements between the upper and lower states. A forbidden transition is denoted with two square brackets – for example,

$$\text{[NII]} 6549.9\text{ Å} {}^3P_1 - {}^1D_2^o.$$

This is a transition between the $\ell = 1s^2 2s^2 2p^2 3P_1$ and $u = 1s^2 2s^2 2p^2 {}^1D_2^o$ levels of NII, with a wavelength $\lambda_{ul} = 6549.9\text{ Å}$ and $A_{ul} = 9.20 \times 10^{-4} \text{ s}^{-1}$. This fails rule 1 (parity is unchanged) and it fails rule 4 (single electron wave functions are unchanged). This is an example of a **magnetic dipole transition**.

Another example of a forbidden transition is the **electric quadrupole transition**

$$\text{[NII]} 5756.2\text{ Å} {}^1D_2 - {}^1S_0^o.$$

This is a transition between the $\ell = 1s^2 2s^2 2p^2 1D_2$ and $u = 1s^2 2s^2 2p^2 1S_0^o$ levels of NII, with a wavelength $\lambda_{ul} = 5756.2 \text{ \AA}$ and $A_{ul} = 1.17 \text{ s}^{-1}$. This fails rules 1 (parity is unchanged) and 4 (single electron wave functions are unchanged) and it fails rules 2 and 3 ($\Delta L = -2$ and $\Delta J = -2$), yet its transition probability is three orders of magnitude larger than the magnetic dipole transition [NII] 6549.9 \AA !

We see then that there is a hierarchy in the transition probabilities: very roughly speaking, inter-system lines are $\sim 10^6$ times weaker than permitted transitions, and forbidden lines are $\sim 10^2 - 10^6$ times weaker than inter-system transitions.

Despite being very “weak,” forbidden transitions are important in astrophysics for the simple reason that every atom and ion has excited states that can only decay via forbidden transitions. At high densities, such excited states would be depopulated by collisions, but at the very low densities of interstellar space, collisions are sufficiently infrequent that there is time for forbidden radiative transitions to take place.

1.5.3 Follow-up Questions

- What are some common examples of forbidden lines?
- When do you get forbidden line absorption?
- What kind of radiation can be absorbed by forbidden line absorption?
- Can we observe any forbidden lines on Earth?
- What is the lifetime of the 21 cm line?
- At what redshifts do CMB photons get absorbed by 21 cm transitions?
- At what redshifts do you get neutral hydrogen?
- How would you estimate the lifetime from the maximum density where the line isn’t washed out?
- Collisions happen much more frequently; why is the 21 cm line still visible?
- Why does forbidden line emission cool the gas?
- Why are they *so good* at cooling the gas?

1.6 Question 6

What is a polytropic equation of state? Give examples of objects for which this is a very good approximation, and explain why it is.

1.6.1 Short answer

Hypothetical stellar models in which the pressure depends on density in the form $P = K\rho^\gamma$ are known as *polytropes*. Some examples include:

- ideal, monatomic gas: $\gamma = 5/3$ (such as radiative main sequence stars);
- non-relativistic, completely degenerate stars: $\gamma = 5/3$ (such as white dwarfs, fully convective stellar cores, and brown dwarfs);
- relativistic, completely degenerate stars: $\gamma = 4/3$ (such as neutron stars).

1.6.2 Additional context

Note that careful inspection of the stellar structure equations show that the mechanical equations of stellar structure could be solved simultaneously without reference to the energy equations only if a simple relationship existed between pressure and density. Unfortunately, such a simple relationship does not generally exist; normally, temperature and chemical composition must also enter into the pressure equation of state, often in a complicated way. However, under certain circumstances, such as for an adiabatic gas, the pressure can be written explicitly in terms of the density alone. Hypothetical stellar models in which the pressure depends on density in the form $P = K\rho^\gamma$ are known as *polytropes*. The discussion of polytropes was originally motivated by the equation of state of an adiabatic gas. For the case of an ideal, monatomic gas, $\gamma = 5/3$. In addition, certain extremely compressed stars in their final stage of evolution known as white dwarfs which are non-relativistic, completely degenerate stars, can also be described by $\gamma = 5/3$. In contrast, relativistic, completely degenerate stars can be described by $\gamma = 4/3$. The development of polytropic models is well worth the effort since their relative simplicity allows us to gain some insight into stellar structure without all of the complications inherent in full-blown numerical models. The famous equation that helps us describe analytical stellar models is referred to as the *Lane-Emden equation*.

Eulerian description: For gaseous, non-rotating, single stars without strong magnetic fields, the only forces acting on a mass element come from pressure and gravity. This results in a spherically symmetric configuration. All functions will then be constant on concentric spheres, and we need only one spatial variable to describe them. It seems natural to use the distance r from the stellar centre as the spatial coordinate, which varies from $r = 0$ at the centre to the total radius $r = R$ at the surface of the star. In addition, the evolution in time t requires a dependence of all functions on t : If we thus take r and t as independent variables, we have a typical “Eulerian” treatment in the sense of classical hydrodynamics. Then all other variables are considered to depend on these two, for example, the density $\rho = \rho(r, t)$.

Lagrangian description: It will turn out that, in the spherically symmetric case, it is often more useful to take a “Lagrangian” coordinate instead of r (i.e., one which is connected to the mass elements). The spatial coordinate of a given mass element then does not vary in time. We choose for this coordinate the above defined m : to any mass element, the value m (which is the mass contained in a concentric sphere at a given moment t_0) is assigned once and for all.

The new independent variables are then m and t ; and all other variables are considered to depend on them; for example, $\rho = \rho(m, t)$. This also includes the radial distance r of our mass element from the centre, which is now described by the function $r = r(m, t)$. Since there is certainly no singularity of ρ at the centre, we have here $m = 0$, while the star’s total mass $m = M$ is reached at the surface (i.e., where $r = R$). This already shows one advantage of the new description for the (normal) case of stars with constant total mass: while the radius R varies strongly in time, the star always extends over the same interval of the independent variable $m : 0 \leq m \leq M$. Although real stars do lose mass, for example, by stellar winds or due to gravitational interaction in binary systems, over short timescales the assumption of constant mass is justified nevertheless. In any case, the range of m never changes by more than a factor of a few.

As just indicated, there will certainly be no problem concerning a unique one-to-one transformation between the two coordinates r and m : We then easily find the connection between the partial derivatives in the two cases from well-known formulae. For any function depending on two variables, one of which is substituted by a new one ($r, t \rightarrow m, t$), the partial derivatives with respect to the new variables are given by

$$\frac{\partial}{\partial m} = \frac{\partial}{\partial r} \cdot \frac{\partial r}{\partial m}$$

$$\left(\frac{\partial}{\partial t} \right)_m = \frac{\partial}{\partial r} \cdot \left(\frac{\partial r}{\partial t} \right)_m + \left(\frac{\partial}{\partial t} \right)_r.$$

Subscripts indicate which of the spatial variables (m or r) is considered constant. The second of these equations reveals the main reason for the choice of the Lagrangian description. Its left-hand side gives the so-called **substantial time derivative** of hydrodynamics. It describes the change of a function in time when following a given mass element (e.g., the change of a physical property of this mass element). The conservation laws for time-dependent spherical stars give very simple equations only in terms of this substantial time derivative. In terms of a local time derivative, $(\partial/\partial t)_r$, the description would become much more complicated since the “convective” terms with the velocity $(\partial r/\partial t)_m$ would appear explicitly. Let us apply the first of these to m . We know that m varies with r as

$$\frac{\partial m}{\partial r} = 4\pi r^2 \rho.$$

Taking the partial derivative $\partial/\partial m$, we obtain

$$\begin{aligned} \frac{\partial}{\partial m} (\partial m) &= \frac{\partial}{\partial m} (4\pi r^2 \rho) \partial r \\ 1 &= \frac{\partial r}{\partial m} (4\pi r^2 \rho) \\ \frac{\partial r}{\partial m} &= \frac{1}{4\pi r^2 \rho}. \end{aligned}$$

This is a differential equation describing the spatial behaviour of the function $r(m, t)$. It replaces $\partial m/\partial r = 4\pi r^2 \rho$ in the Lagrangian description. Introducing this differential equation into the partial differential equation we just found for $\partial/\partial m$, we find the general recipe for the transformation between the two operators:

$$\frac{\partial}{\partial m} = \frac{1}{4\pi r^2 \rho} \frac{\partial}{\partial r}.$$

In the deep interiors of stars the gases are fully ionized (i.e., for each hydrogen nucleus there also exists a free electron, while for each helium nucleus there are two free electrons). We therefore have a mixture of two gases, that of the nuclei (which in itself can consist of more than one component) and that of the free electrons. The mixture can be treated similarly to a one-component gas, if all single components obey the ideal gas law.

Let's consider a mixture of fully ionized nuclei. The chemical composition can be described by specifying all X_i ; the weight fractions of nuclei of type i which have molecular weight μ_i and charge number Z_i . If we have n_i nuclei per volume and a “partial density” ρ_i , then $X_i = \rho_i/\rho$ and

$$n_i = \frac{\rho_i}{\mu_i m_H} = \frac{\rho}{m_H} \frac{X_i}{\rho_i}.$$

By L_r we define the net energy per second passing outward through a sphere of radius r . The function L_r is zero at $r = 0$, since there can be no infinite energy source at the centre, while L_r reaches the total luminosity L of the star at the surface. In between, L_r can be a complicated function, depending on the distribution of the sources and sinks of energy.

The function L_r comprises the energies transported by radiation, conduction, and convection. Not included is a possible energy flux by neutrinos, which normally have negligible interaction with the stellar matter. Also included in L_r are only those fluxes which require a temperature gradient.

Consider a spherical mass shell of radius r ; thickness dr ; and mass dm . The energy per second entering the shell at the inner surface is L_r , while $L_r + dL_r$ is the energy per second leaving it through the outer surface. The surplus power dL_r can be provided by nuclear reactions, by cooling, or by compression or expansion of the mass shell.

We first consider a stationary case in which dL_r is due to the release of energy from nuclear reactions only. Let ϵ be the **nuclear energy produced per gram of stellar material**; then

$$\begin{aligned} dL_r &= 4\pi r^2 \rho \epsilon dr = \epsilon dm \\ \frac{dL_r}{dm} &= \epsilon [\text{eV g}^{-1}]. \end{aligned}$$

In general ϵ depends on the temperature and density, and on the abundance of the different nuclear species that react.

Let us denote the types of reacting particles, X and a , by indices j and k respectively. Suppose there is one particle of type j moving with a velocity v relative to all particles of type k . Its cross section σ for reactions with k sweeps over a volume σv per second. The number of reactions per second will then be $n_k \sigma v$ if there are n_k particles of type k per unit volume. For n_j particles per unit volume the total number of reactions per units of volume and time is

$$\tilde{r}_{jk} = n_j n_k \sigma v.$$

This product may also be interpreted by saying that $n_j n_k$ is the number of pairs of possible reaction partners, and σv gives the reaction probability per pair and second. This indicates what we have to do in the case of reactions between identical particles ($j = k$). Then the number of pairs that are possible reaction partners is $n_j(n_j - 1)/2 \approx n_j^2/2$ for large particle numbers. This has to replace the product $n_j n_k$ so that we can generally write

$$\tilde{r}_{jk} = \frac{1}{1 + \delta_{jk}} n_j n_k \sigma v$$

The full set of equations: Collecting the basic differential equations for a spherically symmetric star in hydrostatic equilibrium, we obtain a full set of differential equations that describe stellar evolution:

$$\begin{aligned} \frac{\partial r}{\partial m} &= \frac{1}{4\pi r^2 \rho} \\ \frac{\partial P}{\partial r} &= -\frac{GM_r}{4\pi r^4} \\ \frac{\partial L}{\partial m} &= \epsilon_n - \epsilon_\nu - c_P \frac{\partial T}{\partial t} + \frac{\delta}{\rho} \frac{\partial P}{\partial t} \\ \frac{\partial T}{\partial m} &= -\frac{GM_r T}{4\pi r^4 P} \nabla \\ \frac{\partial X_i}{\partial t} &= \frac{m_i}{\rho} \left(\sum_j j_{ji} - \sum_k r_{ik} \right), \quad i = 1, \dots, I. \end{aligned}$$

The second equation has an additional term $-\partial^2 r / \partial t^2 (4\pi r^2)^{-1}$ in case the assumption of hydrostatic equilibrium is not fulfilled. In the last equation, we have a set of I equations (one of which may be replaced by the normalization $\sum_i X_i = 1$) for the change of mass fractions X_i of the relevant nuclei $i = 1, \dots, I$ having masses m_i . In the third equation, $\delta \equiv -(\partial \ln \rho / \partial \ln T)_P$, and in the second last equation $\nabla \equiv d \ln T / d \ln P$. If the energy transport is due to radiation (and conduction), then ∇ has to be replaced by ∇_{rad} given by

$$\nabla_{\text{rad}} = \frac{3}{16\pi acG} \frac{\kappa LP}{mT^4} [\text{KP}^{-1}].$$

If the energy is carried by convection, then ∇ has to be replaced by a value obtained from a proper theory of convection; this may be ∇_{ad} in the deep interior given by

$$\nabla_{\text{ad}} = \left(1 - \frac{1}{\gamma}\right) \frac{\mu m_H}{k} \frac{GM_r}{r^2} [\text{K m}^{-1}],$$

or one obtained from a solution of the cubic equation for super-adiabatic convection in the outer layers. The boxed equations contain functions which describe properties of the stellar material such as $\rho, \epsilon_n, \epsilon_\nu, \kappa, c_P$, and the reaction rates r_{ij} . We shall assume them to be known functions of P, T , and the chemical composition described by the functions $X_i(m, t)$. We therefore have an equation of state

$$\rho = \rho(P, T, X_i) [\text{g cm}^{-3}],$$

and equations for the other thermodynamic properties of the stellar matter

$$\begin{aligned} c_P &= c_P(P, T, X_i) \\ \delta &= \delta(P, T, X_i) \\ \nabla_{\text{ad}} &= \nabla_{\text{ad}}(P, T, X_i), \end{aligned}$$

as well as the Rosseland mean of the opacity (including conduction)

$$\kappa = \kappa(P, T, X_i),$$

and the nuclear reaction rates and the energy production and energy loss via neutrinos:

$$\begin{aligned} r_{jk} &= r_{jk}(P, T, X_i) \\ \epsilon_\nu &= \epsilon_\nu(P, T, X_i) \\ \epsilon_\nu &= \epsilon_\nu(P, T, X_i). \end{aligned}$$

Polytropic equations: The temperature does not appear explicitly in the first two boxed equations. Under certain circumstances this provides the possibility of separating them from the “thermo-energetic part” of the equations. For the following it is convenient to introduce the gravitational potential Φ . We here treat stars in hydrostatic equilibrium, which requires

$$\frac{dP}{dr} = -\frac{d\Phi}{dr}\rho,$$

together with Poisson’s equation

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Phi}{dr} \right) = 4\pi G\rho.$$

We have replaced the partial derivatives by ordinary ones since only time-independent solutions shall be considered.

In general the temperature appears in these equations if the density is replaced by an equation of state of the form $\rho = \rho(P, T)$. However, if ρ does not depend on T (i.e., $\rho = \rho(P)$), then this relation can be introduced, which become a system of two equations for P and Φ and can be solved without the other structure equations. An example is the completely degenerate gas of non-relativistic electrons for which $\rho \sim P^{3/5}$.

We shall deal here with similar cases and assume that there exists a simple relation between P and ρ of the form

$$P = K\rho^\gamma \equiv K\rho^{1+1/n} [P],$$

where K , γ , and n are constants. A relation of this form is called a **polytropic equation of state**. K is the **polytropic constant** and γ is the **polytropic exponent** (which we have to distinguish from the adiabatic exponent γ_{ad}). One often uses, instead of γ , the **polytropic index** n , which is defined by

$$n = \frac{1}{\gamma - 1} \text{ [dimensionless].}$$

For a completely degenerate gas the equation of state in its limiting cases has the polytropic form as seen above. In the non-relativistic limit we have $\gamma = 5/3$, $n = 3/2$, while for the relativistic limit $\gamma = 4/3$, $n = 3$. For such cases, where the equation of state has a polytropic form, the polytropic constant K is fixed and can be calculated from natural constants. There are also examples of the polytropic equation of state where K is a free parameter which is constant within a particular star but can have different values from one star to another.

In a star that is completely convective the temperature gradient (except for that in a region near the surface, which we shall ignore) is given, to a very good approximation, by $\nabla = (d \ln T / d \ln P)_{ad} = \nabla_{ad}$. If radiation pressure can be ignored and the gas is completely ionized, we have $\nabla_{ad} = 2/5$. This means that throughout the star $T \sim P^{2/5}$, and for an ideal gas with $\mu = \text{constant}$, $T \sim P/\rho$, and therefore $P \sim \rho^{5/3}$. This again is a polytropic relation with $\gamma = 5/3$, $n = 3/2$. But now K is not fixed by natural constants; it is a free parameter in the sense that it can vary from star to star.

The homogeneous gaseous sphere can also be considered a special case of the polytropic relation. Let us write the polytropic equation of state in the form

$$\rho = K_1 P^{1/\gamma} [\text{g cm}^{-3}].$$

Then $\gamma = \infty$ (or $n = 0$) gives $\rho = K_1 = \text{constant}$.

These examples have shown that we can have two reasons for a polytropic relation in a star:

1. The equation of state is of the simple form $P = K\rho$ with a fixed value of K ;
2. The equation of state contains T (as for an ideal gas), but there is an additional relation between T and P (like the adiabatic condition) that together with the equation of state yields a polytropic relation; then K is a free parameter.

On the other hand, if we assume a polytropic relation for an ideal gas, this is equivalent to adopting a certain relation $T = T(P)$. This means that one fixes the temperature stratification instead of determining it by the thermo-energetic equations of stellar structure. For example, a polytrope with $n = 3$ does not necessarily have to consist of relativistic degenerate gases but can also consist of an ideal gas and have $\nabla = 1/(n + 1) = 0.25$.

Polytropic stellar models: With the polytropic relation (independent of whether K is a free parameter or a constant with a fixed value), the equation of hydrostatic equilibrium can be written as

$$\frac{d\Phi}{dr} = -\gamma K \rho^{\gamma-2} \frac{d\rho}{dr}.$$

If $\gamma \neq 1$ (the case $\gamma = 1, n = \infty$ corresponding to the isothermal model, this can be integrated:

$$\rho = \left(\frac{-\Phi}{(n+1)K} \right)^n [\text{g cm}^{-3}],$$

where we have made use of $n = 1/(\gamma-1)$ and chosen the integration constant to give $\Phi = 0$ at the surface ($\rho = 0$). Note that in the interior of our model, $\Phi < 0$, giving there $\rho > 0$. If we introduce this form of ρ into the right-hand side of the Poisson equation, we obtain an ordinary differential equation for Φ :

$$\frac{d^2\Phi}{dr^2} + \frac{2}{r} \frac{d\Phi}{dr} = 4\pi G \left(\frac{-\Phi}{(n+1)K} \right)^n.$$

We now define dimensionless variables z, w by

$$z = Ar \text{ [dimensionless]}, \quad A^2 = \frac{4\pi G}{(n+1)^n K^n} (-\Phi_c)^{n-1} = \frac{4\pi G}{(n+1)K} \rho_c^{(n-1)/n} \text{ [dimensionless]},$$

$$w = \frac{\Phi}{\Phi_c} = \left(\frac{\rho}{\rho_c} \right)^{1/n} \text{ [dimensionless]},$$

where the subscript c refers to the centre and where the relation between ρ and Φ is taken from $\rho = [-\Phi/(n+1)K]^n$. At the centre ($r = 0$) we have $z = 0, \Phi = \Phi_c, \rho = \rho_c$, and therefore $w = 1$. Then the ordinary differential equation for Φ can be written as

$$\frac{d^2w}{dz^2} + \frac{2}{z} \frac{dw}{dz} + w^n = 0$$

$$\frac{1}{z^2} \frac{d}{dz} \left(z^2 \frac{dw}{dz} \right) + w^n = 0.$$

This is the famous **Lane-Emden equation** (named after J.H. Lane and R. Emden). We are only interested in solutions that are finite at the centre, $z = 0$. This equation shows that we then have to require $dw/dz \equiv w' = 0$. Let us assume we have a solution $w(z)$ of the Lane-Emden equation that fulfills the central boundary conditions $w(0) = 1$ and $w'(0) = 0$; then according to the definition of w the radial distribution of the density is given by

$$\rho(r) = \rho_c w^n \text{ [g cm}^{-3}], \quad \rho_c = \left[\frac{-\Phi_c}{(n+1)K} \right]^n \text{ [g cm}^{-3}].$$

For the pressure we obtain from the polytropic equation that $P(r) = P_c w^{n+1}$, where $P_c = K \rho_c^\gamma$.

1.7 Question 7

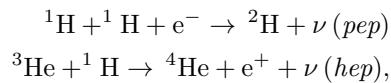
What was the Solar neutrino problem, and how was it resolved?

1.7.1 Short answer

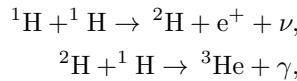
The Solar neutrino problem arose when the number of detected Solar neutrinos was lower by a factor of ~ 3 compared to theoretical predictions (2.5 SNU instead of 7.5 SNU). If this was correct, this could have implied that the Solar interior was cooler than they thought and that their Solar models were wrong. The solution to this problem came with the discovery that neutrinos could oscillate between three different “flavours”: the electron, muon, and tau neutrino. Original measurements were only sensitive to neutrinos with energies above 0.814 MeV, and thus were not sensitive to two of the neutrino flavours (explaining the detection deficit). In order for the neutrinos to be able to oscillate between flavours, the neutrinos must therefore have non-zero rest mass.

1.7.2 Additional context

Solar neutrinos: Some of the nuclear reactions of the pp chain, as well as of the CNO cycle, produce neutrinos. In addition, there are also neutrinos due to the very rare *pep* and *hep* reactions



the latter one being the trivial way to produce ${}^4\text{He}$ after the **pp-chain**,



but it is occurring in only 10^{-8} of all cases. However, the energy of the emitted neutrino is close to 10 MeV, and it is therefore necessary to consider this reaction. The neutrinos leave the star practically without interacting with the stellar matter. The energy spectrum of neutrinos from β decay is continuous, since the electrons can take part of the energy away, while neutrinos released after an inverse β decay are essentially monochromatic. Therefore most reactions of the pp chain have a continuous spectrum, while the pep-reaction and the electron capture on ${}^7\text{Be}$ have a line spectrum. Since ${}^7\text{Be}$ can decay into ${}^7\text{Li}$ either in the ground state or in an excited state, this reaction gives two spectral lines. The neutrino spectrum of the Sun as predicted from the reactions of the pp chain, computed from our standard solar model, is given in Figure 15.

Since the solar neutrinos can leave the Sun almost unimpeded they can in principle be measured in terrestrial laboratories and thus be used to learn directly about conditions in the innermost solar core. This difficult task indeed has been undertaken since 1964, when John Bahcall and Raymond Davies began to plan for an underground neutrino detector in a mine in Homestead, North Dakota. Forty years later the experiments finally have confirmed the standard solar model, and R. Davies received the Nobel Prize for his work. The time in between, however, was characterized by the **solar neutrino problem**.

The solar neutrino problem consisted in the fact that since the first results from the so-called chlorine experiment by Davies there was a lack of neutrinos compared to solar model predictions. The chlorine experiment is sensitive to neutrinos with energies above 0.814 MeV and therefore, as can be seen in Figure 15 mainly to the ${}^8\text{B}$ neutrinos, with some contribution from pep, hep, and ${}^7\text{Be}$ neutrinos. The experiment is based on the reaction ${}^{37}\text{Cl} + \nu \rightarrow {}^{37}\text{Ar}$, where the decays of radioactive argon nuclei are counted. The rate of neutrino captures is commonly measured in solar neutrino units (SNU). One SNU corresponds to 10^{-36} captures per second and per target nucleus. The predicted counts amount to 7.5 SNU for the chlorine experiment, the measurements averaged over several decades to only 2.5 ± 0.2 SNU. The deficit could indicate that the solar centre is cooler than in the models.

To improve the experimental evidence, additional experiments were started. First, another kind of radiochemical detector using gallium in the detector fluid measured, due to a much lower energy threshold, the majority of neutrinos, including those from the pp-reaction. Later, electron-scattering detectors were developed, which are sensitive to the highest energies only, but which provide directional information about the neutrino source (for these detectors the *hep*-neutrinos have to be taken into account.). All experiments confirmed that the solar neutrino flux was of the right order of magnitude, and therefore that indeed the Sun shines by the nuclear fusion of hydrogen, but they also consistently measured a deficit of neutrinos. This deficit, however, varied between different kinds of detectors.

With more and more experimental data it became evident that even hypothetical changes to the solar centre cannot solve the problem and that the solution is most likely to be found in the properties of neutrinos. All nuclear reactions emit electron neutrinos, and these are the only ones that were measured

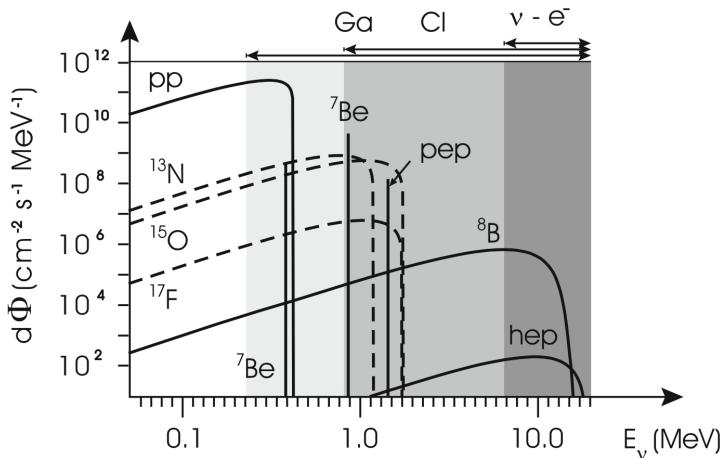


Figure 15: The neutrino spectrum of the Sun as predicted from the theoretical standard solar model. The solid lines belong to reactions of the pp chain while the broken lines are due to reactions of the CNO cycle. The neutrinos from most of the reactions have continuous spectra, while mono-energetic neutrinos come from ${}^7\text{Be}$ and from the pep-reaction. The flux for the continuum sources is given in $\text{cm}^2 \text{s}^{-1} \text{MeV}^{-1}$ and for the line sources in $\text{cm}^2 \text{s}^{-1}$. The sensitivity of the three types of neutrino experiments is indicated above the figure and by the shaded regions. Figure taken from Kippenhahn, Weigert & Weiss (2012).

in terrestrial experiment, with the exception of the electron-scattering Sudbury Neutrino Observatory (SNO) experiment in Canada, where heavy water (with a high percentage of deuterium isotopes) was used as the detector. Here also reactions with the two other types (flavours) of neutrinos, muon and tau neutrinos can be detected. Summing these and the electron neutrinos up, the total number of detections is completely consistent with the solar model prediction, within a few percent. What created the apparent solar neutrino deficit is the fact that neutrinos can change their flavour, both while travelling through vacuum and more efficiently in the presence of electrons in the solar interior. A similar effect was also confirmed for muon neutrinos arising in the Earth's upper atmosphere from high-energy cosmic radiation, when measured before or after they have travelled through the Earth's interior. The modelling of the solar interior, together with sophisticated experiments, has therefore resulted in new knowledge about fundamental properties of neutrinos. In particular, these so-called **neutrino oscillations** are possible only if neutrinos have mass.

The evolution of neutrino astronomy: The following is from the brief Millennium Essay for the PASP on the evolution of neutrino astronomy from the personal perspective of John Bahcall and Raymon Davis, Jr (2000).

The possibility of observing solar neutrinos began to be discussed seriously following the 1958 experimental discovery by H. P. Holmgren and R. Johnston that the cross section for production of the isotope ${}^7\text{Be}$ by the fusion reaction ${}^3\text{He} + {}^4\text{He} \rightarrow {}^7\text{Be} + \gamma$ was more than a thousand times larger than was previously believed. This result led Willy Fowler and Al Cameron to suggest that ${}^8\text{B}$ might be produced in the Sun in sufficient quantities by the reaction ${}^7\text{Be} + p \rightarrow {}^8\text{B} + \gamma$ to produce an observable flux of high energy neutrinos from ${}^8\text{B}$ beta-decay.

We begin our story in 1964, when we published back-to-back papers in Physical Review Letters arguing that a 100,000 gallon detector of perchloroethylene could be built which would measure the Solar neutrino capture rate on chlorine. Our motivation was to use neutrinos to look into the interior of the Sun and thereby test directly the theory of stellar evolution and nuclear energy generation in stars. The particular development that made us realize that the experiment could be done was the theoretical discovery (by John in late 1963) that the principal neutrino absorption cross section on chlorine was 20 times larger than previously estimated due to a super-allowed nuclear transition to an excited state of argon.

If you have a good idea today, it likely will require many committees, many years, and many people in order to get the project from concept to observation. The situation was very different in 1964. Once the decision to go ahead was made, a very small team designed and built the experiment; the entire team consisted of Ray, Don Harmer (on leave from Georgia Tech), and John Galvin (a technician who worked part-time on the experiment). Kenneth Hoffman, a (then) young engineer, provided expert advice on technical questions. The money came out of the chemistry budget at Brookhaven National Laboratory. Neither of us remember a formal proposal ever being written to a funding agency. The total capital expenditure to excavate the cavity in the Homestake Gold Mine in South Dakota, to build the tank, and to purchase the liquid was \$0.6 million dollars (in 1965 dollars).

During the period 1964-1967, Fred Reines and his group worked on three Solar neutrino experiments in which recoil electrons produced by neutrino interactions would be detected by observing the associated light in an organic scintillator. Two of the experiments, which exploited the elastic scattering of neutrinos by electrons, were actually performed and led to a (higher than predicted) upper limit on the ${}^8\text{B}$ solar

neutrino flux. The third experiment, which was planned to detect neutrinos absorbed by ^{7}Li , was abandoned after the initial chlorine results showed that the solar neutrino flux was low. These three experiments introduced the technology of organic scintillators into the arena of Solar neutrino research, a technique that will only finally be used in 2001 when the BOREXINO detector will begin to detect low-energy Solar neutrinos. Also during this period, John investigated the properties of neutrino electron scattering and showed that the forward peaking from ^{8}B neutrinos is large, a feature that was incorporated two and half decades later in the Kamiokande (and later Super-Kamiokande) water Cerenkov detectors. The first results from the chlorine experiment were published in 1968, again in a back-to-back comparison (in Physical Review Letters) between measurements and standard predictions. The initial results have been remarkably robust; the conflict between chlorine measurements and standard Solar model predictions has lasted over three decades. The main improvement has been in the slow reduction of the uncertainties in both the experiment and the theory. The efficiency of the Homestake chlorine experiment was tested by recovering carrier solutions, by producing ^{37}Ar in the tank with neutron sources, and by recovering ^{36}Cl inserted in a tank of perchloroethylene. The Solar model was verified by comparison with precise helioseismological measurements.

For much more than two decades, the best estimates for the observational and for the theoretical prediction have remained essentially constant. The discrepancy between the standard Solar model predictions and the chlorine observations became widely known as the **Solar neutrino problem**.

Very few people worked on Solar neutrinos during the period 1968-1988. The chlorine experiment was the only solar neutrino experiment to provide data in these two decades. It is not easy for us to explain why this was the case; we certainly tried hard to interest others in doing different experiments, and we gave many joint presentations about what came to be known as the solar neutrino problem. Each of us had one principal collaborator during this long period, Bruce Cleveland (experimental) and Roger Ulrich (Solar models). A large effort to develop a chlorine experiment in the Soviet Union was led by George Zatsepin, but it was delayed by the practical difficulties of creating a suitable underground site for the detector. Eventually, the effort was converted into a successful gallium detector, SAGE, led by Vladimir Gavrin and Tom Bowles, that gave its first results in 1990.

Only 1 year after the first (1968) chlorine results were published, Vladimir Gribov and Bruno Pontecorvo proposed that the explanation of the Solar neutrino problem was that neutrinos oscillated between the state in which they were created and a more difficult to detect state. This explanation, which is the consensus view today, was widely disbelieved by nearly all the particle physicists we talked to in those days. In the form in which Solar neutrino oscillations were originally proposed by Gribov and Pontecorvo, the process required that the mixing angles between neutrino states be much larger than the quark mixing angles, something which most theoretical physicists believed, at that time, was unlikely. Ironically, a flood of particle theory papers explained, more or less “naturally,” the large neutrino mixing angle that was decisively demonstrated 30 years later in the Super-Kamiokande atmospheric neutrino experiment.

One of the most crucial events for early Solar neutrino research occurred in 1968 while we were relaxing in the Sun after a swim at the Caltech pool. Gordon Garmire (now a PI for the Chandra X-ray Observatory) came up to Ray, introduced himself, and said he had heard about the chlorine experiment. He suggested to Ray that it might be possible to reduce significantly the background by using pulse rise-time discrimination, a technique used for proportional counters in space experiments. The desired fast-rising pulses from ^{37}Ar Auger electrons are different from the slower rising pulses from a background gamma or cosmic ray. Ray went back to Brookhaven and asked the local electronic experts if it would be possible to implement this technique for the very small counters he used. The initial answer was that the available amplifiers were not fast enough to be used for this purpose with the small Solar neutrino counters. But, in about a year three first-class electronic engineers at Brookhaven, Veljko Radeca, Bob Chase, and Lee Rogers, were able to build electronics fast enough to be used to measure the rise time in Ray’s counters. This “swimming pool” improvement was crucial for the success of the chlorine experiment and the subsequent radio-chemical gallium solar neutrino experiments, SAGE, GALLEX, and GNO. Measurements of the rise time as well as the pulse energy greatly reduce the background for radio-chemical experiments. The backgrounds can be as low as one event in 3 months.

In 1978, after a decade of disagreement between the Homestake neutrino experiment and standard Solar model predictions, it was clear to everyone that the subject had reached an impasse and a new experiment was required. The chlorine experiment is, according to standard Solar model predictions, sensitive primarily to neutrinos from a rare fusion reaction that involves ^{8}B neutrinos. These neutrinos are produced in only 2 of every 10^4 terminations of the basic p-p fusion chain. In the early part of 1978, there was a conference of interested scientists who got together at Brookhaven to discuss what to do next. The consensus decision was that we needed an experiment that was sensitive to the low-energy neutrinos from the fundamental p-p reaction.

The only remotely practical possibility appeared to be another radio-chemical experiment, this time with ^{71}Ga (instead of ^{37}Cl) as the target. But, a gallium experiment (originally proposed by the Russian

theorist V. A. Kuzmin in 1965) was expensive; we needed about 3 times the world's annual production of gallium to do a useful experiment. In an effort to generate enthusiasm for a gallium experiment, we wrote another Physical Review Letters paper, this time with a number of interested experimental colleagues. We argued that a gallium detector was feasible and that a gallium measurement, which would be sensitive to the fundamental p-p neutrinos, would distinguish between broad classes of explanations for the discrepancy between prediction and observation in the ^{37}Cl experiment. Over the next 5 or 6 years, the idea was reviewed a number of times in the United States, always very favorably. The Department of Energy (DOE) appointed a blue ribbon panel headed by Glen Seaborg that endorsed enthusiastically both the experimental proposal and the theoretical justification.

To our great frustration and disappointment, the gallium experiment was never funded in the United States, although the experimental ideas that gave rise to the Russian experiment (SAGE) and the German-French-Italian-Israeli-US experiment (GALLEX) originated largely at Brookhaven. Physicists strongly supported the experiment and said the money should come out of the astronomy budget; astronomers said it was great physics and should be supported by the physicists. DOE could not get the nuclear physics and the particle physics sections to agree on who had the financial responsibility for the experiment. In a desperate effort to break the deadlock, John was even the PI of a largely Brookhaven proposal to the National Science Foundation (which did not support proposals from DOE laboratories). A pilot experiment was performed with 1.3 tons of gallium by an international collaboration (Brookhaven, University of Pennsylvania; Max Planck Institute, Heidelberg; Institute for Advanced Study, Princeton; and the Weizmann Institute) which developed the extraction scheme and the counters eventually used in the GALLEX full-scale experiment.

In strong contrast to what happened in the United States, Moissey Markov, the head of the Nuclear Physics Division of the Russian Academy of Sciences, helped establish a neutrino laboratory within the Institute for Nuclear Research, participated in the founding of the Baksan neutrino observatory, and was instrumental in securing 60 tons of gallium free to Russian scientists for the duration of a Solar neutrino experiment. The Russian-American gallium experiment (SAGE) went ahead under the leadership of Vladimir Gavrin, George Zatsepin (Institute for Nuclear Research, Russia), and Tom Bowles (Los Alamos), and the mostly European experiment (GALLEX) was led by Till Kirsten (Max Planck Institute, Germany). Both experiments had a strong but not primary US participation.

The two gallium experiments were performed in the decade of the 1990s and gave very similar results, providing the first experimental indication of the presence of p-p neutrinos. Both experiments were tested by measuring the neutrino rate from an intense laboratory radioactive source.

There were two dramatic developments in the Solar neutrino saga, one theoretical and one experimental, before the gallium experiments produced observational results. In 1985, two Russian physicists proposed an imaginative solution of the Solar neutrino problem that built upon the earlier work of Gribov and Pontecorvo and, more directly, the insightful investigation by Lincoln Wolfenstein (of Carnegie Mellon University). Stanislav Mikheyev and Alexei Smirnov showed that if neutrinos have masses in a relatively wide range, then a resonance phenomenon in matter (now universally known as the **MSW effect**) could convert efficiently many of the electron-type neutrinos created in the interior of the Sun to more difficult to detect muon and tau neutrinos. The MSW effect can work for small or large neutrino mixing angles. Because of the elegance of the theory and the possibility of explaining the experimental results with small mixing angles (analogous to what happens in the quark sector), physicists immediately began to be more sympathetic to particle physics solutions to the Solar neutrino problem. More importantly, they became enthusiasts for new Solar neutrino experiments.

The next big breakthrough also came from an unanticipated direction. The Kamiokande water Cerenkov detector was developed to study proton decay in a mine in the Japanese Alps; it set an important lower limit on the proton lifetime. In the late 1980s, the detector was converted by its Japanese founders, Masatoshi Koshiba and Yoji Totsuka, together with some American colleagues (Gene Beier and Al Mann of the University of Pennsylvania), to be sensitive to the lower energy events expected from Solar neutrinos. With incredible foresight, these experimentalists completed in late 1986 their revisions to make the detector sensitive to Solar neutrinos, just in time to observe the neutrinos from SN 1987a emitted in the LMC 170,000 years earlier. (Supernova and Solar neutrinos have similar energies, ~ 10 MeV, much less than the energies that are relevant for proton decay.) In 1996, a much larger water Cerenkov detector (with 50,000 tons of pure water) began operating in Japan under the leadership of Yoji Totsuka, Kenzo Nakamura, Yoichiro Suzuki (from Japan), and Jim Stone and Hank Sobel (from the United States).

So far, five experiments have detected solar neutrinos in approximately the numbers (within a factor of 2 or 3) and in the energy range (< 15 MeV) predicted by the standard Solar model. This is a remarkable achievement for Solar theory since the ^8B neutrinos that are observed primarily in three of these experiments (chlorine, Kamiokande, and its successor Super-Kamiokande) depend upon approximately the 25th power of the central temperature. The same set of nuclear fusion reactions that are hypothesized to produce the Solar luminosity also give rise to Solar neutrinos. Therefore, these experiments establish

empirically that the Sun shines by nuclear fusion reactions among light elements in essentially the way described by Solar models.

Nevertheless, all of the experiments disagree quantitatively with the combined predictions of the standard Solar model and the standard theory of electroweak interactions (which implies that nothing much happens to the neutrinos after they are created). The disagreements are such that they appear to require some new physics that changes the energy spectrum of the neutrinos from different fusion sources.

Solar neutrino research today is very different from what it was three decades ago. The primary goal now is to understand the neutrino physics, which is a prerequisite for making more accurate tests of the neutrino predictions of Solar models. Solar neutrino experiments today are all large international collaborations, each typically involving of order 10^2 physicists. Nearly all of the new experiments are electronic, not radio-chemical, and the latest generation of experiments measure typically several thousand events per year (with reasonable energy resolution), compared to rates that were typically 25 – 50 per year for the radio-chemical experiments (which have no energy resolution, only an energy threshold). Solar neutrino experiments are currently being carried out in Japan (Super-Kamiokande, in the Japanese Alps), in Canada (SNO, which uses a kiloton of heavy water in Sudbury, Ontario), in Italy (BOREXINO, ICARUS, and GNO, each sensitive to a different energy range and all operating in the Gran Sasso Underground Laboratory), in Russia (SAGE, in the Caucasus region), and in the United States (Homestake chlorine experiment). The SAGE, chlorine, and GNO experiments are radio-chemical; all the others are electronic. Since 1985, the chlorine experiment has been operated by the University of Pennsylvania under the joint leadership of Ken Lande and Ray Davis. Lande and Paul Wildenhain have introduced major improvements in the extraction and measurement systems, making the chlorine experiment a valuable source of new precision data.

The most challenging and important frontier for Solar neutrino research is to develop experiments that can measure the energies of individual low-energy neutrinos from the basic p-p reaction, which constitutes (we believe) more than 90% of the Solar neutrino flux.

Solar neutrino research is a community activity. Hundreds of experimentalists have collaborated to carry out difficult, beautiful measurements of the elusive neutrinos. Hundreds of other researchers helped redefine the Solar model predictions, measuring accurate nuclear and Solar parameters and calculating input data such as opacities and equation of state.

Three people played special roles. Hans Bethe was the architect of the theory of nuclear fusion reactions in stars, as well as our mentor and hero. Willy Fowler was a powerful and enthusiastic supporter of each new step, and his keen physical insight motivated much of what was done in Solar neutrino research. Bruno Pontecorvo opened everyone's eyes with his original insights, including his early discussion of the advantages of using chlorine as a neutrino detector and his suggestion that neutrino oscillations might be important.

In the next decade, neutrino astronomy will move beyond our cosmic neighborhood and, we hope, will detect distant sources. The most likely candidates now appear to be gamma-ray bursts. If the standard fireball picture is correct and if gamma-ray bursts produce the observed highest energy cosmic rays, then very high energy (10^{15} eV) neutrinos should be observable with a km^2 detector. Experiments with the capability to detect neutrinos from gamma-ray bursts are being developed at the South Pole (AMANDA and ICECUBE), in the Mediterranean Sea (ANTARES, NESTOR), and even in space.

Looking back on the beginnings of solar neutrino astronomy, one lesson appears clear to us: if you can measure something new with reasonable accuracy, then you have a chance to discover something important. The history of astronomy shows that very likely what you will discover is not what you were looking for. It helps to be lucky.

1.7.3 Follow-up Questions

- Can neutrinos account for all of the dark matter in the Universe?
- If there was a fourth neutrino flavour, how would we detect it?
- Why is the first step of the p-p chain not where we looked for neutrino signatures? (i.e., not energetic enough)

1.8 Question 8

Why is nuclear fusion stable inside a main-sequence star? Under what conditions is nuclear fusion unstable? Give examples of actual objects.

1.8.1 Short answer

Answer.

1.8.2 Additional context

Hydrostatic equilibrium: The gravitational force is always attractive, implying that an opposing force must exist if a star is to avoid collapse. This force is provided by pressure. For a star to be supported against its own self-gravity, it must satisfy **hydrostatic equilibrium**:

$$\frac{dP}{dr} = -\frac{GM_r\rho}{r^2} = -\rho g \text{ [P m}^{-1}\text]}, \quad g \equiv \frac{GM_r}{r^2} \text{ [m s}^{-2}\text]},$$

where g is the local acceleration of gravity at radius r . This equation represents one of the fundamental equations of stellar structure for spherically symmetric objects under the assumption that accelerations are negligible. It clearly indicates that in order for a star to be static, a *pressure gradient* dP/dr must exist to counteract the force of gravity. It is not the pressure that supports a star, but the change in pressure with radius. Furthermore, the pressure must *decrease* with increasing radius; the pressure is necessarily larger in the interior than it is near the surface.

The equation of mass conservation: A second relationship involving mass, radius, and density also exists called the **mass conservation equation**:

$$\frac{dM_r}{dr} = 4\pi r^2 \rho \text{ [kg m}^{-1}\text]},$$

which dictates how the interior of the mass of a star must change with distance from the center. This is the second of the fundamental equations of stellar structure.

Pressure equation of state: So far, no information has been provided about the origin of the pressure term required by the equation of hydrostatic equilibrium. To describe this macroscopic manifestation of particle interactions, it is necessary to know the pressure **equation of state** of the material. Such an equation of state relates the dependence of pressure on other fundamental parameters of the material. One well-known example of a pressure equation of state is the **ideal gas law**, often expressed as

$$PV = Nk_B T \text{ [P m}^{-3}\text]},$$

where V is the volume of the gas, N is the number of particles, T is the temperature, and k_B is Boltzmann's constant.

In astrophysical applications, it is often convenient to express the ideal gas law in alternative form. We know that n is the particle number density which must be related to the mass density of the gas. Allowing for a variety of particles of different masses, it is then possible to express n as

$$n = \frac{\rho}{\bar{m}} = \frac{\rho}{\mu m_H} \text{ [m}^{-3}\text]},$$

where \bar{m} is the average mass of a gas particle. Substituting into the ideal gas law, we get

$$P = \frac{\rho k_B T}{\bar{m}} \text{ [P].}$$

We can define a new quantity, the **mean molecular weight**, as

$$\mu \equiv \frac{\bar{m}}{m_H} = \frac{\rho}{nm_H} \text{ [dimensionless]},$$

where $m_H = 1.67 \times 10^{-27} \text{ kg}$ is the mass of the hydrogen atom. *The mean molecular weight is just the average mass of a free particle in the gas, in units of the mass of hydrogen.* The ideal gas law can now be written in terms of the mean molecular weight as

$$P = \frac{\rho k_B T}{\mu m_H} \text{ [P].}$$

The mean molecular weight depends on the composition of the gas as well as on the state of ionization of each species. The level of ionization enters because free electrons must be included in the average mass per particle \bar{m} . This implies that a detailed analysis of the Saha equation is necessary to calculate the relative numbers of ionization states. When the gas is either completely neutral or completely ionized, the calculation simplifies significantly.

For a completely neutral gas,

$$\bar{m}_n = \frac{\sum_j N_j m_j}{\sum_j N_j} [\text{kg}],$$

where m_j and N_j are, respectively, the mass and total number of atoms of type j that are present in the gas, and the sums are assumed to be carried out over all types of atoms. Dividing by m_{H} yields

$$\mu_n = \frac{\sum_j N_j A_j}{\sum_j N_j} [\text{dimensionless}]$$

where $A_j \equiv m_j/m_{\text{H}}$. Similarly, for a completely ionized gas,

$$\mu_i \simeq \frac{\sum_j N_j A_j}{\sum_j N_j (1 + z_j)} [\text{dimensionless}]$$

where $(1 + z_j)$ accounts for the nucleus plus the number of free electrons that result from completely ionizing an atom of type j . (Do not confuse z_j with Z , the mass fraction of metals.)

By inverting the expression for \bar{m} , it is possible to write alternative equations for μ in terms of mass fractions. Recalling that $\bar{m} = \mu m_{\text{H}}$, for a neutral gas we get

$$\begin{aligned} \frac{1}{\mu_n m_{\text{H}}} &= \frac{\sum_j N_j}{\sum_j N_j m_j} \\ &= \frac{\text{total number of particles}}{\text{total mass of gas}} \\ &= \sum_j \frac{\text{number of particles from } j}{\text{mass of particles from } j} \cdot \frac{\text{mass of particles from } j}{\text{total mass of gas}} \\ &= \sum_j \frac{N_j}{N_j A_j m_{\text{H}}} X_j \\ &= \sum_j \frac{1}{A_j m_{\text{H}}} X_j [\text{dimensionless}], \end{aligned}$$

where X_j is the mass fraction of atoms of type j . Solving for $1/\mu_n$, we have

$$\frac{1}{\mu_n} = \sum_j \frac{1}{A_j} X_j [\text{dimensionless}].$$

Thus, for a neutral gas,

$$\frac{1}{\mu_n} \simeq X + \frac{1}{4} Y + \left\langle \frac{1}{A} \right\rangle_n Z [\text{dimensionless}].$$

$\langle 1/A \rangle_n$ is a weighted average of all elements in the gas heavier than helium. For solar abundances, $\langle 1/A \rangle_n \sim 1/15.5$.

The mean molecular weight of a completely ionized gas may be determined in a similar way. It is necessary only to include the *total* number of particles contained in the sample, both nuclei and electrons. For instance, each hydrogen atom contributes one free electron, together with its nucleus, to the total number of particles. Similarly, one helium atom contributes two free electrons plus its nucleus. Therefore, for a completely ionized gas we have

$$\frac{1}{\mu_i} = \sum_j \frac{1 + z_j}{A_j} X_j [\text{dimensionless}].$$

Including hydrogen and helium explicitly, we have

$$\frac{1}{\mu_i} \simeq 2X + \frac{3}{4}Y + \left\langle \frac{1+z}{A} \right\rangle_i Z [\text{dimensionless}].$$

For elements heavier than helium, $1 + z_j \simeq z_j$, where $z_j \gg 1$ represents the number of protons (or electrons) in an atom of type j . It also holds that $A_j \simeq 2z_j$, the relation being based on the facts that sufficiently massive atoms have approximately the same number of protons and neutrons in their nuclei and that protons and neutrons have very similar masses. Thus

$$\left\langle \frac{1+z}{A} \right\rangle_i \simeq \frac{1}{2} [\text{dimensionless}].$$

If we assume that $X = 0.70$, $Y = 0.28$, and $Z = 0.02$, a composition typical of younger stars, then with these expressions for the mean molecular weight, $\mu_n = 1.30$ and $\mu_i = 0.62$.

The average kinetic energy per particle: It is possible to combine the ideal gas law with the pressure integral to find the average kinetic energy per particle:

$$\frac{1}{2}mv^2 = \frac{3}{2}k_B T \text{ [J].}$$

The factor of 3 arises from averaging particle velocities over three coordinate directions (i.e., degrees of freedom). Thus the average kinetic energy of a particle is $(1/2)k_B T$ per degree of freedom.

The contribution due to radiation pressure: Because photons possess momentum $p = h\nu/c$, they are capable of delivering an impulse to other particles during absorption or reflection. Consequently, EM radiation results in another form of pressure called the **radiation pressure**:

$$P_{\text{rad}} = \frac{1}{3}aT^4 \text{ [P],}$$

where $a = 7.57 \times 10^{-15} \text{ erg cm}^{-3} \text{ K}^{-4}$ is the **radiation density constant**.

In many astrophysical situations, the pressure due to photons can actually exceed by a significant amount the pressure produced by the gas. In fact, it is possible that the magnitude of the force due to radiation pressure can become sufficiently great that it surpasses the gravitational force, resulting in an overall expansion of the system.

Combining both the ideal gas and radiation pressure terms, the total pressure becomes

$$P_t = P_g + P_{\text{rad}}$$

$$P_t = \frac{\rho k_B T}{\mu m_H} + \frac{1}{3}aT^4 \text{ [P].}$$

Gravitation potential energy: One likely source of stellar energy is gravitational potential energy. Recall that the gravitational potential of a system of two particles is given by

$$E_g = -G \frac{Mm}{r} \text{ [J].}$$

As the distance between M and m diminishes, the gravitational potential energy becomes *more negative*, implying that energy must have been converted to other forms, such as kinetic energy. If a star can manage to convert its gravitational potential energy into heat and then radiate that heat into space, the star may be able to shine for a significant period of time. However, we must also remember that by the **virial theorem**, the total energy of a system of particles in equilibrium is one-half of the system's potential energy. Therefore, only one-half of the change in gravitational potential energy of a star is actually available to be radiated away.

The total gravitational potential energy can be calculated by integrating over all mass shells from the center to the surface of the star

$$E_g = -4\pi G \int_0^R M_r \rho r dr \text{ [J],}$$

where R is the radius of the star.

An exact calculation of E_g requires knowledge of how ρ and consequently M_r depend on r . Nevertheless, an approximate value can be obtained by assuming that ρ is constant and equal to its average value,

$$\rho \sim \bar{\rho} = \frac{M}{\frac{4}{3}\pi R^3} \text{ [kg m}^{-3}\text{]},$$

M being the total mass of the star. Now we may also approximate M_r as

$$M_r \sim \frac{4}{3}\pi r^3 \bar{\rho} M \text{ [M}_\odot\text{].}$$

Substituting this into our equation for the gravitational potential energy, we get

$$E_g \sim -\frac{16\pi^2}{15} G \bar{\rho}^2 R^5 \sim -\frac{3}{5} \frac{GM^2}{R} \text{ [J].}$$

Lastly, applying the virial theorem, we can obtain the total mechanical energy of the star

$$E \sim -\frac{3}{10} \frac{GM^2}{R} \text{ [J].}$$

Nuclear reaction rates: Writing the nuclear reaction rates in the form of a power law centered at a particular temperature, the reaction rate becomes

$$r_i \simeq r_0 X_i X_x \rho^{\alpha'} T^{\beta},$$

where r_0 is a constant, X_i and X_x are the mass fractions of the two particles, and α and β are determined from the power law expansion of the reaction rate equations. Usually, $\alpha = 2$ for a two-body collision, and β can range from near unit to 40 or more.

By combining the reaction rate equation with the amount of energy released per reaction, we can calculate the amount of energy released per second in each kilogram of stellar material. If ϵ_0 is the amount of energy released per reaction, the amount of energy liberated per kilogram of material per second becomes

$$\epsilon_i = \left(\frac{\epsilon_0}{\rho} \right) r_i,$$

or, in the form of a power law,

$$\epsilon_i = \epsilon_0 X_i X_x \rho^{\alpha} T^{\beta},$$

where $\alpha = \alpha' - 1$, ϵ has units of W kg^{-1} , and the sum of ϵ_i for all reactions is the total nuclear energy generation rate.

The luminosity gradient equation: To determine the luminosity of a star, we must consider all of the energy generated by stellar material. The interior luminosity due to all of the energy generated within the star's interior out to some radius r is given by the **luminosity gradient equation**

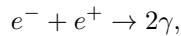
$$\frac{dL_r}{dr} = 4\pi r^2 \rho \epsilon [\text{erg s}^{-1} \text{ m}^{-1}],$$

where ϵ is the *total* energy released per kilogram per second by *all* nuclear reactions and by gravity, or $\epsilon = \epsilon_{\text{nuclear}} + \epsilon_{\text{gravity}}$. It is worth noting that $\epsilon_{\text{gravity}}$ could be negative if the star is expanding. This equation for the luminosity gradient is another one of the fundamental stellar structure equations.

Stellar nucleosynthesis and conservation laws: The process by which one element is converted into another is a process known as **nucleosynthesis**. It is highly unlikely that four hydrogen nuclei are directly converted into helium (i.e., all nuclei hitting simultaneously). For hydrogen fusion to occur, the final product must be created by a chain of reactions, each involving more probable two-body interactions. The process by which a chain of nuclear reactions leads to the final product cannot happen in a completely arbitrary way, however; a series of **particle conservation laws** must be obeyed. In particular, during every reaction it is necessary to conserve:

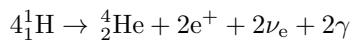
- electric charge,
- the number of nucleons, and
- the number of leptons.

The term **lepton** means a “light thing” and includes electrons, positrons, neutrinos, and anti-neutrinos. Although anti-matter is extremely rare in comparison with matter, it plays an important role in subatomic physics, including nuclear reactions. Antimatter particles are identical to their matter counterparts but have opposite attributes, such as electric charge. Anti-matter also has the characteristic that a collision with its matter counterpart results in complete annihilation of both particles, accompanied by the production of energetic photons. For instance,



where e^- , e^+ , and γ denote an electron, positron, and photon, respectively. Note that two photons are required to conserve both momentum and energy simultaneously. Since electrons and positrons have charges equal in magnitude to that of a proton, these leptons will contribute to the charge conservation requirement while their total **lepton number** must also be conserved. Note that in counting the number of leptons involved in a nuclear reaction, we treat matter and anti-matter differently. Specifically, the total number of matter leptons *minus* the total number of anti-matter leptons must remain constant.

The proton-proton chains: Applying the conservation laws, one chain of reactions that can convert hydrogen into helium is the first **proton-proton chain** (PPI). It involved a reaction sequence that ultimately results in



through the intermediate production of **deuterium** ${}^2_1\text{H}$ and helium-3 ${}^3_2\text{He}$. The entire PPI reaction chain is

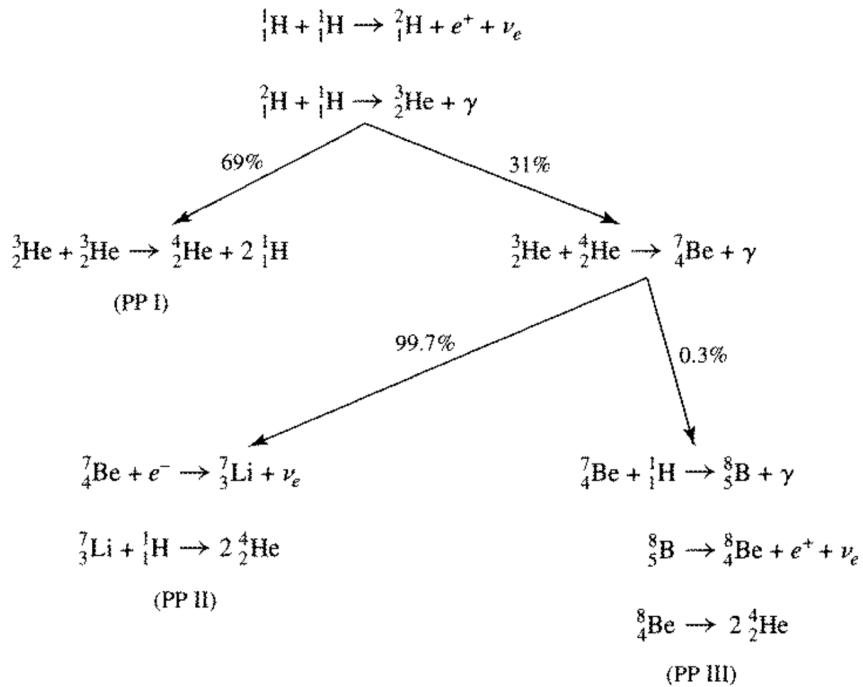
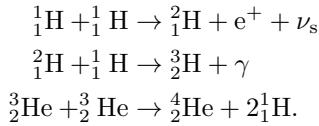
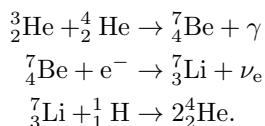


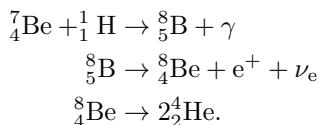
Figure 16: The three branches of the pp chain along with the branching ratios appropriate for conditions in the core of the Sun. Figures taken from Carroll & Ostlie (2007).



Each step of the PPI chain has its own reaction rate, since different Coulomb barriers and cross sections are involved. The slowest step in the sequence is the initial one, because it involves the *decay* of a proton into a neutron via $p^+ \rightarrow n + e^+ + \nu_e$. Such a decay involves the **weak force**. The production of helium-3 nuclei in the PPI chain also provides for the possibility of their interaction directly with helium-4 nuclei, resulting in a second branch of the proton-proton chain. In an environment characteristic of the center of the Sun, 69% of the time a helium-3 interacts with another helium-3 in the PPI chain, whereas 31% of the time the **PPII** chain occurs:



Yet another branch, the **PPIII** chain, is possible because the capture of an electron by the beryllium-7 nucleus in the PPII chain competes with the capture of a proton (a proton is captured only 0.3% of the time in the center of the Sun):



The three branches of the proton-proton chain, along with their branching ratios, are summarized in Figure 16.

The nuclear energy generation rate for the combined pp chain is

$$\epsilon_{pp} = 0.241\rho X^2 f_{pp} \psi_{pp} C_{pp} T_6^{-2/3} e^{-33.80T_6^{-1/3}} [\text{W kg}^{-1}],$$

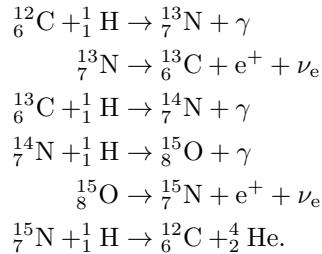
where T_6 is a dimensionless expression of temperature in units of 10^6 K (i.e., $T_6 \equiv T/10^6 \text{ K}$), $f_{pp} = f_{pp}(X, Y, \rho, T) \simeq 1$ is the **pp chain screening factor**, $\psi_{pp} = \psi_{pp}(X, Y, T) \simeq 1$ is a correction factor that accounts for the simultaneous occurrence of PPI, PPII, and PPIII, and $C_{pp} \simeq 1$ involves higher-order correction terms.

When written as a power law near $T = 1.5 \times 10^7$ K, the energy generation rate has the form

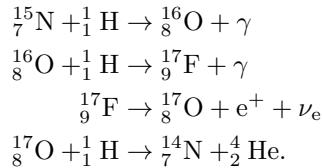
$$\epsilon_{pp} \simeq \epsilon_{0,pp} \rho X^2 f_{pp} \psi_{pp} C_{pp} T_6^4 [\text{W kg}^{-1}],$$

where $\epsilon_{0,pp} = 1.08 \times 10^{-12}$ W m³ kg⁻². The power law form of the energy generation rate demonstrates a modest temperature dependence of T^4 near $T_6 = 15$.

The CNO cycle: A second, independent cycle also exists for the production of helium-4 from hydrogen. In the **CNO cycle**, carbon, nitrogen, and oxygen are used as catalysts, being consumed and then regenerated during the process. Just as with the pp chain, the CNO cycle has competing branches. The first branch **CNO-1** culminates with the production of carbon-12 and helium-4:



The second branch occurs only about 0.04% of the time and arises when the last reaction produces oxygen-16 and a photon, rather than carbon-12 and helium-4:



The energy generation rate for the CNO cycle is given by

$$\epsilon_{\text{CNO}} = 8.67 \times 10^{20} \rho X X_{\text{CNO}} C_{\text{CNO}} T_6^{-2/3} e^{-152.28 T_6^{-1/3}} [\text{W kg}^{-1}],$$

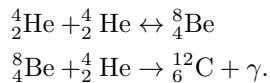
where C_{CNO} is the total mass fraction of carbon, nitrogen, and oxygen, and C_{CNO} is a higher-order correction term. When written as a power law centered about $T = 1.5 \times 10^7$ K, this energy equation becomes

$$\epsilon_{\text{CNO}} \simeq \epsilon_{0,\text{CNO}} \rho X X_{\text{CNO}} C_{\text{CNO}} T_6^{19.9} [\text{W kg}^{-1}],$$

where $\epsilon_{0,\text{CNO}} = 8.24 \times 10^{-31}$ W m³ kg⁻². As shown by the power law dependence, the CNO cycle is much more strongly temperature-dependent than the pp chain is. This property implies that low-mass stars, which have smaller central temperatures, are dominated by the pp chain during their “hydrogen burning” evolution, whereas more massive stars, with their higher central temperatures, convert hydrogen to helium by the CNO cycle. The transition in stellar mass between stars dominated by the pp chain and those dominated by the CNO cycle occurs for stars slightly more massive than our Sun. This difference in nuclear reaction processes plays an important role in the structure of stellar interiors.

When hydrogen is converted into helium by either the pp chain or the CNO cycle, the mean molecular weight μ of the gas increases. If neither the temperature nor the density of the gas changes, the ideal gas law predicts that the central pressure will necessarily decrease. As a result, the star will no longer be in hydrostatic equilibrium and would begin to collapse. This collapse has the effect of actually raising both the temperature and the density to compensate for the increase in μ . When the temperature and density become sufficiently high, helium nuclei can overcome their **Coulomb repulsion** and begin to “burn”.

Triple alpha process of helium burning: The reaction sequence by which helium is converted into carbon is known as the **triple alpha process**. The process takes its name from the historical result that the mysterious alpha particles detected in some types of radioactive decay were shown by Rutherford to be helium-4 nuclei. The triple alpha process is as follows:



In the triple alpha process, the first step produces an unstable beryllium nucleus that will rapidly decay back into two separate helium nuclei if not immediately struck by another alpha particle. As a result, this reaction may be thought of as a three-body interaction, and therefore, the reaction rate depends on $(\rho Y)^3$. The nuclear energy generation rate is given by

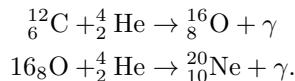
$$\epsilon_{3\alpha} = 50.9 \rho^2 Y^3 T_8^{-3} f_{3\alpha} e^{-44.027 T_8^{-1}} [\text{W kg}^{-1}],$$

where $T_8 \equiv T/10^8 \text{ K}$ and $f_{3\alpha}$ is the screening factor for the triple alpha process. Written as a power law centered on $T = 10^8 \text{ K}$, it demonstrates a very dramatic temperature dependence:

$$\epsilon_{3\alpha} = \epsilon_{0,3\alpha} \rho^2 Y^3 f_{3\alpha} T_8^{41.0} [\text{W kg}^{-1}].$$

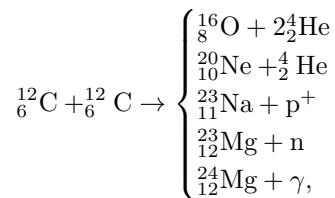
With such a strong dependence, even a small increase in temperature will produce a large increase in the amount of energy generated per second. For instance, an increase of only 10% in temperature raises the energy output rate by more than 50 times!

Carbon and oxygen burning: In the high-temperature environment of helium burning, other competing processes are also at work. After sufficient carbon has been generated by the triple alpha process, it becomes possible for carbon nuclei to capture alpha particles, producing oxygen. Some of the oxygen in turn can capture alpha particles to produce neon.

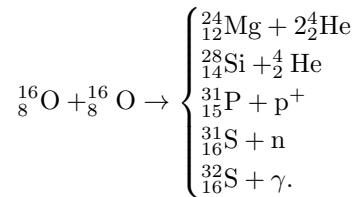


At helium-burning temperatures, the continued capture of alpha particles leading to progressively more massive nuclei quickly becomes prohibitive due to the ever higher Coulomb barrier.

If a star is sufficiently massive, still higher temperatures can be obtained and many other nuclear products become possible. Examples of available reactions include carbon burning reactions near $6 \times 10^8 \text{ K}$,



and oxygen burning reactions near 10^9 K ,



1.8.3 Follow-up Questions

- Can main sequence stars have unstable nuclear fusion?
- If I suddenly increase the core temperature of a star by 20%, how quickly does the star respond?
- What happens when you ignite fusion in a white dwarf? Explain using explicit thermodynamics what happens when you turn up the temperature or reaction rate in the core of the star (i.e. explain adiabatic expansion and contraction).

1.9 Question 9

Why do neutrons inside a neutron star not decay into protons and electrons?

1.9.1 Short answer

Neutrons inside a neutron star do not decay because all lower energy levels are already occupied since they are supported by neutron degeneracy pressure. If the neutrons underwent β -decay ($n \rightarrow p^+ + e^- + \nu_e$), the resulting electron would have an energy exceeding the Fermi energy which is energetically unfavourable.

1.9.2 Additional context

The Pauli exclusion principle and electron degeneracy: Any system (whether an atom of hydrogen, an oven filled with blackbody photons, or a box filled with gas particles) consists of quantum states that are identified by a set of quantum numbers. Just as the oven is filled with standing waves of EM radiation that are described by three quantum numbers (specifying the number of photons of wavelength λ traveling in the x -, y -, and z - directions), a box of gas particles is filled with standing **de Broglie** waves that are also described by three quantum numbers (specifying the particle's component of momentum in each of three directions). If the gas particles are fermions (such as electrons or neutrons), then the **Pauli exclusion principle** allows at most one fermion in each quantum state because no two fermions can have the same set of quantum numbers.

In an everyday gas at standard temperature and pressure, only one of every 10^7 quantum states is occupied by a gas particle, and the limitations imposed by the Pauli exclusion principle become insignificant. Ordinary gas has a *thermal* pressure that is related to its temperature by the ideal gas law. However, as energy is removed from the gas and its temperature falls, an increasingly large fraction of the particles are forced into the lower energy states. If the gas particles are fermions, only one particle is allowed into each state; thus all the particles cannot crowd into the ground state. Instead, as the temperature of the gas is lowered, the fermions will fill up the lowest available unoccupied states, starting with the ground state, and then successively occupy the excited states with the lowest energy. Even in the limit $T \rightarrow 0\text{ K}$, the vigorous motion of all the fermions in excited states produces a pressure in the fermion gas. At zero temperature, *all* of the lower energy states and *none* of the higher energy states are occupied. Such a fermion gas is said to be completely **degenerate**.

The Fermi energy: The maximum energy E_F of any electron in a completely degenerate gas at $T = 0\text{ K}$ is known as the **Fermi energy** given by

$$E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3} [\text{J}],$$

where m is the mass of the fermion and $n \equiv N/L^3$ is the number of fermions per unit volume. For electrons, the average energy per fermion at zero temperature of $(3/5)E_F$.

The condition for degeneracy: At any temperature above absolute zero, some of the states with an energy less than E_F will become vacant as fermions use their thermal energy to occupy other, more energetic states. Although the degeneracy will not be precisely complete when $T > 0\text{ K}$, the assumption of complete degeneracy is a good approximation at the densities encountered in the interior of e.g., white dwarfs. All but the most energetic particles will have an energy less than the Fermi energy.

The condition for degeneracy is the following:

$$\frac{T}{\rho^{2/3}} < \mathcal{D} [\text{K m}^2 \text{kg}^{-2/3}], \quad \mathcal{D} \equiv 1261 \text{ K m}^2 \text{kg}^{-2/3}.$$

The smaller the value of $T/\rho^{2/3}$, the more degenerate the gas. Figure 17 shows the degree of degeneracy in the Sun's center as it evolves.

If the initial mass of the star on the MS was not too large (perhaps $M_{\text{ZAMS}} < 25 M_\odot$, the remnant in the inner core will stabilize and become a **neutron star** supported by **degeneracy pressure**².

Neutron degeneracy: Because neutron stars are formed when the degenerate core of an aging supergiant star nears the **Chandrasekhar limit** and collapses, we take the **Chandrasekhar mass** to be that of a typical neutron star. A $1.4 M_\odot$ neutron star would consist of $1.4 M_\odot/m_n \approx 10^{57}$ neutrons – in effect, a huge nucleus with a mass number of $A \approx 10^{57}$ that is held together by gravity and supported by **neutron degeneracy pressure**.

The radius of a neutron star is given by

²The upper mass limit of the progenitor that results in the formation of a neutron star depends on how metal-rich the original star was. A sufficiently meta-rich star may form a neutron star even if its initial mass is much greater than $25 M_\odot$.

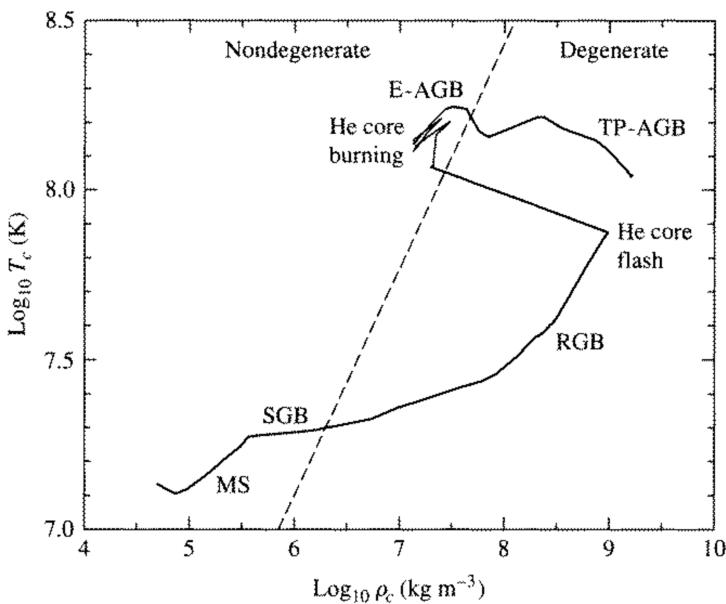


Figure 17:
Degeneracy in the Sun's center as it evolves. (Data from Mazzitelli and D'Antona, Ap. J., 311, 762, 1896.) Figure taken from Carroll & Ostlie (2007).

$$R_{\text{ns}} \approx \frac{(18\pi)^{2/3}}{10} \frac{\hbar^2}{GM_{\text{ns}}^{1/3}} \left(\frac{1}{m_{\text{H}}} \right)^{8/3} [\text{km}].$$

For $M_{\text{ns}} = 1.4 M_{\odot}$, this yields a value of $R_{\text{ns}} = 4.4 \text{ km}$.

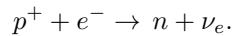
This estimate is too small by a factor about about 3 – that is, the actual radius of a $1.4 M_{\odot}$ neutron star lies roughly between 10 and 15 km.

Density of a neutron star: This incredibly compact stellar remnant would have an average density of $6.65 \times 10^{17} \text{ km m}^{-3}$, greater than the typical density of an atomic nucleus, $\rho_{\text{nuc}} \approx 2.3 \times 10^{17} \text{ kg m}^{-3}$. In some sense, the neutrons in a neutron star must be “touching” one another. At the density of a neutron star, all of Earth’s human inhabitants would be crowded into a cube 1.5 cm on each side.

The pull of gravity at the surface of a neutron star is fierce. For a $1.4 M_{\odot}$ neutron star with a radius of 10 km, $g = 1.86 \times 10^{12} \text{ m s}^{-2}$, 190 billion times stronger than the acceleration of gravity on Earth’s surface. An object dropped from a height of one meter would arrive at the star’s surface with a speed of $1.93 \times 10^6 \text{ m s}^{-1}$.

The equation of state: To appreciate the exotic nature of the material constituting a neutron star and the difficulties involved in calculating the equation of state, imagine compressing the mixture of iron nuclei and degenerate electrons that make up an iron white dwarf at the center of a massive supergiant star. Because the mechanical and thermal properties of degenerate matter are independent of one another, we will assume for convenience that $T = 0 \text{ K}$. The iron nuclei are then arranged in a crystalline lattice. We are interested in the equilibrium configuration of 10^{57} nucleons (protons and neutrons), together with enough free electrons to provide zero net charge. The equilibrium arrangement is the one that involves the least energy.

Initially, at low densities the nucleons are found in iron nuclei. This is the outcome of the minimum-energy compromise between the repulsive Coulomb force between the protons and the attractive nuclear force between all of the nucleons. However, when $\rho \approx 10^9 \text{ kg m}^{-3}$ the electrons become relativistic. Soon thereafter, the minimum-energy arrangement of protons and neutrons changes because the energetic electrons can convert protons in the iron nuclei into neutrons by the process of **electron capture**,



Because the neutron mass is slightly greater than the sum of the proton and electron masses, and the neutrino’s rest-mass energy is negligible, the electron must supply the kinetic energy to make up the difference in energy: $m_n c^2 - m_p c^2 - m_e c^2 = 0.78 \text{ MeV}$.

A careful calculation that takes into account the surrounding nuclei and relativistic degenerate electrons, as well as the complexities of nuclear physics, reveals that the density must exceed $10^{12} \text{ kg m}^{-3}$ for the protons in iron nuclei to capture electrons. At still higher densities, the most stable arrangement of nucleons is one where the neutrons and protons are found in a lattice of increasingly neutron-rich nuclei so as to decrease the energy due to the Coulomb repulsion between protons. This process is known as **neutronization** and produces a sequence of nuclei such as $^{56}\text{Fe}, ^{62}\text{Ni}, ^{64}\text{Ni}, ^{66}\text{Ni}, ^{86}\text{Kr}, \dots, ^{118}\text{Kr}$. Ordinarily, these supernumerary neutrons would revert to protons via the standard β -decay process,

Transition density (kg m ⁻³)	Composition	Degeneracy pressure
$\approx 1 \times 10^9$	iron nuclei, nonrelativistic free electrons	electron
	electrons become relativistic	
$\approx 1 \times 10^{12}$	iron nuclei, relativistic free electrons	electron
	neutronization	
$\approx 4 \times 10^{14}$	neutron-rich nuclei, relativistic free electrons	electron
	neutron drip	
$\approx 4 \times 10^{15}$	neutron-rich nuclei, free neutrons, relativistic free electrons	electron
	neutron degeneracy pressure dominates	
$\approx 2 \times 10^{17}$	neutron-rich nuclei, superfluid free neutrons, relativistic free electrons	neutron
	nuclei dissolve	
$\approx 4 \times 10^{17}$	superfluid free neutrons, superconducting free protons, relativistic free electrons	neutron
	pion production	
	superfluid free neutrons, superconducting free protons, relativistic free electrons, other elementary particles (pions, ...?)	neutron

Table 1: Composition of neutron star material. Table taken from Carroll & Ostlie (2007).

$$n \rightarrow p^+ + e^- + \bar{\nu}_e.$$

However, under the condition of complete electron degeneracy, there are no vacant states for an emitted electron to occupy, so the neutrons cannot decay back into protons. In *isolation*, a neutron decays into a proton in about 10.2 minutes.

When the density reaches about $4 \times 10^{14} \text{ kg m}^{-3}$, the minimum-energy arrangement is one in which some of the neutrons are found *outside* the nuclei. The appearance of these free neutrons is called **neutron drip** and marks the start of a three-component mixture of a lattice of neutron-rich nuclei, non-relativistic degenerate free neutrons, and relativistic degenerate electrons. The fluid of free neutrons has the striking property that it has no viscosity. This occurs because a spontaneous pairing of the degenerate neutrons has taken place. The resulting combination of two fermions (the neutrons) is a boson and so it is not subject to the restrictions of the Pauli exclusion principle. Because degenerate bosons can *all* crowd into the lowest energy state, the fluid of paired neutrons can lose no energy. It is a **superfluid** that flows without resistance. Any whirlpools or vortices in the fluid will continue to spin forever without stopping. As the density increases further, the number of free neutrons increases as the number of electrons declines. The neutron degeneracy pressure exceeds the electron degeneracy pressure when the density reaches roughly $4 \times 10^{15} \text{ kg m}^{-3}$. As the density approaches ρ_{nuc} , the nuclei effectively dissolve as the distinction between neutrons inside and outside of nuclei becomes meaningless. This results in a fluid mixture of free neutrons, protons, and electrons dominated by neutron degeneracy pressure, with both the neutrons and protons paired to form superfluids. The fluid of pairs of positively charged protons is also **superconducting**, with zero electrical resistance. As the density increases further, the ratio of neutrons:protons:electrons approaches a limiting value of 8 : 1 : 1, as determined by the balance between the competing process of electron capture and β -decay inhibited by the presence of degenerate electrons. The properties of the neutron star material when $\rho > \rho_{\text{nuc}}$ are still poorly understood. A complete theoretical description of the behaviour of a sea of free neutrons interacting via the strong nuclear force in the presence of protons and electrons is not yet available, and there is little experimental data on the behaviour of matter in this density range. A further complication is the appearance of sub-nuclear particles such as **pions** (π) produced by the decay of a neutron into a proton and a negatively charged pion, $n \rightarrow p^+ + \pi^-$, which occurs spontaneously in neutron stars with $\rho > 2\rho_{\text{nuc}}$. Nevertheless, these are the values of the density encountered in the interiors of neutron stars, and the difficulties mentioned are the primary reasons for uncertainties in the structure calculated for model neutron stars.

Table 1 summarizes the composition of the neutron star material at various densities.

1.9.3 Follow-up Questions

- What sort of products come from NS-NS mergers?
- What are typical photon energies emitted from each of these scenarios?

1.10 Question 10

What is the typical temperature of matter accreting on a star, a white dwarf, a neutron star, a stellar mass black hole, and a supermassive black hole? In what wavelength range would one best find examples of such sources?

1.10.1 Short answer

Answer.

1.10.2 Additional context

Additional context.

1.10.3 Follow-up Questions

- How does the accretion process work?
- If you don't assume a blackbody, and just dump a bunch of material onto a star and it can't radiate away, what is the maximum temperature it can reach?

1.11 Question 11

You don't usually need to cool down the detectors for short wavelength (e.g., X-ray) observations, but it's critical to cool down the detectors in long wavelength (e.g., far-IR) observations. Why is this, and why is it not necessary for radio observations?

1.11.1 Short answer

IR detectors use CCDs which are made up of semi-conducting silicon material; these detectors, when hit with a photon of the 'right' energy, will release a photoelectron which is later read by the CCD. The required energy for an incoming photon to release a photoelectron must be greater than silicon's band gap energy ($> 1.14 \text{ eV}$). Silicon has a useful photoelectric effect range of $1.1 - 10 \text{ eV}$, covering the range of near-IR to soft x-ray. While CCDs are generally more sensitive at higher temperatures, this results in a dark current of photoelectrons which cannot be distinguished from astrophysical sources. Thus, cooling detectors for short wavelength observations is necessary to reduce these dark currents specifically because they are sensitive to the IR whereas short wavelength observations are not. This is not necessary for radio observations because these observations use bolometers (instead of CCDs) which detect electromagnetic fields (instead of photons).

1.11.2 Additional context

Why use CCDs?: Most astronomical detectors in use today at professional observatories, as well as with many amateur telescopes, are CCDs. This fact alone gives an impression that there must be something very special or useful about CCDs; otherwise why all the fuss? CCDs have revolutionized modern astronomy. They will take their place in astronomical history along with other important discoveries such as the telescope, photographic plates, prisms, and spectroscopy. The contribution to our knowledge of the heavens brought about by CCDs is astounding, even more so when one considers that they have been in use for only about thirty years.

First introduced as electronic analogs to magnetic bubble memory at Bell labs, CCDs provided their first astronomical image in 1975 when scientists from the Jet Propulsion Laboratory imaged the planet Uranus at a wavelength of 8900\AA . This observation used the University of Arizona 61" telescope atop Mt. Lemmon and a prototype version of a CCD made by Texas Instruments Corporation as part of a development project for NASA spacecraft missions.

During the past ten years, tremendous progress has been made in the manufacturing process and therefore in the properties of the CCD itself. These improvements have allowed much lower noise properties for CCDs, thereby increasing their overall efficiency in astronomy. In addition, larger format devices have been produced and the readout times are much shorter, approaching 1 – 2 seconds even for arrays as large as 1024 pixels square. This latter advance is mainly due to the availability of high-speed, low-power and low-noise CCD controllers. The driving technology for CCD manufacturing is for items such as copy machines, TV cameras, and digital cameras, but the requirements for low noise, excellent pixel cosmetics, and nearly perfect performance is still firmly rooted in astronomy. We outline below two of the important reasons why CCDs are considered as essentially the perfect imaging device.

Noise properties: The usefulness of a detector is very often determined by the amount of inherent noise within the device itself. Suffice it to say that modern astronomical CCDs are almost noise free. The original line of photosensitive electronic array detectors, such as television-type imaging detectors, vidicons, silicon intensified targets, and image-dissector scanners, all had very high noise properties. For comparison, silicon intensified target imagers (SITs) had a noise level upon readout of 800 electrons per picture element. Some very good systems of this type could be produced with read noise values of only 200 electrons. The first CCDs had readout noise levels similar to this latter value, while modern CCDs have noise values of ten down to two electrons per pixel per readout. The large noise levels present in early array detectors not only limited the SNR obtainable for a given measurement, they also severely limited the total dynamic range available to the camera. Another "feature" of older, higher noise CCDs was the decision an astronomer had to make about co-addition of frames. Since the read noise adds as its square to the total noise budget, adding two frames resulted in a much higher read noise contribution. Today, with typical read noise values of 2 – 5 electrons, co-addition is essentially equal to a single exposure of longer integration time.

CCD operation: The simplest and very understandable analogy for the operation of a CCD is also one that has been used numerous times for this purpose. This is the "water bucket" idea in which buckets represent pixels on the CCD array, and a rainstorm provides the incoming photons (rain drops). Imagine a field covered with buckets aligned neatly in rows and columns throughout the entirety of the area (Figure 18). After the rainstorm (CCD integration), each bucket is transferred in turn and metered to

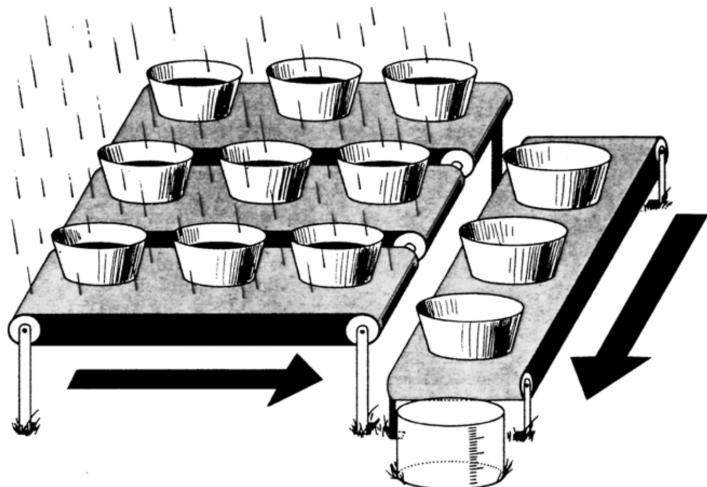


Figure 18:
CCDs can be likened to an array of buckets that are placed in a field and collect water during a rainstorm. After the storm, each bucket is moved along conveyor belts until it reaches a metering station. The water collected in each field bucket is then emptied into the metering bucket within which it can be measured. From Janesick & Blouke (1987). Figure taken from Howell (2006).

determine the amount of water collected. A written record (final CCD image) of the amount of water in each bucket will thus provide a two-dimensional record of the rainfall within the field.

Referring to the actual mechanisms at work within a CCD, we start with the method of charge generation within a pixel: the photoelectric effect.³ Incoming photons strike the silicon within a pixel and are easily absorbed if they possess the correct wavelength (energy). Silicon has a band gap energy of 1.14 eV, and so it easily absorbs light of energy 1.1 to 4 eV (3000 to 11000 Å).⁴ Photon absorption causes the silicon to give up a valence electron and move it into the conduction band. Photons of energy 1.1 eV to near 4 or so eV generate single **electron-hole pairs**, whereas those of higher energy produce multiple pairs. Left to themselves, these conduction band electrons would recombine back into the valence level within approximately 100 μ s. Silicon has a useful photoelectric effect range of 1.1 to about 10 eV, which covers the near-IR to soft X-ray region. Above and below these limits, the CCD material appears transparent to the incoming photons.

Once electrons have been freed to the conduction band of the silicon, they must be collected and held in place until readout occurs. The details of the actual construction of each pixel within a CCD (that is, the formation of the MIS capacitor with its doped silicon, layers of silicon dioxide, etc.) are beyond the scope of this discussion, but suffice it to say that each pixel has a structure allowing applied voltages to be placed on sub-pixel sized electrodes called **gates**. These gate structures provide each pixel with the ability to collect the freed electrons and hold them in a potential well until the end of the exposure. In a typical arrangement, each pixel has associated with it three gates, each of which can be set to a different voltage potential. The voltages are controlled by clock circuits with every third gate connected to the same clock. Figure 19 illustrates this clocking scheme for a typical three-phase device.

We note in Figure 19 that, when an exposure ends, the clock voltages are manipulated such that the electrons that have been collected and held in each pixel's +10 volt potential well by clock voltage V3 can now be shifted within the device. Note that electrons created anywhere within the pixel during the exposure (where each pixel has a surface area equal to the total area under all three gates) will be forced to migrate toward the deepest potential well. When the exposure is terminated and CCD readout begins, the voltages applied to each gate are cycled (this process is called **clocking the device**) such that the charge stored within each pixel during the integration is electronically shifted. A simple change in the voltage potentials (V3 goes to +5 volts, while V1 becomes +10 volts and so on) allows the charge to be shifted in a serial fashion along columns from one CCD pixel to another throughout the array. The transfer of the total charge from location to location within the array is not without losses. Each charge transfer (one of which occurs for each voltage change or clock cycle) has an associated efficiency. This efficiency value is the percent of charge transferred compared with that which was actually collected. Modern values for the charge transfer efficiency (CTE) are approaching 0.999999 (i.e., 99.9999% efficient) for each transfer.

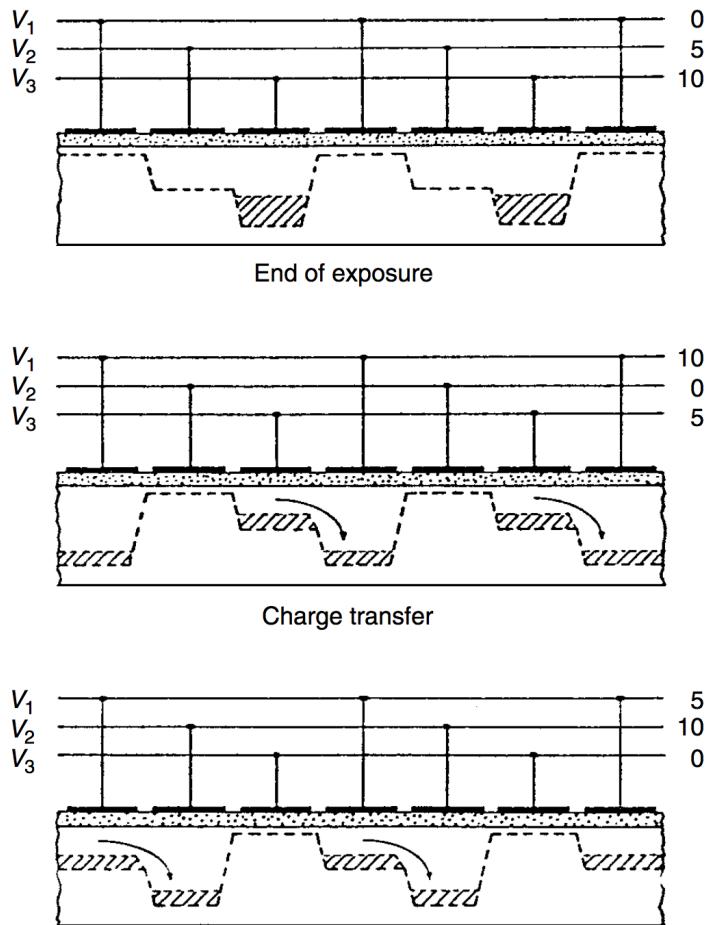
Each column in the array is connected in parallel and thus each pixel shift is mimicked throughout the entire array simultaneously. One clock cycle moves each row of pixels up one column, with the top row being shifted off the array into what is called the **output shift register** or **horizontal shift register**. This register is simply another row of pixels hidden from view (i.e., not exposed to incident light) and serves as the transition between active rows on the array and the output of the device. Once an entire

³Albert Einstein received his Nobel Prize mainly for his work on the photoelectric effect, not, as many think, for relativity.

⁴The energy of a photon of a given wavelength (in electron volts) is given by $E_{\text{eV}} = 12407/\lambda(\text{\AA})$.

Figure 19:

Schematic voltage operation of a typical three-phase CCD. The clock voltages are shown at three times during the readout process, indicating their clock cycle of 0, 10, and 5 volts. One clock cycle causes the stored charge within a pixel to be transferred to its neighboring pixel. CCD readout continues until all the pixels have had their charge transferred completely out of the array and through the A/D converter. From Walker (1987). Figure taken from Howell (2006).



row is shifted into the output register, and before any further row shifts on the active area occur, each pixel in the output register is shifted out one at a time (in a similar manner as before) into the output electronics. Here, the charge collected within each pixel is measured as a voltage and converted into an output digital number. Each pixel's collected charge is sensed and amplified by an **output amplifier**. CCD output amplifiers are designed to have low noise and are built directly into the silicon circuitry; thus they are often referred to as **on-chip amplifiers**. These amplifiers must work with extremely small voltages and are rated, as to their sensitivity, in volts per electron. Typical values are in the range of 0.5 to 4 microvolts per electron.

The output voltage from a given pixel is converted to a **digital number** (DN) and is typically discussed from then on as either **counts** or **analog-to-digital units** (ADUs). The amount of voltage needed (i.e., the number of collected electrons or received photons) to produce 1 ADU is termed the **gain** of the device. A typical CCD gain might be 10 electrons/ADU, which means that for every 10 electrons collected within a pixel, the output from that pixel will produce, on average, a count or DN value of 1. For example, with this gain value if a pixel collects 1000 electrons (photons), the output pixel value stored in the computer would be 100 ADUs. For 1500 electrons 150 ADUs would be produced and for 17234 electrons, the output pixel value would be 1723 ADUs (note, not 1723.4). Digital output values can only be integer numbers and it is clear already that the discrimination between different pixel values can only be as good as the resolution of the gain and digital conversion of the device.

Conversion of the output voltage signal into a DN is performed within a device called an **analog-to-digital converter** (A/D or ADC). There is an intimate connection between the number of digital bits available in the A/D and the value of the gain that can or should be used for the CCD. The output DNs are usually stored initially in computer memory and then moved to disk for storage and later manipulation. The process of shifting each entire CCD row into the output register, shifting each pixel along within this register, and finally performing the voltage conversion of each pixel's stored charge by the A/D to produce a DN value is continued until the entire array of pixels has been read out. For large-format CCD arrays, this process can take upwards of a few minutes to complete a single read out of the entire device. Note that for a 2048×2048 CCD, the charge collected in the last pixel to be read out has to be transferred over four thousand times. However, most modern large-format CCDs or mosaic cameras containing many large CCDs use a few tricks to readout faster. Single monolithic CCDs usually have 2

or 4 output amplifiers available (one in each corner) and given the proper electronic setup, these large chips are often read out from 2 or 4 corners simultaneously, thus decreasing the total readout time by 2 – 4. For a mosaic of CCDs, this same process can read the entire array (using multiple amplifiers on each CCD) much faster than even one single large CCD.

The array size of a single CCD, as well as the size of a given pixel on a device, is controlled by the current limitations of manufacturing. How large one can make a good quality, large-scale integrated circuit and how small one can make an MIS capacitor, both of which have demanding requirements for near perfect operation, set the scale of CCD and pixel sizes that are available. CCDs as large as 5040×10080 and 7168×9216 pixels and pixels as small as $2 - 10 \mu\text{m}$ have been successfully produced.

Modern CCDs have much higher processing standards than even five years ago. Items such as multi-layer registration on the silicon wafer on the photomasks used in the production of the CCD integrated circuit and the ability to make smaller electrical component parts on the wafers (such as output amplifiers) lead to much lower noise characteristics, better pixel charge manipulation, and the ability for faster readout speeds with lower noise. For example, better alignment of the CCD layers in each pixel allow lower clock voltages to be used (as low as 2 volts has been demonstrated) leading to lower overall power consumption. This fact, in turn, allows for items such as plastic packaging instead of ceramic, reducing overall packaging costs, a cost that often rivals that of the CCD itself.

As you might imagine, astronomy is not the driving force for CCD manufacturing. Video devices, cell phones, security cameras, Xerox machines, etc. are the global markets boosting the economy of CCD makers. The trend today is to produce CCDs with small pixels ($10 - 12 \mu\text{m}$ for astronomy down to $\sim 2 \mu\text{m}$ for other applications) in order to increase image resolution. Small pixels (and small CCDs) have lower cost and higher yield but the small pixels have shallow well depths. This is somewhat compensated for using fast readout techniques and/or microlens arrays, which focus light from an incoming source onto each small CCD pixel. Not all CCD pixels are desired to have shallow wells. The CCDs produced by E2V for the NASA Kepler Discovery mission have $27 \mu\text{m}$ pixels with well depths of nearly 1 million electrons each and a capacity of > 50000 electrons per pixel is quite common in astronomy. Even CCDs with built-in electronic shutters are being experimented with. Each pixel contains a $p^+ - n - p^-$ vertical overflow drain (VOD) photodiode structure on its top through which the incoming light passes. The absorption of incoming light when the “shutter” is open is minimal and, within a few hundred nanoseconds, the electronic shutter can be biased and become opaque.

Quantum efficiency: The composition of a CCD is essentially pure silicon. This element is thus ultimately responsible for the response of the detector to various wavelengths of light. The wavelength dependence of silicon can be understood in an instant by glancing at Figure 20. Shown here is the length of silicon needed for a photon of a specific wavelength to be absorbed. Absorption length is defined as the distance for which 63% ($1/e$) of the incoming photons will be absorbed. Figure 20 clearly shows that, for light outside the range of about 3500 to over 8000Å, the photons (1) pass right through the silicon, (2) get absorbed within the thin surface layers or gate structures, or (3) simply reflect off the CCD surface. At short wavelengths, 70% or more of the photons are reflected, and for very short wavelengths (as for long wavelengths) the CCD becomes completely transparent. Thus the quantum efficiency of a typical CCD device will approximately mirror the photon absorption curve for silicon. Shortward of $\sim 2500\text{\AA}$ (for thinned devices) or about $\sim 25\text{\AA}$ (for thick devices) the detection probability for photons increases again. However, owing to their much higher energy, these photons lead to the production of multiple electron-hole pairs within the silicon and may also produce damage to the CCD itself.

CCD quantum efficiencies are therefore very dependent on the thickness of the silicon that intercepts the incoming photons. This relation between absorption probability and CCD thickness is why front-side illuminated (thick) devices are more red sensitive (the photons have a higher chance of absorption) and why they have lower overall (blue) QE (since the gate structures can be close to or even exceed the necessary absorption depths of as small as only a few atomic layers). A few front-side CCDs have been produced with special gate structures that are transparent to incoming blue and UV photons. In thinned devices, the longest wavelength photons are likely to pass right through the CCD without being absorbed at all.

Figure 21 (left) shows the quantum efficiencies for various imaging devices. Note that the y scale is logarithmic and the much superior QE provided by CCDs over previous detectors. Figure 21 (right) shows a selection of modern CCD QE. The large difference in QE that used to exist between thinned and thick CCDs is now mostly eliminated due to manufacturing processes and coatings although other differences (such as location of peak QE, cosmic ray detection, etc.) remain. Quantum efficiency or QE curves allow one quickly to evaluate the relative collecting power of the device as a function of wavelength. Measured QE curves, such as in Figure 21 (right) and those shown in the literature, are generally assumed to be representative of each and every pixel on the device, that is, all pixels of a given device are assumed to work identically and have the same wavelength response. This is almost true, but it is the “almost” that makes flat fielding of a CCD necessary. In addition, the QE curves shown or delivered with a

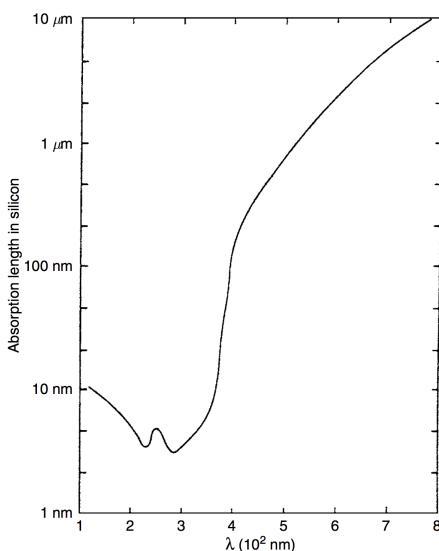


Figure 20:
The photon absorption length in silicon is shown as a function of wavelength in nanometers. From Reicke (1994). Figure taken from Howell (2006).

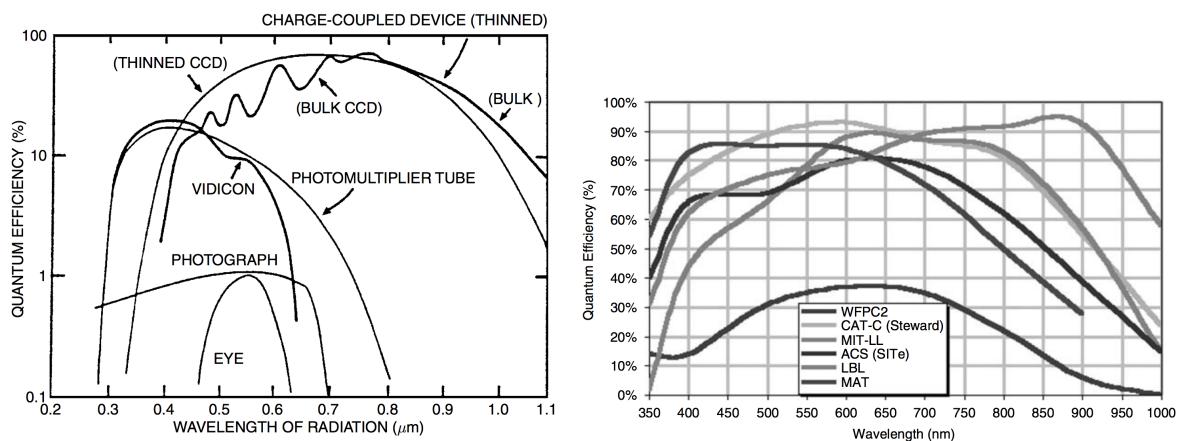


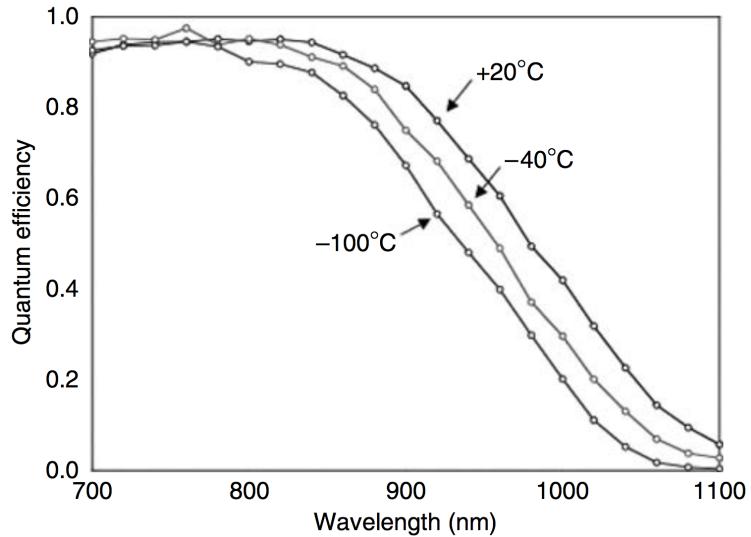
Figure 21: (left): QE curves for various devices, indicating why CCDs are a quantum leap above all previous imaging devices. The failure of CCDs at optical wavelengths shorter than about 3500\AA has been essentially eliminated via thinning or coating of the devices. (right): QE curves for a variety of CCDs. WFPC2 is the second generation wide-field/planetary camera aboard HST, CAT-C is a new generation SITe CCD used in a mosaic imager at the University of Arizona's 90'' telescope on Kitt Peak, MIT-LL is a CCD produced at the MIT Lincoln Laboratories and optimized for red observations, ACS is the Hubble Space Telescope Advanced Camera for Surveys SITe CCD, LBL is a Lawrence Berkeley Lab high resistivity, "deep depletion" CCD with high red QE, and MAT is a front-side, processed CCD showing high blue QE. Figures taken from Howell (2006).

particular device may only be representative of a “typical” device of the same kind, but they may not be 100% correct for the exact device of interest.

The quantum efficiency of a CCD is temperature sensitive especially in the red wavelength region. It has long been known that measurement of the QE at room temperature is a poor approximation to that which it will have when operated cold. Thus QE curves should be measured at or near the operating temperature at which the CCD will be used. As an example of the temperature sensitivity of the efficiency of a CCD, Figure 22 shows three QE measurements of the same CCD for temperatures of $+20^\circ \text{C}$ (\sim room temperature), -40°C , and -100°C (\sim operating temperature). Note that the variations are small at 8000\AA but increase to 20% at $9000 - 10000\text{\AA}$. The curves in this plot would lead one to the conclusion that operating a CCD at room temperature is the best thing to do. However, this is *not* the case and a compromise between operating temperature (i.e., dark current) and delivered QE must be used.

Charge diffusion: Once an electron is captured in a CCD pixel, the voltages applied during integration attempt to hold it in place. However, situations arise within a CCD pixel that provide a finite possibility for any given electron to wander out of its collection pixel and into a neighboring pixel. This process is called **charge diffusion** and until recently it was noted but of low significance compared with other noise and readout issues. Today CCDs are of excellent quality and have very low readout noise, good pixel registration on the array, and reside in high quality optical systems. These facts mean that CCD imaging now has the ability to show great detail of any optical aberrations and geometric distortions. Even items such as better mechanical tolerances in instrumentation can reveal noticeable focus variations

Figure 22:
 Sensitivity of the quantum efficiency of a MIT/LL CCD for three operating temperatures. The blue sensitivity is little affected by a change in operating temperature but the red QE can change by a factor of two. The use of such devices requires a balance of higher operating temperature and keeping the dark current under control. Figure taken from Howell (2006).



as materials breathe with thermal changes. Given CCDs with deep pixel wells, large format front-side illuminated thinned devices, and the related improvements to modern astronomical instrumentation, the effects of charge diffusion on the point-spread function are noticeable.

A few ways in which charge diffusion can occur may be useful to discuss. Imagine a deep (geometrically long) pixel modeled after that which is shown in Figure 23. Electrons produced by long wavelength photons are captured in the front-side illuminated pixel near the bottom, far from the applied voltages in the front gates. Thus the potential well for these electrons is more like a potential dip. Given the right circumstances, an electron can move into a neighboring pixel. Another example would be impurities in the silicon material the CCD was manufactured from. These variations in the Si lattice can slightly bend or slightly redirect the potential within a pixel and provide weak spots from which electron escape is possible. Ways to mitigate electron loss are the use of higher potential voltages (although this can lead to other issues such as logic glow or shorting on the array), higher resistivity Si to more tightly hold the electrons (the Si lattice) in place, or to use small pixels (but these have lower red QE and small well depths). Again, we see that compromise and specific application come into play and can be tuned into the CCD as a part of its production process.

Charge diffusion often varies significantly across a CCD, especially thinned devices as they are not equal thickness everywhere. For example, the HST ACS wide field camera thinned CCD shows a variation in the core of the PSF, caused by charge diffusion, across the field of view. The variation is 30 – 40% at 5000Å with larger variations at shorter wavelengths. The effects of charge diffusion are not to be taken lightly. The ACS/WFC suffers a loss of about 0.5 magnitudes in its limiting magnitude at short wavelengths and near 0.2 magnitudes elsewhere. Charge diffusion is especially important in thinned devices that under-sample the PSF.

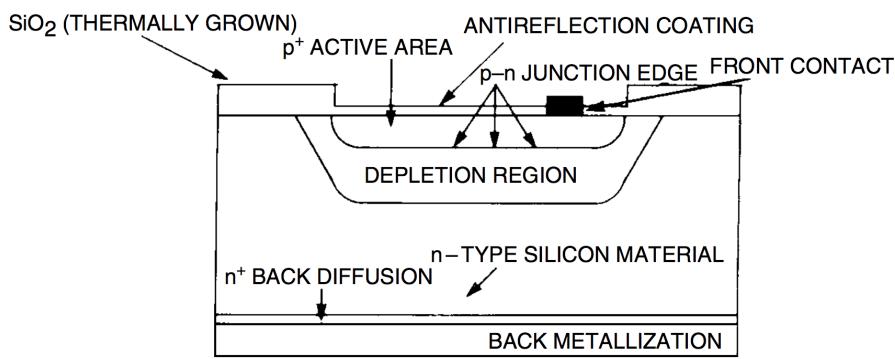
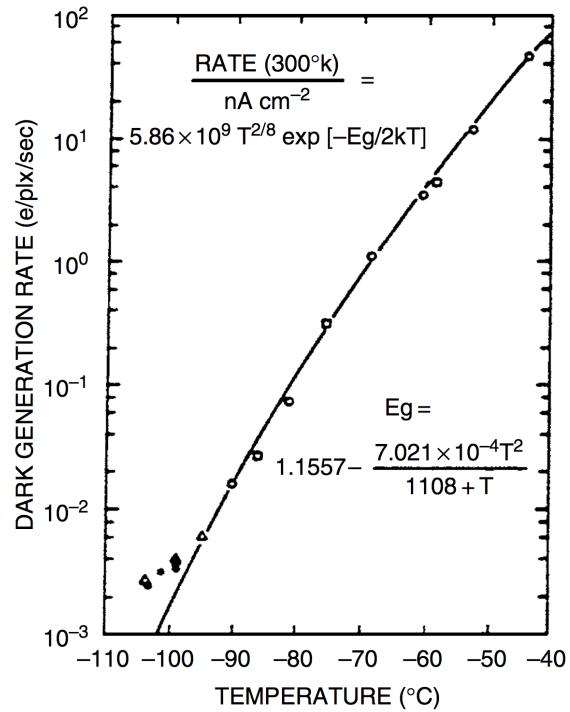


Figure 23:
 Schematic view of a single front-side illuminated CCD pixel. The square labeled “front contact” is a representation of part of the overall gate structure. The letters “p” and “n” refer to regions within the pixel consisting of silicon doped with phosphorus and boron respectively. Figure taken from Howell (2006).

Charge transfer efficiency: Charge transfer efficiency or CTE is a measure of the fraction of the charge that is successfully transferred for each pixel transfer. CTE values of 0.999995 or more are typical in good modern CCDs. For a CCD with 1024×1024 pixels, the charge collected in the last pixel readout has shifted 2048 times thus the CTE must be nearly 100% in order to preserve the charge in each pixel during readout. CTI (charge transfer inefficiency) is 1-CTE or numerically near $10 - 5$ or $10 - 6$ in value. CTI can be and usually is different in the vertical and horizontal directions. The loss in charge from a CCD

Figure 24:
 Experimental (symbols) and theoretical (line) results for the dark current generated in a typical three-phase CCD. The rate of dark current, in electrons generated within each pixel every second, is shown as a function of the CCD operating temperature. E_g is the band gap energy for silicon. From Robinson (1988a). Figure taken from Howell (2006).



pixel containing N electrons that is shifted 1024 times vertically and 1024 times horizontally is given by $L(e) = N(1024 \times \text{CTI}(H) + 1024 \times \text{CTI}(V))$, or, if a single CTI value is given, $L(e) = 2048 \times N \times \text{CTI}$. CCDs with poor CTE generally show charge tails in the direction opposite readout for bright stars. These tails are the charge left behind as the image is shifted out.

The standard method for measuring CTE is to use X-ray stimulation of a CCD with a Fe^{55} source. CCDs are good X-ray detectors and for a specific X-ray source such as Fe^{55} , each X-ray photon collected produces ~ 1620 electrons.⁵ A CCD is exposed to X-ray illumination and the resulting image readout. An X-ray transfer plot is made of signal in DN (y -axis) vs. running pixel number on the x -axis. Often hundreds of rows are summed together to increase the signal generated by the weak X-ray source. If the CCD has good CTE, a horizontal line will be seen at 1620 electrons (assuming a gain of 1.0). If the CTE is poor, this line starts at 1620 electrons (for rows close to the output amplifier) but tilts toward lower signal values for pixels further away from the output amplifier. This behavior indicates a loss of charge being transferred due to poor CTE. The CTE of a given CCD will generally degrade rapidly with decreasing operating temperature and is also a function of the clocking pulse shape and speed.

Readout noise: CCDs can be thought of as having three noise regimes: **read noise**, **shot noise**, and **fixed pattern noise**. In astronomy, we speak of these as **read noise limited**, **photon noise limited**, and **flat field uncertainties**. A plot of the log of the standard deviation of the signal (y -axis) vs. the log of the signal itself (x -axis) for a CCD is called the **photon transfer curve**. Read noise (or any noise independent of signal level) sets a noise floor for a device. Upon illumination, photon or Poisson noise raises the sigma measured following a \sqrt{N} slope. Finally, for large signal values, pixel to pixel variations due to processing errors and photomask misalignment begin to dominate. This latter noise is proportional to the signal and rises with a slope of 1.0. Full well sets in at some very high illumination and the slope of the photon transfer curve turns over or breaks.

Readout noise, or just read noise, is usually quoted for a CCD in terms of the number of electrons introduced per pixel into your final signal upon readout of the device. Read noise consists of two inseparable components. First is the conversion from an analog signal to a digital number, which is not perfectly repeatable. Each on-chip amplifier and A/D circuit will produce a statistical distribution of possible answers centered on a mean value.⁶ Thus, even for the hypothetical case of reading out the same pixel twice, each time with identical charge, a slightly different answer may be produced. Second, the electronics themselves will introduce spurious electrons into the entire process, yielding unwanted random fluctuations in the output. These two effects combine to produce an additive uncertainty in the final output value for each pixel. The average (1σ) level of this uncertainty is the read noise and is limited by the electronic properties of the on-chip output amplifier and the output electronics.

⁵Remember that for optical photon detection, one photon collected produces one photoelectron, regardless of its wavelength. For the much higher energy X-rays, a single photon collected produces multiple electrons in proportion to the photon's energy.

⁶The distribution of these values is not necessarily Gaussian

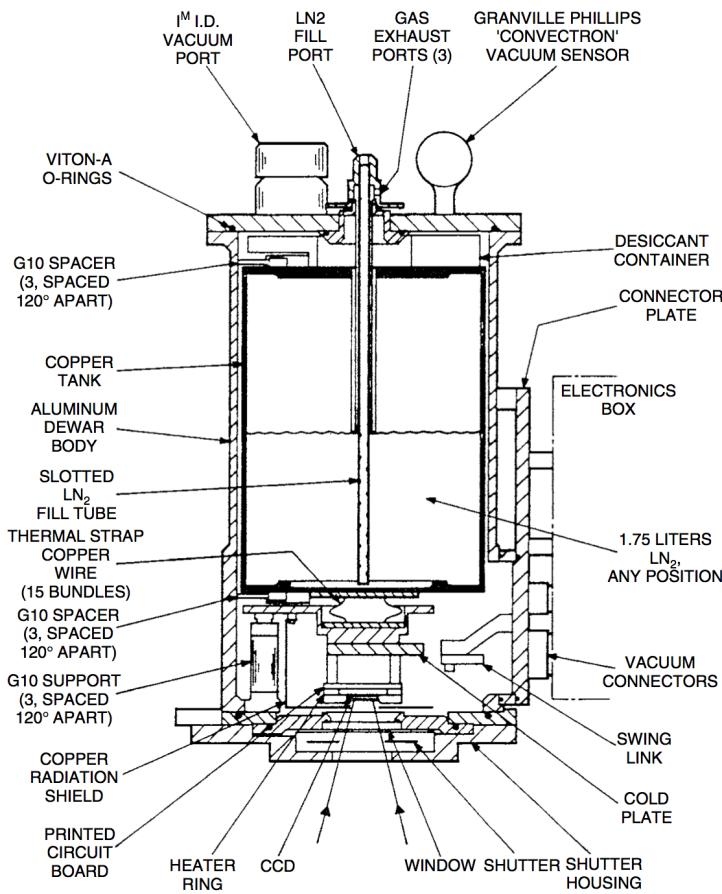


Figure 25:
A typical CCD dewar. This is the Mark-II Universal dewar originally produced in 1984 at Kitt Peak National Observatory. The dewar held 1.75 liters of liquid nitrogen providing a CCD operating time of approximately 12 hours between fillings. This dewar could be used in up-looking, down-looking, and side-looking orientations. From Brar (1984). Figure taken from Howell (2006).

In the output CCD image, read noise is added into every pixel every time the array is readout. This means that a CCD with a read noise of 20 electrons will, on average, contain 20 extra electrons of charge in each pixel upon readout. High read noise CCDs are thus not very good to use if co-addition of two or more images is necessary. The final resultant image will not be quite as good as one long integration of the same total time, as each co-added image will add in one times the read noise to every pixel in the sum. However, for modern CCDs, read noise values are very low and are hardly ever the dominant noise with which one must be concerned. Good read noise values in today's CCDs are in the range of 10 electrons per pixel per read or less. These values are far below read noise levels of ten years ago, which were as high as 50 – 100 electrons, and even those are well down from values of 300 – 500 or more electrons/pixel/read present in the first astronomical CCDs.

Dark current: Every material at a temperature much above absolute zero will be subject to thermal noise within. For silicon in a CCD, this means that when the thermal agitation is high enough, electrons will be freed from the valence band and become collected within the potential well of a pixel. When the device is read out, these dark current electrons become part of the signal, indistinguishable from astronomical photons. Thermal generation of electrons in silicon is a strong function of the temperature of the CCD, which is why astronomical use generally demands some form of cooling. Figure 24 shows a typical CCD dark current curve, which relates the amount of thermal dark current to the CCD operating temperature. Within the figure the theoretical relation for the rate of thermal electron production is given.

Dark current for a CCD is usually specified as the number of thermal electrons generated per second per pixel or as the actual current generated per area of the device (i.e., picoamps cm^{-2}). At room temperature, the dark current of a typical CCD is near 2.5×10^4 electrons/pixel/second. Typical values for properly cooled devices range from 2 electrons per second per pixel down to very low levels of approximately 0.04 electrons per second for each pixel. Although 2 electrons of thermal noise generated within a pixel every second sounds very low, a typical 15 minute exposure of a faint astronomical source would include 1800 additional (thermal) electrons within each CCD pixel upon readout. These additional charges cannot, of course, be uniquely separated from the photons of interest after readout. The dark current produced in a CCD provides an inherent limitation on the noise floor of a CCD. Because dark noise has a Poisson distribution, the noise actually introduced by thermal electrons into the signal is proportional to the square root of the dark current.

Cooling of CCDs is generally accomplished by one of two methods. The first, and usually the one used for scientific CCDs at major observatories, is via the use of liquid nitrogen (or in some cases liquid air). The CCD and associated electronics (the ones on or very near the actual CCD itself, called the head electronics) are encased in a metal dewar under vacuum. Figure 25 shows a typical astronomical CCD dewar. The liquid nitrogen (LN2) is placed in the dewar and, although not in direct physical contact with the CCD, cools the device to temperatures of near -100°C . Since LN2 itself is much colder than this, CCDs are generally kept at a constant temperature $\pm 1^{\circ}\text{C}$ with an on-board heater. In fact, the consistency of the CCD temperature is very important as the dark current is a strong function of temperature (Figure 24) and will vary considerably owing to even modest changes in the CCD temperature.

A less expensive and much less complicated cooling technique makes use of **thermoelectric cooling** methods. These methods are employed in essentially all “off-the-shelf” CCD systems and allow operation at temperatures of -20°C to -50°C or so, simply by plugging the cooler into an electrical outlet. Peltier coolers are the best known form of thermoelectric cooling devices. CCD operation and scientific quality imaging at temperatures near -30°C is possible, even at low light levels, due to advances in CCD design and manufacturing techniques and the use of multipinned phase operation.

The amount of dark current a CCD produces depends primarily on its operating temperature, but there is a secondary dependence upon the bulk properties of the silicon used in the manufacture. Even CCDs produced on the same silicon wafer can have slightly different dark current properties. Today’s CCDs are made from high purity epi wafers produced with low occurrences of integrated circuit error. These factors have greatly reduced many of the sources of dark current even at warmer temperatures. As with most of the noise properties of a given CCD, custom tailoring the CCD electronics (such as the bias level and the readout rate) can produce much better or much worse overall dark current and noise performance.

1.12 Question 12

Compare the S/N ratios between the following two cases where photon noise is dominant (assume an unresolved point source): [A] 1-minute exposure with a 10-m telescope; [B] 10-minute exposure with a 1-m telescope.

1.12.1 Short answer

Since we're in the photon-noise regime, the SNR can be written as:

$$\text{SNR} = \frac{N}{\sqrt{N}} = \frac{N}{\sqrt{N}} \frac{\sqrt{N}}{\sqrt{N}} = \frac{N\sqrt{N}}{N} = \sqrt{N}.$$

To get the total number of photons N , we can divide the flux by the photon energy and multiply by the product of collecting area and integration time:

$$N = \left(\frac{F}{h\nu} \right) A\Delta t = \left(\frac{F}{h\nu} \right) (\pi R^2)\Delta t \text{ [dimensionless].}$$

Plugging this back into our equation for the SNR, we get:

$$\text{SNR} = \sqrt{\frac{F\pi R^2 \Delta t}{h\nu}} \text{ [dimensionless]}$$

We can now use this to find the SNR ratio of a one-minute expose with a 10 m telescope to that of a ten-minute exposure with a 1 m telescope:

$$\begin{aligned} \frac{\text{SNR}_{1\text{ min}}}{\text{SNR}_{10\text{ min}}} &= \frac{\sqrt{\frac{F_{1\text{ min}}\pi R_{1\text{ min}}^2 \Delta t_{1\text{ min}}}{h\nu_{1\text{ min}}}}}{\sqrt{\frac{F_{10\text{ min}}\pi R_{10\text{ min}}^2 \Delta t_{10\text{ min}}}{h\nu_{10\text{ min}}}}} = \frac{\sqrt{R_{1\text{ min}}^2 \Delta t_{1\text{ min}}}}{\sqrt{R_{10\text{ min}}^2 \Delta t_{10\text{ min}}}} = \frac{\sqrt{(10\text{ m})^2(1\text{ min})}}{\sqrt{(1\text{ m})^2(10\text{ min})}} \\ &= \frac{10}{\sqrt{10}} = \frac{10}{\sqrt{10}} \left(\frac{\sqrt{10}}{\sqrt{10}} \right) = \frac{10\sqrt{10}}{10} = \sqrt{10} \text{ [dimensionless].} \end{aligned}$$

Thus, the SNR of a one-minute exposure with a 10 m telescope is $\sqrt{10}$ that of a ten-minute exposure with a 1 m telescope.

1.13 Question 13

Describe linear and circular polarizations of electromagnetic waves and give examples of their relevance to astronomical observations.

1.13.1 Short answer

An electromagnetic wave is said to be linearly polarized if it oscillates in a single plane; circularly polarized electromagnetic waves, on the other hand, are composed of two plane waves of equal amplitude with a phase difference of 90° . Circular polarization can either be right- or left-handed circularly polarized with the electric field vector rotating in a clockwise or counter-clockwise direction, respectively. Astrophysical examples of linear polarization include synchrotron radiation, dust polarization, and starlight polarization; those of circular polarization include starlight polarization and dust polarization (though generally very weak).

1.13.2 Additional context

Electromagnetic waves can be characterized by four properties: intensity, frequency, direction of propagation, and polarization state. Figure 26 shows why polarization is considered an important characteristic of an EM wave. The oscillating electric and magnetic field vectors will have a particular orientation at a given instant in time (Figure 26a). Figure 26b shows how the polarization state of a wave changes its orientation over time. If the wave oscillates in a single plane, the wave is said to be **linearly polarized** (see Figure 27b), while a wave composed of two plane waves of equal amplitude but differing in phase by 90° is **circularly polarized** (Figure 27a). If two plane waves have differing amplitudes or a phase difference that is not 90° , then the resultant wave is **elliptically polarized** (Figure 27c).

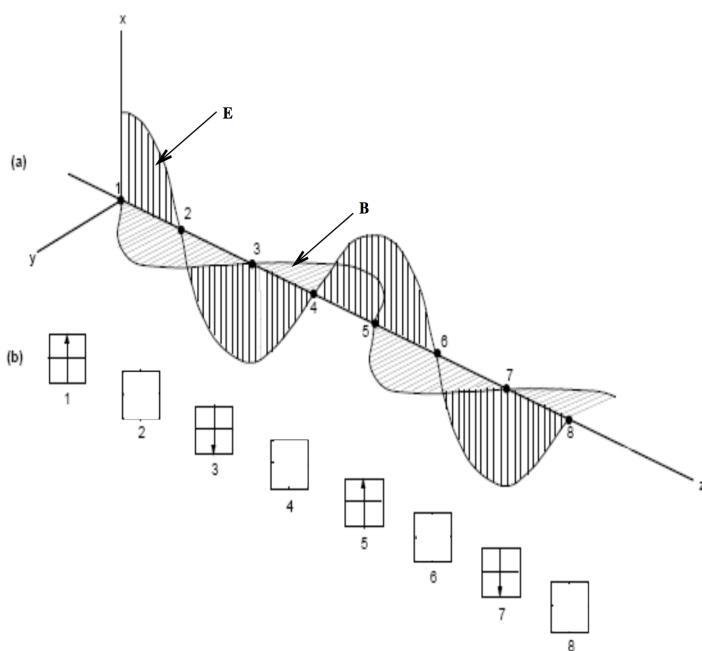


Figure 26:
This figure shows (a) an EM wave and (b) how its polarization changes as it propagates along the z -axis. At position 1, the electric vector is pointing in the x -direction, but at point 2, some time in the future, the electric vector is pointing in the y -direction. One complete revolution has been completed by point 4, and points 5 – 8 repeat the cycle. The wave is circularly polarized as there are equal amplitude components in both the x and y planes. This figure is reproduced from <http://astronomyonline.org>. Figure taken from Newton-McGee (2009).

There are many ways of describing the polarization state of an EM wave. If we begin with a quasi-monochromatic wave propagating in the z -direction, it can be resolved into three Cartesian components (x, y, z):

$$\begin{aligned}E_x &= a_1(t) \exp[i(\phi_1(t) - 2\pi\nu t)] \text{ [N C}^{-1}], \\E_y &= a_2(t) \exp[i(\phi_2(t) - 2\pi\nu t)] \text{ [N C}^{-1}], \\E_z &= 0 \text{ [N C}^{-1}],\end{aligned}$$

where E is the magnitude of the electric field vector, a_1 and a_2 are the amplitudes, and ϕ_1 and ϕ_2 are the phases of the component waves.

The polarization states of the wave are described by the **Stokes parameters** are:

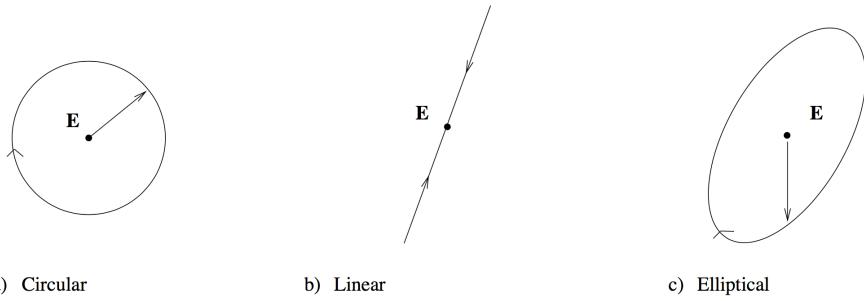


Figure 27: End on view of the trace of the tip of the electric vector for: (a) circularly polarized radiation, (b) linearly polarized radiation, and (c) elliptically polarized radiation. This figure is adapted from Cotton (1999). Figure taken from Newton-McGee (2009).

$$\begin{aligned} I &= \langle a_1^2 \rangle + \langle a_2^2 \rangle [\text{N C}^{-1}] \\ Q &= \langle a_1^2 \rangle - \langle a_2^2 \rangle [\text{N C}^{-1}] \\ U &= 2\langle a_1 a_2 \cos \phi \rangle [\text{N C}^{-1}] \\ V &= 2\langle a_1 a_2 \sin \phi \rangle [\text{N C}^{-1}]. \end{aligned}$$

where $\langle \rangle$ represents the time averaged quantity. **Stokes \mathbf{I}** is a measurement of the total intensity of the wave, **Stokes \mathbf{Q}** and **\mathbf{U}** together describe the linearly polarized components while **Stokes \mathbf{V}** represents the circularly polarized component of the wave. If $V > 0$ then the wave is right-handed circularly polarized (RCP), while if $V < 0$ then the polarization is left-handed circularly polarized (LCP). From correctly calibrated Stokes images, useful polarization images can be produced. Customarily, three linear polarization parameters are constructed: the **polarized intensity** (P) image, the **polarization angle** (χ) image, and the **polarization fraction** (p) image. These are formed from the Stokes Q and U images in the following manner:

$$\begin{aligned} P &= \sqrt{Q^2 + U^2} [\text{Jy}] \\ \chi &= \frac{1}{2} \arctan \left(\frac{U}{Q} \right) [\text{rad}] \\ p &= \frac{\sqrt{Q^2 + U^2}}{I} [\text{dimensionless}]. \end{aligned}$$

The polarization of electromagnetic radiation is fundamental to many physical processes including: how bees view the world, why rainbows exist, and many astrophysical phenomena including:

- synchrotron radiation (linear polarization);
- CMB polarization (linear and circular polarization);
- dust polarization (linear polarization);
- starlight polarization (linear polarization);
- astrophysical magnetic fields (linear polarization);
- Faraday rotation (linear polarization).

1.14 Question 14

What's the field of view of a $2K \times 2K$ CCD camera on a 5 m telescope with $f/16$ focal ratio? The pixel size of the CCD is $20\ \mu\text{m}$. Now, lets bring this to a 10 m telescope with the same focal ratio. Explain how the field of view changes on the 10 m telescope (compared to that of the 5 m telescope) based on the Etendue conservation rule.

1.14.1 Short answer

The FOV of a telescope operating with a CCD is given as follows:

$$\text{FOV} = \frac{1 \text{ rad}}{\text{focal length}} \cdot (\# \text{ pixels}) \cdot (\text{pixel size}) = \frac{206265''}{f \# D} N \Delta x,$$

where we have used the fact that $1 \text{ rad} = 206265''$ and that the focal length is given by $f = f \# D$, where D is the telescope diameter.

Using this to find the FOV for a $2K \times 2K$ CCD camera on a 5 m telescope with $f/16$ focal ratio:

$$\text{FOV} = \frac{206265''}{(16)(5 \text{ m})} (2000)(20 \times 10^{-6} \text{ m}) \sim 100''.$$

We can now use the Etendue conservation rule to find the FOV at a 10 m telescope. Etendue is a property of light which characterizes how “spread out” the light is in an area and solid angle, and the Etendue conservation rule says that this property is conserved following the equation

$$\epsilon = \text{FOV}^2 \cdot D^2 = \text{const} \text{ [dimensionless].}$$

This allows us to equate the etendue of the first telescope setup to that of the second in order to solve for the new FOV:

$$\begin{aligned}\epsilon_{5 \text{ m}} &= \epsilon_{10 \text{ m}} \\ \text{FOV}_{5 \text{ m}}^2 \cdot D_{5 \text{ m}}^2 &= \text{FOV}_{10 \text{ m}}^2 \cdot D_{10 \text{ m}}^2 \\ \text{FOV}_{10 \text{ m}} &= \left(\frac{D_{5 \text{ m}}}{D_{10 \text{ m}}} \right) \text{FOV}_{5 \text{ m}} \\ \text{FOV}_{10 \text{ m}} &= \left(\frac{5 \text{ m}}{10 \text{ m}} \right) 100'' \\ \text{FOV}_{10 \text{ m}} &= \frac{1}{2} 100'' \\ \text{FOV}_{10 \text{ m}} &= 50''.\end{aligned}$$

Thus, using Etendue conservation, we can see that the FOV is reduced by a factor of 2 when brought to a telescope twice the diameter.

1.14.2 Follow-up Questions

- If you wanted a smaller FOV with the same size telescope, what would you change?

1.15 Question 15

Sketch and give the equations for each of the following distributions: 1. Gaussian (Normal distribution); 2. Poisson distribution; 3. Log-normal distribution. Give two examples from astrophysics where each of these distributions apply.

1.15.1 Short answer

The Gaussian (normal) distribution:

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

The Poisson distribution:

$$f(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}$$

The log-normal distribution:

$$f(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right]$$

1.15.2 Additional context

Accuracy versus precision: It is important to distinguish between the terms *accuracy* and *precision*. The **accuracy** of an experiment is a measure of how close the result of the experiment is to the true value; the **precision** is a measure of how well the result has been determined, without reference to its agreement with the true value. The precision is also a measure of the **reproducibility** of the result in a given experiment. The distinction between *accuracy* and *precision* is illustrated by the two sets of measurements in Figure 28 where the straight line on each graph shows the expected relation between the dependent variable y and the independent variable x . In both graphs, the scattered of the data points is a reflection of uncertainties in the measurements, consistent with the error bars on the points. The data in Figure 28a have been measured to a high degree of *precision* as illustrated by the small error bars, and are in excellent agreement with the expected variation of y with x , but are clearly *inaccurate*, deviating from the line by a constant offset. On the other hand, the data points in Figure 28b are rather *imprecise* as illustrated by the large error bars, but are scattered around the predicted distribution.

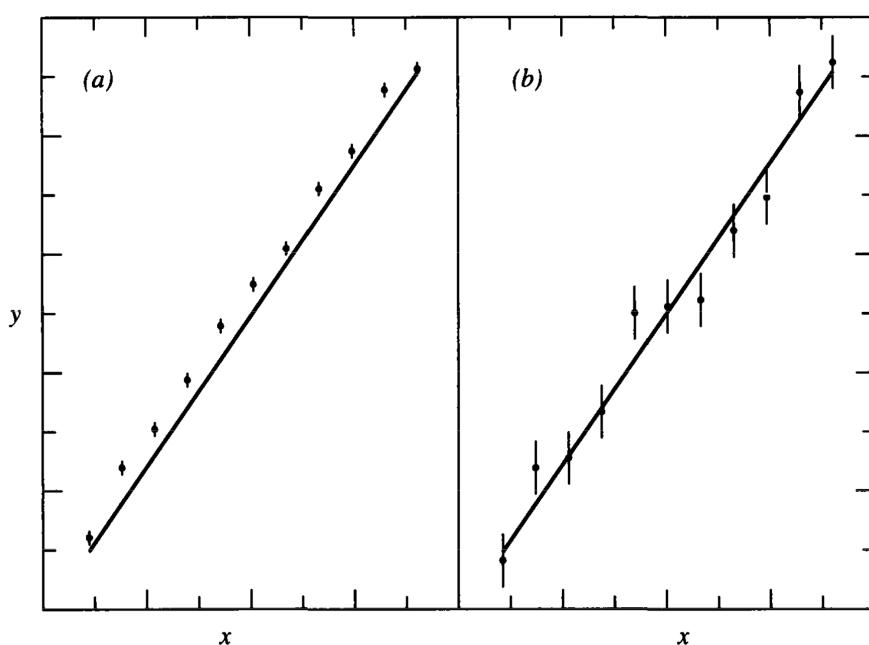


Figure 28:
Illustration of the difference between precision and accuracy. (a) Precise but inaccurate data. (b) Accurate but imprecise data. True values are represented by the straight lines. Figure taken from Bevington & Robinson (2003).

It is obvious that we must consider the accuracy and precision simultaneously for any experiment. It would be a waste of time and energy to determine a result with high precision if we knew that the result would be highly inaccurate. Conversely, a result cannot be considered to be extremely accurate if the precision is low. In general, when we know the **uncertainty** or **error** of an experimental result, we are

referring to the precision with which that result has been determined. **Absolute precision** indicates the magnitude of the uncertainty in the result in the same units as the result; **relative precision** indicates the uncertainty in terms of a fraction of the value of the result.

Systematic errors: The accuracy of an experiment, as we have defined it, is generally dependent on how well we can control or compensate for **systematic errors**, or errors that will make our result different from the “true” value with reproducible discrepancies. Errors of this type are not easy to detect and not easily studied by statistical analysis. They may result from faulty calibration of equipment or bias on the part of the observer. They must be estimated from an analysis of the experimental conditions and techniques. A major part of the planning of an experiment should be devoted to understanding and reproducing sources of systematic errors.

Random errors: The precision of an experiment depends upon how well we can overcome **random errors**, or fluctuations in observations that yield different results each time the experiment is repeated, and thus requires repeated experimentation to yield precise results. A given accuracy implies an equivalent precision and, therefore, also depends to some extent on random errors.

The problem of reducing random errors is essentially one of improving the experimental method and refining the techniques, as well as simply repeating experiment. If the random errors result from instrumental uncertainties, they may be reduced by using more reliable and more precise measuring techniques. If the random errors result from statistical fluctuations in a limited number of measurements, they may be reduced by making more measurements. There are, however, practical limits to these improvements.

Mean, median, and mode: The **mean** \bar{x} of an experimental distribution is given as the sum of N determinations of x_i of the quantity x divided by the number of determinations

$$\bar{x} \equiv \frac{1}{N} \sum x_i \text{ [function]},$$

and the **mean of the parent population** μ is defined as the limit

$$\mu \equiv \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum x_i \right) \text{ [function]}.$$

The mean is therefore equivalent to the **centroid** or **average** value of the quantity x .

Deviations: The deviation d_i of any measurement x_i from the mean μ of the parent population is defined as the difference between x_i and μ :

$$d_i \equiv x_i - \mu \text{ [function]}.$$

For computational purposes, deviations are generally defined with respect to the mean, rather than the median or most probable value. If μ is the true value of the quantity, d_i is also the true error in x_i .

The average of the deviations d_i must vanish by virtue of the definition of the mean:

$$\lim_{N \rightarrow \infty} \bar{d}_i = \lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum (x_i - \mu) \right] = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum x_i \right) - \mu = 0 \text{ [function]}.$$

The **average deviation** α , therefore, is defined as the absolute values of the deviations:

$$\alpha \equiv \lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum |x_i - \mu| \right] \text{ [function]}.$$

The average deviation is a measure of the **dispersion** of the expected observations about the mean. The presence of the absolute value signs makes its use inconvenient for statistical analysis.

A parameter that is easier to use analytically and that can be justified fairly well on theoretical grounds to be a more appropriate measure of the dispersion of the observations is the **standard deviation** σ . The **variance** σ^2 is defined as the limit of the average of the squares of the deviations from the mean μ :

$$\sigma^2 \equiv \lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum (x_i - \mu)^2 \right] = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum x_i^2 \right) - \mu^2 \text{ [function]}$$

and the standard deviation σ is the square root of the variance. Note that the second form of the variance shown above is often described as the “average of the squares minus the square of the average”. The standard deviation of the **root mean square** of the deviation, and is associated with the **second moment** of the date about the mean. The corresponding expression for the variance s^2 of the sample population is given by

$$s^2 \equiv \frac{1}{N} \sum (x_i - \bar{x})^2 \text{ [function]}$$

where the factor of $N - 1$, rather than N , is required in the denominator to account for the fact that the parameter \bar{x} has been determined from the data and not independently. We note that the symbol

σ (instead of s) is often used to represent the best estimate of the standard deviation of the parent distribution determined from a sample distribution.

Significance: The mean μ and the standard deviation σ , as well as the median, the most probable value, and the average deviation, are all parameters that characterize the information we are seeking when we perform an experiment. Often we wish to describe our distribution in terms of just the mean and standard deviation. The mean may not be exactly equal to the datum in question if the parent distribution is not symmetrical about the mean, but it should have the same characteristics. If a more detailed description is desired, it may be useful to compute **higher moments about the mean**.

In general, the best we can say about the mean is that it is one of the parameters that specifies the probability distribution: it has the same units as the “true” value and, in accordance with convention, we shall consider it to be the best estimate of the “true” value under the prevailing experimental conditions. The variance s^2 and the standard deviation s characterize the uncertainties associated with our experimental attempts to determine the “true” values. For a given number of observations, the uncertainty in determining the mean of the parent distribution is proportional to the standard deviation of that distribution. The standard deviation s is, therefore, an appropriate measure of the uncertainty due to fluctuations in the observations in our attempt to determine the “true” value.

Although, in general, the distribution resulting from purely statistical errors can be described well by the two parameters, the mean and the standard deviation, we should be aware that, at distances of a few standard deviations from the mean of an experimental distribution, non-statistical errors may dominate. In especially severe cases, it may be preferable to describe the spread of the distribution in terms of the average deviation, rather than the standard deviation, because the latter tends to de-emphasize measurements that are far from the mean. There are also distributions for which the variance does not exist. The average deviation or some other quantity must be used as a parameter to indicate the spread of the distribution in such cases.

Probability distributions: Of the many probability distributions that are involved in the analysis of experimental data, three play a fundamental role: the binomial distribution, the Poisson distribution, and the Gaussian distribution. Of these, the **Gaussian, or normal error, distribution**, is undoubtedly the most important in statistical analysis of data. Practically, it is useful because it seems to describe the distribution of random observations for many experiments, as well as describing the distributions obtained when we try to estimate the parameters of most other probability distributions.

The **Poisson distribution** is generally appropriate for counting experiments where the data represents the number of items or events observed per unit time interval. It is important in the study of random processes such as those associated with the radioactive decay of elementary particles or nuclear states, and is also applied to data that have been sorted into ranges to form a frequency table or a histogram.

The **binomial distribution** is generally applied to experiments in which the result is one of a small number of possible final states, such as the number of “heads” or “tails” in a series of coin tosses, or the number of particles scattered forward or backwards relative to the direction of the incident particle in a particle physics experiment. Because both the Poisson and the Gaussian distribution can be considered as limiting cases of the binomial distribution, we shall devote some attention to the derivation of the binomial distribution from basic considerations.

Binomial distribution: Suppose we toss a coin in the air and let it land. There is a 50% probability that it will land heads up and a 50% probability that it will land tails up. By this, we can see that if we continue tossing a coin repeatedly, the fraction of times that it lands with heads up will asymptotically approach 1/2, indicating that there was a probability of 1/2 in doing so. For any given coin toss, the probability cannot determine whether or not it will land heads up; it can only describe how we should expect a large number of tosses to be divided into these two possibilities.

Suppose we toss two coins at a time. There are now four different possible permutations of the way in which they can land: both heads up, both tails up, and two mixtures of heads and tails depending on which one is heads up. Because each of these permutations is equally probable, the probability for any choice of them is 1/4 or 25%. To find the probability for obtaining a particular mixture of heads and tails, without differentiating between the two kinds of mixtures, we must add the probabilities corresponding to each possible kind. Thus, the total probability of finding either heads up and the other tails up is 1/2. Note that the sum of the probabilities for all possibilities ($1/4 + 1/4 + 1/4 + 1/4$) is always equal to 1 because *something* is bound to happen.

Let us extrapolate these ideas to the general case. Suppose we toss n coins into the air, where n is some integer. Alternatively, suppose that we toss one coin n times. What is the probability that exactly x of these coins will land heads up, without distinguishing which of the coins actually belongs to which group? We can consider the probability $P(x; n)$ to be a function of the number of coins tossed n and of the number of coins that land heads up x . For a given experiment in which n coins are tossed, this probability $P(x; n)$ will vary as a function of x . Of course, x must be an integer for any physical experiment, but we can consider the probability to be smoothly varying with x as a continuous variable for mathematical

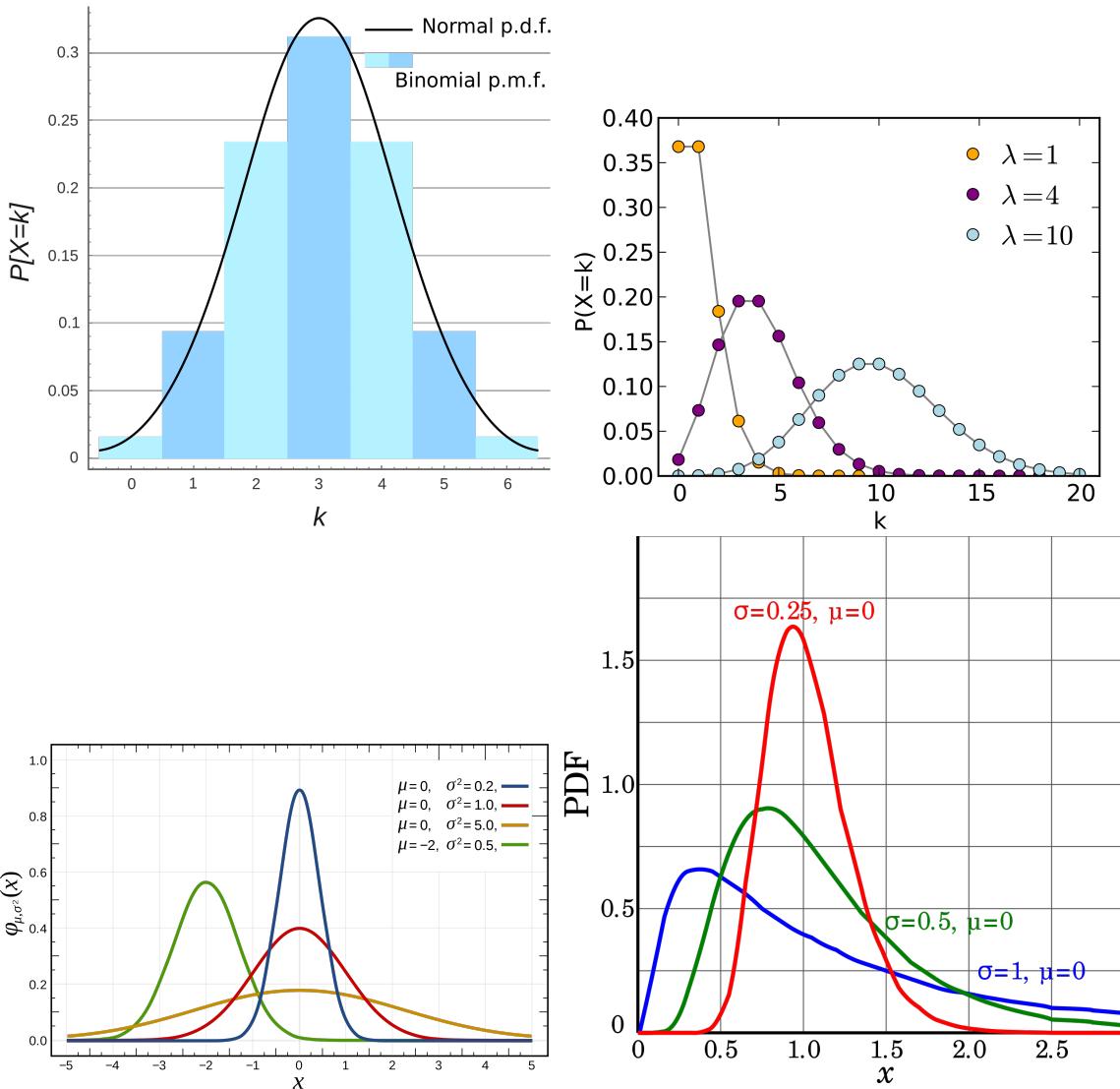


Figure 29: (Top left): Binomial distribution. (Top right): Poisson distributions. (Bottom left): Gaussian distributions. (Bottom right): Log-normal distributions. Figures taken from Wikipedia.

purposes.

Permutations and combinations: If n coins are tossed, there are 2^n different possible ways in which they can land. This follows from the fact that the first coin has two possible orientations, for each of these the second coin also has two such orientations, for each of these the third coin also has two, and so on. Because each of these possibilities is equally probable, the probability for any one of these possibilities to occur at any toss of n coins has to be $1/2^n$.

How many of these possibilities will contribute to our observation of x coins with heads up? Imagine two boxes, one labeled “heads” and divided into x slots, and the other labeled “tails”. We shall consider first the question of how many permutations of the coins result in the proper separation of x in one box and $n - x$ in the other; then we shall consider the question of how many combinations of these permutations should be considered different from each other.

In order to enumerate the number of **permutations** $Pm(n, x)$, let us pick up the coins one at a time from the collection of n coins and put x of them into the “heads” box. We have a choice of n coins for the first one we pick up. For our second selection, we can choose from the remaining $n - 1$ coins. The range of choice is diminished until the last selection of the x th coin can be made from only $n - x + 1$ remaining coins. The total number of choices for coins to fill the x slots in the “heads” box is the product of the numbers of individual choices:

$$Pm(n, x) = n(n - 1)(n - 2) \cdots (n - x + 2)(n - x + 1).$$

This expansion can be expressed more easily in terms of factorials:

$$Pm(n, x) = \frac{n!}{(n-x)!} \text{ [dimensionless].}$$

So far we have calculated the number of permutations $Pm(n, x)$ that will yield x coins in the “heads” box and $n - x$ coins in the “tails” box, with the provision that we have identified which coin was placed in the “heads” box first, which was placed in the second, and so on. That is, we have *ordered* the x coins in the “heads” box. In our computation of 2^n different possible permutations of the n coins, we are only interested in which coins landed heads up or heads down, not which landed first. Therefore, we must consider contributions different only if there are different coins in the two boxes, not if the x coins within the “heads” box are permuted into different time orderings.

The number of different **combinations** $C(n, x)$ of the permutations in the preceding enumeration results from combining the $x!$ different ways in which x coins in the “heads” box can be permuted within the box. For every $x!$ permutations, there will be only one new combination. Thus, the number of different combinations $C(n, x)$ is the number of permutations $Pm(n, x)$ divided by the degeneracy factor $x!$ of the permutations:

$$C(n, x) = \frac{Pm(n, x)}{x!} = \frac{n!}{x!(n-x)!} = \binom{n}{x} \text{ [dimensionless]}$$

where $q = 1 - p$. The name for the binomial distribution comes from the fact that the coefficients $P_B(x; n, p)$ are closely related to the binomial theorem for the expansion of a power of a sum. According to the **binomial theorem**,

$$(p+q)^n = \sum_{x=0}^n \left[\binom{n}{x} p^x q^{n-x} \right] \text{ [dimensionless].}$$

This is the number of different possible combinations of n items taken x at a time, commonly referred to as $\binom{n}{x}$ or “ n over x ”.

Probability: The probability $P(x; n)$ that we should observe x coins with heads up and $n - x$ tails up is the product of different combinations $C(n, x)$ that contribute to that set of observations multiplied by the probability for each of the combinations to occur, which we have found to be $(1/2)^n$.

Actually, we should separate the probability for each combination into two parts: one part is the probability $p^x = (1/2)^x$ for x coins to be heads up; the other part is the probability $q^{n-x} = (1 - 1/2)^{n-x} = (1/2)^{n-x}$ for the other $n - x$ coins to be tails up. For symmetrical coins, the product of these two parts $p^x q^{n-x} = (1/2)^n$ is the probability of the combination with x coins heads up and $n - x$ coins tails up. In the general case, the probability p of success for each item is not equal in magnitude to the probability $q = 1 - p$ for failure. For example, when tossing a die, the probability that a particular number will show up is $p = 1/6$, while the probability of it now showing is $q = 1 - 1/6 = 5/6$ so that $p^x q^{n-x} = (1/6)^x \times (5/6)^{n-x}$.

With these definitions of p and q , the probability $P_B(x; n, p)$ for observing x of the n items to be in the state with probability p is given by the **binomial distribution**:

$$P_B(x; n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \text{ [dimensionless].}$$

Figure 29 (top left) shows an example binomial distribution.

Poisson distribution: The Poisson distribution represents an approximation to the binomial distribution for the special case where the average number of successes is much smaller than the possible number; that is, when $\mu \ll n$ because $p \ll 1$. For such experiments the binomial distribution correctly describes the probability $P_B(x; n, p)$ of observing x events per unit time interval out of n possible events, each of which has a probability p of occurring, but the large number of n possible events makes exact evaluation from the binomial distribution impractical. Furthermore, neither the number n of possible events nor the probability p for each is usually known. What may be known instead is the average number of events μ expected in each time interval or its estimate \bar{x} . The Poisson distribution provides an analytical form appropriate to such investigations that describes the probability distribution in terms of just the variable x and the parameter μ .

Let us consider the binomial distribution in the limiting case of $p \ll 1$. We are interested in its behaviour as n becomes infinitely large while the mean $\mu = np$ remains constant. The equation for the probability function of the binomial distribution can be written as

$$P_B(x; n, p) = \frac{1}{x!} \frac{n!}{(n-x)!} p^x (1-p)^{n-x} (1-p)^n \text{ [function].}$$

If we expand the second factor,

$$\frac{n!}{(n-x)!} = n(n-1)(n-2) \cdots (n-x-2)(n-x-1) \text{ [function]}$$

we can consider it to be the product of x individual factors, each of which is very nearly equal to n because $x \ll n$ in the region of interest. The second factor in $P_B(x; n, p)$ above thus asymptotically approaches n^x . The product of the second and third factors then becomes $(np)^x = \mu^x$. The fourth factor is approximately equal to $(1+px)$ which tends to 1 as p tends to 0.

The last factor can be rearranged by substituting μ/p for n and expanding the expression to show that it asymptotically approaches $e^{-\mu}$:

$$\lim_{p \rightarrow 0} (1-p)^x = \lim_{p \rightarrow 0} [(1-p)^{1/p}]^\mu = \left(\frac{1}{e}\right)^\mu = e^{-\mu} \text{ [function].}$$

Combining these approximation, we find that the binomial distribution probability function $P_B(x; n, p)$ asymptotically approaches the **Poisson distribution** $P_P(x; \mu)$ as p approaches 0:

$$\lim_{p \rightarrow 0} P_B(x; n, p) = P_P(x; \mu) \equiv \frac{\mu^x}{x!} e^{-\mu} \text{ [function].}$$

Because this distribution is an approximation to the binomial distribution for $p \ll 1$, the distribution is asymmetric about its mean μ . Note that $P_P(x; \mu)$ does not become 0 for $x = 0$ and is not defined for negative values of x . This restriction is not troublesome for counting experiments because the number of counts per unit time interval can never be negative.

Derivation: The Poisson distribution can also be derived for the case where the number of events observed is small compared to the total possible number of events. Assume that the average rate at which events of interest occur is constant over a given interval of time and that event occurrences are randomly distributed over that interval. Then, the probability dP of observing no events in a time interval dt is given by

$$dP(0; t\tau) = -P(0; t, \tau) \frac{dt}{\tau} \text{ [function]}$$

where $P(x; t, \tau)$ is the probability of observing x events in the time interval dt , τ is a constant proportionality factor that is associated with the mean time between events, and the minus sign accounts for the fact that increasing the differential time interval dt decreases the probability proportionality. Integrating this equation yields the probability of observing no events within a time t to be

$$P(0; t, \tau) = P_0 e^{-t/\tau} \text{ [function]}$$

where P_0 , the constant of integration, is equal to 1 because $P(0; t, \tau) = 1$ at $t = 0$.

The probability $P(x; t, \tau)$ for observing x events in the time interval τ can be evaluated by integrating the differential probability

$$d^x P(x; t, \tau) = \frac{e^{-t/\tau}}{x!} \prod_{i=1}^x \frac{dt_i}{\tau} \text{ [function]}$$

which is the product of the probabilities of observing each events in a different interval dt_i and the probability $e^{-t/\tau}$ of not observing any other events in the remaining time. The factor of $x!$ in the denominator compensates for the ordering implicit in the probabilities $dP_i(l, t, \tau)$.

Thus, the probability of observing x events in the time interval t is obtained by integration

$$P_P(x; \mu) = P(x; t, \tau) = \frac{e^{-t/\tau}}{x!} \left(\frac{t}{\tau}\right)^x \text{ [function]}$$

or

$$P_P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} \text{ [function]}$$

which is the expression for the Poisson distribution, where $\mu = t/\tau$ is the average number of events observed in the time interval t . This equation represents a **normalized probability function**; that is, the sum of the function evaluated at each of the allowed values of the variable x is unity:

$$\sum_{x=0}^{\infty} P_P(x, \mu) = \sum_{x=0}^{\infty} \frac{\mu^x}{x!} e^{-\mu} = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = e^{-\mu} e^{\mu} = 1.$$

Mean and standard deviation: The Poisson distribution, like the binomial distribution, is a *discrete* distribution. That is, it is only defined only at integral values of the variable x , although the parameter μ is a positive, real number. The mean of the Poisson distribution is actually the parameter μ that appears in the probability function $P_P(x; \mu)$. To verify this, we can evaluate the expectation value $\langle x \rangle$ of x :

$$\langle x \rangle = \sum_{x=0}^{\infty} \left(x \frac{\mu^x}{x!} e^{-\mu} \right) = \mu e^{-\mu} \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!} = \mu e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{y!} = \mu \text{ [dimensionless].}$$

To find the standard deviation σ , the expectation value of the square of the deviations can be evaluated:

$$\sigma^2 = \langle (x - \mu)^2 \rangle = \sum_{x=0}^{\infty} \left[(x - \mu)^2 \frac{\mu^x}{x!} e^{-\mu} \right] = \mu \text{ [function].}$$

Thus, the standard deviation σ is equal to the square root of the mean μ and the Poisson distribution has only a single parameter, μ .

Computation of the Poisson distribution by

$$P_P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} \text{ [function]}$$

can be limited by the factorial function in the denominator. The problem can be avoided by using logarithms or by using the recursive relations

$$P(0; \mu) = e^{-\mu}, \quad P(x; \mu) = \frac{\mu}{x} (x-1; \mu) \text{ [dimensionless].}$$

This form has the disadvantage that, in order to calculate the function for particular values of x and μ , the function must be calculated at all lower values of x as well. However, if the function is to be summed from $x = 0$ to some upper limit to obtain the summed probability or to generate the distribution for a Monte Carlo calculation, the function must be calculated at all lower values of x anyway.

Summed probability: We may want to know the probability of obtaining a sample value of x between limits x_1 and x_2 from a Poisson distribution with mean μ . This probability is obtained by summing the values of the function calculated at the integral values of x between the two integral limits x_1 and x_2 ,

$$S_P(x_1, x_2; \mu) = \sum_{x_1}^{x_2} P_P(x; \mu) \text{ [dimensionless].}$$

More likely, we may want to find the probability of recording n or more events in a given interval when the mean number of events is μ . This is just the sum

$$S_P(n, \infty; \mu) = \sum_{x=n}^{\infty} P_P(x; \mu) = 1 - \sum_{x=0}^{n-1} P_P(x; \mu) = 1 - e^{-\mu} \sum_{x=0}^{n-1} \frac{\mu^x}{x!} \text{ [function].}$$

So, this distribution has the special property that the mean is equal to the variance.

The first four moments of the Poisson distribution are given by:

1. Mean = μ
2. Variance = μ
3. Skewness = μ
4. Kurtosis = $3\mu^2 + \mu$.

For large values of μ , the probability sum may be approximated by an integral of the Gaussian function. Figure 29 (top right) shows an example Poisson distribution.

Application: A steady celestial source of constant luminosity produces on average a finite number of counts in a detector every second. The photons do not arrive with equal intervals. The average rate of arrival is controlled by a fixed probability of an event occurring in some fixed interval of time. There is some chance of an event occurring in every millisecond, but no certainty that the event will indeed arrive in any specific millisecond. This randomness leads to a variation in the number of counts detected in successive intervals.

Gaussian (or normal error) distribution: The Gaussian distribution is an approximation to the binomial distribution for the special limiting case where the number of possible different observations n becomes infinitely large and the probability of success for each is finitely large so $np \gg 1$. It is also the limiting case for the Poisson distribution as μ becomes large.

There are several derivations of the Gaussian distribution from first principles, none of them as convincing as the fact that the distribution is reasonable, that it has a fairly simple analytic form, and that it is accepted by convention and experimentation to be the most likely distribution for most experiments. In addition, it has the satisfying characteristic that the most probable estimate of the mean μ from a random sample of observations x is the average of those observations \bar{x} .

Characteristics: The Gaussian probability density is defined as

$$p_G = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \text{ [function].}$$

This distribution is **symmetrical** and **mesokurtic**. It is a continuous function describing the probability of obtaining the value x in a random observation from a parent distribution with parameters μ and σ , corresponding to the mean and standard deviation, respectively. Because the distribution is continuous, we must define an interval in which the value of the observation x will fall. The probability density function is properly defined such that the probability that the value of a random observation $dP_G(x; \mu, \sigma)$ will fall within an interval dx around x is given by

$$dP_G(x; \mu, \sigma) = p_G(x; \mu, \sigma) rmdx \text{ [function],}$$

considering dx to be an infinitesimal differential, and the probability density function to be normalized, so that

$$\int_{x=-\infty}^{x=\infty} dP_G(x; \mu, \sigma) = \int_{x=-\infty}^{x=\infty} p_G(x; \mu, \sigma) dx \text{ [function].}$$

The width of the curve is determined by the value of σ , such that for $x = \mu + \sigma$, the height of the curve is reduced to $e^{-1/2}$ of its value at the peak:

$$p_G(x; \mu \pm \sigma, \sigma) = e^{-1/2} p_G(\mu; \mu, \sigma) \text{ [function].}$$

We can characterize a distribution by its **full-width at half maximum** (FWHM) Γ , often referred to as the **half-width**, defined as the range of x between values at which the probability $p_G(x; \mu, \sigma)$ is half its maximum value:

$$p_G(\mu \pm \frac{1}{2}\Gamma, \mu, \sigma) = \frac{1}{2} p_G(\mu; \mu, \sigma) \text{ [function].}$$

With this definition, we can determine from the Gaussian probability density p_G that

$$\Gamma = 2\sqrt{2 \ln 2}\sigma \approx 2.354\sigma \text{ [function].}$$

Tangents drawn along a portion of steepest descent of the curve intersect the curve at the $e^{-1/2}$ points $x = \mu \pm \sigma$ and intersect the x axis at the points $x = \mu \pm 2\sigma$.

The first four moments of the Gaussian distribution are given by:

1. Mean = μ
2. Variance = σ^2
3. Skewness = 0
4. Kurtosis = $3\sigma^4$.

Figure 29 (bottom left) shows an example Gaussian distribution.

Regardless of statistical considerations, there are a number of reasons why the normal distribution is much better known than the log-normal. A major one appears to be symmetry, one of the basic principles realized in nature as well as in our culture and thinking. Thus, probability distributions based on symmetry may have more inherent appeal than skewed ones. Two other reasons relate to simplicity. First, as Aitchison and Brown states, “Man has found addition an easier operation than multiplication, and so it is not surprising that an additive law of errors was the first to be formulated.” Second, the established, concise description of a normal sample is handy, well-known, and sufficient to represent the underlying distribution, which made it easier, until now, to handle normal distributions than to work with log-normal distributions. Another reason relates to the history of distributions: The normal distribution has been known and applied more than twice as long as its log-normal counterpart. Finally, the very notion of “normal” conjures more positive associations for non-statisticians than does “log-normal.” For all of these reasons, the normal or Gaussian distribution is far more familiar than the log-normal distribution is to most people.

Application: Gaussian function approximates the shapes of many observables in astronomy, such as the profiles of seeing disks, the width of spectral lines and the distribution of noise in radio receivers. In error analysis, the Gaussian distribution is often used to determine the significance of a measurement in the presence of noise.

Log-normal distribution: Many measurements show a more or less skewed distribution. Skewed distributions are particularly common when mean values are low, variance large, and values cannot be

negative. Such skewed distributions often closely fit the log-normal distribution. What is the difference between normal and log-normal variability? Both forms of variability are based on a variety of forces acting independently of one another. A major difference, however, is that the effects can be additive or multiplicative, thus leading to normal or log-normal distributions, respectively.

Some basic principles of additive and multiplicative effects can easily be demonstrated with the help of two ordinary dice with sides numbered from 1 to 6. Adding the two numbers, which is the principle of most games, leads to values of 2 to 12, with a mean of 7, and a symmetrical frequency distribution. The total range can be described as 7 plus or minus 5 (i.e., 7 ± 5); note that in this case, 5 is not the standard deviation. Multiplying the two numbers, however, leads to values between 1 and 36 with a highly skewed distribution. The total variability can be described as 6 multiplied or divided by 6. In this case, the symmetry has moved to the multiplicative level.

Although these examples are neither normal nor log-normal distributions, they do clearly indicate that additive and multiplicative effects give rise to different distributions. Thus, we cannot describe both types of distributions in the same way. Unfortunately, however, common belief has it that quantitative variability is generally bell shaped and symmetrical. The current practice in science is to use symmetrical error bars in graphs to indicate standard deviations or errors, and the sign \pm to summarize data, even though the data or underlying principles may suggest skewed distributions.

Log-normal distributions are usually characterized in terms of the log-transformed variable, using as parameters the expected value, or mean, of its distribution and its standard deviation. This characterization can be advantageous as, by definition, log-normal distributions are symmetrical again at the log level. Unfortunately, the widespread aversion to statistics becomes even more pronounced as soon as logarithms are involved. This may be the major reason that the log-normal distributions are so little understood in general, which leads to frequent misunderstandings and errors. Plotting the data can help, but graphs are difficult to communicate orally. In short, current ways of handling log-normal distributions are often unwieldy. To get an idea of a sample, most people prefer to think in terms of the original rather than the log-transformed data. This conception is indeed feasible and advisable for log-normal data too, because the familiar properties of the normal distribution have their analogies in the log-normal distribution.

A set of observations x is said to have a log-normal distribution if $\ln(x)$ or $\log(x)$ is *normally distributed*. Here, the range of x is exclusively positive (i.e., $0 < x < \infty$) and skewed towards the left. It is also known as a **Galton distribution**. A variable might be modelled as log-normal if it can be thought of as the multiplicative product of many independent random variables each of which is positive. This distribution is **positively skewed** and **leptokurtic**.

The log-normal probability distribution is defined as

$$p_L = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \text{ [function]}$$

which has the following descriptors:

1. Mean = $e^{\mu+\sigma^2/2}$
2. Median = e^μ
3. Mode = $e^{\mu-\sigma^2}$
4. Variance = $(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$
5. Skewness = $(e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$.

Figure 29 (bottom right) shows an example log-normal distribution.

Application:

Summary:

- **Log-normal distribution:** if $\ln(x)$ or $\log(x)$ is *normally distributed*:

$$p_L = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \text{ [function]}$$

- **Poisson distribution:** Limiting case of the binomial distribution for large n and constant μ ; appropriate for describing small samples from large populations:

$$p_P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} \quad \sigma^2 = \mu$$

- **Gaussian distribution:** Limiting case of the binomial distribution for large n and finite μ ; appropriate for smooth symmetric distributions:

$$p_G(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

(FWHM) $\Gamma = 2\sqrt{2\ln 2}\sigma \approx 2.354\sigma$

1.15.3 Follow-up Questions

- What sort of noise do you expect for these?
- What extra sources of noise can you get from a CCD?
- Where does the noise come from, and what does the variance/standard deviation actually tell you?
- What property of the Gaussian distribution makes it particularly useful? (i.e., Central limit theorem.)
- How do we get a log-normal distribution?
- Why do we quote n-sigma probabilities and confidence levels with a Gaussian assumption when the actual probability distribution might not be Gaussian?
- If I'm doing a cosmological survey and I say I assume a Poisson distribution for radio galaxy sources and a Gaussian distribution for dust sources, why would I use two different distributions?

1.16 Question 16

You are trying to determine a flux from a CCD image using aperture photometry, measuring source(+sky) within a 5-pixel radius, and sky within a 20-25 pixel annulus. Assume you find 10000 electrons inside the aperture and 8100 electrons in the sky region, and that the flux calibration is good to 1%. What is the fractional precision of your measurement? (Ignore read noise.) More generally, describe how you propagate uncertainties, what assumptions you implicitly make, and how you might estimate errors if these assumptions do not hold.

1.16.1 Short answer

Answer.

1.16.2 Additional context

Signal-to-noise ratio: The equation for the S/N of a measurement made with a CCD is given by

$$\frac{S}{N} = \frac{N_*}{\sqrt{N_* + n_{\text{pix}}(N_s + N_D + N_R^2)}} \text{ [dimensionless].}$$

the “signal” term, N_* , is the total number of photons (i.e., signal) collected from the object of interest. N_* may be from one pixel (if determining the S/N of a single pixel is done for a background measurement), or N_* may be from several pixels, such as all those contained within a stellar profile (if determining the S/N for the measurements of a star), or N_* may even be from say a rectangular area of X by Y pixels (if determining the S/N in a portion of the continuum of a spectrum).

The “noise” terms in the above equation are the square roots of N_* , plus n_{pix} (the number of pixels under consideration for the S/N calculation) times the contributions from N_s (the total number of photons per pixel from the background or sky), N_D (the total number of dark current electrons per pixel), and N_R^2 (the total number of electrons per pixel resulting from the read noise. Note that this noise source is not a Poisson noise source but a **shot noise**; therefore it enters into the noise calculation as the value itself, not the square root of the value as Poisson noise sources do.

The CCD Equation above provides the formulation for a S/N calculation given typical conditions and a well-behaved CCD. For some CCD observations, particularly those that have high background levels, faint sources of interest, poor spatial sampling, or large gain values, a more complete version of the error analysis is required. We can write the complete S/N equation as

$$\frac{S}{N} = \frac{N_*}{\sqrt{N_* + n_{\text{pix}} \left(1 + \frac{n_{\text{pix}}}{n_B}\right) (N_s + N_D + N_R^2 + G^2 \sigma_f^2)}} \text{ [dimensionless].}$$

This form of the S/N equation is essentially the same as that given above, but two additional terms have been added. The first term, $(1 + n_{\text{pix}}/n_B)$, provides a measure of the noise incurred as a result of any error introduced in the estimation of the background level on the CCD image. The term n_B is the total number of background pixels used to estimate the mean background (sky) level. One can see that small values of n_B will introduce the largest error as they will provide a poor estimate of the mean level of the background distribution. Thus, very large values of n_B are to be preferred but clearly some trade-off must be made between providing a good estimate of the mean background level and the use of pixels from areas on the CCD image that are far from the source of interest or possibly of a different character. The second new term added into the complete S/N equation accounts for the error introduced by the digitization noise within the A/D converter. The error introduced by this process can be considerable if the CCD gain has a large value. In this term, $G^2 \sigma_f^2$, G is the gain of the CCD (in electrons/ADU) and σ_f is an estimate of the 1 sigma error introduced within the A/D converter and has a value of approximately 0.289. The parameter σ_f and its value depend on the actual internal electrical workings of a given A/D converter.

In practice for most CCD systems in use and for most observational projects, the two additional terms in the complete S/N equation are often very small error contributors and can be ignored. In the instances for which they become important – for example, cases in which the CCD gain has a high value (e.g., 100 electrons/ADU), the background level can only be estimated with a few pixels (e.g., less than 200), or the CCD data are of poor pixel sampling – ignoring these additional error terms will lead to an overestimation of the S/N value obtained from the CCD data.

1.17 Question 17

Suppose you measure the brightness of a star ten times (in a regime where source-noise dominates. (1) How do you calculate the mean, median, and mode and standard deviation? (2) How can you tell if any points are outliers? Say some points are outliers, what do you do now (i.e., how does this impact the calculation of the quantities in part 1)?

1.17.1 Short answer

Answer.

1.17.2 Additional context

Additional context.

1.17.3 Follow-up Questions

- How would you go about fitting a model to your data? (i.e., Bayesian approach.)
- Why is using a gaussian generally a good choice?

1.18 Question 18

Suppose you do an imaging search for binaries for a sample of 50 stars, and that you find companions in 20 cases. What binary fraction do you infer? Suppose a binary-star fraction of 50% had been found previously for another sample (which was much larger, so you can ignore its uncertainty). Determine the likelihood that your result is consistent with that fraction.

1.18.1 Short answer

Answer.

1.18.2 Additional context

Additional context.

1.19 Question 19

What are the primary wavelength bands at which searches for gravitational waves are conducted? What techniques are used to search in each band? What are the sources of gravitational waves in each band? What can we learn from detections (or non-detections)?

1.19.1 Short answer

Four gravitational-wave frequency bands are being explored experimentally:

- The high-frequency band (HF; $\nu \sim 1$ Hz to $\nu \sim 10^4$ Hz) is the domain of Earth-based gravitational wave detectors or laser interferometers and resonant mass antennas.
- The low-frequency band (LF; 1 Hz to 10^4 Hz) is the domain of detectors flown in space (Earth orbit or interplanetary orbit).
- The very-low frequency band (VLF; 10^7 Hz is the domain of millisecond pulsar timing.
- The extremely-low-frequency band (ELF; 10^{15} Hz to 10^{18} Hz) is the domain of quadrupolar anisotropies in the CMB.

1.19.2 Additional context

In 1916, the year after the final formulation of the field equations of general relativity, Albert Einstein predicted the existence of gravitational waves. He found that the linearized weak-field equations had wave solutions: transverse waves of spatial strain that travel at the speed of light, generated by time variations of the mass quadrupole moment of the source. Einstein understood that gravitational-wave amplitudes would be remarkably small; moreover, until the Chapel Hill conference in 1957 there was significant debate about the physical reality of gravitational waves.

Also in 1916, Schwarzschild published a solution for the field equations that was later understood to describe a black hole, and in 1963 Kerr generalized the solution to rotating black holes. Starting in the 1970s theoretical work led to the understanding of black hole quasi-normal modes, and in the 1990s higher-order post-Newtonian calculations preceded extensive analytical studies of relativistic two-body dynamics. These advances, together with numerical relativity breakthroughs in the past decade, have enabled modeling of binary black hole mergers and accurate predictions of their gravitational waveforms. While numerous black hole candidates have now been identified through EM observations, black hole mergers have not previously been observed.

The discovery of the binary pulsar system PSR B1913+16 by Hulse and Taylor in 1975 and subsequent observations of its energy loss by Taylor and Weisberg in 1982 demonstrated the existence of gravitational waves. This discovery, along with emerging astrophysical understanding, led to the recognition that direct observations of the amplitude and phase of gravitational waves would enable studies of additional relativistic systems and provide new tests of general relativity, especially in the dynamic strong-field regime.

Experiments to detect gravitational waves began with Weber and his resonant mass detectors in the 1960s, followed by an international network of cryogenic resonant detectors. Interferometric detectors were first suggested in the early 1960s and the 1970s. A study of the noise and performance of such detectors, and further concepts to improve them, led to proposals for long-baseline broadband laser interferometers with the potential for significantly increased sensitivity. By the early 2000s, a set of initial detectors was completed, including TAMA 300 in Japan, GEO 600 in Germany, the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the United States, and Virgo in Italy. Combinations of these detectors made joint observations from 2002 through 2011, setting upper limits on a variety of gravitational-wave sources while evolving into a global network. In 2015, Advanced LIGO became the first of a significantly more sensitive network of advanced detectors to begin observations.

A century after the fundamental predictions of Einstein and Schwarzschild, the first direct detection of gravitational waves and the first direct observation of a binary black hole system merging to form a single black hole was discovered in January of 2016. This discovery provided unique access to the properties of space-time in the strong-field, high-velocity regime and confirmed predictions of general relativity for the nonlinear dynamics of highly disturbed black holes.

There is an enormous difference between gravitational waves and the EM waves on which our present knowledge of the universe is based:

- EM waves are oscillations of the EM field that propagate through spacetime; gravitational waves are oscillations of the “fabric” of spacetime itself.

- Astronomical EM waves are almost always incoherent superpositions of emission from individual electrons, atoms, or molecules. Cosmic gravitational waves are produced by coherent, bulk motions of huge amounts of mass-energy – either material mass or the energy of vibrating, nonlinear spacetime curvature.
- Since the wavelengths of EM waves are small compared to their sources (gas clouds, stellar atmospheres, accretion disks, etc.), from the waves we can make pictures of the sources. The wavelengths of cosmic gravitational waves are comparable to or larger than their coherent, bulk-moving sources, so we cannot make pictures from them. Instead, the gravitational waves are like sound; they carry, in two independent waveforms, a stereophonic, symphony-like description of their sources.
- EM waves are easily absorbed, scattered, and dispersed by matter. Gravitational waves travel nearly unscathed through all forms and amounts of intervening matter.
- Astronomical EM waves have frequencies that begin at $\nu = 10^7$ Hz and extend upward by roughly 20 orders of magnitude. Astronomical gravitational waves should begin at $\nu \sim 10^4$ Hz (1000-fold lower than the lowest-frequency astronomical EM waves) and should extend downward from there by roughly 20 orders of magnitude.

These enormous differences make it likely that:

- The information brought to us by gravitational waves will be very different from (almost “orthogonal to”) that carried by EM waves; gravitational waves will show us details of the bulk motion of dense concentrations of energy, whereas electromagnetic waves show us the thermodynamic state of optically thin concentrations of matter.
- Most (but not all) gravitational-wave sources that our instruments detect will not be seen electromagnetically, and conversely, most objects observed electromagnetically will never be seen gravitationally. Typical EM sources are stellar atmospheres, accretion disks, and clouds of interstellar gas – none of which emit significant gravitational waves, while typical gravitational-wave sources are the cores of supernova (which are hidden from electromagnetic view by dense layers of surrounding stellar gas), and colliding black holes (which emit no EM waves at all).
- Gravitational waves may bring us great surprises. In the past, when a radically new window has been opened onto the Universe, the resulting surprises have had a profound, indeed revolutionary, impact. For example, the radio universe, as discovered in the 1940s, '50s, and '60s, turned out to be far more violent than the optical Universe; radio waves brought us quasars, pulsars, and the CMB, and with them, our first direct observational evidence for black holes, neutron stars, and the heat of the Big Bang. It is reasonable to hope that gravitational waves will bring a similar “revolution.”

Gravitational waves: The essence of GR is that mass and energy produce a curvature of four-dimensional space-time, and that matter moves in response to this curvature. The Einstein field equations prescribe the interaction between mass and spacetime curvature, much as Maxwell’s equations prescribe the relationship between electric charge and electromagnetic fields. Just as EM waves are time-dependent vacuum solutions to Maxwell’s equations, GWs are time-dependent vacuum solutions to the field equations. GWs are oscillating perturbations to a flat, or **Minkowski**, space-time metric, and can be thought of equivalently as an oscillating strain in space-time or as an oscillating tidal force between free test masses.

As with electromagnetic waves, GWs travel at the speed of light and are transverse in character, i.e. the strain oscillations occur in directions orthogonal to the direction in which the wave is propagating. Whereas EM are **dipolar** in nature, GWs are **quadrupolar**: the strain pattern contracts space along one transverse dimension, while expanding it along the orthogonal direction in the transverse plane (see Figure 30). Gravitational radiation is produced by oscillating multipole moments of the mass distribution of a system. The principle of mass conservation rules out monopole radiation, and the principles of linear and angular momentum conservation rule out gravitational dipole radiation. Quadrupole radiation is the lowest allowed form and is thus usually the dominant form. In this case, the GW field strength is proportional to the second time derivative of the quadrupole moment of the source, and it falls off in amplitude inversely with distance from the source. The tensor character of gravity (i.e., the hypothetical graviton is a spin-2 particle) means that the transverse strain field comes in two orthogonal polarizations. These are commonly expressed in a linear polarization basis as the + polarization (depicted in Figure 30) and the \times polarization, reflecting the fact that they are rotated 45° relative to one another. An astrophysical GW will, in general, be a mixture of both polarizations.

GWs differ from electromagnetic waves in that they propagate essentially unperturbed through space, as they interact only very weakly with matter. Furthermore, GWs are intrinsically non-linear, because

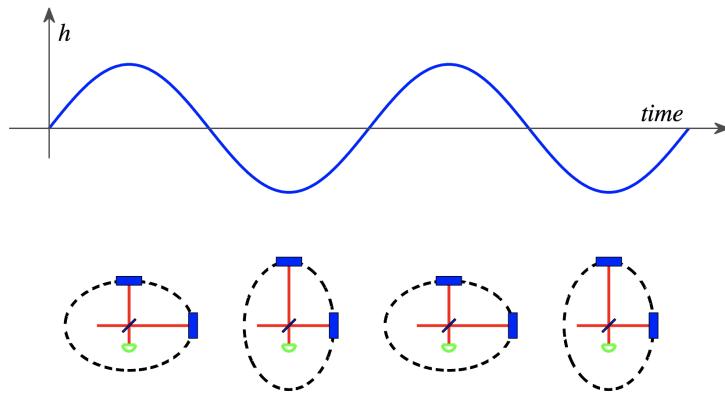


Figure 30: A GW traveling perpendicular to the plane of the diagram is characterized by a strain amplitude h . The wave distorts a ring of test particles into an ellipse, elongated in one direction in one half-cycle of the wave, and elongated in the orthogonal direction in the next half-cycle. This oscillating distortion can be measured with a Michelson interferometer oriented as shown. The length oscillations modulate the phase shifts accrued by the light in each arm, which are in turn observed as light intensity modulations at the photo-detector (green semi-circle). This depicts one of the linear polarization modes of the GW. Figure taken from Abbott et al. (2009).

the wave energy density itself generates additional curvature of space-time. This phenomenon is only significant, however, very close to strong sources of waves, where the wave amplitude is relatively large. More usually, GWs distinguish themselves from EM waves by the fact that they are very weak. One cannot hope to detect any waves of terrestrial origin, whether naturally occurring or man-made; instead one must look for very massive compact astrophysical objects, moving at relativistic velocities. For example, strong sources of GWs that may exist in our Galaxy or nearby galaxies are expected to produce wave strengths on Earth that do not exceed strain levels of one part in 10^{21} . Finally, it is important to appreciate that GW detectors respond directly to GW amplitude rather than GW power; therefore the volume of space that is probed for potential sources increases as the cube of the strain sensitivity.

Frequency bands, sources, and detection methods: Four gravitational-wave frequency bands are being explored experimentally: the high-frequency band (HF; $\nu \sim 1$ Hz to $\nu \sim 10^4$ Hz), the low-frequency band (LF; 1 Hz to 10^4 Hz), the very-low frequency band (VLF; 10^7 Hz to 10^9 Hz), and the extremely-low-frequency band (ELF; 10^{15} Hz to 10^{18} Hz).

HF band: 1 Hz to 10^4 Hz: A gravitational-wave source of mass M cannot be much smaller than its gravitational radius, $2GM/c^2$, and cannot emit strongly at periods much smaller than the light-travel time $4GM/c^3$ around this gravitational radius. Correspondingly, the frequencies at which it emits are

$$\nu \lesssim \frac{1}{4\pi GM/c^3} \sim 10^4 \frac{M_\odot}{M} [\text{Hz}],$$

where M_\odot is the mass of the Sun, and G and c are Newton's gravitation constant and the speed of light, respectively. To achieve a size of the same order as its gravitational radius and thereby emit near this maximum frequency, an object presumably must be heavier than the **Chandrasekhar limit**, about the mass of the sun, M_\odot . Thus, the highest frequency expected for strong gravitational waves is $\nu_{\max} \sim 10^4$ Hz. This defines the upper edge of the high-frequency gravitational-wave band.

The HF band is the domain of the Earth-based gravitational-wave detectors of laser interferometers and resonant mass antennas. At frequencies below about 1 Hz, Earth-based detectors face nearly insurmountable noise from:

1. fluctuating Newtonian gravity gradients (due, e.g., to the gravitational pulls of inhomogeneities in the Earth's atmosphere which move overhead with the wind); and
2. Earth vibrations (which are extremely difficult to filter out mechanically below ~ 1 Hz).

This defines the 1 Hz lower edge of the high-frequency band; to detect waves below this frequency, one must fly one's detectors in space.

A number of interesting gravitational-wave sources fall in the HF band: the stellar collapse to a neutron star or black hole in our Galaxy and distant galaxies, which sometimes triggers SNe; the rotation and vibration of neutron stars (pulsars) in our Galaxy; the coalescence of neutron-star and stellar-mass black-hole binaries ($M < 1000 M_\odot$) in distant galaxies; and possibly such sources of stochastic background as vibrating loops of cosmic string, phase transitions in the early Universe, and the Big Bang in which the Universe was born.

LF band: 10^{-4} Hz to 1 Hz: The LF band is the domain of detectors flown in space (in Earth orbit or in interplanetary orbit). The most important of these are the Doppler tracking of spacecraft via microwave signals sent from Earth to the spacecraft and there transponded back to Earth (a technique that NASA has pursued since the early 1970s), and optical tracking of spacecraft by each other (laser interferometry in space, a technique now under development for possible flight in \sim 2014 or sooner).

The 1 Hz upper edge of the LF band is defined by the gravity gradient and seismic cutoffs on Earth-based instruments; the 10^{-4} Hz lower edge is defined by expected severe difficulties at lower frequencies in isolating spacecraft from the buffeting forces of fluctuating solar radiation pressure, solar wind, and cosmic rays.

The LF band should be populated by waves from short-period binary stars in our own Galaxy (MS binaries, cataclysmic variables, white-dwarf binaries, neutron-star binaries, etc.); from white dwarfs, neutron stars, and small black holes spiraling into massive black holes ($M \sim 3 \times 10^5 M_\odot$ to $3 \times 10^7 M_\odot$) in distant galaxies; and from the inspiral and coalescence of super-massive black-hole binaries ($M \sim 100 M_\odot$ to $10^8 M_\odot$). The upper limit, $10^8 M_\odot$, on the masses of black holes that can emit in the LF band is set by

$$\nu \lesssim \frac{1}{4\pi GM/c^3} \sim 10^4 \frac{M_\odot}{M} [\text{Hz}],$$

with $\nu > 10^4$ Hz. There should also be a LF stochastic background from such early-Universe processes as vibrating cosmic strings, phase transitions, and the Big Bang itself.

VLF band: 10^{-9} Hz to 10^{-7} Hz: Joseph Taylor and others have achieved a remarkable gravity-wave sensitivity in the VLF band by the timing of millisecond pulsars. When a gravitational wave passes over the Earth, it perturbs our rate of flow of time and hence the ticking rates of our clocks relative to clocks outside the wave. Such perturbations will show up as apparent fluctuations in the times of arrival of the pulsar's pulses. If no fluctuations are seen at some level, we can be rather sure that neither Earth nor the pulsar is being bathed by gravitational waves of the corresponding strength. If fluctuations with the same time evolution are seen simultaneously in the timing of several different pulsars, then the cause could well be gravitational waves bathing the Earth.

By averaging the pulses' times of arrival over long periods of time (months to tens of years), a very high timing precision can be achieved, and correspondingly, tight limits can be placed on the waves bathing the Earth or the pulsar. The upper edge of the VLF band, $\sim 10^7$ Hz, is set by the averaging time of a few months needed to build up high accuracy; the lower edge, $\sim 10^9$ Hz, is set by the time of ~ 20 years, since very steady millisecond pulsars were first discovered.

Strong gravitational-wave sources are generally compact, not much larger than their own gravitational radii. The only compact bodies that can radiate in the VLF band or below (i.e., at $\nu < 10^7$ Hz) are those with $M > 10^{11} M_\odot$. Conventional astronomical wisdom suggests that compact bodies this massive do not exist, and that therefore, the only strong waves in the VLF band and below are a stochastic background produced by the same early-Universe processes as might radiate at low and high frequencies: cosmic strings, phase transitions, and the Big Bang.

Of course, conventional wisdom could be wrong. Nevertheless, it is conventional to quote measurement accuracies in the VLF band and below in the language of a stochastic background: the fraction $\Omega_g(\nu)$ of the energy required to close the Universe that lies in a bandwidth $\Delta\nu = \nu$ centered on frequency ν . The current 95% confidence limit on $\Omega_g(\nu)$ from pulsar timing in the VLF band is $\Omega_g(4 \times 10^9 \text{ Hz}) < 6 \times 10^{-8} H^{-2}$, where H is the Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. This is a sufficiently tight limit that is beginning to cast doubt on the (not terribly popular) suggestion that the Universe contains enough vibrating loops of cosmic string for their gravitational pulls to have seeded galaxy formation.

ELF band: 10^{-18} Hz to 10^{-15} Hz: Gravitational waves in the ELF band should produce anisotropies in the CMB. The tightest limit from microwave observations comes from the lower edge of the ELF band $\nu \sim 10^{-18}$ Hz, where the gravitational wavelength is about π times the Hubble distance, and the waves, by squeezing all of the space inside our cosmological horizon in one direction and stretching it all in another, should produce a quadrupolar anisotropy in the CMB. The quadrupolar anisotropy measured by the COBE satellite, if due primarily to gravitational waves (which it could be), corresponds to an energy density $\Omega_g(10^{-18} \text{ Hz}) \sim 10^{-9}$.

Wave polarizations, waveforms, and how an interferometer works: According to general relativity theory (which we shall assume to be correct), a gravitational wave has two linear polarizations, conventionally called + (plus) and \times (cross). Associated with each polarization, there is a gravitational wave field, hP_+ or $h\times$, which oscillates in time and propagates with the speed of light. Each wave field produces tidal forces (stretching and squeezing forces) on any object or detector through which it passes. If the object is small compared to the waves' wavelength (as is the case for ground-based interferometers and resonant mass antennas), then relative to the object's center, the forces have the quadrupolar patterns shown in Figure 31. The names "plus" and "cross" are derived from the orientations of the axes

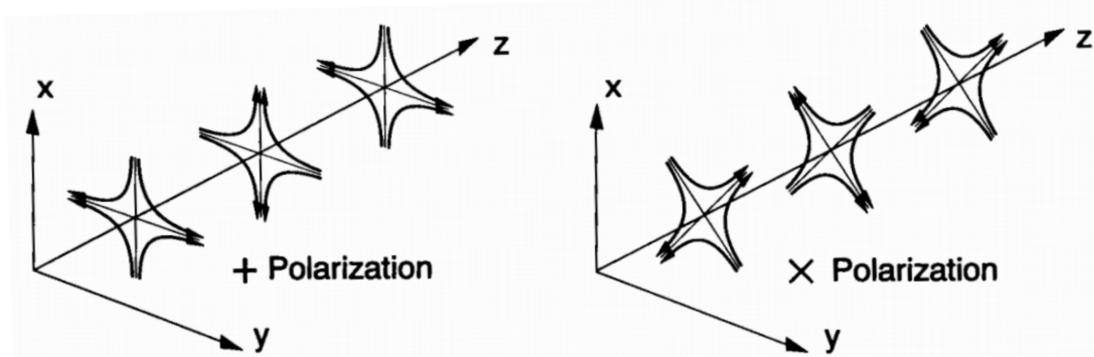


Figure 31: The lines of force associated with the two polarizations of a gravitational wave. Figure taken from Thorne (1994).

that characterize the force patterns.

A laser interferometer gravitational wave detector (**interferometer** for short) consists of four masses that hang from vibration-isolated supports as shown in Figure 32, and the indicated optical system for monitoring the separations between the masses. Two masses are near each other, at the corner of an “L,” and one mass is at the end of each of the L’s long arms. The arm lengths are nearly equal, $L_1 \simeq L_2 = L$. When a gravitational wave, with high frequencies compared to the masses’ ~ 1 Hz pendulum frequency, passes through the detector, it pushes the masses back and forth relative to each other as though they were free from their suspension wires, thereby changing the arm-length difference, $\Delta L \equiv L_1 - L_2$. That change is monitored by laser interferometry in such a way that the variations in the output of the photodiode (the interferometer’s output) are directly proportional to $\Delta L(t)$.

If the waves are coming from overhead or underfoot, and the axes of the + polarization coincide with the arms’ directions, then it is the wave’s + polarization that drives the masses, and $\Delta L(t)/L = h_+(t)$. More generally, the interferometer’s output is a linear combination of the two wave fields:

$$\frac{\Delta L}{L} = F_+ h_+(t) + F_\times h_\times(t) \equiv h(t) \text{ [dimensionless].}$$

The coefficients F_+ and F_\times are of order unity and depend in a quadrupolar manner on the direction to the source and the orientation of the detector. The combination $h(t)$ of the two h ’s is called the **gravitational-wave strain** that acts on the detector; and the time evolutions of $h_+(t)$, $h_\times(t)$, and $h(t)$ are sometimes called **waveforms**.

Interferometer test masses at present are made of transparent fused silica, though other materials might be used in the future. The masses’ inner faces (shown white in Figure 32) are covered with high-reflectivity dielectric coatings to form the indicated “mirrors,” while the masses’ outer faces are covered with anti-reflection coatings. The two mirrors facing each other on each arm form a **Fabry-Perot** cavity. A **beam splitter** splits a carefully prepared laser beam in two and directs the resulting beams down the two arms. Each beam penetrates through the anti-reflection coating of its arm’s corner mass, through the mass, and through the dielectric coating (the mirror); and thereby – with the length of the arm’s Fabry-Perot cavity adjusted to be nearly an integral number of half wavelengths of light – the beam gets trapped in the cavity. The cavity’s end mirror has much higher reflectivity than its corner mirror, so the trapped light leaks back out through the corner mirror and then hits the beam splitter where it recombines with light from the other arm. Most of the recombined light goes back toward the laser (where it can be returned to the interferometer by a **light-recycling mirror** labeled R), but a tiny portion goes toward the photodiode.

When a gravitational wave hits the detector and moves the masses, thereby changing the lengths L_1 and L_2 of the two cavities, it shifts each cavity’s resonant frequency slightly relative to the laser frequency, and thereby changes the phase of the light in the cavity and the phase of the light that exits from the cavity toward the beam splitter. Correspondingly, the relative phase of the two beams returning to the splitter is altered by an amount $\Delta\Phi \propto \Delta L$, and this relative phase shift causes a change in the intensity of the recombined light at the photodiode, $I_{pd} \propto \Delta L \propto h(t)$. Thus, the change of photodiode output current is directly proportional to the gravitational-wave strain $h(t)$. This method of monitoring $h(t)$, which was invented by Ronald Drever as a modification of Rainer Weiss’s seminal concept for such an interferometer, is capable of very high sensitivity.

Wave strengths and interferometer arm lengths: The strengths of the waves from a gravitational-wave source can be estimated using the “Newtonian quadrupole” approximation to the Einstein field

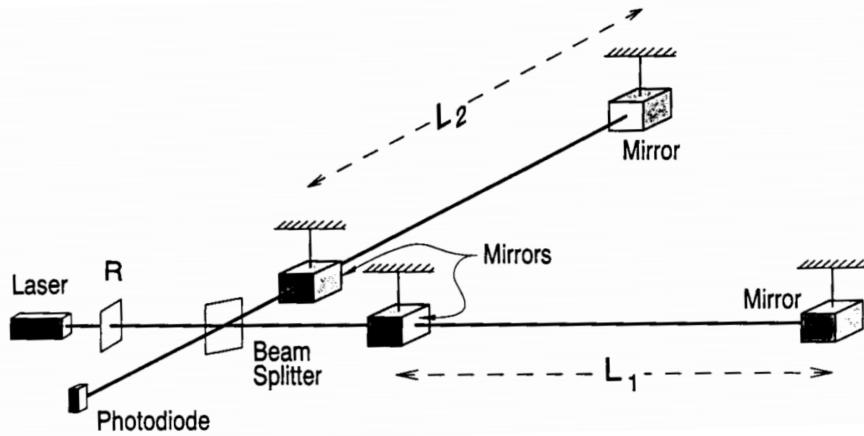


Figure 32: Schematic diagram of a laser interferometer gravitational wave detector. Figure taken from Thorne (1994).

equations. This approximation says that $h \simeq (G/c^4)\ddot{Q}/r$, where \ddot{Q} is the second time derivative of the source's quadrupole moment, r is the distance of the source from Earth, and G and c are Newton's gravitation constant and the speed of light, respectively. The strongest sources will be highly non-spherical, and thus, will have $Q \simeq ML^2$ where M is their mass and L their size, and correspondingly, will have $\ddot{Q} \simeq 2Mv^2 \simeq 4E_{\text{kin}}^{\text{ns}}$ where v is their internal velocity and $E_{\text{kin}}^{\text{ns}}$ is the non-spherical part of their internal kinetic energy. This provides us with the estimate

$$h \sim \frac{1}{c^2} \frac{4G(E_{\text{kin}}^{\text{ns}}/c^2)}{r} \text{ [dimensionless]},$$

that is, h is about four times the gravitational potential produced at Earth by the mass-equivalent of the source's non-spherical, internal kinetic energy, made dimensionless by dividing by c^2 . Thus, in order to radiate strongly, the source must have a very large, non-spherical internal kinetic energy.

The best known way to achieve a huge internal kinetic energy is via gravity; and by energy conservation (or the virial theorem), any gravitationally-induced kinetic energy must be of the order of the source's gravitational potential energy. A huge potential energy, in turn, requires that the source be very compact, not much larger than its own gravitational radius. Thus, the strongest gravity-wave sources must be highly compact, dynamical concentrations of large amounts of mass (e.g., colliding and coalescing black holes and neutron stars).

Such sources cannot remain highly dynamical for long; their motions will be stopped by energy loss to gravitational waves and/or the formation of an all-encompassing black hole. Thus, the strongest sources should be transient. More-over, they should be very rare; so rare that to see a reasonable event rate will require reaching out through a substantial fraction of the Universe. Thus, just as the strongest radio waves arriving at Earth tend to be extragalactic, so also the strongest gravitational waves are likely to be extragalactic.

For highly compact, dynamical objects that radiate in the high-frequency band (e.g., colliding and coalescing neutron stars and stellar-mass black holes), the internal, non-spherical kinetic energy $E_{\text{kin}}^{\text{ns}}/c^2$ is of the order of the mass of the Sun; and, correspondingly, $h \sim 10^{-22}$ for such sources at the Hubble distance (3000 Mpc, i.e., 10^{10} light years); $h \sim 10^{-21}$ at 200 Mpc (a best-guess distance for several neutron-star coalescences per year), $h \sim 10^{-20}$ at the Virgo cluster of galaxies (15 Mpc); and $h \sim 10^{-17}$ in the outer reaches of our own Milky Way galaxy (20 kpc). These numbers set the scale of sensitivities that ground-based interferometers seek to achieve: $h \sim 10^{-21}$ to 10^{-22} .

When one examines the technology of laser interferometry, one sees good prospects to achieve measurement accuracies $\Delta L \sim 10^{-16}$ cm (1/1000 the diameter of the nucleus of an atom). With such an accuracy, an interferometer must have an arm length $L = \Delta L/h \sim 1$ to 10 km in order to achieve the desired wave sensitivities of 10^{-21} to 10^{-22} .

LIGO and the worldwide detector network: As illustrated in Figure 30, the oscillating quadrupolar strain pattern of a GW is well matched by a Michelson interferometer, which makes a very sensitive comparison of the lengths of its two orthogonal arms. LIGO utilizes three specialized Michelson interferometers, located at two sites (see Figure 33): an observatory on the Hanford site in Washington houses two interferometers, the 4 km-long H1 and 2 km-long H2 detectors; and an observatory in Livingston Parish, Louisiana, houses the 4 km-long L1 detector. Other than the shorter length of H2, the three interferometers are essentially identical. Multiple detectors at separated sites are crucial for rejecting instrumental and environmental artifacts in the data, by requiring coincident detections in the analy-



Figure 33: Aerial photograph of the LIGO observatories at Hanford, Washington (**left**) and Livingston, Louisiana (**right**). The lasers and optics are contained in the large corner buildings. From each corner building, evacuated beam tubes extend at right angles for 4 km in each direction (the full length of only one of the arms is seen in each photo); the tubes are covered by the arched, concrete enclosures seen here. Figure taken from Abbott (2009).

sis. Also, because the antenna pattern of an interferometer is quite wide, source localization requires triangulation using three separated detectors.

The initial LIGO detectors were designed to be sensitive to GWs in the frequency band $40 - 7000$ Hz, and capable of detecting a GW strain amplitude as small as 10^{-21} . With funding from the National Science Foundation, the LIGO sites and detectors were designed by scientists and engineers from the California Institute of Technology and the Massachusetts Institute of Technology, constructed in the late 1990s, and commissioned over the first 5 years of this decade. From November 2005 to September 2007, they operated at their design sensitivity in a continuous data-taking mode. The data from this science run, known as S5, were analyzed for a variety of GW signals by a group of researchers known as the LIGO Scientific Collaboration. At the most sensitive frequencies, the instrument rms strain noise has reached an unprecedented level of 3×10^{-22} in a 100 Hz band.

Although in principle LIGO can detect and study GWs by itself, the potential to do astrophysics can be quantitatively and qualitatively enhanced by operation in a more extensive network. For example, the direction of travel of the GWs and the complete polarization information carried by the waves can only be extracted by a network of detectors. Such a global network of GW observatories has been emerging over the past decade. In this period, the Japanese TAMA project built a 300 m interferometer outside Tokyo, Japan; the German-British GEO project built a 600 m interferometer near Hanover, Germany; and the European Gravitational Observatory built the 3 km-long interferometer Virgo near Pisa, Italy. In addition, plans are underway to develop a large scale GW detector in Japan sometime during the next decade.

Early in its operation LIGO joined with the GEO project; for strong sources the shorter, less sensitive GEO 600 detector provides added confidence and directional and polarization information. In May 2007 the Virgo detector began joint observations with LIGO, with a strain sensitivity close to that of LIGO's 4 $rmkm$ interferometers at frequencies above ~ 1 kHz. The LIGO Scientific Collaboration and the Virgo Collaboration negotiated an agreement that all data collected from that date would be analyzed and published jointly.

LIGO instrument performance and noise: The challenge is to make the instrument sufficiently sensitive: at the targeted strain sensitivity of 10^{-21} , the resulting arm length change is only $\sim 10^{-18}$ m, a thousand times smaller than the diameter of a proton. Meeting this challenge involves the use of special interferometry techniques, state-of-the-art optics, highly stable lasers and multiple layers of vibration isolation. And of course a key feature of the detectors is simply their scale: the arms are made as long as practically possible to increase the signal due to a GW strain.

The response of the interferometer output is a function of GW frequency and is derived in length in other papers. In the long-wavelength approximation, where the wavelength of the GW is much longer than the size of the detector, the response R of a Michelson-Fabry-Perot interferometer is approximated by a single-pole transfer function:

$$R(\nu) \propto \frac{1}{1 + i\nu/\nu_p},$$

where the pole frequency ν_p is related to the arm cavity storage time by $\nu_p = 1/4\pi\tau_s$. Above the pole frequency ($\nu_p = 85$ Hz for the LIGO 4 km interferometers), the amplitude response drops off as $1/\nu$. The measurement noise above the pole frequency has a white (flat) spectrum, and so the strain sensitivity decreases proportionally to frequency in this region. The single-pole approximation is quite accurate,

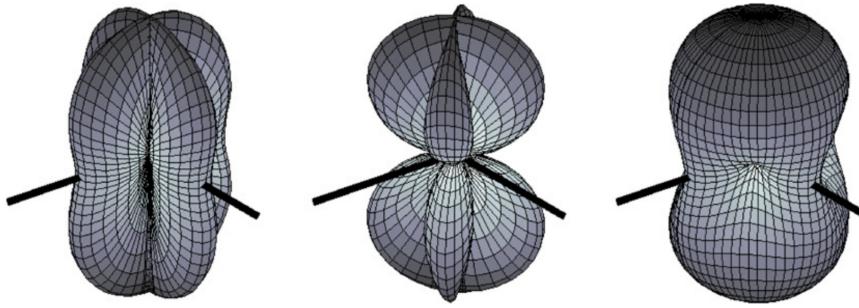


Figure 34: Antenna response pattern for a LIGO GW detector, in the long-wavelength approximation. The interferometer beamsplitter is located at the center of each pattern, and the thick black lines indicate the orientation of the interferometer arms. The distance from a point of the plot surface to the center of the pattern is a measure of the GW sensitivity in this direction. The pattern on the left is for $+$ polarization, the middle pattern is for \times polarization, and the right-most one is for unpolarized waves. Figure taken from Abbott (2009).

differing from the exact response by less than a percent up to ~ 1 kHz.

In the long-wavelength approximation, the interferometer directional response is maximal for GWs propagating orthogonally to the plane of the interferometer arms and linearly polarized along the arms. Other angles of incidence or polarizations give a reduced response, as depicted by the antenna patterns shown in Figure 34. A single detector has blind spots on the sky for linearly polarized GWs.

To complete the LIGO detector, the interferometers were supplemented with a set of sensors to monitor the local environment. Seismometers and accelerometers measure vibrations of the ground and various interferometer components; microphones monitor acoustic noise at critical locations; magnetometers monitor fields that could couple to the test masses or electronics; radio receivers monitor RF power around the modulation frequencies. These sensors are used to detect environmental disturbances that can couple to the GW channel.

Since the interferometers detect GW strain amplitude, their performance is typically characterized by an amplitude spectral density of detector noise (the square root of the power spectrum), expressed in equivalent GW strain. Typical high-sensitivity strain noise spectra are shown in Figure 35. Over the course of S5 the strain sensitivity of each interferometer was improved, by up to 40% compared with the beginning of the run through a series of incremental improvements to the instruments.

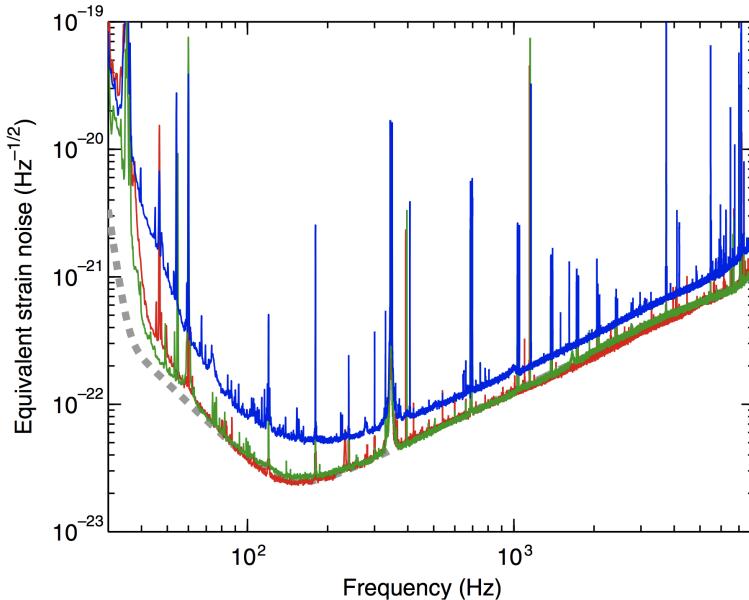


Figure 35:
Strain sensitivities, expressed as amplitude spectral densities of detector noise converted to equivalent GW strain. The vertical axis denotes the rms strain noise in 1 Hz of bandwidth. Shown are typical high sensitivity spectra for each of the three interferometers (red: H1; blue: H2; green: L1), along with the design goal for the 4 km detectors (dashed gray). Figure taken from Abbott et al (2009).

The primary noise sources contributing to the H1 strain noise spectrum are shown in Figure 35. Understanding and controlling these instrumental noise components has been the major technical challenge in the development of the detectors. The noise terms can be broadly divided into two classes: displacement noise and sensing noise. **Displacement noises** cause motions of the test masses or their mirrored surfaces. **Sensing noises**, on the other hand, are phenomena that limit the ability to measure those motions; they are present even in the absence of test mass motion. The strain noises shown in Figure 35 consist of spectral lines superimposed on a continuous broadband noise spectrum. The majority of the

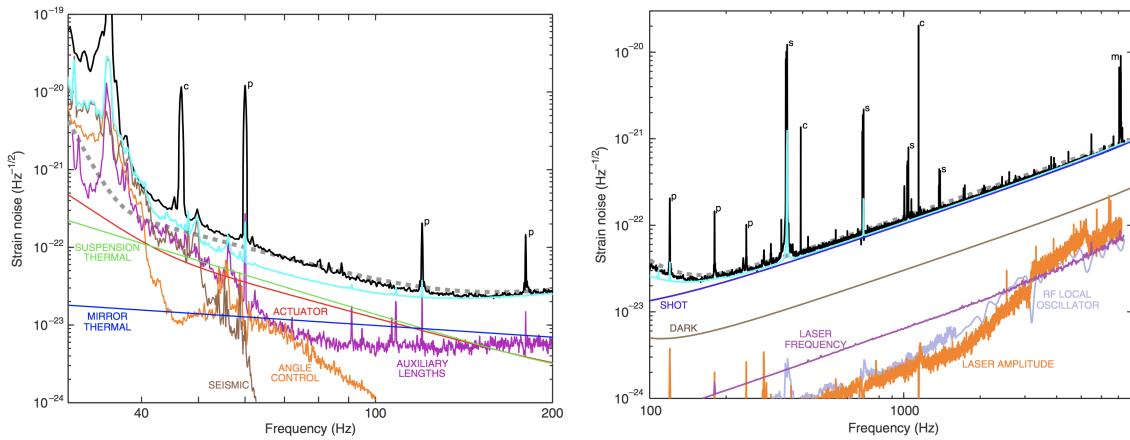


Figure 36: Primary known contributors to the H1 detector noise spectrum. The left panel shows the displacement noise components, while the right panel shows sensing noises (note the different frequency scales). In both panels, the black curve is the measured strain noise (same spectrum as in Figure 35), the dashed gray curve is the design goal and the cyan curve is the root-square-sum of all known contributors (both sensing and displacement noises). The known noise sources explain the observed noise very well at frequencies above 150 Hz, and to within a factor of 2 in the 40 – 100 Hz band. Spectral peaks are identified as follows: c, calibration line; p, power line harmonic; s, suspension wire vibrational mode; m, mirror (test mass) vibrational mode. Figure taken from Abbott (2009).

lines are due to power lines (60, 120, 180 Hz, etc.), “violin mode” mechanical resonances (350, 700 Hz, etc) and calibration lines (55, 400 and 1100 Hz). These narrow noise lines are easily excluded from analysis; the broadband noise dominates the instrument sensitivity.

The left panel of Figure 36 shows the displacement noise. At the lowest frequencies the largest such noise is **seismic noise** – motions of the Earth’s surface driven by wind, ocean waves, human activity and low-level earthquakes; filtered by the isolation stacks and pendulums. The seismic contribution is estimated using accelerometers to measure the vibration at the isolation stack support points, and propagating this motion to the test masses using modeled transfer functions of the stack and pendulum. The seismic wall frequency, below which seismic noise dominates, is approximately 45 Hz, a bit higher than the goal of 40 Hz, as the actual environmental vibrations around these frequencies are ~ 10 times higher than what was estimated in the design.

Mechanical thermal noise is a more fundamental effect, arising from finite losses present in all mechanical systems, and is governed by the **fluctuation-dissipation theorem**. It causes arm length noise through thermal excitation of the test mass pendulums (**suspension thermal noise**), and thermal acoustic waves that perturb the test mass mirror surface (**test mass thermal noise**). Most of the thermal energy is concentrated at the resonant frequencies, which are designed (as much as possible) to be outside the detection band. Away from the resonances, the level of thermal motion is proportional to the mechanical dissipation associated with the motion. Designing the mirror and its pendulum to have very low mechanical dissipation reduces the detection-band thermal noise. It is difficult, however, to accurately and unambiguously establish the level of broadband thermal noise in situ; instead, the thermal noise curves in Figure 36 are calculated from models of the suspension and test masses, with mechanical loss parameters taken from independent characterizations of the materials.

The right panel of Figure 36 shows the sensing noise. By design, the dominant noise source above 100 Hz is **shot noise**, as determined by the Poisson statistics of photon detection. The ideal shot-noise limited strain noise density, $\bar{h}(f)$, for this type of interferometer is

$$\bar{h}(f) = \sqrt{\frac{\pi\hbar\lambda}{\eta P_{\text{BS}}c}} \frac{\sqrt{1 + (4\pi\nu\tau_s)^2}}{4\pi\tau_s} \quad [\text{dimensionless}],$$

where λ is the laser wavelength, \hbar is the reduced Planck constant, c is the speed of light, τ_s is the arm cavity storage time, ν is the GW frequency, P_{BS} is the power incident on the beamsplitter, and η is the photo-detector quantum efficiency. For the estimated effective power of $\eta P_{\text{BS}} = 0.9 \times 250 \text{ W}$, the ideal shot-noise limit is $\tilde{h} = 1.0 \times 10^{-23} \text{ Hz}^{-1/2}$ at 100 Hz. The shot-noise estimate in Figure 36 is based on measured photo-currents in the AS port detectors and the measured interferometer response. The resulting estimate, $\tilde{h}(100 \text{ Hz}) = 1.3 \times 10^{-23} \text{ Hz}^{-1/2}$, is higher than the ideal limit due to several inefficiencies in the heterodyne detection process.

1.20 Question 20

Self-similarity is a useful idealization of many astrophysical systems. Explain what self-similarity means, when it works, and why it is so useful, and provide two examples from any field.

1.20.1 Short answer

Answer.

1.20.2 Additional context

Additional context.

1.21 Question 21

Explain why diffraction-limited detectors tend to have sidelobes, and how sidelobes can be suppressed in optical and radio observations.

1.21.1 Short answer

Answer.

1.21.2 Additional context

Additional context.

1.22 Resources

- Astronomical Statistics, Taylor (2004)
- Statistical Methods for Astronomical Data Analysis, Chattopadhyay & Chattopadhyay (2014)
- Data Reduction and Error Analysis for the Physical Sciences, Bevington & Robinson (2003)
- Bayesian Logical Data Analysis for the Physical Sciences, Gregory (2005)
- Magnetic Fields in Diffuse Media, Lazarian (2015)
- Optics, Hecht (2002)
- Astronomical Optics, Schroeder (1987)
- Handbook of CCD Astronomy, Howell (2006)
- Gravitational Waves, Thorne (1994)
- Gravitational Bending of Light, Edwards (2007)
- Gravitational Lensing, Abdo
- Principles of Interferometry, Jackson (2008)
- Astrophysics of the Interstellar Medium; Maciel (2013)
- Radio Polarimetry as a Probe of Interstellar Magnetism; Newton-McGee (2009)
- Synchrotron Radiation; Mobilio et al. (2015)
- Cosmic Magnetobremsstrahlung (Synchrotron Radiation); Ginzburg & Syrovatskii (1965)
- The Evolution of Neutrino Astronomy; Bahcall & Davies (2000)
- Observation of Gravitational Waves from a Binary Black Hole Merger; Abbott et al. (2016)
- LIGO: the Laser Interferometer Gravitational-Wave Observatory; Abbott et al. (2009)
- Self-similar Collapse of Isothermal Spheres and Star Formation; Shu (1977)
- Beyond the Diffraction Limit via Optical Amplification; Kellerer & Ribak (2016)
- Log-Normal Distributions Across the Sciences: Keys and Clues; Limpert et al. (2001)
- Microlensing; Encyclopedia of Astronomy and Astrophysics (2001)
- Interferometry and Synthesis in Radio Astronomy; Thompson et al. (2017)
- Essential Radio Astronomy; Condon & Ransom (2016)
- Principles of Interferometry; Jackson