1 Astro notes 2018/08/22 - Wed - Celestial Mechanics 1: fundamental equations, limitations

1.1 Intro to course

Topics schedule on syllabus

homework: about 5 homeworks, 2 per each of 3 parts. Talk to each other, but write up individually. Your goal is to communicate to me (or the grader) that you understand the exercise. Less about getting the correct answer than communicating with the grader. Solutions will be handed out soon after due date, so if you have an excuse please let me know right away.

exams: 2 during semester (in-class + take-home) plus final exam

in-class: Will call randomly in class for some exercises and for some homework discussion before HW is due

1.2 Context

The orbit problem formed the basis of the first physics, and arguably is the birth of modern physics, and certainly mechanics.

There are basically 4 stages in the development of our understanding of the orbit problem:

Not physics:

0. Earth-centered motion - clearly the Sun moves across the sky - no, only apparently
1. Kepler's empirical laws - sun-centered - no, center of mass is real center

Physics:

2. Newton's laws - motion about center of mass - no, but we will use this
3. Einstein's laws (General relativity) - orbits don't actually close, light bends despite having no mass, simultaneity isn't well defined, information propogates at finite speed, etc - maybe this is correct (so far it is)

1.3 Fundamental equations for Newton's theory

We wish to consider how the positions of two particles change in time. The first thing we do is construct frame of reference (a rectilinear 3D coordinate system, and a time coordinate):

$$\vec{r}_1(t), \vec{r}_2(t)$$

Newtonian laws of motion and gravity tell us how these behave. The first thing to do is to define an instantaneous state of the objects' motion:

$$\vec{v}_1 \equiv \frac{d\vec{r}_1}{dt}, \quad \vec{v}_2 = \frac{d\vec{r}_2}{dt}$$

Then we introduce the concept of a force and the definition of mass such that

$$\vec{F} = \frac{dm_i v_i}{dt}$$

But Newton said that this is actually nonsense by itself. Forces always come in pairs. In this case there is a force on each object directed towards the other of common magnitude Gm_1m_2/d^2 where d is the separation distance. So we must write two equations to specify the motion:
(ask student)

$$\frac{dm_1v_1}{dt} = \frac{Gm_1m_2}{|\vec{r}_2 - \vec{r}_1|^2}\frac{\vec{r}_2 - \vec{r}_1}{\sqrt{|\vec{r}_2 - \vec{r}_1|^2}}$$

$$\frac{dm_2v_2}{dt} = \frac{Gm_1m_2}{|\vec{r}_2 - \vec{r}_1|^2}\frac{\vec{r}_1 - \vec{r}_2}{\sqrt{|\vec{r}_2 - \vec{r}_1|^2}}$$

At this point we're basically done. These are the equations of motion of the system.

(another student):
Identify: how many functions of time, what are they? how many equations, and what are they? what are initial conditions?

We have 12 time-dependent unknown functions, the components of $\vec{r}_1$, $\vec{r}_2$, v_1, and v_2, and 12 equations, (6 from the definition of $\vec{v}_i$ and 6 from Newton's laws) we just need to set the initial values for the $\vec{r}_i$ and $\vec{v}_i$ and integrate.

Noting some unspoken assumptions...
- Information travels at infinite speed. The position of each object is known when evaluating the force on the other, even though they are separated in space.
- Nominally any reference frame moving at constant velocity is equivalently good.
- Time and space are clearly separated and all simultaneous events are well defined and don't depend on coordinate choice.
All these are actually wrong.

Next time will make equations more convenient...

2 Astro notes 2018/08/24 - Fri - Celestial Mechanics 2: simplifying to reduced mass

Noting some unspoken assumptions...
- Information travels at infinite speed. The position of each object is known when evaluating the force on the other, even though they are separated in space.
- Nominally any reference frame moving at constant velocity is equivalently good.
- Time and space are clearly separated and all simultaneous events are well defined and don't depend on coordinate choice.
All these are actually wrong.

2.1 Adding some convenience

But these equations aren't in a very convenient form. Let's start by first of all changing our coordinate system so that the overall system is at rest. This has a good physical basis because, since there are no "outside" forces, this frame will not change in time. Summing the accelerations we see that they cancel out so that

$$\frac{d}{dt}(m_1 v_1 + m_2 v_2) = 0$$

Thus if we define a place

$$\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}, \quad \frac{d\vec{R}}{dt} = constant = \vec{V}$$

So if we find uniform motion uninteresting, we can subtract it out, and define new coordinates. Doing so gives exactly the same equations except that now

$$\vec{R} = 0 \implies \vec{V} = 0 \implies \vec{r}_1 = -\frac{m_2}{m_1}\vec{r}_2$$

and these can also be expressed easily in terms of the difference vector

$$\vec{r} = \vec{r}_2 - \vec{r}_1 = (1 + \frac{m_2}{m_1})\vec{r}_2 = \frac{m_1 + m_2}{m_1}\vec{r}_2$$

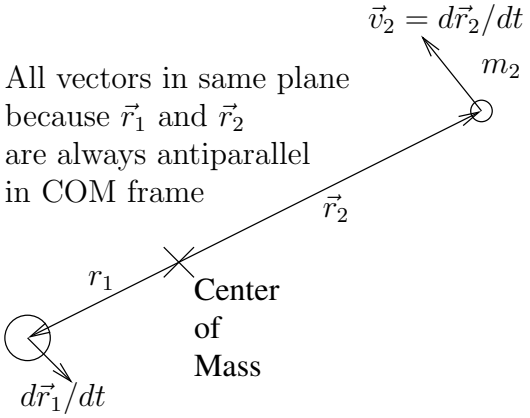
so that, with $M = m_1 + m_2$,

$$\vec{r} = \frac{M}{m_1}\vec{r}_2 = -\frac{M}{m_2}\vec{r}_1$$

(Make sure you can invert this process!) Then the the dynamical equation for body 1 is

$$\frac{dm_1\vec{v}_1}{dt} = \frac{Gm_1m_2}{r^2}(\widehat{\vec{r}_2 - \vec{r}_1}) \quad \implies \quad \frac{dm_1\vec{v}_1}{dt} = -\frac{Gm_1m_2^3}{M^2r_1^2}\hat{r}_1$$

Which, notably, now only involves $\vec{r}_1$. we'll come back to the constants, just consider the directions for a moment.



Since $\vec{r}_1$ and $\vec{r}_2$ are antiparallel to each other and $\vec{v}_1$ and $\vec{v}_2$ are always antiparallel, there are effectively only two vectors involved, and all the vectors are in the plane formed by, for example, $\vec{r}_1$ and $\vec{v}_1$.

Knowing the relative direction of the change in $\vec{v}_1$ allows us to further constrain the continuing *motion* of our problem to a plane.

Since $\vec{r}$ (the direction of the force) and $\vec{r}_1$ are parallel, the angular momentum is conserved since

$$\vec{L}_1 = \vec{r}_1 \times m_1 \vec{v}_1, \quad \frac{dL_1}{dt} = \frac{d\vec{r}_1}{dt} \times m_1 \vec{v}_1 + \vec{r}_1 \times m_1 \frac{d\vec{v}_1}{dt} \propto \vec{v}_1 \times \vec{v}_1 + \vec{r}_1 \times \hat{r}_1 = 0$$

Note also that, since $\vec{r}_1$ and $\vec{v}_1$ are both therefore perpendicular to $\vec{L}_1$, all the motion lies in the plane defined by them. Note also that since $\vec{r}_2 \propto -\vec{r}_1$ this is the same plane for body 2. The total angular momentum works out to be

$$L = \mu r v_\theta.$$

At this point things get even nicer if we switch to the reduced mass problem

$$r_1 = -\frac{m_2}{M}\vec{r}, \quad \vec{r}_2 = \frac{m_1}{M}\vec{r}$$

And the equations of motion are both now

$$\frac{d\mu\vec{v}}{dt} = -\frac{GM\mu}{r^2}\hat{r}$$

where $\mu = m_1 m_2 / M$.

2.2 "Kepler's Laws"

By forming the product L^2, which we know is constant, in a particular way, we can get an equation for the orbit. We know that L is constant, so consider

$$L^2 = \mu(\vec{r} \times \vec{v}) \cdot \vec{L} = \mu \left[\vec{r} \cdot (\vec{v} \times \vec{L})\right]$$

It turns out that from the vectors in the EOM it is not too hard to show

$$\frac{d}{dt}(\vec{v} \times \vec{L}) = \frac{d}{dt}(GM\mu\hat{r})$$

and as a result

$$\vec{v} \times \vec{L} = GM\mu\hat{r} + \vec{D}$$

where $\vec{D}$ is an integration constant. This vector is directed toward perihelion (point of closest approach) and related to the eccentricity.

We now have

$$\frac{L^2}{\mu} = \vec{r} \cdot (GM\mu\hat{r} + \vec{D}) = GM\mu r + rD\cos\theta$$

defining $e = D/GM\mu$ and solving for r this becomes

$$r = \frac{L^2/\mu^2}{GM(1 + e\cos\theta)}$$

which is the equation of an ellipse with θ being the angle from perihelion.

Note that this is not the orbit, since this must be multiplied by some factor to get that the orbit of either the more or less massive body.

Getting the equal area law is significantly easier. The swept out area is just $\int_0^r r dr d\theta = \frac{1}{2}r^2 d\theta$ so that

$$\frac{dA}{dt} = \frac{1}{2}r^2\dot{\theta} = \frac{1}{2}\frac{L}{\mu}$$

since $L = \mu r v_\theta$ and $v_\theta = r\dot{\theta}$.

Integrating that and comparing the the area of an ellipse gives

$$P^2 = \frac{4\pi^2}{GM}a^3$$

3 Astro notes 2018/08/27 - Mon - Light1: Parallax, Basic Flux, Magnitude

First a couple of things to finish up the 2-body problem...

3.1 Counting Parameters

It is useful to count independent parameters, since there are several forms of them floating around. We started with 12 initial conditions ($\vec{r}_1$, $\vec{r}_2$, $\vec{v}_1$ and $\vec{v}_2$). The position and uniform motion of the center of mass, which we removed, used 6 of these. Two parameters set the direction of the normal to the orbital plane, and another sets its orientation in space (the direction of perihelion, for example). This leaves 3 parameters, along with the masses, to describe the nature of the two orbits. One parameter describes the current state (where in the orbit the bodies are). The masses set the relative scale of the orbits of the two bodies as well as the overall scale. So there are two parameters, along with the masses, that set the shape of the orbits. There are several combinations of parameters that can be used:
- a, e for one orbit or the reduced mass problem
- a, b for one orbit or the reduced mass problem
- E, L

Specifying any two of these, along with the masses, sets the shape and size of the orbits and their period.

3.2 Virial Theorem

With these expressions for orbit shape, it can be shown that the orbital energy (which is consant) is

$$E = -\frac{Gm_1m_2}{2a}$$

This is a form of the virial theorem,

$$E = \frac{1}{2}\langle U\rangle \quad or \quad -2\langle K\rangle = \langle U\rangle$$

where K is the kinetic energy and U is the total potential energy. This relation is true for much more complex systems, as long as they are in some kind of equilibrium.

The virial theorem can be derived with some generality by considering time derivatives of the moment of inertia $I = \sum m_i r_i^2$. If this is expected to be fairly static on average, constraints can be placed on the energy of the system. Consider derivatives of I,

$$\frac{1}{2}\frac{dI}{dt} = \sum_i m_i\vec{v}_i\cdot\vec{r}_i = \sum_i \vec{p}_i\cdot\vec{r}_i$$

Taking the time derivative of this gives

$$\frac{1}{2}\frac{d^2I}{dt^2} = \sum_i m_i\vec{v}_i\cdot\vec{v}_i + \sum_i m_i\frac{d\vec{v}_i}{dt}\cdot\vec{r}_i = 2K + \sum_i \vec{F}_i\cdot\vec{r}_i$$

This latter term can be summed by expanded

$$\sum_i \vec{F}_i\cdot\vec{r}_i = \sum_i\sum_{i\neq j}\vec{F}_{ij}\cdot\vec{r}_i$$

Then writing r_i as $(\vec{r}_i+\vec{r}_j)/2 + (\vec{r}_i-\vec{r}_j)/2$ this becomes

$$= \frac{1}{2}\sum_i\sum_{i\neq j} F_{ij}\cdot(\vec{r}_i+\vec{r}_j) + \frac{1}{2}\sum_i\sum_{i\neq j}\vec{F}_{ij}\cdot(\vec{r}_i-\vec{r}_j)$$

THe first term cancels to zero by symmetry since $\vec{F}_{ij} = -\vec{F}_{ji}$. The second term turns out to be surprisingly simple for our force law since it includes $\hat{r}_{ij}$:

$$= -\frac{1}{2}\sum_i\sum_{j\neq i}\frac{m_im_j}{r_{ij}^2}r_{ij} = -\frac{1}{2}\sum_i\sum_{j\neq i}\frac{m_im_j}{r_{ij}} = U$$

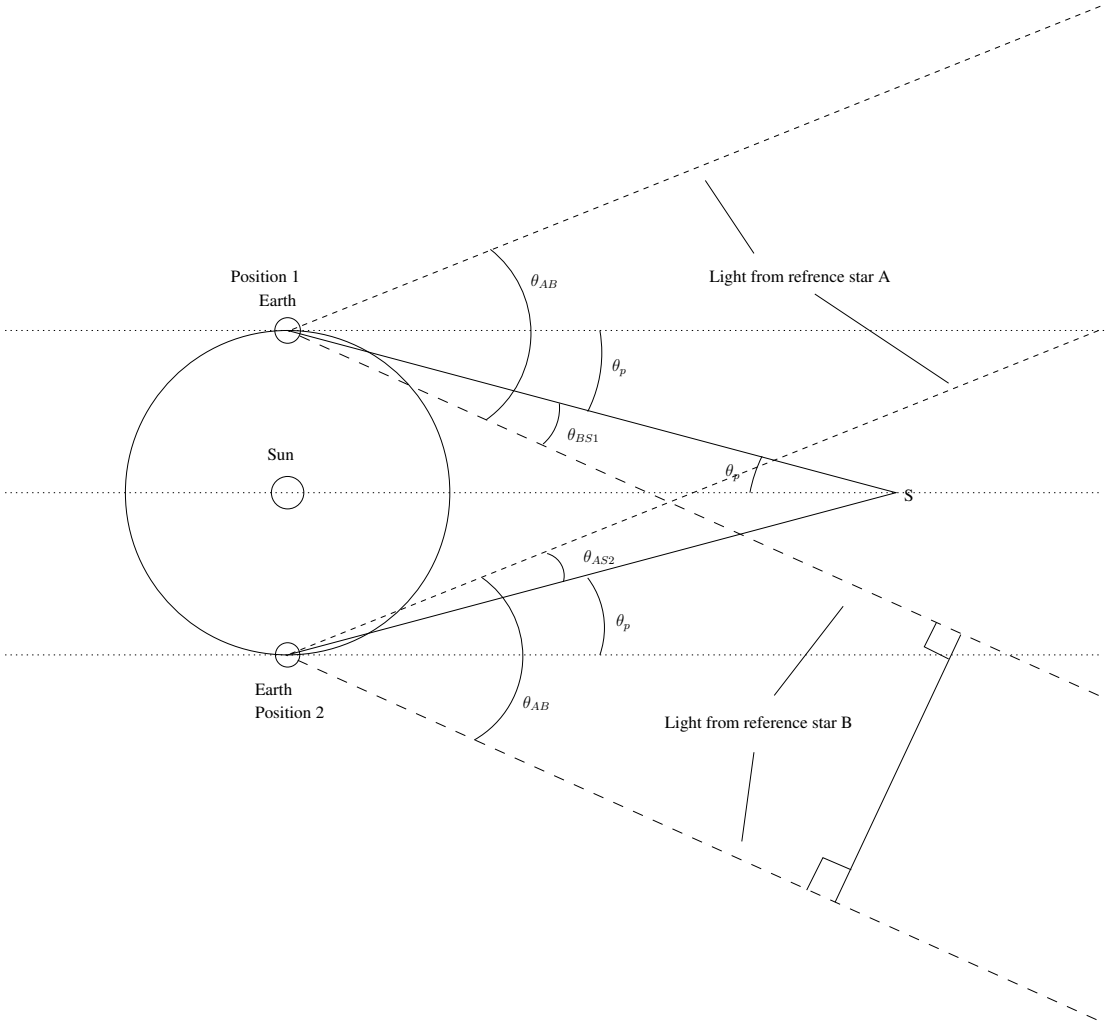
the total potential energy of the system.

By averaging everything

$$\frac{1}{2}\langle\frac{d^2I}{dt}\rangle = 2\langle K\rangle + \langle U\rangle = 0$$

approximately as long as dI/dt is roughly periodic, so that its average derivative is zero. This is commonly true of systems in equilibrium, since they are by definiton static in time on average.

3.3 Measuring distance with parallax



Parallax is determined by comparing the direction of the light of a star with light from other more distant objects at different times of the year (as earth moves through space).

Referencing the diagram above the parallax angle of the star is actually

$$\theta_p = \frac{1}{2}(\theta_{AB} - \theta_{AS2} - \theta_{BS1})$$

Where θ_{AB} is the angle between stars A and B, which is the same regardless of earth position, and θ_{AS2}, for example, is the angle between reference star A and the target star when Earth is in position 2.

See also a real parallax measurement, Thorstensen (2003):
http://adsabs.harvard.edu/abs/2003AJ....126.3017T
Notably as can be seen in the first figure, one must also account for the proper motion of the object, which will give an overall motion in some direction across the sky (called "proper motion") on which the back-and-forth (actually quasi-circular) motion due to the motion of the Earth is superimposed.

Also note the discussion of Lutz-Kelker bias. Parallax measurements are biased toward larger parallaxes (nearer distances) because errors tend to increase parallaxes. This necessitates shifting parallaxes slightly to correct them. This can also lead to unexpected ambiguities if the nature of the object is not well known - i.e. error can be mistaken for parallax in a very distant object. But if something about the object is known, it can have different implications, as discussed in Thorstensen (2003).

Also take a look at the Gaia mission, recently measured distances out to several thousand parsecs – including around a billion stars, about 1% of Milky Way stars.

A parsec is then defined such that a star with a parallax of 1 arcsecond is at 1 parsec distance so that

$$d = \frac{1}{\theta_p}\text{pc}$$

when θ_p is given in arcseconds. Using

$$d \times 1arsec = 1AU, \qquad d = \frac{180 \times 60 \times 60}{\pi} 1.5 \times 10^{11} m = 3.1 \times 10^{16} \text{ m}$$

One parsec is about 3.26 light years

Note that typical seeing at a telescope (the approximate size of the image of a star) is between 0.8 and 1 arcsecond typically, down to a few tenths at a good site. Even the closest stars are hard to measure distances to even with modern instruments.

3.4 Flux and magnitude

As a first introduction to flux density, consider how the energy emitted by a star like the sun spreads out in space. The energy emitted by the star (or other object) is measured by its luminosity, L, in Watts or $L_\odot$. This has *dimensions* of energy/time. But what we generally measure with a telescope is the flux density

$$F = \frac{L}{4\pi d^2} \quad \text{has units of} \quad \text{Watt m}^{-2}$$

or dimensions of energy per time per area, that is, the energy passing through a unit area perpendicular to the direction with the light is travelling.

The magnitude scales are constructed to be logarithmic such that a 5 magnitudes represents a factor of 100 in brightness. That is:

$$\frac{F_2}{F_1} = 100^{(m_1 - m_2)/5}$$

Note that here we are discussing bolometric fluxes, but this type of scale applies equally well to the flux in some wavelength range, which we will come to later.

We must choose a reference flux, which nominally would get into the messy details of the various magnitude scales of different wavelength ranges – they are actually standardized on a particular star, Vega – but it easier at this point to just declare that the sun observed from Earth has an apparent bolometric magnitude of -26.83.

In-class exercise: Derive an expression for the apparent magnitude directly from the flux. i.e. what is $m(F)$?

one should get

$$m_2(F_2) = m_1 - 2.5\log(F_2/F_1)$$

or

$$m(F) = m_\odot - 2.5\log(F/F_\odot)$$

Absolute magnitude is simply defined as the apparent magnitude at 10 pc. (student:) Insert $F = L/4\pi d^2$, with the Sun as reference.

$$M_{bol}(L) = M_{bol,\odot} - 2.5\log(L/L_\odot)$$

and

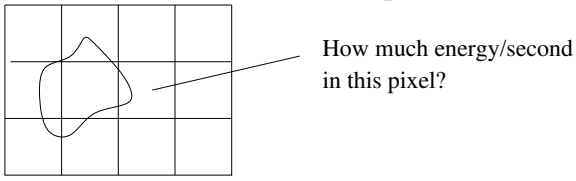
$$m(d) = M + 5\log(d/10\text{ pc})$$

The absolute magnitude of the Sun is 4.74.

Note this is deceptive because it means that while apparent magnitude is a measure of *flux* in energy/time/area, absolute magnitude is actually a mesaure of just luminosity, in energy/time.

4 Astro notes 2018/8/29 - Wed - Light 2 - EM radiation, general flux density

One way to consider the question we will be addressing is: how do we compute the energy received from a cloud that fills part of a CCD pixel?



Similarly if we are observing a star, the scattered light from the sky will contribute to the energy measured. How do we quantify this? It will be uniform over the sky, but how much ends up in a single pixel in addition to the starlight?

4.1 Electromagnetic radiation

Electromagnetic radiation follows from the equations for electrodynamics. i.e. Maxwell's equations

$$\vec{\nabla}\cdot\vec{E} = 4\pi\rho, \qquad \vec{\nabla}\times\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t}$$

$$\vec{\nabla}\cdot B = 0, \qquad \vec{\nabla}\times\vec{B} = \frac{1}{c}\left(4\pi\vec{J} + \frac{\partial E}{\partial t}\right)$$

where ρ is the charge density and $\vec{J}$ is the current density. Note this is Gaussian units (cgs). Note force on a charge then given by

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

For Gaussian units, charge is measured in "electrostatic units" (statcoulombs) rather than Coulombs. This causes a different constant to appear in Gauss' Law (the first equation above) and therefore the typical equation for the electric field of a point particle. In SI $\vec{\nabla} \cdot \vec{E} = \rho/\epsilon_0$.

But we want to investigate the situation in empty space where there are no charges (and therefore no currents), so that $\rho = 0$ and $\vec{J} = 0$. Then we have

$$\vec{\nabla} \cdot \vec{E} = 0 \qquad \vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \qquad \vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t}$$

Taking $\vec{\nabla}\times$ of the second equations, and using the identity

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{X}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{X}) - \nabla^2 \vec{X}$$

gives

$$\nabla \times \nabla \times \vec{E} = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = \nabla \times (-\frac{1}{c}\frac{\partial B}{\partial t})$$

or

$$\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0$$

and similarly for $\vec{B}$

$$\frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0$$

These clearly indicate waves moving at speed c through space. Solutions look like $\vec{E}_0 e^{i(\vec{k}\cdot\vec{x}-\omega t)}$. Such that

$$\omega^2 - c^2 k^2 = 0, \quad \text{or} \quad c = \lambda\nu$$

where λ is the wavelength ($k = 2\pi/\lambda$) and ν is the frequency in cycles per time ($\omega = 2\pi\nu$).

Note that for $\vec{\nabla} \cdot E = 0$, $\vec{k} \cdot \vec{E} = 0$, so that the waves are **transverse**, and the same holds for $\vec{B}$.

Note also that, $\vec{E}$ and $\vec{B}$, are perpendicular to each other as well.

Many important aspects of light follow from the fact that the above equations for propagating light are linear, constant-coefficient wave equations. This means that any two solutions can form a simple **superposition**, i.e. they can just be added and they continue to form a solution and do not influence each other. So when light waves cross they continue to propagate independently. They can interfere, if measured where they overlap, but each propagates independently.

Solutions at different frequencies, ν, and moving in different directions, $\vec{k}$, are independent.

Thus it is not enough just to know how much radiation there is in some region of space, but also what direction it is moving in, i.e. $\vec{k}$.

4.2 Poynting Vector, actual flux density

A straightforward way to characterize energy transfer of a radiation field is by its Poynting vector (in mks units):

$$\vec{S} = \frac{1}{\mu_0}\vec{E} \times \vec{B}$$

This gives both the magnitude and direction of energy transfer. However, this is not uniform in space (due to the wave) so it is typically easier to discuss $\langle \vec{S} \rangle$.

Now that we know that light of different wavelengths and directions is indepentent, and that it makes sense to speak of it carrying energy in some direction through space, we will state the fully **general flux density.** The energy passing through a surface (dA) is

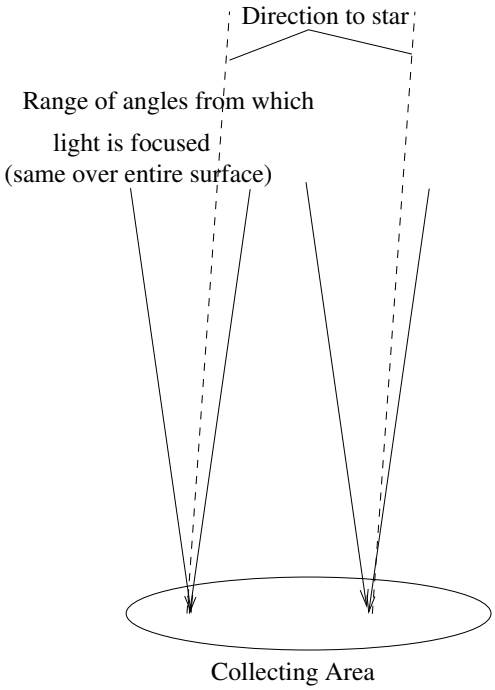
$$dE = F(\lambda, \hat{n}, \vec{x}) dA d\lambda d\Omega dt$$

so F has dimensions of energy per wavelength, per solid angle, per area, and depends on wavelength, direction and position. For example, consider a star in space at position $\vec{x}_s$. Far away from the star, its Flux density looks approximately like

$$F = \frac{B(T, \lambda) R^2}{|\vec{x} - \vec{x}_s|^2} \delta(\hat{n} - \widehat{\vec{x} - \vec{x}_s})$$

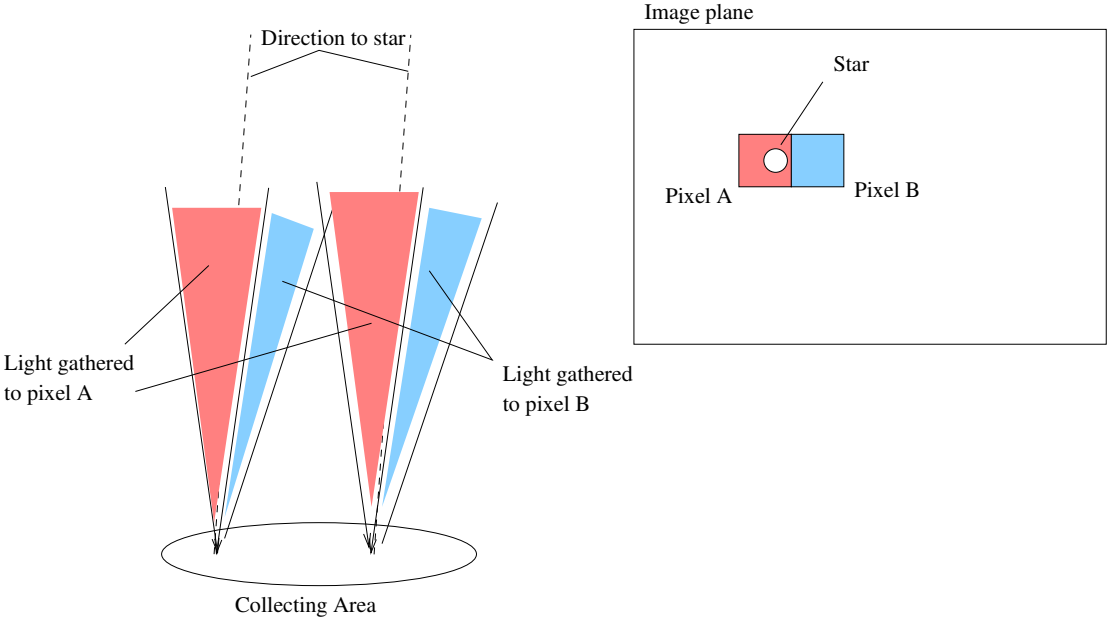
where $B(T, \lambda)$ is the plank function spectrum.

Performing a measurement generally corresponds to integrating F (sometimes called J in this form) over some collecting area *and* over some field of view. The field of view specifies the directionality of photons being collected, and, by design, is very well defined for focusing optics.



For an extended source, the delta function might go away and be replaced by defining F for some value in some portion of the sky. Thus the power measured (energy received per time) is dependent both on how big the collecting area is, and how much of the source light is collected from. If F is approximately independent of position on the sky over some range, then the power collected is just proportional to Ω, the solid angle observed. i.e. the collected flux is the product $F\Omega$. This product having dimensions of energy per time per area per wavelength.

A diagram demonstrating this for neighboring pixes in a detector:



5 Astro notes 2018/08/31 - Fri - Light 3: Thermal spectrum, Colors

Last time we introduced the flux density, which is a function F such that energy can be added up using

$$dE = F_\lambda(\lambda, \hat{n}, \vec{x}) dA d\lambda d\Omega dt$$

As a result, F has dimensions of energy per time per area per wavelength per solid angle. Usually the units are ergs/ (s-cm^2-Å-steradian).

The simplist example is for a uniform background (no anglular dependence), with a pixel subtending a solid angle Ω the flux collected is $f_\lambda = F_\lambda \Omega$.

For a star, the full flux density looks like

$$F_\lambda = \frac{B_\lambda(T, \lambda) R^2}{|\vec{x} - \vec{x}_s|^2} \delta(\hat{n} - \widehat{\vec{x} - \vec{x}_s})$$

So that if the angles subtended by the pixel include the direction toward the star, we recover the previous expression for flux $f = L/4\pi d^2$. But we really need a bit more info about the blackbody function...

5.1 Thermalized radiation (Blackbody)

Without any other assumptions, there is no reason to necessarily believe that the wavelengith dependence of flux, $F(\lambda)$, would be simple. Independent wavelengths can propagate indepently. However it is possible to reduce the wavelength dependence to a single parameter, T, the temperature if we assume **thermal equilibrium** in the radiation field. This can be posed in several ways, the most common being a box with a small hole to the outside, where the photon field has ample opportunity to equilibriate inside the box. A fully ionized, optically thick medium, such as the interior of a star, also has a photon field which is thermalized, and therefore follows the same distribution.

The construction of the Plank function comes from assuming that creating light occurs by creating photons whose energy is related to their wavelength,

$$E_\gamma = h\nu = \frac{hc}{\lambda}$$

where h is Planck's constant. It turns out that this is a deep meaningful statement about the nature of radiation that *does not follow from Maxwell's equations.* By doing this, it is possible to determine statistically what the **occuption of different wavelengths** should be in thermal equilibrium from which Planck's function is obtained:

$$B_\lambda(T) = \frac{2hc^2/\lambda^5}{e^{hc/\lambda kT} - 1}$$

where k is Boltzmann's constant. Note this is energy/s per wavelength interval per area per steradian. So flux in power is $\int B_\lambda(T)d\lambda$. This means that if we want the frequency-specific power, we must convert to $B_\nu(T)d\nu$, so that

$$B_\nu(T) = \frac{2h\nu^3/c^2}{e^{h\nu/kT}-1} \ .$$

Fluxes must also be given either per wavelenth or per frequency, thus you will see F_λ and F_ν to denote the difference.

There are two interesting limits to take. The first is long wavelength. In this case we can expand $e^{hc/\lambda kT} \simeq 1 + hc/\lambda kT$. This gives

$$B_\lambda(T) \simeq \frac{2hc^2}{\lambda^5}\frac{\lambda kT}{hc} = \frac{2ckT}{\lambda^4}$$

which is the **Rayleigh-Jeans Law**.

Also if λ small, or, equivalently, ν is large, we find

$$B_\nu(T) \simeq \frac{2h\nu^3}{c^2}e^{-h\nu/kT}$$

which follows naturally from classical thermal physics.

Wein's Law Finally it is straight forward to derive the maximum of $B_\lambda(T)$ in wavelength just by taking derivatives

$$\frac{dB_\lambda}{d\lambda} = 0 = B_\lambda\frac{-5}{\lambda} + \frac{B_\lambda}{e^{hc/\lambda kT}-1}e^{hc/\lambda kT}\frac{hc}{\lambda^2 kT}$$

or

$$\frac{hc}{\lambda kT}\frac{e^{hc/\lambda kT}}{e^{hc/\lambda kT}-1} - 5 = 0$$

The solution to this equation is $hc/\lambda kT = 4.97...$ giving

$$\lambda_{max}T = 0.002898 \text{ m K}$$

5.2 Effective temperature

The thermal flux law can be used to obtain the flux from the surface of a star. The flux at the surface of a blackbody is

$$F_\lambda = B_\lambda \cos\theta$$

where θ is the angle from the normal. At every point on the surface of the star we must integrate over the upper half-volume above the surface

$$F_{surf} = \int_{\lambda=0}^{\infty}\int_{\theta=0}^{\pi/2}\int_{\phi=0}^{2\pi} B_\lambda \cos\theta \ \sin\theta \, d\theta \, d\phi = \sigma_{\rm SB}T^4$$

where σ_{SB} is the Stefan-Bolztmann constant. This is often generalized so that even if F_λ is not a blackbody, an "effective" temperature can be defined such that $F_{\rm surf} = \sigma_{\rm SB} T_{\rm eff}^4$.

But so far we have only integrated F over direction (with respect to the surface) and wavelength. We still must integrate over the surface to find the total energy release:

$$L = \int_{surf} \sigma_{\rm SB} T^4 = 4\pi R^2 \sigma_{\rm SB} T^4$$

where R is the stellar radius.

6 Astro notes 2018/9/5 - Wed - colors, special relativity

6.1 Filters and color

We have been assuming the wavelength response of our detector is performing the following integral of our flux field

$$F_{\rm bol} = \int_0^\infty F_\lambda d\lambda$$

This corresponds to measuring the bolometric flux, the energy at all wavelengths. But real detectors generally only measure over a limited range of wavelengths. This is partially *on purpose* because it allows one to get a gross idea of the spectral shape (and therefore the temperature) by just measuring the fux in a few wavelength ranges. So for a real instrument we do something like:

$$F_X = \int_0^\infty F_\lambda(\lambda)\, \mathcal{S}_X(\lambda)\, d\lambda$$

where $\mathcal{S}$ is the sensitivity and X is some label for the "pass band".

We usually select a **filter** that is constructed to reproduce, in combination with our detector, one of several standard sensitivity functions. (Though typically only approximately.)

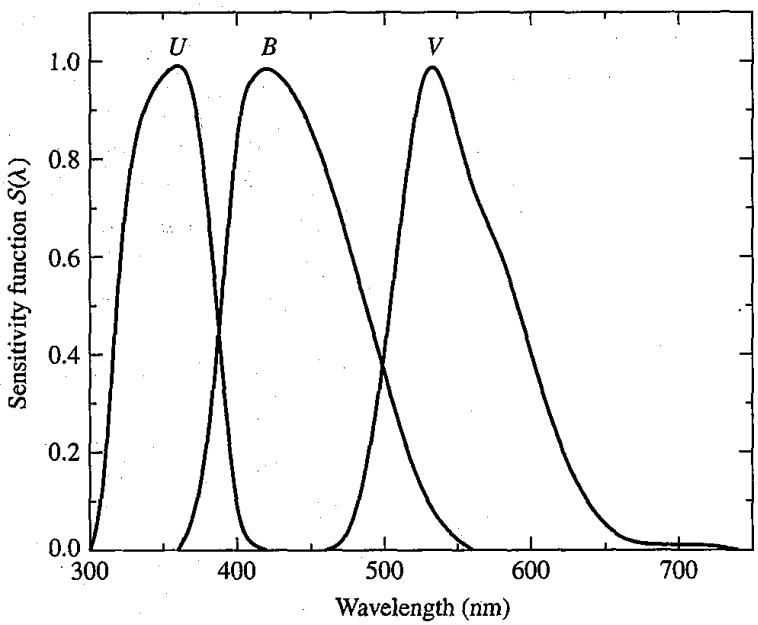


FIGURE 3.10 Sensitivity functions $\mathcal{S}(\lambda)$ for U, B, and V filters. (Data from Johnson, *Ap. J.*, *141*, 923, 1965.)

Filter	center λ (nm)	width $\Delta\lambda$ (nm)
U	365	68
B	440	98
V	550	89

There are also many others, R, I, J, etc and entire other systems, e.g. *ugriz*, which are those from the Sloan Digital Sky Survey (SDSS).

The magnitudes are define generally like so:

$$V = -2.5 \log_{10} \left(\frac{\int F_\lambda \mathcal{S}_V d\lambda}{\int F_{\lambda,\text{Vega}} \mathcal{S}_V d\lambda} \right)$$

Where here the star Vega is used as a reference. (not quite really, but that is the historical reference. Now often "AB" magnitudes, based on single-frequency standards.) And similarly for B, V, etc. Thus Vega has $V = B = U = 0$ by definition, though it is certainly not equally bright in all these wavelengths.

A VERY rough way to characterize the spectral shape is the flux ratio obtained effectively by differencing two magnitudes:

$$B - V = -2.5 \log_{10} \left(\frac{\int F_\lambda \mathcal{S}_B d\lambda}{\int F_\lambda \mathcal{S}_V d\lambda} \right) + 2.5 \log_{10} \left(\frac{\int F_{\lambda,\text{Vega}} \mathcal{S}_B d\lambda}{\int F_{\lambda,\text{Vega}} \mathcal{S}_V d\lambda} \right)$$

$$= -2.5 \log_{10} \left(\frac{\int F_\lambda \mathcal{S}_B d\lambda}{\int F_\lambda \mathcal{S}_V d\lambda} \right) + C_{B-V}$$

Also called the **color index**. Here the constant C is set again by the reference flux distribution, that of the star Vega. So this just characterizes the ratio between fluxes at different wavelengths, compared to that ratio in Vega. Approximately:

$$B - V \simeq -2.5 \log_{10} \left(\frac{F_\lambda(\lambda_{B,center}) \Delta\lambda_B}{F_\lambda(\lambda_{V,center}) \Delta\lambda_V} \right) + C_{B-V}$$

which is a simple way to estimate color index.

Note that, due to the negative, larger $B - V$ means that F_B is smaller than F_V i.e. the object is redder.

Note that if the spectrum is characterized by a single parameter, like T for a thermal (or T_{eff} for a stellar) spectrom, there is a simple mapping from color index to that parameter.

Mapping from color index $B - V$ to T_{eff} for luminosity class V stars (main sequence stars). (From Binney & Merrifield, *Galactic Astronomy*, tables 3.7 and 3.10.)

spectral type	$B - V$	T_{eff} (K)
O3	-0.33	52,500
O5	-0.33	44,500
O8	-0.32	35,500
B0	-0.30	30,000
B3	-0.20	18,700
B5	-0.17	15,400
B8	-0.11	11,900
A0	-0.02	9,520
A5	0.15	8,200
F0	0.30	7,200
F5	0.44	6,440
G0	0.58	6,030
G5	0.68	5,770
K0	0.81	5,250
K5	1.15	4,350
M0	1.40	3,850
M5	1.64	3240

Why no change at high T? in Rayleigh-Jeans tail, so ratio indepdent of T

6.2 Discussion of homework

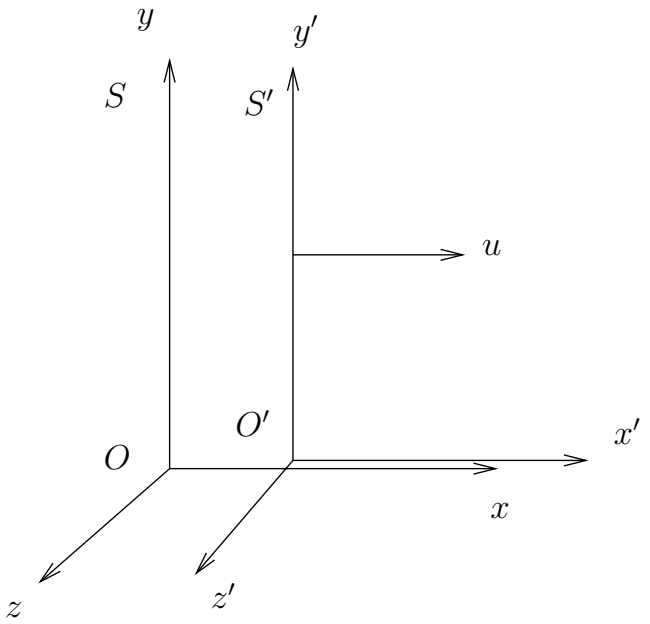
Some discussion of homework questions, particularly problem 4 on homework 1a.

7 Astro notes 2018/9/7 - Fri - special relativity

7.1 Intro Special Relativity

Galilean transformation stems from assuming that it makes sense to speak of the global structure of space and time. i.e. there are coordinates where it is well defined to measure

the position, x, y, z, and time t of *events*. One can then imagine two such coordinate systems that are in relative motion at constant velocity. We call the second coordinate system the "prime" system, which is moving at speed u in the x direction.



Then we can easily convert events (posititon, time) sets from the base coordinates to the prime coordinates, assuming that O and O' are coincident at $t = 0$. That is $x = x'$ at $t = t' = 0$.

Galilean:

$$x' = x - ut$$
$$y' = y, \quad z' = z, \quad t' = t$$

Then if we have a particle path $\vec{x}(t)$, its velocity $\vec{v}(t)$ can be computed in prime coordinates, giving

$$v'_x = v_x - u, \quad v'_y = v_y, \quad v'_z = v_z$$

Newtons laws, phrased in terms of $d^2\vec{x}/dt^2$ are then unchanged.

Problems arise when one considers that there is some reference frame in which the speed of light is the value obtained from Maxwell's equations. This is equivilent to there being some correct definition of electric vs. magnetic field. Recall that static charges generate electric fields, and moving charges generate magnetic fields, so which field is present appears to depend on how the chosen reference frame moves.

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho, \qquad \vec{\nabla} \times \vec{B} = \frac{1}{c}\left(4\pi\vec{J} + \frac{\partial \vec{E}}{\partial t}\right)$$

(not only does $\vec{J}$ depend on the speed of the frame, but so does the time dependence of ρ, which enters via $\partial\vec{E}/\partial t$.) This seems, at first glance to imply that there is some favored reference frame, though it turns out that it implies exactly the opposite!

This was first shown experimentally by attempting to measure the speed of light in opposite directions. If there is some favored reference frame, it should then be easy to find ones relative motion with respect to it. It turned out, despite trying, no difference in the speed of light was ever measured.

7.2 Lorentz transformations

It turns out the concept of a well-defined global coordinate system is what is flawed. Local connections between events based on the propagation speed of light, however, does provide a way to construct a workable concept of spacetime. Einstein expressed this in two postulates:

- The laws of physics are the same in all inertial reference frames.
- Light moves through vacuum at a constant speed c independent of source.

These two basically tell you how to construct valid reference frames for Newton's laws (inertial) and how they are connected (all measure the same speed for light).

Consider the general coodinate transformation:

$$x' = a_{11}x + a_{12}y + a_{13}z + a_{14}t$$

etc. This is a fully general linear transformation. A transformation of similar form, but corresponding to a different operation, is that of a rotated reference frame:

$$x' = (\cos\theta)x - (\sin\theta)y \qquad y' = (\sin\theta)x + (\cos\theta)y$$

The Galilean transform is also of this form.

We seek a transform between frames where the prime frame is moving at constant speed in the x direction as before. The transverse directions are nominally unchanged, so that

$$y' = y, \quad z' = z$$

The symmetry with respect to y or z also implies that the time transformation should not depend on them. So

$$t' = a_{41}x + a_{44}t$$

Finally for the x coordinates, we have defined the relative motion of the coordinate system origins. They coincide at $t = 0$ and the origin of the prime coordinate system ($x' = 0$) is moving with speed u. Thus the coordinates of the origin of the prime coordinates are $x' = 0$ and $x = ut$.

(Student:) What does this tell us?

$$0 = a_{11}ut + a_{12}y + a_{13}z + a_{14}t$$

since this holds on the yz plan for any y, z, $a_{12} = a_{13} = 0$. So we are left with $a_{11}u = -a_{14}$. So far we have

$$x' = a_{11}(x - ut), \qquad t' = a_{41}x + a_{44}t$$

To set the final coefficients, consider the propagation of a light pulse started at the origin at time zero. Invoking the second of Einstein's postulates this gives the location in both frames

(Student:) How is this written in each coordinate system?

$$x^2 + y^2 + z^2 = c^2t^2, \qquad x'^2 + y'^2 + z'^2 = c^2t'^2$$

Putting in the above we get

$$a_{11}^2(x^2 - 2xut + u^2t^2) + y^2 + z^2 = c^2(a_{41}^2x^2 + 2a_{41}a_{44}xt + a_{44}^2t^2)$$

or

$$(a_{11}^2 - c^2a_{41}^2)x^2 + y^2 + z^2 = (a_{11}^2u + c^2a_{41}a_{44})2xt + (c^2a_{44}^2 - a_{11}^2u^2)t^2$$

Matching up, the t^2 term tells us

$$c^2a_{44}^2 - a_{11}^2u^2 = c^2$$

$$a_{11}^2 - c^2a_{41}^2 = 1$$

$$a_{11}^2u + c^2a_{41}a_{44} = 0$$

or

$$a_{11}^4u^2 = c^2a_{41}^2c^2a_{44}^2 = (a_{11}^2 - 1)(c^2 + a_{11}^2u^2) = a_{11}^2c^2 - c^2 + a_{11}^4u^2 - a_{11}^2u^2$$

so that

$$a_{11}^2 = \frac{1}{1 - u^2/c^2}$$

and then

$$a_{44} = a_{11} \qquad a_{41} = -ua_{11}/c^2$$

so that our transformations are

$$x' = \gamma(x - ut), \qquad t' = \gamma(t - ux/c^2), \qquad y' = y, z' = z$$

where

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

is called the Lorentz factor.

7.3 Summary

we obtained the Lorentz transformation, the relation between the coordinates of events in a reference frame and one moving at speed u with respect to it in the x direction.

$$t' = \gamma(t - ux/c^2)$$

$$x' = \gamma(-ut + x)$$

$$y' = y, \qquad z' = z$$

where $\gamma = 1/\sqrt{1 - u^2/c^2}$, and note that $\gamma \geq 1$ always.

We will also make use of the inverse transforms which are most easily obtained by swapping u to $-u$ and the primes to non-primes

$$t = \gamma(t' + ux'/c^2)$$

$$x = \gamma(ut' + x')$$

7.4 Simultaneity

Using these transformations we can compute the time, in the moving frame, between two events which are simultaneous (equal t) in the base frame.

$$t'_2 - t'_1 = -\frac{(x_2 - x_1)u/c^2}{\sqrt{1 - u^2/c^2}}$$

thus the events are not simultaneous in the prime frame. Also which one is first actually depends on relative sign of u compared to $x_2 - x_1$, i.e. the direction of the motion.

7.5 Homework problem on simultaneity (or, rather, lack thereof)

Handed out:

Consider a spaceship that is traveling past the solar system toward a star 10 light-years away at $0.5c$. A signal is emitted from Earth when the spaceship is halfway there, and therefore received just as the spaceship reaches the star.

Draw, to scale as much as possible, a spacetime diagram in both the solar system rest frame and a frame moving with the spaceship that shows the path of the Sun, the distant star, the spaceship and the signal. How far does the signal travel in the frame of the spaceship? By working back from the time the signal was received using the speed of light and the varying distance to the solar system, what fraction of the distance to the star would the occupants of the spaceship say that they had covered at the time, according to their coordinates, that the signal was emitted? (That is, what is the time coordinate of the emission event in the spaceship's frame, and how does this compare to the time coordinate of the spaceship arriving at the star?)

You should explicitly give the coordinates in both the solar system and spaceship frames of the followin events:

- Departure of spaceship from solar system
- Arrival of spaceship at star
- Emission of signal
- Arrival of signal at star

You must also compute and show the positions of both the solar system and the distant star at the beginning and end of the trip in both frames.

8 Astro notes 2018/9/10 - Mon - special relativity

8.1 Summary

we obtained the Lorentz transformation, the relation between the coordinates of events in a reference frame and one moving at speed u with respect to it in the x direction.

$$t' = \gamma(t - ux/c^2)$$

$$x' = \gamma(-ut + x)$$

$$y' = y, \qquad z' = z$$

where $\gamma = 1/\sqrt{1 - u^2/c^2}$, and note that $\gamma \geq 1$ always.

We will also make use of the inverse transforms which are most easily obtained by swapping u to $-u$ and the primes to non-primes

$$t = \gamma(t' + ux'/c^2)$$

$$x = \gamma(ut' + x')$$

8.2 Time dilation

Consider the ticking of a clock that is in a moving reference frame. We can compute the time between ticks as determined by an observer that is not moving with the clock (note this is an "inverse" transformation from prime to non-prime, so we negate u):

$$t_2 - t_1 = \gamma[(t'_2 - t'_1) + (x'_2 - x'_1)u/c^2]$$

or, since the object is assumed to be at $x' = 0$ at both times,

$$\Delta t_{\text{moving}} = \gamma \Delta t_{\text{rest}}$$

where here "moving" is considered with respect to the clock being considered (i.e. an "at rest" frame is moving with respect to a moving clock).

Thus it appears to take longer, from the point of view of the observer not moving with the clock, for the a tick to elapse than it does in the clocks rest frame.

Note that this is not a "apparent" effect. If the ticks are measured by measuring pulses of light, the discrepancy is that present after accounting for the light travel time.

8.3 Length Contraction

Now consider a bar with ends at x'_2 and x'_1 such that $L' = x'_2 - x'_1$. The lack of simultaneity makes the definition of "measuring" this bar's length in a moving reference frame non-trivial. It must be defined in terms of events. We will take its length in the non-prime frame (which is moving with respect to the bar) to be the distance between the positions that the ends of the bar pass at the same time t.

$$x_2 - x_1 = \gamma(u(t'_2 - t'_1) + x'_2 - x'_1) = \gamma^2[u(t_2 - t_1) - (x_2 - x_1)u^2/c^2] + \gamma(x'_2 - x'_1)$$

since $t_2 - t_1 = 0$, by our assumption,

$$(1 + \gamma^2 u^2/c^2)(x_2 - x_1) = \gamma(x'_2 - x'_1)$$

The left hand side of this is

$$\frac{1 - u^2/c^2 + u^2/c^2}{1 - u^2/c^2}(x_2 - x_1) = \gamma^2(x_2 - x_1)$$

so

$$\gamma^2(x_2 - x_1) = \gamma(x'_2 - x'_1)$$

or

$$L = L'\sqrt{1 - u^2/c^2}$$

where L' is the actual length of the bar in its own rest frame, and L is the length perceived by a moving observer.

8.4 Homework problem on simultaneity (or, rather, lack thereof)

Consider a spaceship that is traveling past the solar system toward a star 10 light-years away at $0.5c$. A signal is emitted from Earth when the spaceship is halfway there, and therefore received just as the spaceship reaches the star.

Draw, to scale as much as possible, a spacetime diagram in both the solar system rest frame and a frame moving with the spaceship that shows the path of the Sun, the distant star, the spaceship and the signal. How far does the signal travel in the frame of the spaceship? By working back from the time the signal was received using the speed of light and the varying distance to the solar system, what fraction of the distance to the star would the occupants of the spaceship say that they had covered at the time, according to their coordinates, that the signal was emitted? (That is, what is the time coordinate of the emission event in the spaceship's frame, and how does this compare to the time coordinate of the spaceship arriving at the star?)

You should explicitly give the coordinates in both the solar system and spaceship frames of the followin events:

- Departure of spaceship from solar system
- Arrival of spaceship at star
- Emission of signal
- Arrival of signal at star

You must also compute and show the positions of both the solar system and the distant star at the beginning and end of the trip in both frames. Verify by comparing the distance and time between the emission and arrival events of the light in the spaceship frame that the light is moving at the speed of light.

9 Astro notes 2018/9/12 - Wed - Special Relativity 3 - Doppler and beaming

Lorentz transforms:

$$t' = \gamma(t - ux/c^2), \qquad x' = \gamma(-ut + x)$$

inverse:

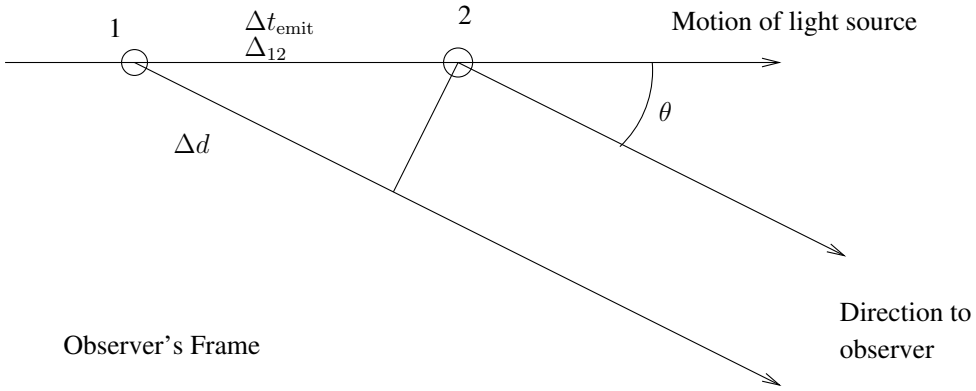
$$t = \gamma(t' + ux'/c^2), \qquad x = \gamma(ut' + x')$$

Note on homework discussion last time: transform of location, $(t, x) = (0, 10)$, of distant star in prime frame make its t' less than zero.

9.1 Relativistic Doppler shift

Context: emission from a jet launched from a black hole, GRB, or young star. For example, see SS433 (a stellar-mass BH or NS) and 3C 273 an AGN jet.

This is much like traditional doppler shift, which depends on the object moving closer to or further from the observer during an emission period, but includes time dilation. Consider the following geometry of motion of a light source in an observer's frame: We consider the time $\Delta t_{\rm rec}$ between reception of consecutive pulses emitted by the source, moving at u, with some $\Delta t_{\rm rest}$ in its own frame.



Assume that the frames were aligned when the first pulse was emitted and the second pulse is emitted at time $\Delta t_{\rm rest}$ in the moving frame from the light source still at the origin in the moving frame, So first pulse is emitted at $(t'_1, x'_1) = (0, 0)$ and second is at $(t'_2, x'_2) = (\Delta t_{\rm rest}, 0)$. We can do an inverse transform to get time and location of the second emission (note $(t_1, x_1) = (0, 0)$):

$$\Delta t_{\rm emit} = t_2 = \gamma \Delta t_{\rm rest}, \qquad \Delta_{12} = x_2 = \gamma u \Delta t_{\rm rest}$$

The light path for pulse 1 is $\Delta d = \Delta_{12} \cos\theta$ shorter than that for pulse 2. So the time difference of receipt is

$$\Delta t_{\rm rec} = \Delta t_{\rm emit} - \frac{\Delta d}{c} = \gamma \Delta t_{\rm rest} - \gamma (u/c) \Delta t_{\rm rest} \cos\theta = \gamma [1 - (u/c)\cos\theta] \Delta t_{\rm rest}$$

Frequencies are then related by

$$\nu_{\rm rec} = \nu_{\rm rest} \frac{\sqrt{1 - u^2/c^2}}{1 - (u/c)\cos\theta} = \nu_{\rm rest} \frac{\sqrt{1 - u^2/c^2}}{1 + (v_r/c)}$$

note that θ here is opposite of the convention in the textbook.

9.2 Velocity transformation

To consider beaming, which involves which direction light is moving, we need a velocity transformation.

We can find how velocities transform directly from the coordinate transformations by considering dt, dx, etc. This is relatively straightforward because u just a parameter, so that this is a simple, constant-coefficient coordinate transformation.

$$dt' = \gamma(dt - (u/c^2)dx)$$

$$dx' = \gamma(-u\,dt + dx)$$

$$dy' = dy, \qquad dz' = dz$$

then

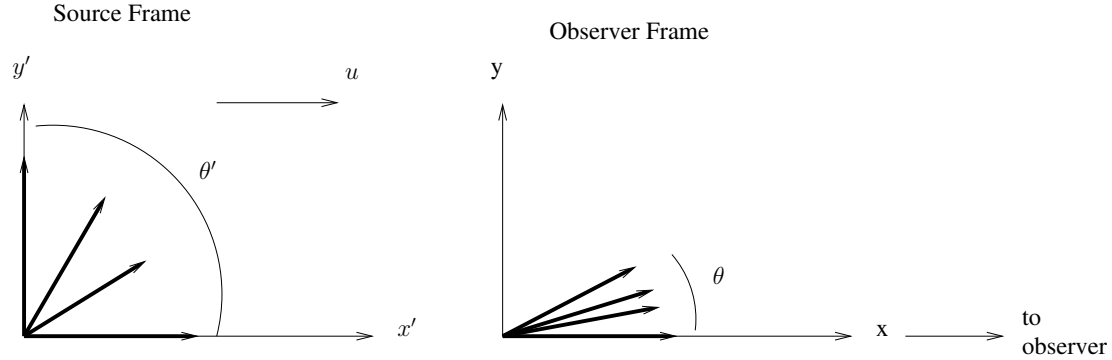
$$v'_x = \frac{dx'}{dt'} = \frac{dx - u\,dt}{dt - (u/c^2)dx} = \frac{dx/dt - u}{1 - (u/c^2)dx/dt} = \frac{v_x - u}{1 - uv_x/c^2}$$

$$v'_y = \frac{dy}{\gamma(dt - (u/c^2)dx} = \frac{v_y\sqrt{1 - u^2/c^2}}{1 - uv_x/c^2}$$

$$v'_z = \frac{dz}{\gamma(dt - (u/c^2)dx} = \frac{v_z\sqrt{1 - u^2/c^2}}{1 - uv_x/c^2}$$

9.3 Beaming

While at first thought it may seem that, since the speed of light is the same in all reference frames, the motion of the source should have little effect on the direction in which light travels. However, this turns out not to be the case because the Lorentz transformation is quite non-isotropic. Consider a light source moving at some speed u that is emitting isotropically in its rest frame. We would like to know into what angle the light emitted in all "forward" directions ends up traveling in what we would consider the observer's frame.



To do this we consider the light ray emitted perpendicular to the direction of travel. This light ray has $v'_y = c$ and $v'_x = 0$. The transform is straightforward:

$$v_x = \frac{0 + u}{1 + 0} = u$$

$$v_y = \frac{c\sqrt{1 - u^2/c^2}}{1 - 0} = c/\gamma$$

Thus all light emitted "forward" in the emitter's frame is directed even more forward in the observer's frame. The angle that this transverse light ray makes can be computed since we know its velocity components (note $v_x^2 + v_y^2 = u^2 + c^2(1 - u^2/c^2) = c^2$).

$$v_y = c \sin\theta \quad \Rightarrow \quad \sin\theta = \frac{v_y}{c} = \gamma^{-1}$$

Thus, for relativistic motion, we say the light is "beamed" into a small angle, with higher u (larger γ) giving more beaming. (smaller θ)

Boosting:

To find the on-axis enhancement of brightness, it is useful to consider the more general case of light in the observer frame a small angle θ from the direction of motion, this will then have been at some angle θ' in the emitter's frame. We want to find how big θ' is for large u. This calculation is much like the above, since $v_x = c\cos\theta$ and $v_y = c\sin\theta$. We actually only need

$$v'_y = \frac{c\sin\theta\sqrt{1 - u^2/c^2}}{1 - (u/c)\cos\theta} = c\sin\theta'$$

For both θ and θ' small, which will be true for the collecting area of a very distant observer,

$$\sin\theta' \approx \theta' \approx \frac{\theta\sqrt{1 - u^2/c^2}}{1 - (u/c)}$$

Now consider large $u = c(1 - \epsilon)$ for small ϵ. Then

$$\gamma = 1/\sqrt{1 - u^2/c^2} \approx 1/\sqrt{1 - (1 - 2\epsilon)} = 1/\sqrt{2\epsilon}$$

In the same limit

$$\theta' = \frac{\theta\sqrt{2\epsilon}}{\epsilon} = \frac{2}{\sqrt{2\epsilon}}\theta = 2\gamma\theta$$

Thus the received radiation for a given collecting area is enhanced by a factor of about 2γ due to the beaming of light from the sources frame.

10 Astro notes 2018/9/14 - Fri - Light and Matter 1

Brief review of doppler boosting approximations from last lecture.

10.1 Photoelectric effect and CCDs

Previously we described the Poynting vector

$$\vec{S} = \frac{1}{\mu_0}\vec{E} \times \vec{B}$$

Which describes the direction of energy transport by radiation. This would naively imply that the energy being carried is just a function of the intensity of the radiation. (the size of the field) but this turns out to be incomplete.

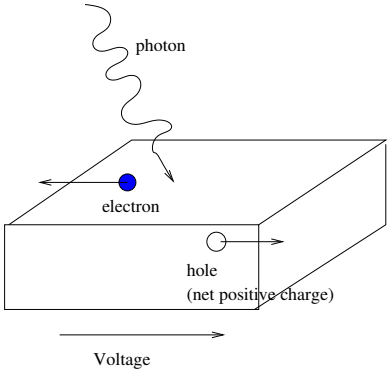
If you shine a light on the surface of a piece of metal, it can cause electrons to be ejected. The odd fact is that the maximum kinetic energy of these electrons only depends on the **frequency** or, equivalently, the wavelength of the radiation NOT on its intensity. At higher intensity, more electrons are produced, but their energy range is the same. It is found that the kinetic energy follows

$$K_{max} = h\nu - \phi$$

where ϕ is the work function characteristic of the particular metal (typically a few eV).

This implies that the energy contained in light is not a continuum, but comes in quanta, at least when you try to measure or convert it.

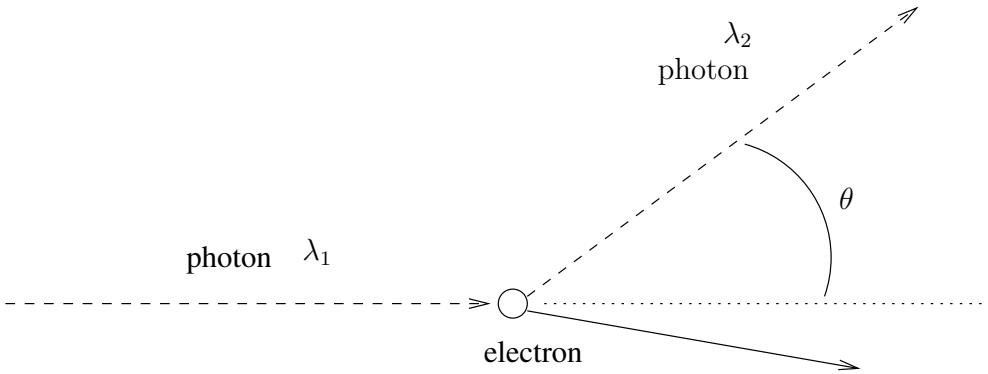
Converting photons to electrons on a semimetal material appears ubiquitously in astrophysics. Much of astrophysics is done using charged-coupled devices (CCDs) to sense light. These work by allowing a photon to create an electron-hole pair on the surfaces of a semiconductor and then sensing the charge by counting the electrons as a proxy for photons.



This process turns out to produce a detector with a very linear response and that can be made to have low noise.

10.2 Compton effect

The particle nature of the photon is also revealed in Compton scattering, in which a photon essentially bounces off of an electron.



By treating the photon as a particle with momentum $p = E/c = h\nu/c$, and performing a simple elastic scattering calculation, we find that

$$\Delta\lambda = \frac{h}{m_e c}(1 - \cos\theta)$$

This experiment can be performed, measuring the wavelength of photons leaving the collision at various angles. This angle-dependent wavelength shift is confirmed, demonstrating that light is made up of some sort of particle that has momentum determined by the wavelength of the light.

10.3 Atomic Levels

When we measure spectra, we will often measure spectral lines, either in emission or absorption. These are a fundamental aspect of the interaction of light with matter, and a direct result of the quantum structure of matter.

It was known that the hydrogen atom exhibited emission and absorption features at wavelengths given by

$$\frac{1}{\lambda} = R_H \left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

Where $R_H = 13.6eV/hc = 1.1 \times 10^7$ m^{-1} is called the Rydberg constant. While this is fundamentially quantum mechanical, it possible to obtain this behavior by insterting quantum mechanics into a classical problem.

Consider the electron orbiting around the proton nucleus under the influence of the force

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}$$

Using the virial theorem, or other means, it is possible to obtain the energy of this system as a function of the orbital distance of the electron

$$E = K + U = -\frac{e^2}{8\pi\epsilon_0}\frac{1}{r}$$

Now we may posit that the angular momentum of this orbit is not free to have any value, but must be quantized

$$L = \mu v r = n\hbar$$

where n = an integer and μ is the reduced mass. Using the virial theorem to relate the potential and kinetic energy we get

$$-E = K = \frac{1}{2}\mu v^2 = \frac{1}{2}\mu \frac{n^2\hbar^2}{\mu^2 r^2}$$

setting equal to the energy found above

$$\frac{n^2\hbar^2}{2\mu r^2} = \frac{e^2}{8\pi\epsilon_0 r}$$

and solving for r_n gives

$$r_n = \frac{4\pi\epsilon_0\hbar^2}{\mu e^2}n^2 = a_0 n^2$$

where a_0 gives the size of the orbit

$$E_n = -\frac{\mu e^4}{32\pi^2\epsilon_0^2\hbar^2}\frac{1}{n^2} = -13.6eV\frac{1}{n^2}$$

and we can get energies of the photon by just differencing

$$h\nu = \frac{hc}{\lambda} = E_n - E_m = hcR_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

where we find that

$$R_H = \frac{\mu e^4}{64\pi^3\epsilon_0^2\hbar^3 c}$$

11 Astro notes 2018/9/17 - Mon - Saha equation

11.1 Ionization state – Chemical Equilibrium and the Saha Equation

Since spectral lines depend on absorption/emission by electrons moving between energy levels, it is critical to know how many electrons are in each level.

One might guess that hydrogen would be ionized at 13.6eV$\approx 10^5$ K. But this is wrong and this is what we want to explain. Quantum mechanincs IS important here because the distinction between ionized and bound hydrogen is essentially quantum mechanical.

The QM distribution function for Fermi dirac is

$$f(p) = \frac{g}{h^3}\frac{1}{\exp[(E_p - \mu)/kT] + 1}$$

This is the number of particles per volume and per momentum bin. p is the momentum and g is the degeneracy factor (usually 2 for a spin 1/2 particle).

A sideline on chemical potential μ: Just like temperature is the thermodynamic potential for entropy (e.g. heat transfer), μ is the thermodynamic potential for particles (motion, diffusion, etc). Consider

$$dE = T\,dS - P\,dV + \mu\,dN$$

This is actually just a statement of the chain rule for an energy function $E(S, V, N)$,

$$dE = \left(\frac{\partial E}{\partial S}\right)_{V,N} dS + \left(\frac{\partial E}{\partial V}\right)_{S,N} dV + \left(\frac{\partial E}{\partial N}\right)_{S,V} dN$$

where the subscripts indicate explicitly which variables are being held fixed in the derivative. This is effectively the definition of the thermodynamic potentials T, P, and μ.

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N}, \quad P = -\left(\frac{\partial E}{\partial V}\right)_{S,N}, \quad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V}.$$

So just as temperature tells us how much energy wants to be moved via heat and pressure tells us how much energy is added when you squish something, chemical potential tells us the energy cost or benefit of adding / moving / removing particles.

Now we want to relate n, the number density in just space, to μ. For $s = \pm 1/2$, $g = 2$, and integrating over all momenta we get

$$n = \frac{g}{h^3}\int_0^\infty 4\pi p^2 dp \frac{1}{\exp(E_p/kT)e^{-\mu/kT} + 1}$$

where

$$E_p = \frac{p^2}{2m} + mc^2$$

Note tha binding energy differences will appear in the rest mass m here. For a nearly ideal gas where $a_{sep} \gg \lambda_{Debroglie}$ then

$$\exp[(E_p - \mu)/kT] \gg 1$$

(had student substitute in E_p and take this limit)
so then

$$n = \frac{g4\pi}{h^3}\exp\left(\frac{\mu - mc^2}{kT}\right)\int p^2 dp e^{-p^2/2mkT}$$

Now define

$$n_Q = \frac{4\pi}{h^3}\int_0^\infty p^2 e^{-p^2/2mkT} dp = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} \sim \frac{1}{\lambda_{de}^3}$$

or restating

$$\exp\left(\frac{\mu - mc^2}{kT}\right) = \frac{n}{gn_Q}$$

were we define which is the density where quantum mechanics becomes important. So we have, for e^-,

$$\mu_e = kT \ln\left(\frac{n_e}{2n_{Q,e}}\right) + m_e c^2$$

We want to solve for chemical equilibrium

$$e^- + p \rightleftharpoons H + \gamma$$

for this to be in chemical equilibrium,

$$\mu_e + \mu_p = \mu_H + \mu_\gamma$$

but note $\mu_\gamma = 0$. So writing out explictly

$$\mu_p = m_p c^2 - kT \ln\left(\frac{g n_{Q,p}}{n_p}\right)$$

and

$$\mu_H = m_H c^2 - kT \ln\left(\frac{g n_{Q,H}}{n_H}\right)$$

Our binding energy appears via

$$m_p c^2 + m_e c^2 - m_H c^2 = 13.6 \text{ eV} \ .$$

Going through the algebra, we get what is typically called the Saha equation

$$\frac{n_e n_p}{n_H} = \frac{g_p g_e}{g_H}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \exp\left(-\frac{13.6eV}{kT}\right)$$

charge balance demands $n_e = n_p$.

For half ionization $n_H = n_p$ then

$$n_e = \frac{g_p g_e}{g_H} n_{Q,e} \exp\left(-\frac{13.6eV}{kT}\right)$$

so really,

$$\ln\left(\frac{n_e}{n_{Q,e}}\right) = \frac{-13.6eV}{kT}$$

so then

$$kT \approx \frac{13.6eV}{\ln(n_{Q,e}/n_e)}$$

So the temperature can be very different from 13.6 eV, since typically $n_e \ll n_{Q,e}$.

Putting in numbers for the half ionization point ($n_p = n_e = n_H$)

$$n_e = 2.4 \times 10^{21} cm^{-3} T_4^{3/2} \exp(-15.76/T_4)$$

where $T_4 = T_{1/2}/10^4$ K.

First example is to use density like a stellar surface: put in $n_p = n_e = 10^{16}$ cm^{-3}. so that $T_{1/2} = 12,400$ K or $kT_{1/2} = 1.1$ eV. Around 10^4 K the H transitions from mostly neutral to mostly ionized. Sharp transition.

12 Astro notes 2018/9/19 - Wed - Saha equation example

12.1 Follow up on half ionization point

Putting in numbers for the half ionization point ($n_p = n_e = n_H$)

$$n_e = 2.4 \times 10^{21} cm^{-3} T_4^{3/2} \exp(-15.76/T_4)$$

where $T_4 = T_{1/2}/10^4$ K.

Last time mentioned example of sun: put in $n_p = n_e = 10^{16}$ cm^{-3}. so that $T_{1/2} = 12,400$ K or $kT_{1/2} = 1.1$ eV. Around 10^4 K the H transitions from mostly neutral to mostly ionized. Sharp transition.

In the early universe, When does the H recombine? here $n = 3 \times 10^4$ cm^{-3} and so we get $T = 4000$ K, which is $z = 1400$.

Places where the Saha equation breaks down - Pressure ionization At high densities the free electrons either start to interact with each other, or $\lambda \simeq$ spacing. (we used the ideal limit.) This becomes important when say $\rho \sim m_p/a_0^3 \simeq$ gram/cc.

12.2 Saha example

(done with student scribes)

Consider a hydrogen gas at $T = 5000$ to 25,000 K and $P_e = 20$ N/m^2. This is example 8.1.4 in the text, but I will do it with my notation.

Assuming H has two bound states at different energies ($n = 1$ and $n = 2$) (i.e. ignore higher states), what is the relative population of these two states and what is the fraction of H that is ionized at several values of the temperature in this range?

I derived the Saha equation as:

$$\frac{n_e n_{i+1}}{n_i} = \frac{g_e g_{i+1}}{g_i} \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} \exp\left(-\frac{\chi_i}{kT} \right)$$

where χ_i is the ionization energy of level i and n_i and n_{i+1} are the number density of atoms in the i and $i+1$ ionization state, for example neutral and singly ionized hydrogen respectively.

If we want to consider bound states in which the electron can be in different energy levels, there is one equation of this form form *each* energy level. So if we keep the index i to indicate the ionization state (number of electrons) we may add indices to indicate the current energy level of each electron.

$$\frac{n_e n_{i+1,j_1,\ldots j_{N-1}}}{n_{i,j_1,\ldots,j_{N-1},j_N}} = \frac{g_e g_{i+1,j_1,\ldots,j_{N-1}}}{g_{i,j_1,\ldots,j_{N-1},j_N}} \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} \exp\left(-\frac{\chi_{i,n_N}}{kT} \right)$$

Where there are N electrons present in ionization state i and $N-1$ present in ionization state $i+1$.

Let us consider 3 species, neutral hydrogen in the ground state $n_{1,1}$, neutral hydrogen in the first excited state $n_{1,2}$ and ionized hydrogen $n_{2,.}$. Note there is no second label on the last. For convenience we will use the following notation instead of two indices:

$$n_+ = n_{2,.}$$

$$n_1 = n_{1,1}$$

$$n_2 = n_{1,2}$$

So that in the notation below, an integer subscript indicates the energy level within the bound H atom.

The Boltzmann equation that the text uses to relate ionization states can be recovered by taking

$$\frac{n_2}{n_1} = \frac{n_+/n_1}{n_+/n_2} = \frac{g_2}{g_1}\exp\left(-(\chi_1-\chi_2)/kT\right) = \frac{g_2}{g_1}\exp\left(-(\chi_1-\chi_2)/kT\right)$$

since all the prefactors cancel, and then $\chi_1 - \chi_2$ is the difference in energy levels, which is also the difference in ionization energies.

If we would like to know the ionization fraction

$$\frac{n_+}{n_{tot}} = \frac{n_+}{n_1+n_2+n_+} = \left(\frac{n_1}{n_+}+\frac{n_2}{n_+}+1\right)^{-1}$$

Where

$$\frac{n_+}{n_1} = \frac{g_e g_+}{n_e g_1}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi_1/kT}$$

Here $g_e = 2$ and $g_1 = 2$, for the electron in each, and g_+ is taken as 1 (actually assume proton is the same in all so it would cancel). Also $n_e = P_e/kT$. so

$$\frac{n_+}{n_1} = \frac{kT}{P_e}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi_1/kT}$$

and with $g_2 = 8$ (1 for 2s, 3 for 2p, times 2 for electron spin)

$$\frac{n_+}{n_2} = \frac{kT}{4P_e}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi_2/kT}$$

Then

$$\frac{n_1}{n_+}+\frac{n_2}{n_+} = \frac{n_1+n_2}{n_+} = \frac{P_e}{kT}\left(\frac{2\pi m_e kT}{h^2}\right)^{-3/2}\left(e^{\chi_1/kT}+4e^{\chi_2/kT}\right)$$

The text assumes $e^{\chi_1/kT} \gg e^{\chi_2/kT}$ for temperatures of interest where $\chi_1 = 13.6$ eV is much bigger than $\chi_2 = 13.6/2^2 = 3.4$ eV, which is a good approximation. But we see this assumption is not really necessary, since everything in the above equation is known.

Plugging in numbers for 8300 K we get $n_+/n_{tot} = 0.05$, so that 5% of the hydrogen is ionized. At 9600 K, we get 0.52 and at 11300 we get 0.95.

Plots from text:

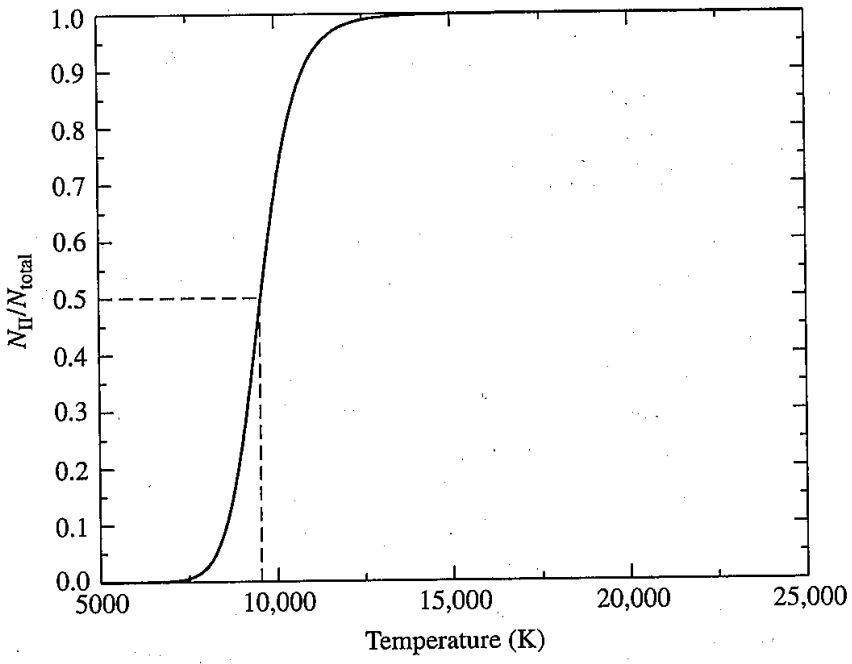


FIGURE 8.8 N_{II}/N_{total} for hydrogen from the Saha equation when $P_e = 20$ N m^{-2}. Fifty percent ionization occurs at $T \simeq 9600$ K.

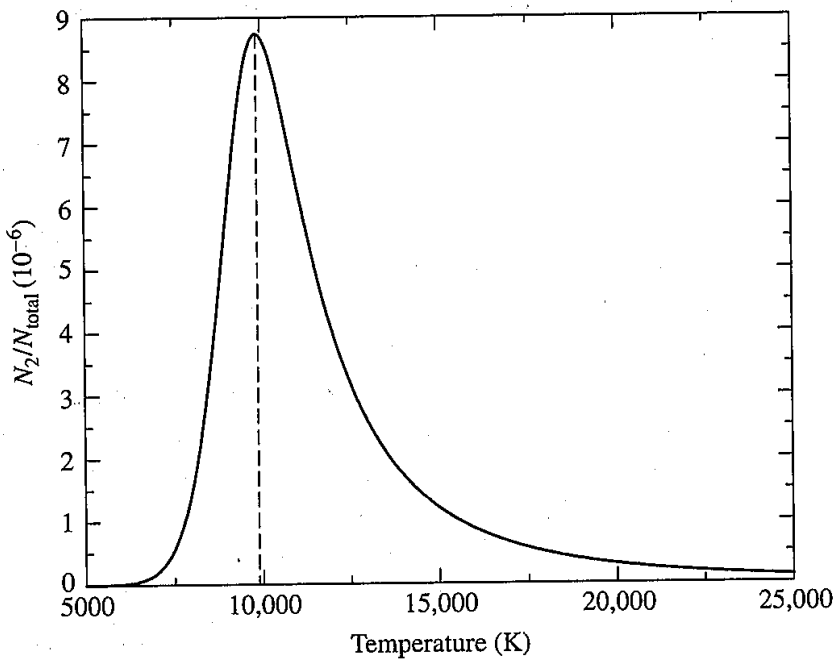


FIGURE 8.9 N_2/N_{total} for hydrogen from the Boltzmann and Saha equations, assuming $P_e = 20$ N m^{-2}. The peak occurs at approximately 9900 K.

13 Astro notes 2018/9/21 - Fri - selection rules; binaries

Binaries will not be on first exam. Just through light and matter.

Some homework things:

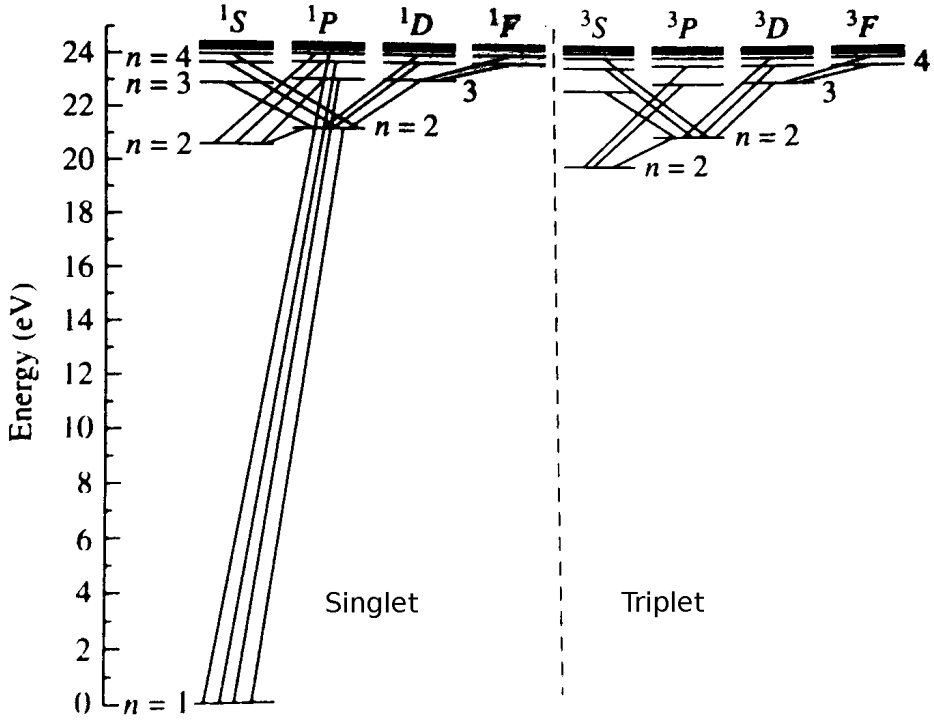
What is meant by integrating an ODE

Briefly revisit relativity problem: not when ship is halfway! Helps to distinguish between "coordinate time", which is different in the two frames, and "elapsed time" which can only be determined for a specific path. Physics problems typically use coordinate time and then compute the time elapsed on a clock moving on some path by doing an integral in those coordinates.

13.1 Selection rules and forbidden/permitted lines

When a transition occurs in an atom, it is necessary for the photon produced to have an angular momentum of 1. This restricts the transitions to energy levels which have a difference in angular momentum of $\Delta\ell = 1$. These are called "permitted" transitions, while the converse are "forbidden" transitions. Forbidden transitions can still occur, but the relevant levels have much longer lifetimes and are therefore "weaker" emitters. Typically an emission is possible via higher-order effects (e.g. a non-dipole photon emission), but is therefore much slower to occur. Also interaction with matter can easily de-excite these lines.

Below is a permitted line transition diagram for helium. Figure 5.14 in text.



13.2 Binary Stars

Several major types:

- resolved
  - "visual binaries" where both stars are seen in motion – both masses can be determined.
  - "astrometric binaries" where one star is too dim to see
- unresolved
  - spectroscopic – doppler shift of one or both spectra observed – can get total mass up to inclination factor
  - eclipsing, spectroscopic – everything determined including stellar radii

Why is this important? Critical to measure mass-luminosity relation in order to understand the stellar main sequence. i.e. to test that stellar luminosity is related to age and initial mass, not just age.

Counting free parameters:
Given observations of a binary, we would like to determine the masses of the stars. But how many free parameters must be determined in order to do that? Kepler's law says

$$P^2 = \frac{4\pi^2}{G(m_1 + m_2)} a^3$$

and so each star forms an ellipse of size $a_1 = (m_2/M)a$ and $a_2 = (m_1/M)a$. If we can measure P, a_1 and a_2 this gives us three equations and three unknowns, m_1, m_2 and a. Said another way, we must measure the orbital period and the orbits of BOTH stars in order to get both masses.

This sounds easy, however there is one significantly complicating factor: inclination, typically denoted i. We will come back to this. Also the binary will be moving with some constant velocity, but that is relatively straightforward to compensate for.

First note that the problem is often rephrased as determining the mass ratio $m_1/m_2 = q$ and the total mass $M = m_1 + m_2$ instead of discussed in terms of m_1 and m_2 directly. This is because these are what have different difficulties of measurement. m_1/m_2 can be measured from

$$q = \frac{m_1}{m_2} = \frac{a_2}{a_1}$$

while, since $a = a_1 + a_2$,

$$M = m_1 + m_2 = \frac{4\pi^2}{GP^2}(a_1 + a_2)^3$$

13.3 Resolved binary, simplest geometry

In the simplest geometry, a resolved binary will have orbits in the plane of the sky. This is defined as $i = 0$. In this case we can measure the angles $\alpha_1 \approx a_1/d$ and $\alpha_2 \approx a_2/d$, where d is the distance to the star. Even if d is not known, the mass ratio can be determined

$$q = \frac{m_1}{m_2} = \frac{\alpha_2}{\alpha_1}$$

(Student: find total mass)

$$M = m_1 + m_2 = \frac{4\pi^2}{GP^2}(\alpha_1 + \alpha_2)^3 d^3$$

The total mass, then, does require the distance to set the overall scale.

But this example is essentially useless, as binaries nearly always have **inclination** from the line of sight. Consider the simplest inclination, which is about the semiminor axis (note that there is actually both an inclination and orientation). In this case the angles measured are based on the projected distances so that

$$\alpha_1 = \frac{a_1 \cos_i}{d} \quad \text{and} \quad \alpha_2 = \frac{a_2 \cos i}{d}$$

(note my notation is a bit different from the book, as α is still the angle observed, but it is just no longer a/d when there is nonzero inclination). The mass ratio is still straightforward:

$$q = \frac{a_2}{a_1} = \frac{\alpha_2 d/\cos i}{\alpha_1 d/\cos_i} = \frac{\alpha_2}{\alpha_1}$$

We will address the total mass next time.

14 Astro notes 2018/9/24 - Mon - binaries 2

hints for line identification problem - consecutive lines, lines should shift "together" i.e. there is a relationship among the lines that is independent of Doppler shift.

Last time we found that the total mass could be obtained from the semimajor axes, a_1 and a_2, of the orbits of two stars orbiting with period P.

$$M = m_1 + m_2 = \frac{4\pi^2}{GP^2}(a_1 + a_2)^3$$

14.1 Inclination

Last time we were considering the simplest inclination, which is about the semiminor axis (note that there is actually both an inclination and orientation). In this case the angles measured are based on the projected distances so that

$$\alpha_1 = \frac{a_1 \cos_i}{d} \quad \text{and} \quad \alpha_2 = \frac{a_2 \cos i}{d}$$

(note my notation is a bit different from the book, as α is still the angle observed, but it is just no longer a/d when there is nonzero inclination). The mass ratio is still straightforward:

$$q = \frac{a_2}{a_1} = \frac{\alpha_2 d/\cos i}{\alpha_1 d/\cos_i} = \frac{\alpha_2}{\alpha_1}$$

but the total mass now contains both the distance and inclination:
(student)

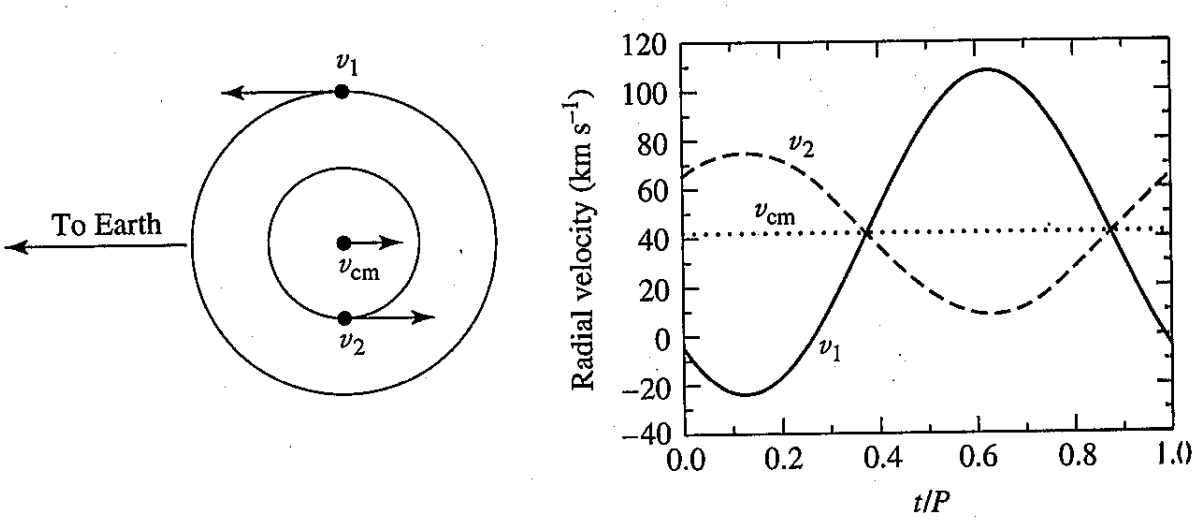
$$M = \frac{4\pi^2}{GP^2}\left(\frac{d}{\cos i}\right)^3 (\alpha_1 + \alpha_2)^3$$

It turns out that the inclination can be determined from the projected orbital shape because the the focus of the projected ellipse is different from the projected focus of the actual ellipse. The latter can be determined from the motion of the stars about the center of mass (which is the projection of the focus of the ellipse.)

Note that an alternative to knowing the distance is to know the radial velocities

14.2 Spectroscopic binaries

In spectroscopic binaries, which are the more common case, the Doppler shifts of the two stars are measered in time.



The circular orbit case is actually fairly common due to tidal dissipation in shorter-period binaries. If the binary has $i = 90°$ then the peak speeds will be determined by
(student)

$$v_1 = \frac{2\pi a_1}{P} \quad \text{and} \quad v_2 = \frac{2\pi a_2}{P}$$

From this the mass ratio is just
(student continued)

$$q = \frac{m_1}{m_2} = \frac{v_2}{v_1}$$

Note that inclinition does not adversely effect this.

Now with arbitrary inclination:

If we define the observed line-of-sight velocity as $v_{1r} = v_1 \sin i$, and similar for v_{2r} we get

$$q = \frac{m_1}{m_2} = \frac{v_{2r}/\sin i}{v_{1r}/\sin i} = \frac{v_{2r}}{v_{1r}}$$

The total mass does require the inclination:

$$M = \frac{P}{2\pi G}\frac{(v_{1r} + v_{2r})^3}{\sin^3 i}$$

While the mass can't be know without i, this is still useful statistically since the orientation of the system to the line of sight is **random**, thus the statistics of $\sin i$ are known, in particular its average value is about 2/3, though one must be careful about selection bias with small radial velocities given by low inclination.

For systems in which the second star is dim, one may only know v_{1r}. In this case it is convenient to eliminate v_2 in favor of m_1/m_2. This gives what is sometimes called the **mass function**

$$\frac{m_2^3}{(m_1 + m_2)^2}\sin^3 i = \frac{P}{2\pi G}v_{1r}^3$$

Note that the right hand side provides a lower limit on m_2 since

$$m_2 \geq m.f.\frac{(m_1 + m_2)^2}{m_2^2 \sin i} \geq m.f. = \frac{P}{2\pi G}v_{1r}^3$$

since $m_1 > 0$ and $\sin i < 1$. This is useful for black holes, for example, in which only the star is visible and can have its radial velocity measured, but this gives a secure lower limit on the mass of the black hole.

Alternatively, if we assume $m_2 \ll m_1$, and that we know m_1, as for an extrasolar planet, we find that

$$m_2 \sin i \approx \left(\frac{P}{2\pi G}v_{1r}^3 m_1^2\right)^{1/3}$$

this again gives a lower limit on the mass, has known statistics for random orientations, and is often quoted for extrasolar planatary systems.

14.3 Discussion of computing radial velocity curve shape

For more realistic systems with eccentricity things are more complicated

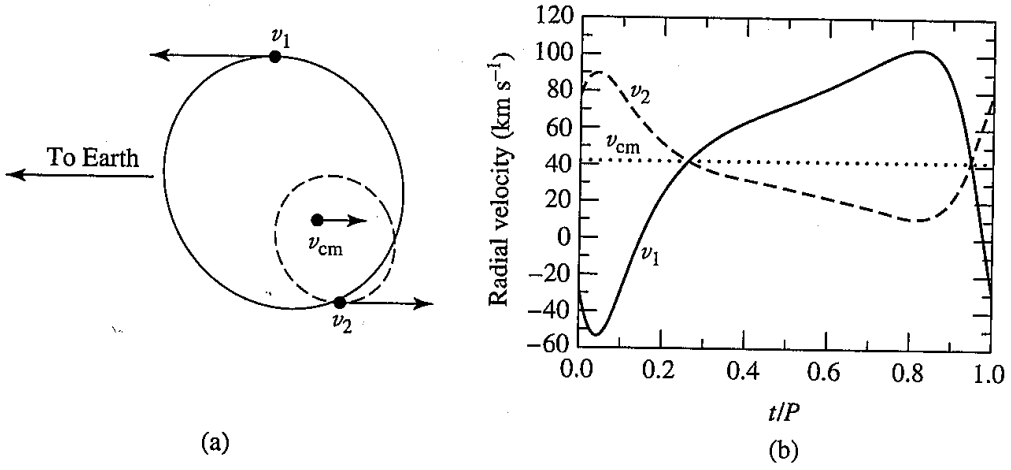


FIGURE 7.6 The orbital paths and radial velocities of two stars in elliptical orbits ($e = 0.4$). As in Fig. 7.5, $M_1 = 1\ M_\odot$, $M_2 = 2\ M_\odot$, the orbital period is $P = 30$ d, and the radial velocity of the center of mass is $v_{cm} = 42$ km s^{-1}. In addition, the orientation of periastron is 45°. v_1, v_2, and v_{cm} are the velocities of Star 1, Star 2, and the center of mass, respectively. (a) The plane of the orbits lies along the line of sight of the observer. (b) The observed radial velocity curves.

Discussion exercise:
Imagine that you were given the task of fitting a bunch of radial velocity data. How would you compute a model radial velocity curve from the information that we know so far? That is imagine we know can compute $\vec{v}_1(t)$ and $\vec{v}_2(t)$ for any parameters, how do we get v_{1r} and v_{2r} to compare to observations?

(set up the geometry and understand how to compute things as a function of time – can pretty much be done by choosing geometry in a concrete way, placing the line-of-sight (toward obsever) in that geometry, and taking dot products to evaluate radial velocities)

15 Astro notes 2018/9/26 - Wed - eclipses; spectral parameters

15.1 Eclipsing

Refer to figures from text for timing definitions.

For exclipsing stars, the inclination is very close to edge-on, $i = 90°$, as long as the separation between stars is more than a few stellar radii, which is the normal case.

In this case it is also possible to measure the **radius** of the stars as well. (subscript s is smaller (in radius) and ℓ is larger) The radius of the smaller star is

$$r_s = \frac{v}{2}\Delta t_{transition}$$

where $v = v_s + v_l$ and $\Delta t_{transition}$ is the time for the transition into eclipse.

The size of the larger star can also be measured from the duration of the transit

$$r_\ell = \frac{v}{2}(\Delta t_{duration})$$

where the duration is measured from the time the smaller star just begins to enter the projected area of the larger, until the smaller just begins to emerge.

Eclipse depth The depth of the eclipses is determined by the relative temperatures of the components, because it depends on their relative surface fluxes. In both cases (smaller in front or in back) the same area is covered up, it is just a question of whether this area is on the hotter or cooler star. (Note that generally for core-hydrogen-fusion stars bigger is higher T. But if one star is giant, this is the opposite.) If we take F_ℓ and F_s to be the surface fluxes of the larger and smaller stars, we can compute the flux outside eclipse:

$$B_0 = k(\pi r_\ell^2 F_l + \pi r_s^2 F_s)$$

Where k is a constant that depends on distance. The primary minimum happens when the hotter star is eclipsed by the cooler, for now lets assume that hotter star is the smaller one, so

$$B_p = k\pi r_\ell^2 F_\ell$$

and

$$B_s = k(\pi r_\ell^2 - \pi r_s^2)F_\ell + k\pi r_s^2 F_s$$

We can then take the ratio of the *depths* of the eclipses

$$\frac{B_0 - B_p}{B_0 - B_s} = \frac{F_s}{F_\ell} = \frac{T_s^4}{T_\ell^4}$$

15.2 Parameters of stellar spectra

We have already learned that temperature can effect both the broadband shape of the spectrum, as in the Planck function, as well as the absorbtion features via the ionization state.

Overall stellar spectra have 2 major parameters:
Flux density, $F = \sigma_{\rm SB} T_{\rm eff}^4 \leftrightarrow$ spectral type OBAFGKM
surface gravity $g \leftrightarrow$ Luminosity class I through V, where I is lowest gravity, largest R
Where the thing on the left is the actual theoretical parameter, and the thing on the right is the corresponding observational classification

Composition is also important, but most stellar photospheres are not too diffferent in composition from the Sun.

In observations, $T_{\rm eff}$ corresponds to the spectral type. In order these are OBAFGKM from highest to lowest temperature.

Figure 8.4 and 8.5 from Carroll and Ostlie show examples of various spectral types:

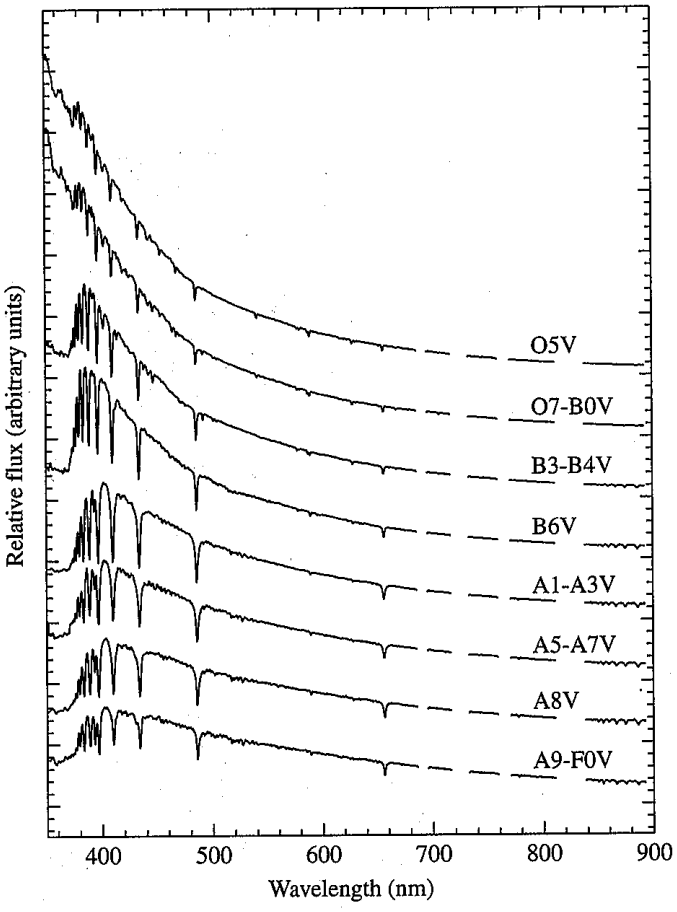


FIGURE 8.4 Digitized spectra of main sequence classes O5–F0 displayed in terms of relative flux as a function of wavelength. Modern spectra obtained by digital detectors (as opposed to photographic plates) are generally displayed graphically. (Data from Silva and Cornell, *Ap. J. Suppl.*, *81*, 865, 1992.)

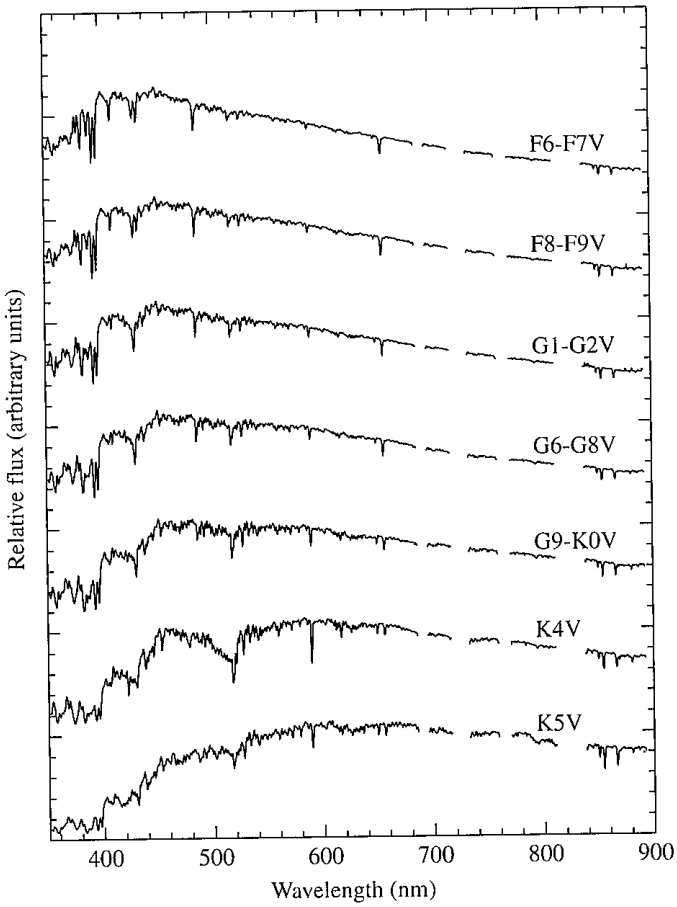


FIGURE 8.5 Digitized spectra of main sequence classes F6–K5 displayed in terms of relative flux as a function of wavelength. (Data from Silva and Cornell, *Ap. J. Suppl.*, *81*, 865, 1992.)

The strength of absorption features of various elements (and molecules at the low temperature end) is strongly dependent on the stellar temperature, making this a good diagnostic of the temperature. The dependence is shown in figure 8.11:

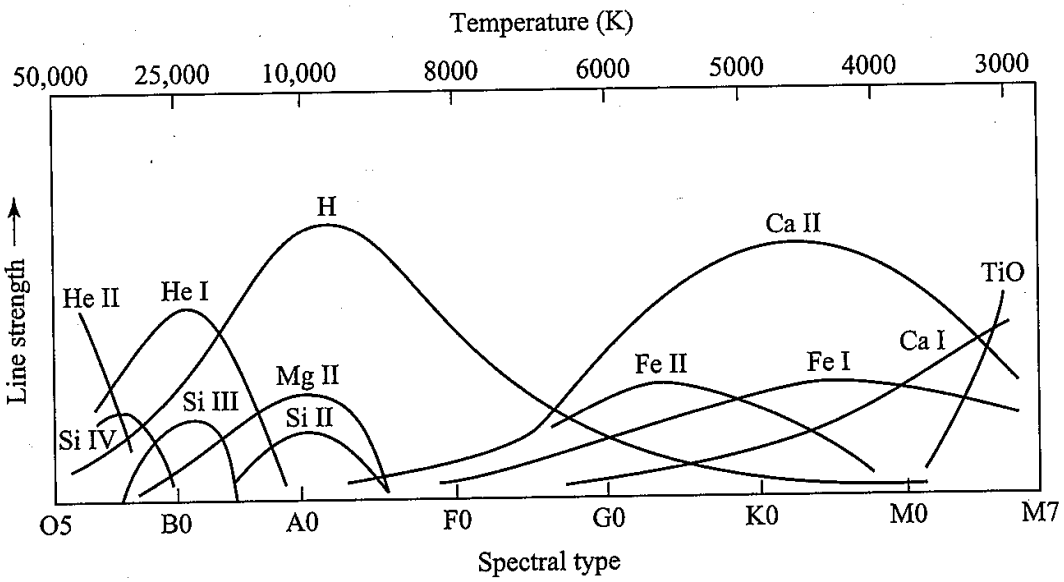


FIGURE 8.11 The dependence of spectral line strengths on temperature.

Strength of absorption features is determined by population of levels in the atoms near the photosphere. The hydrogen absorbtion lines in the optical are Balmer, and so based off of $n = 2$. At high temperatures, most of the H is ionized, so the lines are weak. At low temperature, most of the atoms' elecrons are in $n = 1$, and so the Balmer lines are weak there as well. Their strength peaks around 8000 K, which corresponds roughly to an A star.

Surface gravity is measured by luminosity classes: V is highest gravity, I is lowest gravity = largest star.

Figure 8.15 shows how different "broadness" of lines can manifest in the traditional spectral display:

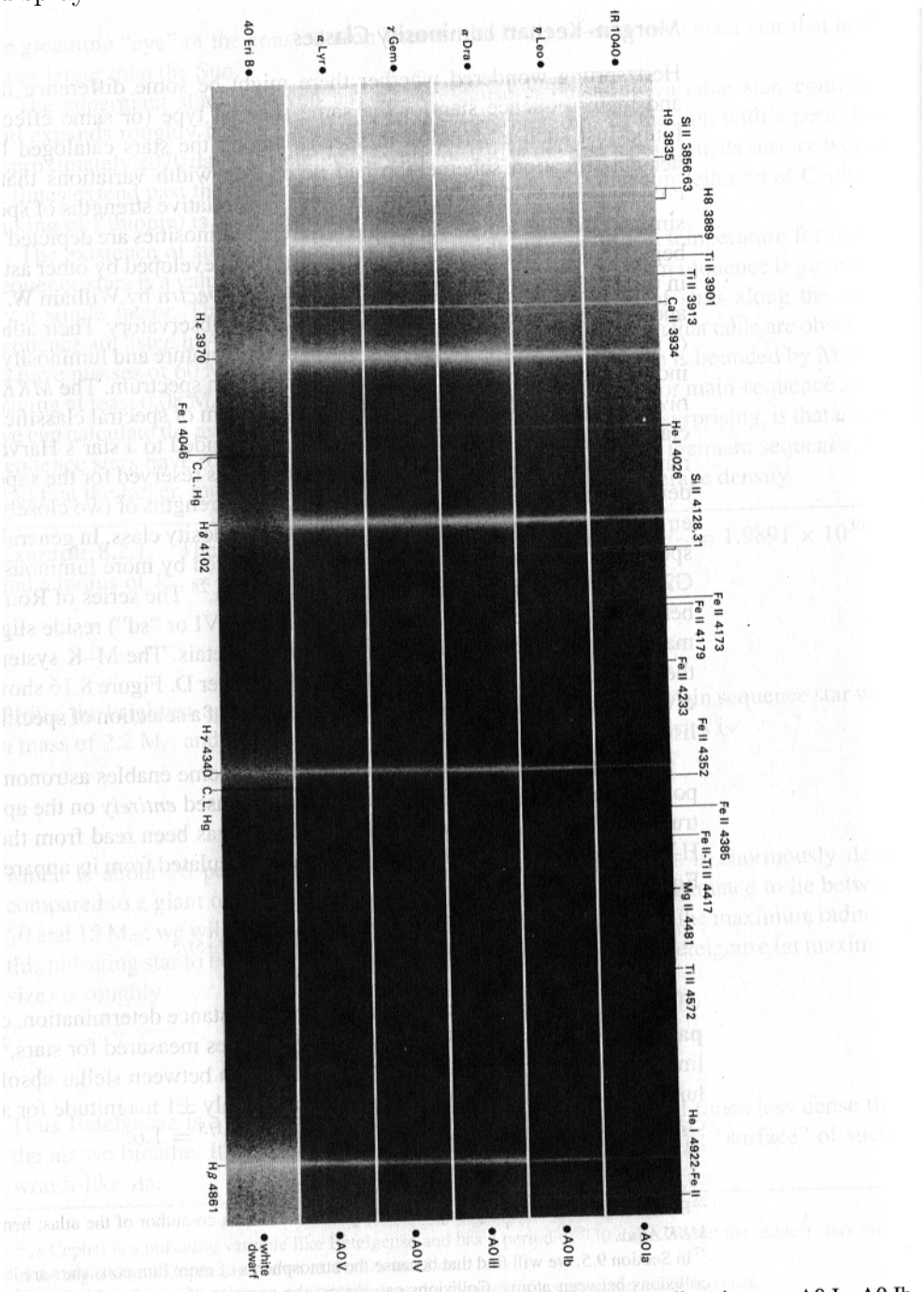


FIGURE 8.15 A comparison of the strengths of the hydrogen Balmer lines in types A0 Ia, A0 Ib, A0 III, A0 IV, A0 V, and a white dwarf, showing the narrower lines found in supergiants. These spectra are displayed as negatives, so absorption lines appear bright. (Figure from Yamashita, Nariai, and Norimoto, *An Atlas of Representative Stellar Spectra*, University of Tokyo Press, Tokyo, 1978.)

Stars are typically placed on an HR diagram - a scatter plot of absolute magnitude (Luminosity) and spectral type (temperature).

Surface gravity also appears on the H-R diagram in which the main sequence is luminosity class V and the giants are lower luminosity class.

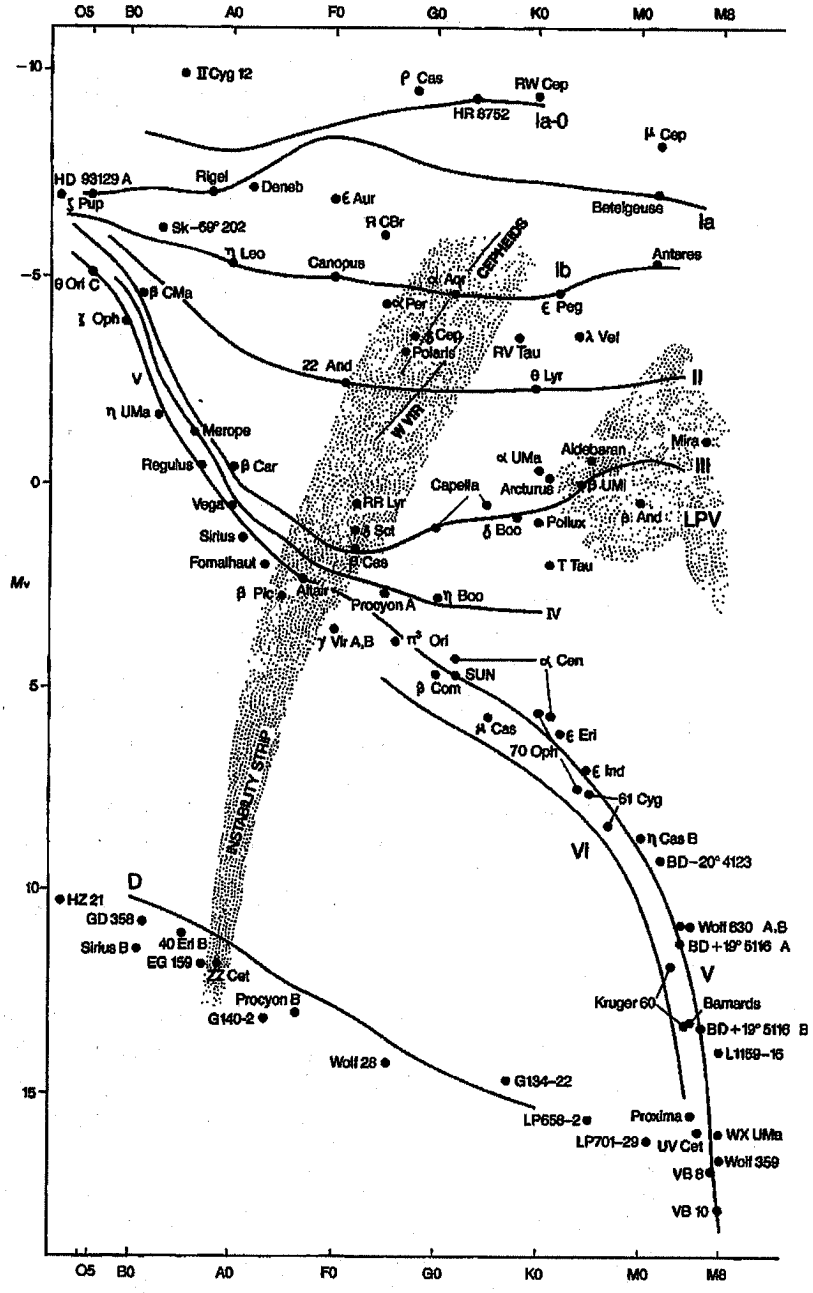


FIGURE 8.16 Luminosity classes on the H–R diagram. (Figure from Kaler, *Stars and Stellar Spectra*, © Cambridge University Press 1989. Reprinted with the permission of Cambridge University Press.)

For higher gravity, absorption lines appear more broad. This is because the density in the atmosphere increases more quickly with depth at higher gravity, so the lines are

broadened more by the pressure. So giant stars, being very large, have low surface gravity due to their large size.

Note tha the H-R diagram (based on stellar observations) is more typically thought of as a Luminosity-Temperature diagram.

Note that this is also an T - R diagram since $L = 4\pi R^2 \sigma T^4$.

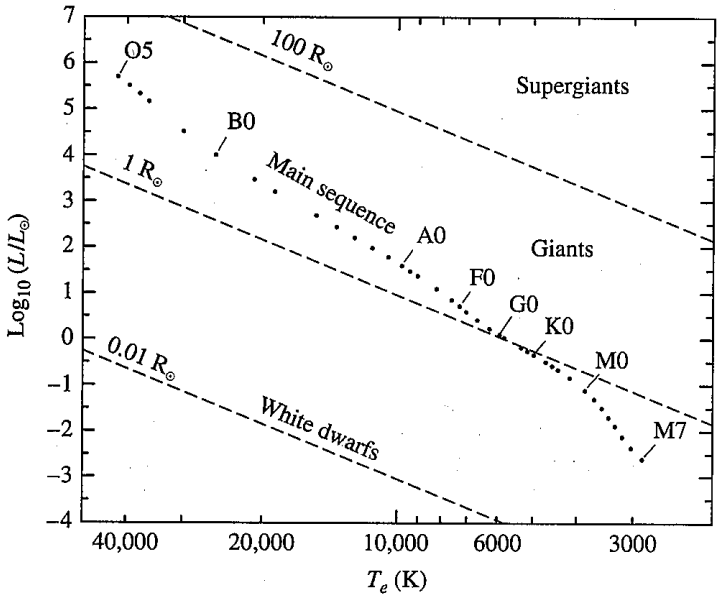


FIGURE 8.14 The theorist's Hertzsprung–Russell diagram. The dashed lines indicate lines of constant radius.

Also can compare to a color-magnitude diagram, shown in 8.13.

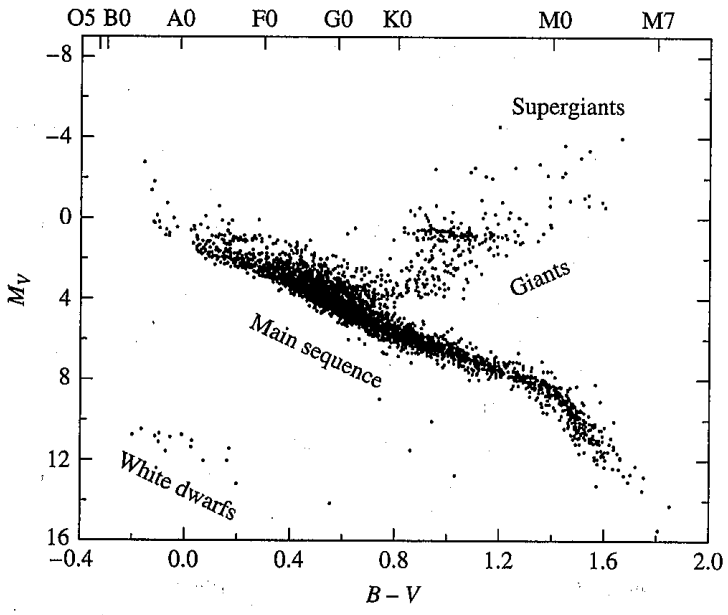


FIGURE 8.13 An observer's H–R diagram. The data are from the Hipparcos catalog. More than 3700 stars are included here with parallax measurements determined to better than 20%. (Data courtesy of the European Space Agency.)

Color, being a ratio of fluxes, indicates the slope of the spectrum, and is therefore fairly monotonically dependent on temperature. It is easier to measure two broad-band fluxes to get color, so it is common to have many stars on color-magnitude diagram. For nearby stars the magnitude is absolute magnitude (since the distance can be measured) while for clusters often just the magnitude in one of the bands used for the color is used, since all the stars are at the same distance.

16 Astro notes 2018/9/28 - Fri - Star formation

Some discussion of exam and homework just turned in.

discussed boost direction in doppler boost question and pressure balance in ISM question.

16.1 HR and CM diagrams

Note tha the H-R diagram (based on stellar observations) is more typically thought of as a Luminosity-Temperature diagram.

Note that this is also a T - R diagram since $L = 4\pi R^2 \sigma T^4$.

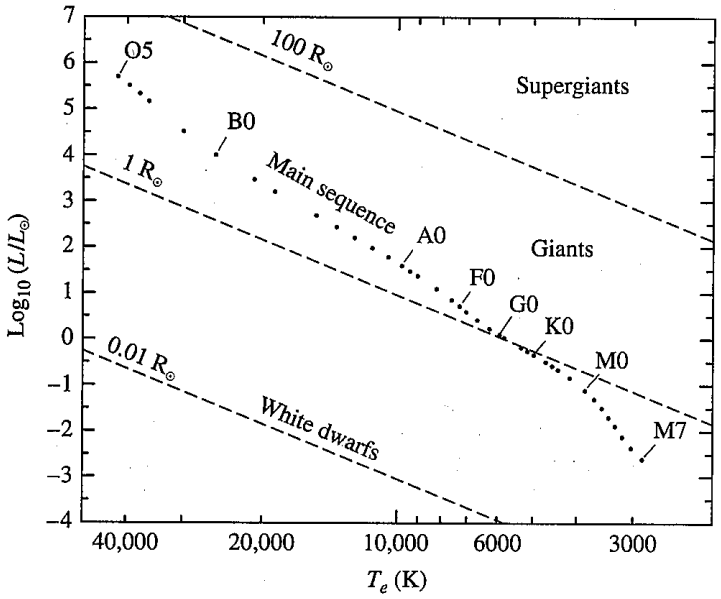


FIGURE 8.14 The theorist's Hertzsprung–Russell diagram. The dashed lines indicate lines of constant radius.

Also can compare to a color-magnitude diagram, shown in 8.13.

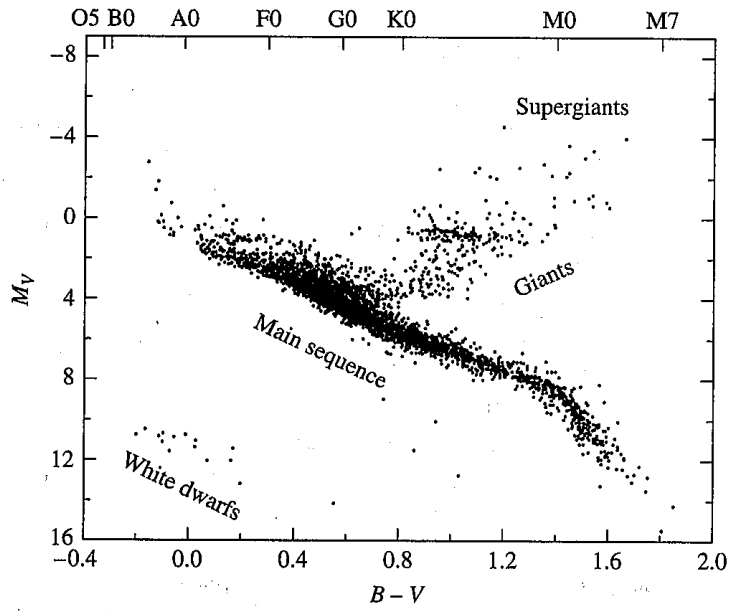


FIGURE 8.13 An observer's H–R diagram. The data are from the Hipparcos catalog. More than 3700 stars are included here with parallax measurements determined to better than 20%. (Data courtesy of the European Space Agency.)

Color, being a ratio of fluxes, indicates the slope of the spectrum, and is therefore fairly monotonically dependent on temperature. It is easier to measure two broad-band fluxes to get color, so it is common to have many stars on color-magnitude diagram. For nearby stars the magnitude is absolute magnitude (since the distance can be measured) while for clusters often just the magnitude in one of the bands used for the color is used, since all the stars are at the same distance.

16.2 Star formation and Jean's mass

Take a piece of the ISM with density ρ and T which has mass M over radius R. the gravitational energy is

$$E_{GR} \sim -\frac{GM^2}{R}$$

Typical densities in a star forming clound are $n \simeq 100/\text{cc}$.

Want to ask: for a known density n and temperature T of a cloud, how big (in mass) does the cloud need to be to collapse under its own gravity?

Total KE content of that region is

$$E_{th} \simeq \frac{3}{2}\frac{M}{m_p}kT$$

critical condition (Jeans length or Jeans mass) set by

$$E_{GR} + E_{th} < 0$$

i.e. when cloud is bound.
(student): put in and eliminate R in favor of $\rho = m_p n$
Putting in the energies this is

$$\frac{GM^2}{R} > \frac{3}{2}\frac{M}{m_p}kT$$

Rewrite in terms of density

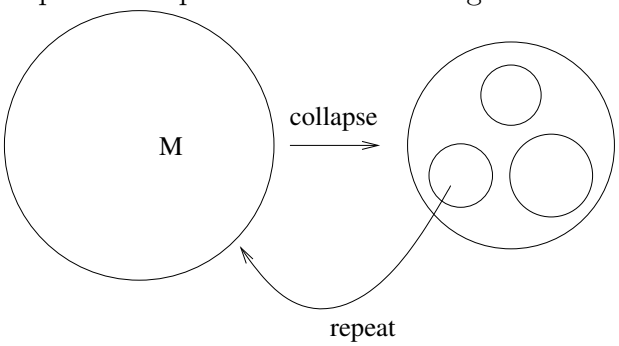
$$M = \frac{4\pi}{3}R^3\rho$$

or

$$M_J = 500M_\odot \left(\frac{T}{10^4\text{ K}}\right)^{3/2}\left(\frac{1\text{ cm}^{-3}}{n}\right)^{1/2}$$

where $n = \rho/m_p$. In a region with T, ρ a mass in excess of this can collapse.

The problem is this is big. How do we get from this to the distribution of stellar masses from this? How does a collapsing mass fragment into $M_\odot$ chunks? We will simply try to answer why would it fragment at all? The Jean's mass scales as $M_J \propto T^{3/2}\rho^{-1/2}$. Imagine that the region keeps it's same temperuture as it collopses. Then M_J decreases during the collapse. Then as the jeans mass decreases this can allow for "fragmentation" or the subsequent collapse of less massive regions.



Next time we'll talk about halting fragmentation to get stars.

17 Astro notes 2018/10/3 - Wed - Star formation, protostars

Mention SPS meeting tonight 5-6pm room 227.

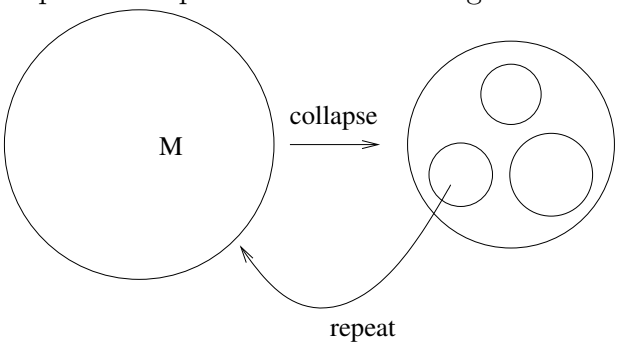
17.1 Star formation and fragmentation

we found the Jeans mass:

$$M_J = 500M_\odot \left(\frac{T}{10^4\text{ K}}\right)^{3/2}\left(\frac{1\text{ cm}^{-3}}{n}\right)^{1/2}$$

where $n = \rho/m_p$. In a region with T, ρ a mass in excess of this can collapse.

The problem is this is big. How do we get from this to the distribution of stellar masses from this? How does a collapsing mass fragment into $M_\odot$ chunks? We will simply try to answer why would it fragment at all? The Jean's mass scales as $M_J \propto T^{3/2}\rho^{-1/2}$. Imagine that the region keeps it's same temperuture as it collopses. Then M_J decreases during the collapse. Then as the jeans mass decreases this can allow for "fragmentation" or the subsequent collapse of less massive regions.



So what halts the fragmentation. As density is rising in the fragments, they eventually become optically thick and so can't necessarily radiate on the collapse time. After it becomes adiabatic. Then $T \propto \rho^{2/3}$. Then the Jeans mass, $M_J \propto \sqrt{\rho}$, which is increasing as collapse continues. This shuts off fragmentation.

Halting of fragmentation is then due to the isothermal-adiabatic transition. Hard because you have to do both the dynamics and radiative transfer.

What's a characteristic timescale? The dynamical time. This can be obtained by considering a simplified version of free-fall. Consider the acceleration of a particle at the surface of the star if pressure support is removed:

$$\frac{d^2R}{dt^2} = a = -\frac{GM}{R^2}$$

If we use this acceleration as if it were constant we can estimate the time it takes for the particle to cross a distance R:

$$\frac{1}{2}at_{\rm dyn}^2 \approx -R \implies t_{\rm dyn}^2 \simeq \frac{R^3}{GM}$$

One conventional way to write this, which is similar up to factors of order unity, is

$$t_{dyn} = \frac{1}{\sqrt{G\rho}} \simeq \frac{10^7 yrs}{(n/100)^{1/2}}$$

17.2 Halting Fragmentation

Want to estimate when fragmentation of the protostellar cloud into smaller and smaller chunks halts. This should determine the approximate mass scale at which individual stars would form, since the cloud could not fragment into any smaller pieces. In reality a wide

range of masses are formed, but there is a characteristic scale. Estimate by finding when the freefall luminosity is similar to what can be emitted by radiative transfer. That is

$$L_{ff} \sim L_{rad}$$

Can make a simple estimate assume energy GM^2/R is released on collapse (free fall) time. This gives
(student)

$$L_{ff} \sim \frac{E_{grav}}{t_{dyn}} \sim \frac{GM^2/R}{1/\sqrt{G\rho}} \sim G^{3/2}\frac{M^{5/2}}{R^{5/2}}$$

The time when the medium becomes optically thick corresponds to when some modest fraction of the flux is carried by radiative diffusion

$$L_{rad} = e4\pi R^2 \sigma T^4$$

where e is some fudge efficiency factor. Then if we equate these two

$$L_{ff} \sim L_{rad}$$

get

$$M^{5/2} \sim R^{9/2} G^{-3/2} e\sigma T^4$$

The last ingredient is that for the collapsing mass

$$\frac{GM^2}{R} \sim \frac{M}{\mu m_p} kT$$

which was used to get the Jeans mass. Here μ is the mean weight of each particle in units of m_p, close to 1 since material is mostly hydrogen. This allows us to relate R to M with a factor of T, $R \propto M/T$. So that

$$M^{5/2} \propto M^{9/2}/T^{1/2}$$

or, with numbers,

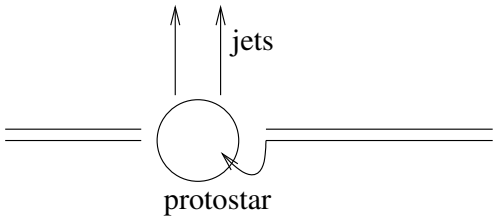
$$M_{iso-adiab} \sim 0.03 M_\odot \left(\frac{1}{e^{1/2}\mu^{9/4}}\right)\left(\frac{T}{10^4 \text{ K}}\right)^{1/4}$$

17.3 Protostars

But on this same timescale there is evidence that the "protostar" accretes more material from the surrounding cloud. The luminosity that such an object has is

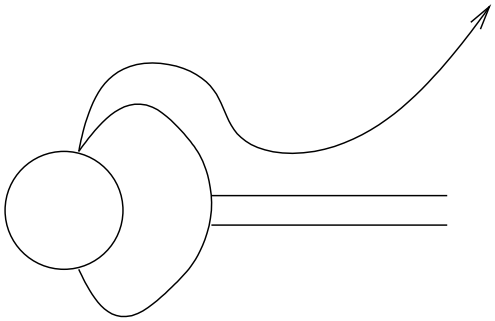
$$L = \dot{M}\frac{GM}{R}$$

just accretion luminosity. The picture is a disk with a jet.



For $M \simeq M_\odot$ the first $\sim (1-5)$ Myrs is this accretion stage.

Why the jets? Two questions: what launches them, and what collimates them. The resulting stars are NOT spinning very fast. The field lines truncate the disk, and then those further out go to infinity.



The problem is figuring out how incoming material chooses between the closed field line and the open one.

See figures in Feigelson & Montmerle 1999, ARA&A, 37, 363

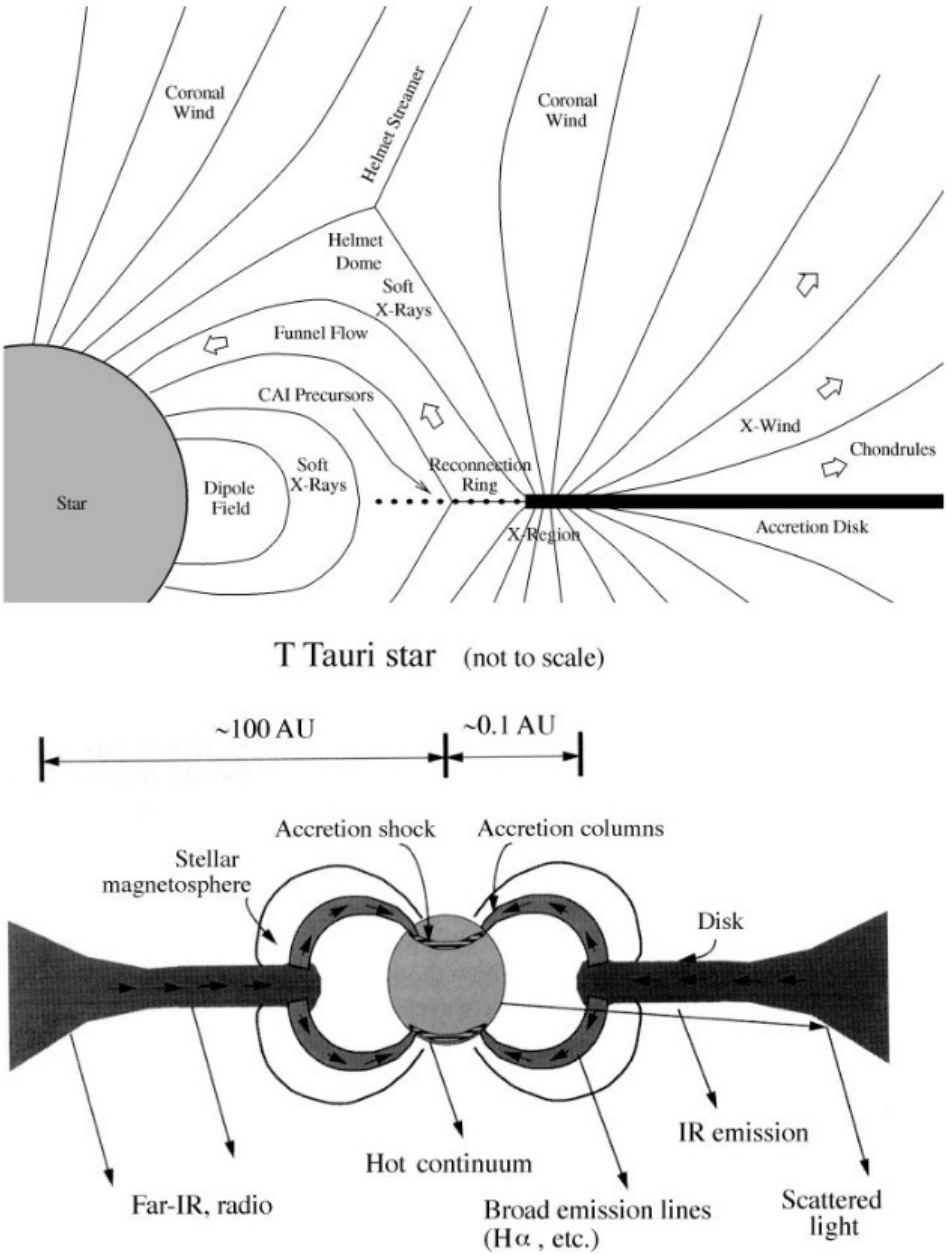


Figure 3 Two contemporary models for Class I–II YSOs, in which magnetic fields play crucial roles: (*top*) the x-wind model of YSOs showing magnetically collimated accretion and outflows with irradiated meteoritic solids (Shu et al 1997); (*bottom*) magnetically funneled accretion streams producing broadened emission lines (Hartmann 1998).

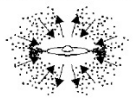
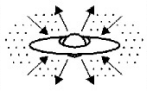
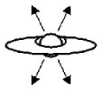
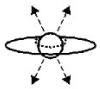
PROPERTIES	*Infalling Protostar*	*Evolved Protostar*	*Classical T Tauri Star*	*Weak-lined T Tauri Star*	*Main Sequence Star*
SKETCH					
AGE (YEARS)	10^4	10^5	10^6 - 10^7	10^6 - 10^7	$> 10^7$
mm/INFRARED CLASS	Class 0	Class I	Class II	Class III	(Class III)
DISK	Yes	Thick	Thick	Thin or Non-existent	Possible Planetary System
X-RAY	?	Yes	Strong	Strong	Weak
THERMAL RADIO	Yes	Yes	Yes	No	No
NON-THERMAL RADIO	No	Yes	No ?	Yes	Yes

Figure 1 The stages of low-mass young stellar evolution. This review chiefly addresses the bottom three rows of the chart. (Adapted from Carkner 1998.)

18 Astro notes 2018/10/5 - Fri - Protostars

18.1 Pre Main sequence contraction

Eventually (1-10 Myr) accretion ceases and we can consider a hydrostatic object of fixed M. These objects are fully convective from core to photosphere. This is necessitated by the boundary conditions, radiative transfer just won't work to build a consistent model.

Hayashi showed that a proper treatment of the outer boundary condition actually mattered (unlike for sun's gross structure).

Because of this objects decrease in luminosity at constant temperature.

on HR diagram:

THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 192:3 (35pp), 2011 January

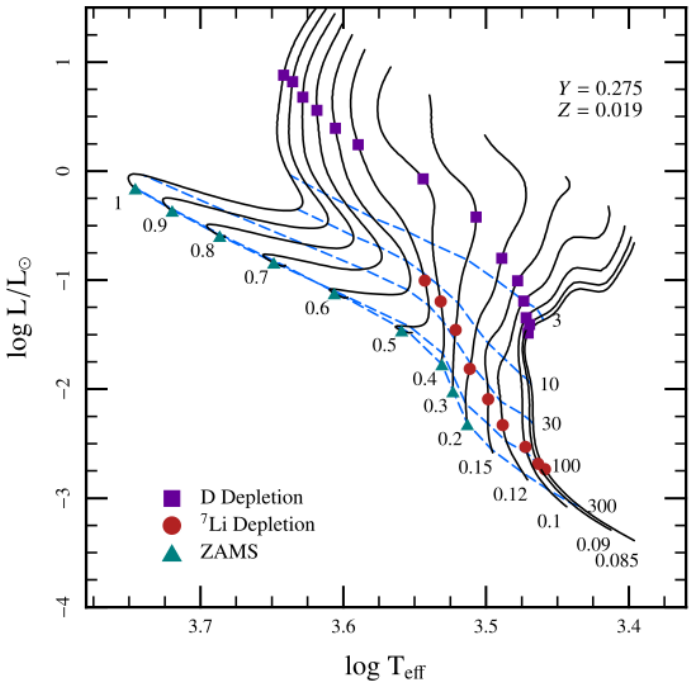


Figure 15. Location in the Hertzsprung-Russell (H-R) diagram for $0.085\,M_\odot < M < 1\,M_\odot$ stars as they arrive at the main sequence for $Y = 0.275$ and $Z = 0.019$. The mass of the star is noted by the values at the bottom of the line. The dashed blue lines are isochrones for ages of 3, 10, 30, 100 and 300 Myr, as noted to the right. The purple squares (red circles) show where D (^{7}Li) is depleted by a factor of 100. The green triangles show the ZAMS.

The temperature scale of around 1500 K is the minimum for photons to be able to be produced efficiently. Makes opacity very steep function of T near 2-3000 K.

18.2 Energy budget and timescale

The protostar lives on contraction energy – gravitational potential energy. So its lifetime will be its gravi-thermal time, also known as the Kelvin-Helmholtz time:

$$t_{g-t} = \frac{E_{grav}}{L}$$

Here $E_{grav} \sim -GM^2/R$.

Another way to write this is
(student)

$$L = -\frac{dE}{dt} \sim -\frac{GM}{R^2}\frac{dR}{dt}$$

We can get L from the characteristic T_{eff}, which we will address in a moment. Since $T_{\rm eff}$ is nearly constant during contraction,
(student)

$$L = (constant) \times R^2$$

This ends up giving an equation that looks like

$$\frac{dR}{dt} = -\frac{R^4}{\tau}$$

where τ is some timescale that depends on mass and not R. The solution is then

$$R(t) \approx R_\odot \left(\frac{t}{\tau}\right)^{-1/3}$$

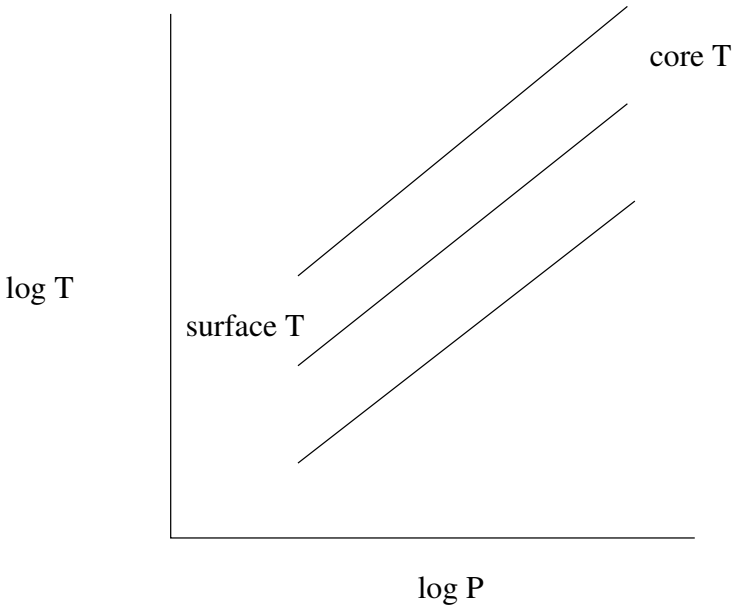
where the timescale is roughly 10^8 years for a solar mass, i.e. around 100 Myr.

18.3 Hayashi contraction phase

Why does contraction occur at nearly constant surface temperature?

We want to find structure of star which is fully convective from surface to interior, in which structure is set by conditions at the photosphere. Hayashi's insight was that the surface boundary condition controls the evolution of these protostellar objects.

For protostellar objects, the inside of the star is fully convective. This means that material is constantly moving from surface to interior, and is therefore well-mixed and at nearly constant entropy. So the T profile can be approximated as an adiabatic profile. For adiabatic compression $T \propto P^{2/5}$. This means the pressure-temperature structure of the star is quite simple:



At the photosphere, where $T \simeq T_{\text{eff}}$, we have

$$P_{ph} = \frac{g}{\kappa}$$

Where κ is the "opacity" and sets how "hard" it is for photons to carry energy through the medium. Note that these are not free-streaming photons, but thermally diffused ones.

The flux transmitted follows

$$F = \frac{1}{3}\frac{c}{\kappa\rho}\frac{d(aT^4)}{dz}$$

where z is some spatial coordinate (we will use radius in a moment). Higher κ means a higher T and higher T gradient is needed to transfer energy at a given rate.

By setting the surface T and P the surface condition sets the power-law relation that the interior of the star follows. It selects among the straight lines in this plot. This determines the central values, T_c and P_c. Thus we know

$$\frac{T_c}{P_c^{2/5}} \sim \frac{T_{ph}}{P_{ph}^{2/5}} = \frac{T_{eff}\kappa^{2/5}}{g^{2/5}}$$

The relation $T \propto P^{2/5}$, the ideal gas law, and the equation of hydrostatic balance allows us to solve for the gross structure of the star. This is gives the following relations for fully convective star:

$$P_c = 0.77\frac{GM^2}{R^4}$$

and

$$T_c = 0.54\frac{GM\mu m_p}{k_B R}$$

this gives that

$$k_B T_{eff} = 0.6\frac{GM\mu m_p}{R}\left[\frac{R^2}{M\kappa_{ph}}\right]^{2/5}$$

So now we NEED to understand the opacity, κ, at the photosphere.

Next time we will put in the opacity to see why we get constant T.

19 Astro notes 2018/10/8 - Mon - Protostars, Dust, Stars

19.1 Hayashi contraction phase continued

Why does contraction occur at nearly constant surface temperature?

Last time we got

$$\frac{T_c}{P_c^{2/5}} \sim \frac{T_{ph}}{P_{ph}^{2/5}} = \frac{T_{eff}\kappa_{ph}^{2/5}}{g^{2/5}}$$

And for fully convective star:

$$P_c = 0.77\frac{GM^2}{R^4}$$

and

$$T_c = 0.54\frac{GM\mu m_p}{k_B R}$$

this gives that

$$k_B T_{eff} = 0.6 \frac{GM\mu m_p}{R} \left[\frac{R^2}{M\kappa_{ph}} \right]^{2/5}$$

So now we NEED to understand the opacity, κ, at the photosphere.

Putting in Thompson scattering – which is the opacity from free electrons also called "electron scattering" opacity – gives a solution that is so cold that electrons can't be free, inconsistent. What Hayashi did is find that the main opacity is from H^-. The last electron is very weakly bound, about 0.75 eV. We'll find that the temperatures are 2000-3000 K. Electrons come from reactive(or low valence energy) elements Na, Li, etc. This process is very temperature sensative, at lower temp, metals not ionized, at higher there is no H^-.

The opacity is from the photo-ionization of H^-, $\gamma + H^- \rightarrow H + e^-$. Figuring out the opacity is hard. The fitting form is

$$\kappa_H = 2.5 \times 10^{-31} \rho^{1/2} T^9 \text{ cm}^2/\text{gr}$$

put this into the above and you get

$$T_{eff} \simeq 2500 \text{ K} \left(\frac{M}{M_\odot} \right)^{1/7} \left(\frac{R}{R_\odot} \right)^{1/49}$$

Almost independent of the radius, so the lines are vertical. Also the increase in effective temperature with mass. You will show in the homework that this relation allows the star to have a wide range of luminosities with nearly the same T_{eff}.

19.2 Dust extinction

observed magnitudes are effected by extinction

$$m_\lambda = M_\lambda + 5 \log_{10} d - 5 + A_\lambda$$

where d is the distance in parsecs and $A_\lambda > 0$ is called the extinction. The amount of extinction depends on wavelength, even for the same amount of dust. The amount of dust is typically measured using optical depth. For some intensity field (parallel flux rays) I_λ,

$$\frac{I_\lambda}{I_{\lambda,0}} = e^{-\tau_\lambda}$$

where τ_λ is called the optical depth. Working this out gives

$$m_\lambda - m_{\lambda,0} = -2.5 \log_{10}(e^{-\tau_\lambda}) = 2.5 \tau_\lambda \log_{10} e = 1.086 \tau_\lambda$$

or

$$A_\lambda = 1.086 \tau_\lambda$$

The optical depth is found by integrating the cross section along the line of sight

$$\tau_\lambda = \int_0^z n_d(z')\sigma_\lambda dz'$$

where n is number density and σ_λ is wavelength-dependent cross section. Larger σ blocks more light. If σ_λ is constant

$$\tau_\lambda = \sigma_\lambda N_d$$

where

$$N_d = \int_0^d n_d(z)dz$$

is called the column depth.

For dust we can express

$$\sigma_\lambda = Q_\lambda \pi a^2$$

where πa^2 is the simply geometric cross section. Q is the extinction coefficent. For large λ, $Q \sim a/\lambda$ so that

$$\sigma_\lambda \propto \frac{a^3}{\lambda} \quad for \quad \lambda \geq a$$

which goes to zero for long wavelengths, and for small

$$\sigma_\lambda \propto a^2 \quad for \quad \lambda \ll a$$

Real dust will have some distribution of a values.

Since the cross section is wavelength dependent, dust changes the *color* of the light. i.e. the relative amount at different wavelengths. This is often given as a color excess:

$$(B-V)_{actual} - (B-V)_{observed} = E(B-V)$$

Note that that $E(B-V)$, since it linear in the log of the flux, is therefore multiplicative in flux, and therefore is the same for any object along this line of sight. Thus you will often see $E(B-V)$ quoted for a cluster since the light from all stars passes through approximately the same column of dust.

19.3 Initial mass function

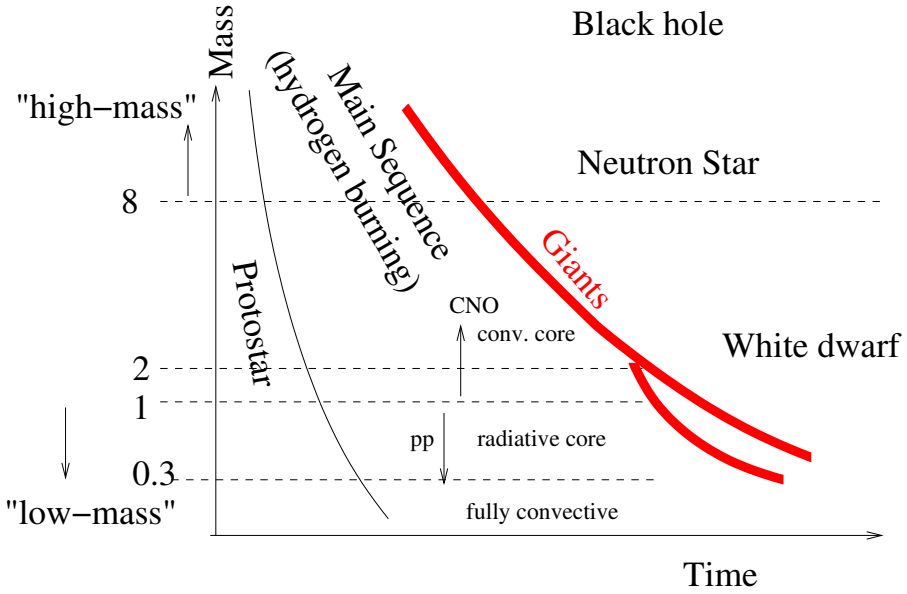
Some comments on initial mass function, i.e. the relative number of stars of different masses at the end of star formation. Peaks just under $0.1 M_\odot$. Decreases steeply to higher masses. In the classic form this decreases as $M^{-2.35}$, so there are hudreds of solar mass stars for each few stars of order 10-20 solar masses.

20 Astro notes 2018/10/10 - Wed - Stars

handed back and discussed exam

20.1 Stellar evolution and the main sequence

Evolution during the hydrogen burning phase.



Main parameters: birth mass, time. Other paramers (not shown) composition, rotation, maybe magnetic field.

Want to understand how the star's structure changes as it evolves. The main reason that this can be done is that there is a hierachy of timescales.

$$t_{\rm nuc} \gg t_{\rm grav-therm} \gg t_{\rm dyn}$$

The time on the main sequence is
(student)

$$t_{MS} = \frac{E_{nuc}}{L} \sim \frac{\Delta E_{nuc,H} M/m_p}{L} \propto \frac{M}{M^3} \propto M^{-2}$$

Thus higher mass stars live much shorter time.

Also since the evolution time in H and He burning phases is much longer than the Kelvin-Helmlotz time, the star evolves from one semi-static solution to another where each solution is in balance, hydrostatic, or even mostly thermal balance.

Evolution during this time is one of changing the interior composition (H→He) from one thermal equilibrium state to another.

20.2 Thresholds in initial mass

There are several thresholds between high- and low-mass stars that depend on the stage of life. Four are shown in the diagram:

1. Remnant type, Neutron star or White dwarf, $\sim 9 M_\odot$.

The most obvious threshold is based on end of life, in that stars about about 8 to 10 $M_\odot$ will have their cores collapse and explode as a supernovae, but less massive stars will just stop fusion and settle into being a white dwarf star.

2. Fusion process, $\sim 1M_\odot$

Also called the upper vs. lower main sequence. This has to do with the how fusion takes place, CNO catalytic fusion for higher masses, p+p for lower masses. This results in the presence or absence of a convective core.

3. Immediate or delayed transition to core helium fusion, $\sim 2.5M_\odot$

i.e. is shell (off-center) H burning required between the end of core H burning and core He burning? Schönberg-Chandrasekhar limit. Below this limit there is a distinct RGB (red giant branch, shell H burning) and AGB (Asymptotic giant branch, shell He,H burning) separated by the HB (horizontal branch, core He burning). Above, there is not really a separate RGB.

4. Fully convective vs. radiative core, $\sim 0.3M_\odot$

Below this threshold star is always fully mixed (i.e. surface and interior abundances evolve together) unlike at higher masses

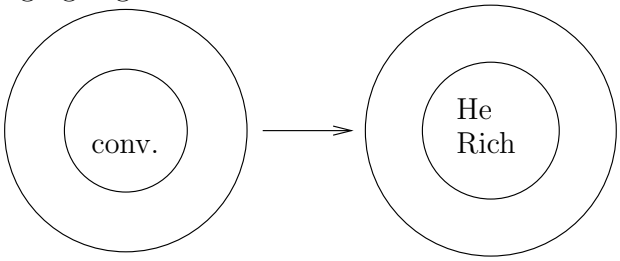
21 Astro notes 2018/10/12 - Fri - Burning processes in stars

21.1 Upper vs lower (mass) main sequence stars

Upper ($M \gtrsim M_\odot$): C/N/O catalyzed H fusion, convective core

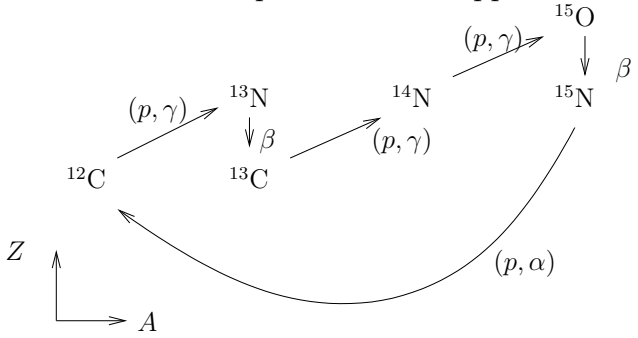
Lower ($M \lesssim M_\odot$): p+p H fusion, stable core

$M > M_\odot$: Convective core, burning via CNO cycle (H to He), has a radiative envelope. What is the main thing that's changing with time? The composition of the interior is changing together.



Lower mass has non-convecting core

What is the fusion process for the upper MS? Carbon, nitrogen, oxygen catalyzed fusion:

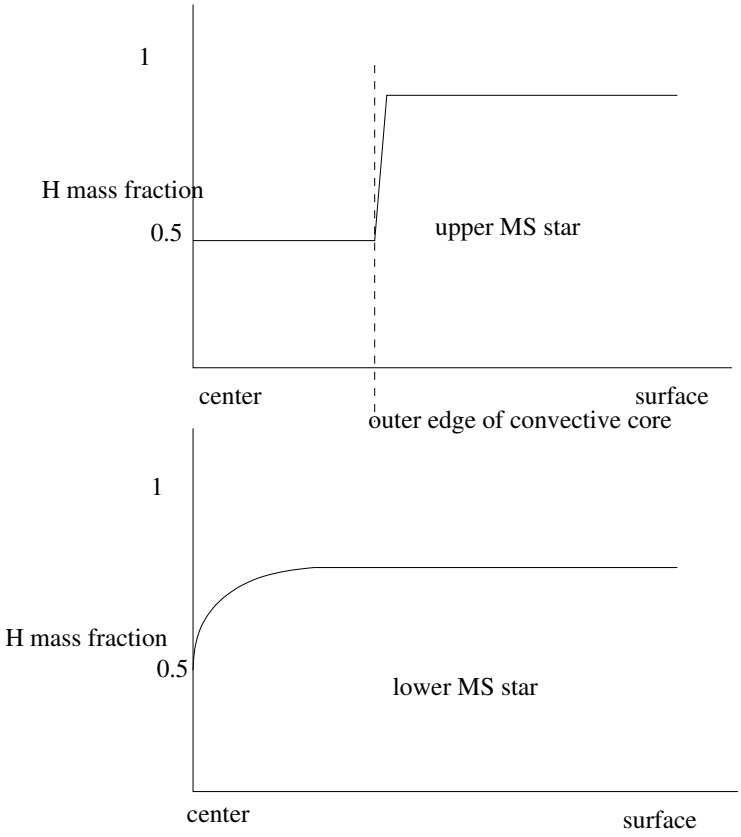


This net turns 4 protons into 2 protons and 2 neutrons in the form of a Helium-4 nucleus. And the carbon seed is recycled.

Contrast with low-mass stars: $M < M_{\odot}$ radiative core, direct proton+proton fusion. in this case 2 things different: no convective core. So depletion is occurs just at the center instead of in whole convective region. Also T dependence of L_{nuc} is different.

Basic process is $p + p \rightarrow^2$ H, $p +^2$ H $\rightarrow^3$ He, ^{3}He $+^3$ He $\rightarrow^4$ He $+ 2p$. Less T sensitive because limiting reaction is between lower-charged things ($p + p$ instead of $p +^{14}$ N.

Can represent profile of interior hydrogen abundance graphically when, for example, $X_{H,center} = 0.5$, i.e. half depletion.



The large core being depleted all together leads to the "Henyey hook" between point B and C in the stellar evolution HR diagram above. Star must contract in order for H fusion to ignite at the surface of the newly made He core. Lower MS star transitions smoothly to off-center fusion.

The star changes slightly during the core hydrogen burning phase (the "main sequence") because converting H to He reduces the number of particles exerting pressure (per unit mass). 8 particles become 3. Star does change as μ increases.

Reference to figure below with HR evolution diagram. You can see increase in luminosity, A to B, as H is converted to He in core. Massive stars evolve on the main sequence by increasing radius and L. Also see MESA paper 1, figures 14 and 22.

Table 26.7
EVOLUTIONARY PHASES IN FIG. 26.10 AND FIG. 26.11

Phase	*Segment of Track or Point(s)*	*Discussed in Sect.*
MS	A–B	26.4a
Central H exhaustion	B–C	26.4b
H burning shell source	C–E	26.4c
Deep convective envelope	D, K	26.4c
Core He burning	E–F	26.4d
Central He exhaustion	F–G	26.4e
He burning shell source	G–H–K	26.4f
Cessation of H burning shell source	H	26.4g

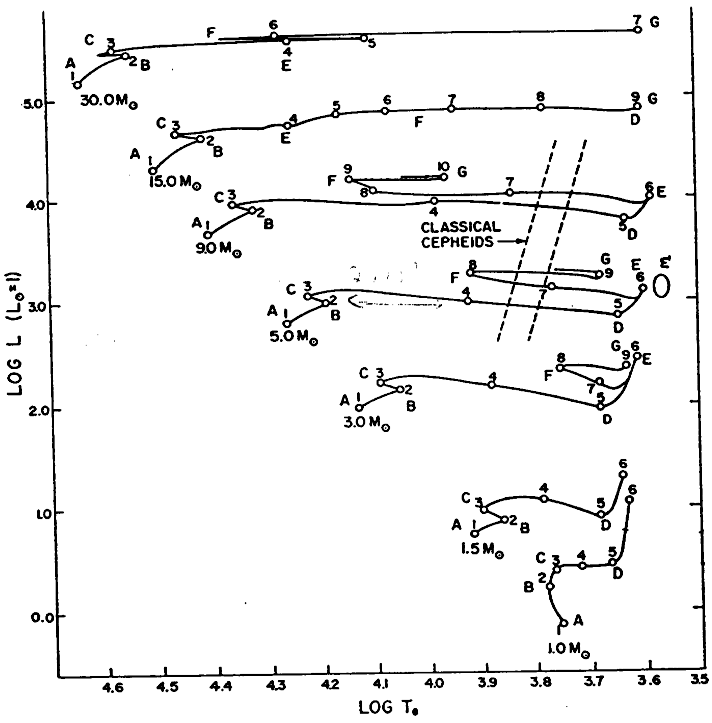


Fig. 26.10 Post-MS evolutionary tracks on the theoretical H-R diagram for stars having masses in the range $1.0 \leq M \leq 30.0$ and, initially, a "Population I" composition ($X = 0.708$, $Z = 0.02$ for all tracks except the $M = 30$ track; for this track, $X = 0.70$, $Z = 0.03$). Numbered points along the tracks are those listed in Table 26.2. The significance of the letters alongside the tracks is summarized in Table 26.7 (*cf.* also Fig. 26.12 below). The tracks for $1.0 \leq M \leq 15.0$ are due to Iben [Ib64]; the track for $M = 30.0$ is due to Stothers [St66d]. The oval labelled E' is the region where the "helium flash" occurs in stars with $M \approx 1.0-1.3$ (see Sect. 26.4c). The approximate location of the classical cepheids is also shown (see Chap. 27).

EVOLUTIONARY LIFETIMES (YEARS)†

(*Initial composition:* $X = 0.708$, $Z = 0.02$ *for* $1.0 \leq M \leq 15.0$; $X = 0.70$, $Z = 0.03$ *for* $M = 30.0$)

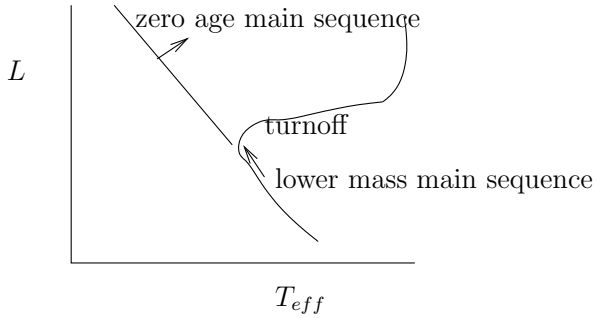
	M (solar units)						
Point	1.0	1.5	3.0	5.0	9.0	15.0	30.0
1	5.016(7)	1.821(7)	2.510(6)	5.760(5)	1.511(5)	6.160(4)	2 (4)
2	8.060(9)	1.567(9)	2.273(8)	6.549(7)	2.129(7)	1.023(7)	4.82(6)
3	9.705(9)	1.652(9)	2.394(8)	6.823(7)	2.190(7)	1.048(7)	4.91(6)
4	1.0236(10)	2.036(9)	2.478(8)	7.019(7)	2.208(7)	1.050(7)	4.92(6)
5	1.0446(10)	2.105(9)	2.488(8)	7.035(7)	2.213(7)	1.149(7)	4.93(6)
6	1.0875(10)	2.263(9)	2.531(8)	7.084(7)	2.214(7)	1.196(7)	5.45(6)
7	–	–	2.887(8)	7.844(7)	2.273(7)	1.210(7)	5.46(6)
8	–	–	3.095(8)	8.524(7)	2.315(7)	1.213(7)	–
9	–	–	3.262(8)	8.782(7)	2.574(7)	1.214(7)	–
10	–	–	–	–	2.623(7)	–	–

* Numbers in parentheses are the powers of ten by which the corresponding entries are to be multiplied.

† From Iben [Ib64] and Stothers [St66d].

Another result: Upper main sequence stars, $M > M_\odot$, evolve perpendicular to the zero-age main sequence line (the birth line), while lower main sequence stars, $M \lesssim M_\odot$ evolve upward parallel to the MS.

This makes dating of Globular clusters difficult. These clusters have roughly ages of 10^{10} years so that the stars leaving the main sequence today have masses of ~ 0.9.



Also see MESA paper 1, figure 14. Current state of the art for ages of Globular clusters from MS. Turnoff provides the age to only ± 2 Gyr.

Show movies of MS phase from 1 and 16 Msun on MESA website. (see course webpage)

21.2 Transition to post-MS - Schönberg-Chandrasekhar limit

In movies of 1 and 16 $M_\odot$ cases.
Pay particular attention to the X (= hydrogen mass fraction) and Y (=helium mass fraction) curves in the upper right. The figure in the lower right is a "Kippenhahn diagram" it shows the extent of the convection zones (hatched lines) and the fusion burning (red shading) in mass (the vertical axis) as a function of "time" (horizontal axis). Note the horizontal axis here is not really time but "model number" in the simulation and though it increases with time, only increments when the structure changes by modest amount. This means that the main sequence (core hydrogen burning phase) though long-lived in time, is only the first little bit in model numbers since it is not complex evolution.

22 Astro notes 2018/10/15 - Mon - End of life of stars

22.1 Drivers of stellar evolution

Two drivers for Stellar evolution:

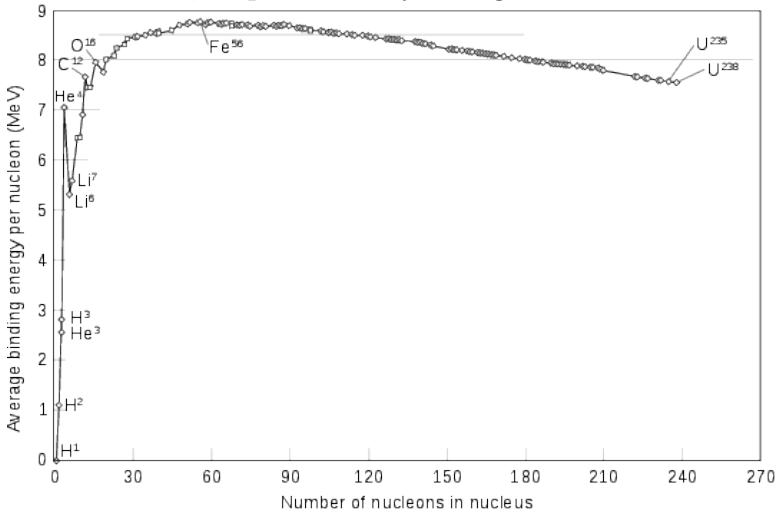
- As a star loses energy, its central temperature will increase.
  (i.e. it has a negative heat capacity!)
  This starts consecutive burning stages. This is halted if the star becomes degenerate – typically the case for low-mass stars.
  Also higher temperature, as in shell burning, leads to higher luminosity.

- Each successive fuel stage releases less energy per mass than the previous.
  While converting H to He releases about 7 MeV per nucleon, later stages release no more than 1 MeV per nucleon each.

More bound means less rest mass energy. So a helium has less rest mass than 2 protons and 2 neutrons. The extra is released when the bound helium is made, and is called the "Binding Energy".

$$E_{\text{bind}} = Zm_pc^2 + Nm_nc^2 - m_ic^2$$

Where m_p is the proton mass, m_n is the neutron mass, and m_i is the mass of the nuclide being considered. Also this is usually given as E_{bind}/A, i.e. "per nucleon", which is what is shown in this plot. This is appropriate because the number of nucleons is conserved during nuclear reactions, so protons may change into neutrons.



(image from binding energy curve entry on wikipedia, but also appearing in many textbooks.)

22.2 H-depleted core formation

$M > M_\odot$: Convective core, burning via CNO cycle (H to He), has a radiative envelope.
$M < M_\odot$: radiative core, pp burning

Kippenhahn diagram. Can see that only the inner regions get burned. also the burning shuts off when the core is converted to helium, which is at 5.6 on the time axis.

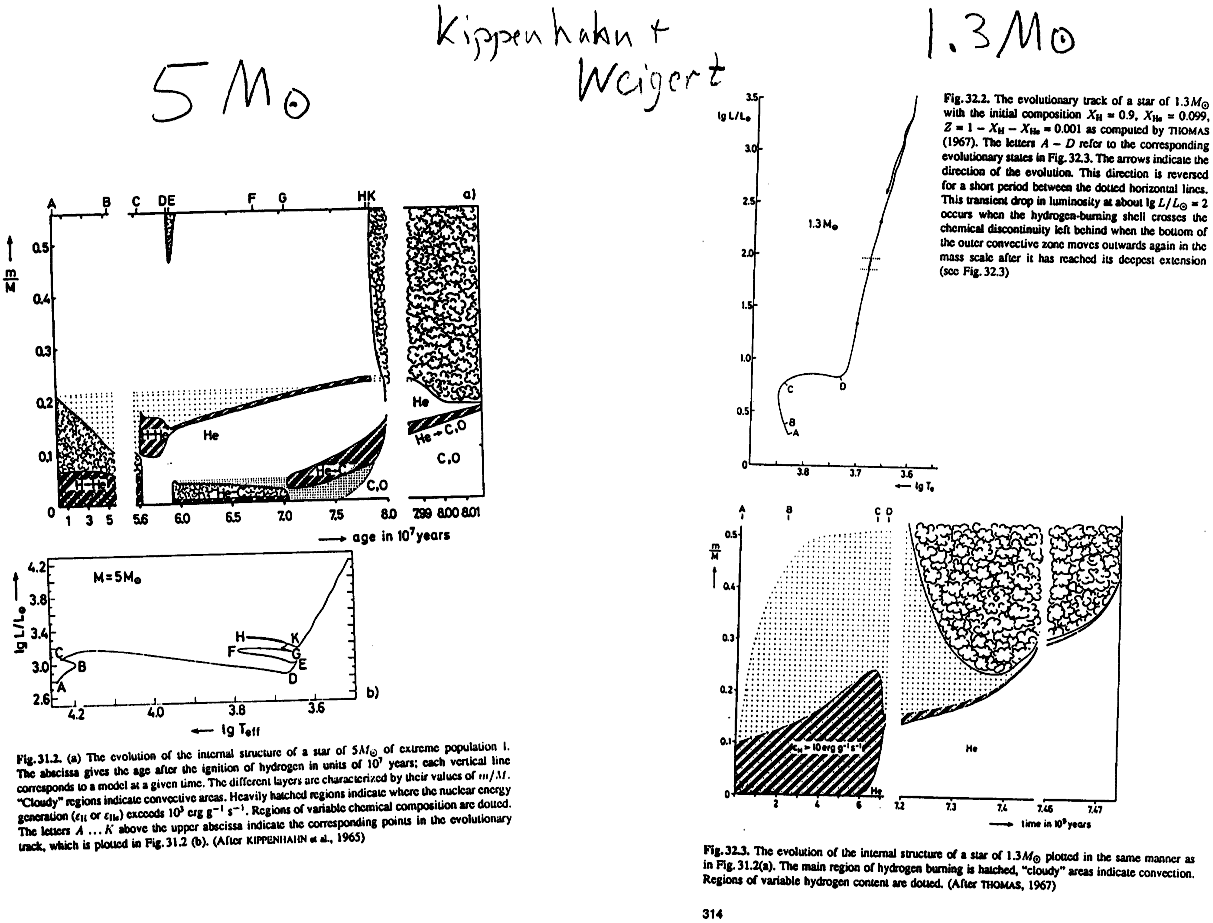


Fig. 31.2. (a) The evolution of the internal structure of a star of $5M_\odot$ of extreme population I. The abscissa gives the age after the ignition of hydrogen in units of 10^7 years; each vertical line corresponds to a model at a given time. The different layers are characterized by their values of m/M. "Cloudy" regions indicate convective areas. Heavily hatched regions indicate where the nuclear energy generation (ε_H or ε_{He}) exceeds 10^3 erg g^{-1} s^{-1}. Regions of variable chemical composition are dotted. The letters $A \ldots K$ above the upper abscissa indicate the corresponding points in the evolutionary track, which is plotted in Fig. 31.2 (b). (After KIPPENHAHN et al., 1965)

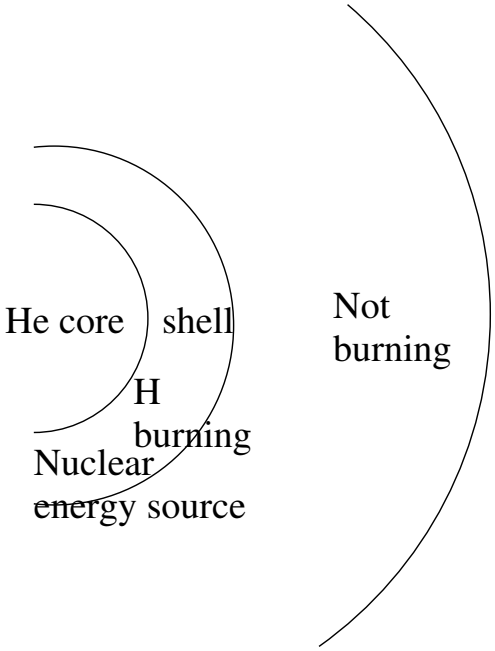
Fig. 32.2. The evolutionary track of a star of $1.3M_\odot$ with the initial composition $X_H = 0.9$, $X_{He} = 0.099$, $Z = 1 - X_H - X_{He} = 0.001$ as computed by THOMAS (1967). The letters $A - D$ refer to the corresponding evolutionary states in Fig. 32.3. The arrows indicate the direction of the evolution. This direction is reversed for a short period between the dotted horizontal lines. This transient drop in luminosity at about $\lg L/L_\odot = 2$ occurs when the hydrogen-burning shell crosses the chemical discontinuity left behind when the bottom of the outer convective zone moves outwards again in the mass scale after it has reached its deepest extension (see Fig. 32.3)

Fig. 32.3. The evolution of the internal structure of a star of $1.3M_\odot$ plotted in the same manner as in Fig. 31.2(a). The main region of hydrogen burning is hatched, "cloudy" areas indicate convection. Regions of variable hydrogen content are dotted. (After THOMAS, 1967)

314

$M > M_\odot$ first. Eventually the temperature does rise as X decreases until the fuel runs out in the core. When this happens the He core gravitationally contracts on t_{KH}. The B to C hook in the HR diagram called the Henyey hook is due to this. Contraction halts once hydrogen burning has ignited in a shell surrounding the helium core.

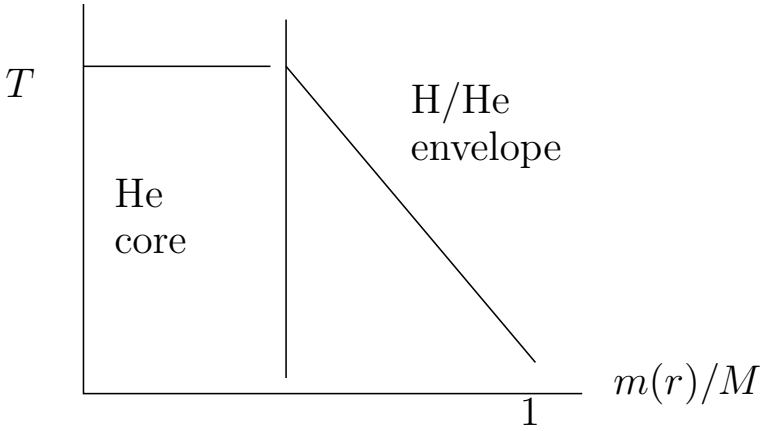
for $M < M_\odot$, T is rising even on the main sequence, so the transition to shell burning is more gradual. Since there is no convective core, there is no sudden depletion, the helium core just slowly grows. Can see this in the Kippenhahn diagrams, as the helium core ramps up at about 6 Gyr.



The helium core mass will be growing in time. Question is can this evolution persist until all H is burned. We will show that the above picture can only be constructed when the helium core is a small fraction of the mass.

22.3 Schönberg-Chandrasekhar Limit

Consider a star after hydrogen burning in the core is done and presume that the He core is non-degenerate.



The envelope exerts pressure on the core and we want to see if there is always a hydrostatic solution. Go back to the virial theorem. Hydrostatic balance:

$$\frac{dP}{dr} = -\rho g = -\frac{Gm(r)\rho(r)}{r^2}$$

multiply both sides by $4\pi r^3$ and integrate (this is what we did to get the virial theorem).

Let R_c be the radius of the core. integrating to here:

$$\int_0^{R_c} 4\pi r^3 \frac{dP}{dr}\, dr = 4\pi r^3 P(r)|_0^{R_c} - 12\pi \int_0^{R_c} r^2 P(r) dr$$

just integrating by parts.

$$= 4\pi R_c^3 P(R_c) - 12\pi \int_0^{R_c} r^2 P(r) dr$$

where now $P(R_c)$ is not zero, since it has the envelope above it. That's the left hand side. The right hand side is the gravitational bining energy which is roughly $-GM_c^2/R_c$. For an isothermal, ideal gas core

$$P(r) = nkT = \left(\frac{\rho}{\mu_c} m_p\right) kT_c = \frac{\rho k T_c}{\mu_c m_p}$$

where μ_c is the mean molecular weight in the core (different from the envelope). Mean molecular weight is approximately the number of protons or neutrons per particle in the (ionized gas). So pure hydrogen has one proton per nucleus and electron so that $\mu = 1/2$, helium has 4 protons and neutrons per 3 particles, one nucleus and two electrons, giving $\mu = 4/3$.

Putting in

$$12\pi \int r^2 \frac{\rho(r) k T_c}{\mu_c m_p} dr = 3 \frac{k T_c}{\mu_c m_p} M_c$$

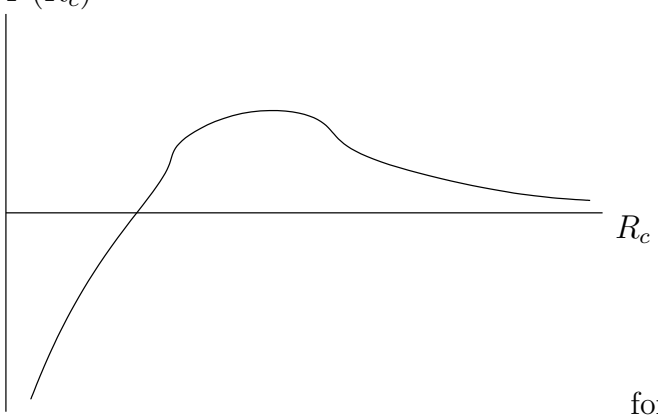
having done all the integrals, we pull it together:

$$4\pi R_c^3 P(R_c) - 3\frac{k T_c}{\mu_c m_p} M_c = -\frac{G M_c^2}{R_c}$$

Solve for $P(R_c)$,

$$P(R_c) = \frac{3}{4\pi} \frac{k T_c}{\mu_c m_p} \frac{M_c}{R_c^3} - \frac{1}{4\pi} \frac{G M_c^2}{R_c^4}$$

fixing everything and only considering dependence on radius, wee see the latter term dominates at small radius and the former at large.



for fixed M_c, T_c there is a $P_{c,max}$. This

gives:

$$R_{c,crit} = \frac{GM_c\mu_c m_p}{kT_c}\frac{4}{9}$$

and

$$P(R_{c,crit}) = \frac{3}{4}\frac{1}{4\pi R_c^3}\frac{M_c}{\mu_c m_p}kT_c = 0.7\left(\frac{kT_c}{\mu_c m_p}\right)^4\frac{1}{G^3M_c^2}$$

(Note the temperature of the core is actually set by the Hydrogen burning shell.)

Now we need to look at the envolope. Pesume that $M = M_{env} \gg M_c$ so that

$$P_{base} = \frac{GM^2}{R^4}$$

what is R here? for the envelope: we would say that

$$kT_e = \frac{GM\mu_e m_p}{R}$$

This is that the hydrogen burning must be distributed over a scale height in order for it to be stable (for the thermostat mechanism of star to work). So then

$$P_{base} = \left(\frac{kT_e}{\mu_e m_p}\right)^4\frac{1}{G^3M^2}$$

To have a solution, we must have $P_{base} < P_{c,max}$. If $T_e = T_c$, then this constraint is just concerned with the relative masses and relative mean molecular weights. (don't believe any coefficients in this derivation) for the scaling we get:

$$\frac{1}{\mu_e^4M^2} < \frac{1}{\mu_c^4M_c^2}$$

or

$$\frac{M_c}{M} < \left(\frac{\mu_e}{\mu_c}\right)^2(0.4)$$

for a stable solution. Here the 0.4 comes from a numerical analysis. Recall that approximately $\mu_e = 0.6$ and $\mu_c = \mu_{He} = 1.33$ and we get

$$\frac{M_c}{M} < 0.08$$

If the Helium core is non-degenerate, then $M_c < 0.08M$ for a stable hydrostatic solution.

Note that this derivation assumes non-degenerate gas – with degeneracy, any pressure can be supported.

Stars that finish core H burning with a core larger than this limit can proceed almost directly to core He burning instead of needing to build the He core more with shell burning first. As a result, some don't even make it to the giant branch before He burning starts.

So we have found the limit for an isothermal He core to support an overlying star:

$$\frac{M_c}{M} < 0.08$$

Above this core contracts directly until He fusion begins. That changes the constant-T core into one with a T gradient, which can be stable.

22.4 With Degenerate core

When the helium core becomes degenerate, then ANY M_c/M value is possible. Won't derive this.

22.5 Helium ignition and the core limit

Three ranges, $M/M_\odot > 6$, $6 > M/M_\odot > 2$ and $2 > M/M_\odot$.

Consider $M > 6M_\odot$. $M_c/M > 0.08$ when the H shell burning ignites. No isothermal solution is possible. Proceeds directly to helium core burning, with no shell phase in between. (that is after burning of H runs out the star "discovers" that it can't support itself, and begins to contract until He ignites) The He core contracts over a time set by the overall energy loss-rate. This is about the Kelvin-Helmholtz time,

$$t = \frac{GM^2/R}{L} \simeq 10^6 \text{ yr at } 6M_\odot$$

or 3×10^4 yr at $30M_\odot$. Refer to first figure of interior of $5M_\odot$ star on Kippenhan figure for borderline case. From C to D is the expansion of the giant, which happens quickly.

This short-lived phase when the He core is contracting is referred to as the Hertzprung Gap. For higher mass stars core burning ignites before reaching giant branch.

Can be seen in upper MS HR diagram, He burning starts at E:

Table 26.7

EVOLUTIONARY PHASES IN FIG. 26.10 AND FIG. 26.11

Phase	*Segment of Track or Point(s)*	*Discussed in Sect.*
MS	A–B	26.4a
Central H exhaustion	B–C	26.4b
H burning shell source	C–E	26.4c
Deep convective envelope	D, K	26.4c
Core He burning	E–F	26.4d
Central He exhaustion	F–G	26.4e
He burning shell source	G–H–K	26.4f
Cessation of H burning shell source	H	26.4g

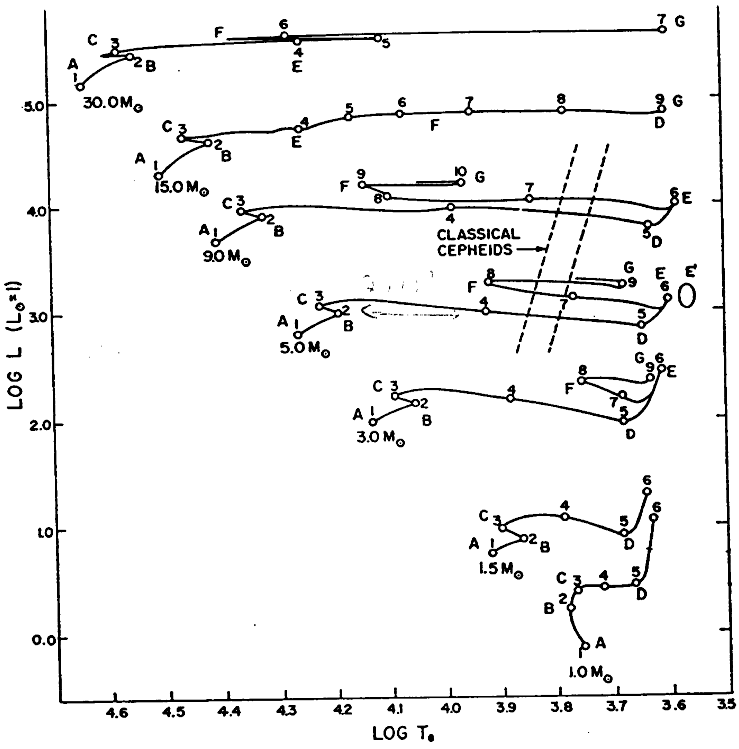


Fig. 26.10 Post-MS evolutionary tracks on the theoretical H-R diagram for stars having masses in the range $1.0 \leq M \leq 30.0$ and, initially, a "Population I" composition ($X = 0.708$, $Z = 0.02$ for all tracks except the $M = 30$ track; for this track, $X = 0.70$, $Z = 0.03$). Numbered points along the tracks are those listed in Table 26.2. The significance of the letters alongside the tracks is summarized in Table 26.7 (*cf.* also Fig. 26.12 below). The tracks for $1.0 \leq M \leq 15.0$ are due to Iben [Ib64]; the track for $M = 30.0$ is due to Stothers [St66d]. The oval labelled E' is the region where the "helium flash" occurs in stars with $M \approx 1.0-1.3$ (see Sect. 26.4c). The approximate location of the classical cepheids is also shown (see Chap. 27).

EVOLUTIONARY LIFETIMES (YEARS)†

(*Initial composition:* $X = 0.708$, $Z = 0.02$ *for* $1.0 \leq M \leq 15.0$; $X = 0.70$, $Z = 0.03$ *for* $M = 30.0$)

	M (solar units)						
Point	1.0	1.5	3.0	5.0	9.0	15.0	30.0
1	5.016(7)	1.821(7)	2.510(6)	5.760(5)	1.511(5)	6.160(4)	2 (4)
2	8.060(9)	1.567(9)	2.273(8)	6.549(7)	2.129(7)	1.023(7)	4.82(6)
3	9.705(9)	1.652(9)	2.394(8)	6.823(7)	2.190(7)	1.048(7)	4.91(6)
4	1.0236(10)	2.036(9)	2.478(8)	7.019(7)	2.208(7)	1.050(7)	4.92(6)
5	1.0446(10)	2.105(9)	2.488(8)	7.035(7)	2.213(7)	1.149(7)	4.93(6)
6	1.0875(10)	2.263(9)	2.531(8)	7.084(7)	2.214(7)	1.196(7)	5.45(6)
7	–	–	2.887(8)	7.844(7)	2.273(7)	1.210(7)	5.46(6)
8	–	–	3.095(8)	8.524(7)	2.315(7)	1.213(7)	–
9	–	–	3.262(8)	8.782(7)	2.574(7)	1.214(7)	–
10	–	–	–	–	2.623(7)	–	–

* Numbers in parentheses are the powers of ten by which the corresponding entries are to be multiplied.

† From Iben [Ib64] and Stothers [St66d].

Now consider $2 < M < 6$. Then the initial value of $M_c/M < 0.08$ when H burning in core is completed. Core can be isothermal during H shell burning, and does not collapse until the H shell burning has forced $M_c > 0.08M$. This happens before the core becomes degenerate. $5M_\odot$ case is example, but core building phase is brief.

Finally $M < 2M_\odot$. Core becomes degenerate before $M_c > 0.08M$. Thus the He core can get as large as it wants until the He ignites. (ends up igniting when it is about $0.4M_\odot$)

22.6 Evolution in density-temperature

final fate of the star is determined by where the curves for burning intersect the degeneracy curves. Refer to plots that show the calculated evolution in this diagram.

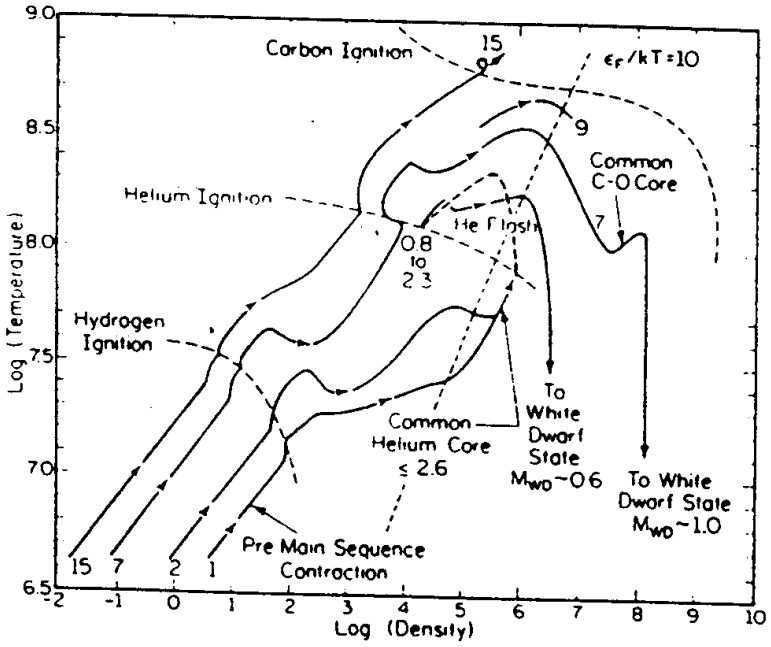


FIGURE 2.7. Central density versus central temperature for evolving stellar models. Reproduced, with permission, from I. Iben Jr. 1985, "The Life and Times of an Intermediate Mass Star," in *Quarterly Journal of the Royal Astronomical Society*, Volume 26, published by Blackwell Scientific Publications.

Also similar figures in MESA paper - Figure 22, 29, 30

22.7 Post-Main-Sequence Evolution

Two important effects - Shell burning and Mass loss

Shell burning leads to giant - on lower main sequence ($\lesssim 2M_\odot$) 2 separate giant branches separated by core helium burning. (Horizontal Branch) Shell burning luminosity depends on the mass of the underlying core, not the total mass. Monotonically increasing.

Mass loss important for both low and high mass stars. For low mass stars Riemers law, calibrated on observations:

$$\dot{M} = -4 \times 10^{-13} \eta \frac{L}{gR} M_\odot/yr = -4 \times 10^{-13} \eta \frac{L}{GM/R} M_\odot/yr$$

where η is an adjustable parameter calibrated on observations.

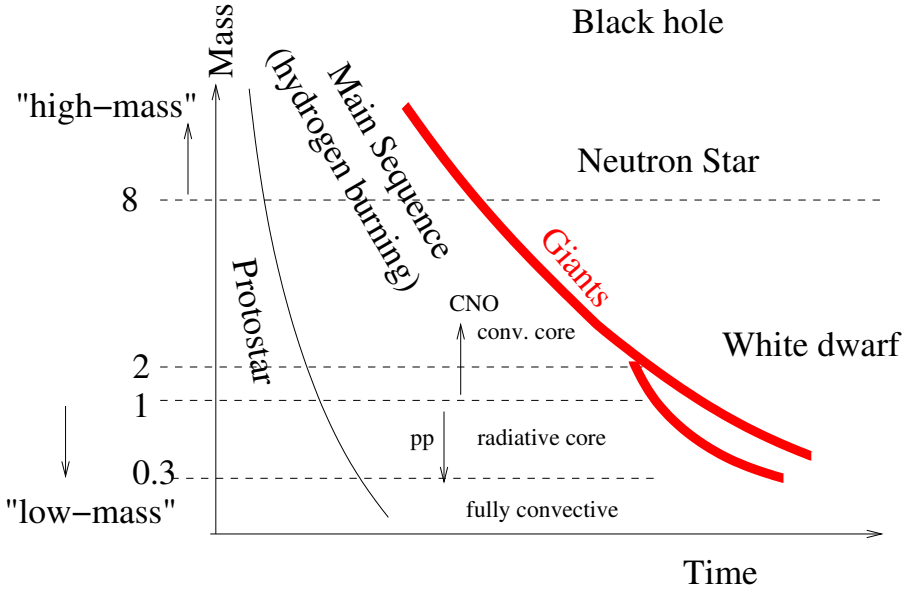
(show movie of low mass star to white dwarf) indicate horizontal branch - only for low mass stars. Mass loss a random variable that gives spread in HB. Was thought to happen during He flash, but now though to just be at tip of RGB.

22.8 end of life

Low mass stars lose mass and end up as white dwarfs. This also forms a planetary nebula.

High mass stars eventually form Fe, which cannot be fused to release energy. The Fe core eventually becomes high enough mass that central pressure causes electron capture $p \rightarrow n$. This removes electron pressure support in a runaway process called core collapse that forms a neutron star. Bounce and neutrinos released powers supernova. (Mesa fig 31)

This is the last of the three major divisions.



22.9 Cluster Color-lum diagrams

Since, in clusters, all stars form at same time, but bright, hot stars evolve in shorter time, the cluster can be dated by distribution of stars in the color-magnitude diagram.

Figures from text:

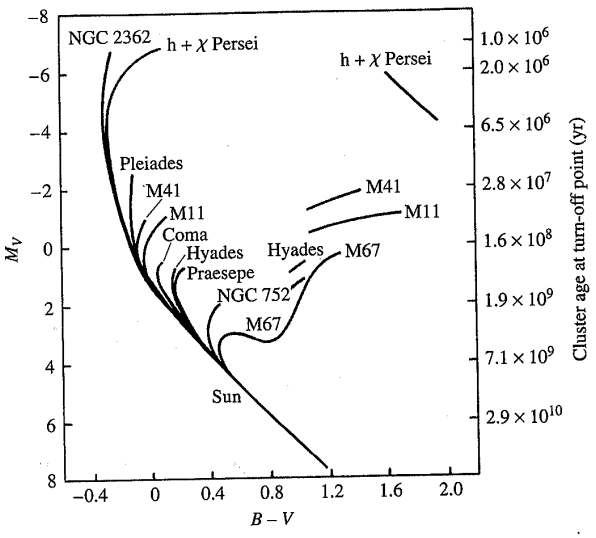


FIGURE 13.19 A composite color–magnitude diagram for a set of Population I galactic clusters

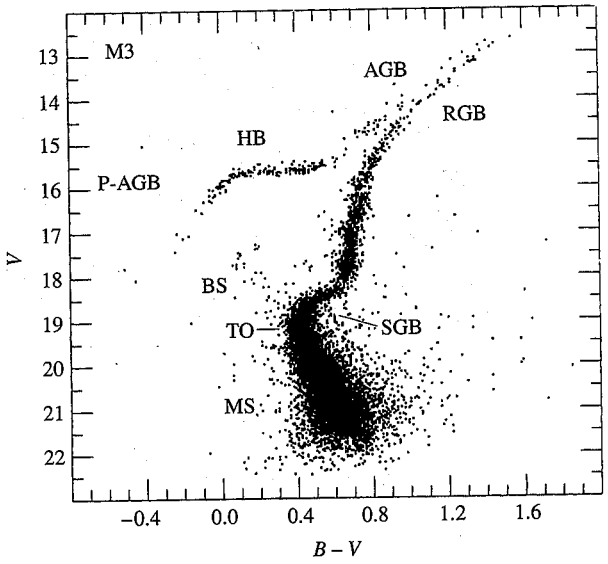


FIGURE 13.17 A color–magnitude diagram for M3, an old globular cluster. The ma

23 Astro notes 2018/10/17 - Wed - Intro to general relativity

-A problem: excess precession of the perihelion of Mercury
-A question: is light bent by gravity if it has no mass?

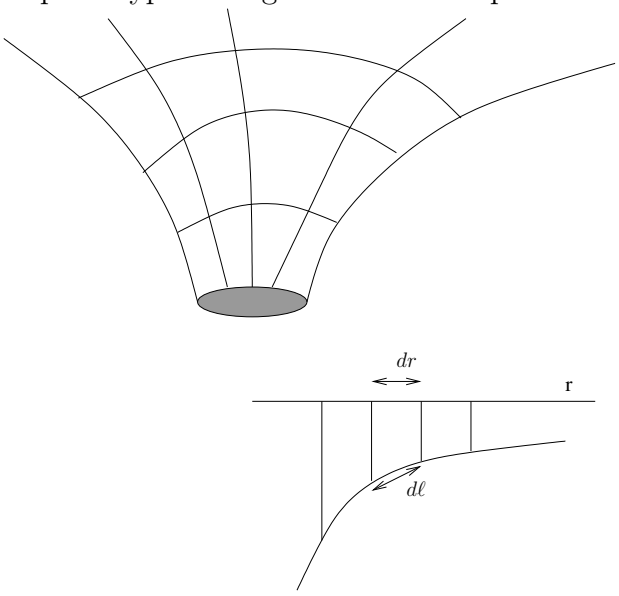
When starting special relativity: had to give up that there is some preferred true frame of reference. But though we said the critical thing was the (local) propagation speed of light, we still defined relation between global coordinate systems moving at some speed. Really need to give up on global (flat) coordinate systems and discuss spacetime by stitching

together its local structure.

i.e. the reference frames we used for special relativity are only effectively valid locally, and if we want to relate ones that are far apart we must use general relativity to stitch together what is between them.

What is curved spacetime? The distance between points depends on path, and objects move on "shortest" path, like straight-line motion in flat space.

Explain typical diagram of curved space:



radius can be measured two ways: distance from center, and circumference/2π. The typical curved spacetime diagram is intended to indicate that these are not the same thing i.e. if we measure one radius as

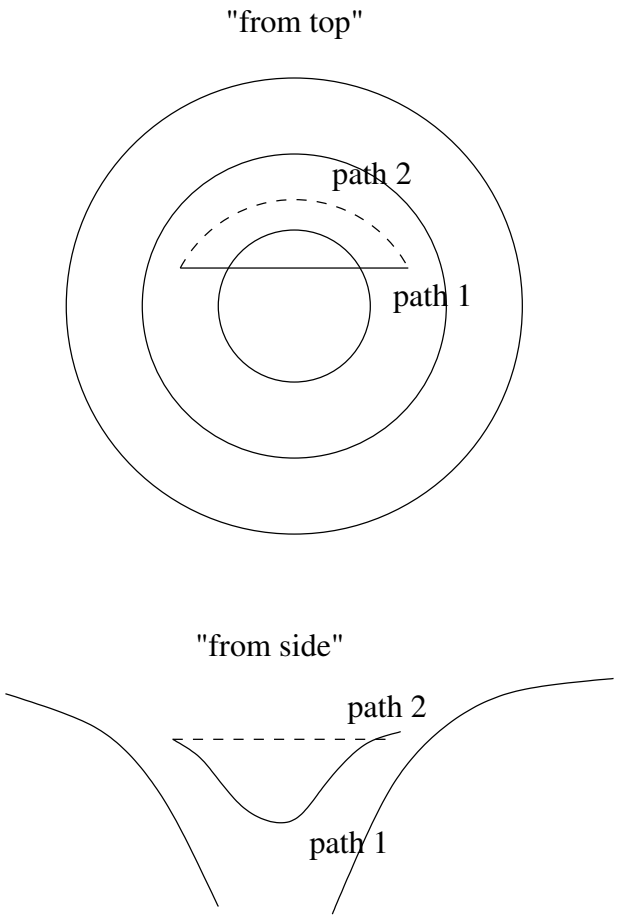
$$r_1 = \frac{s_1}{2\pi}$$

where s_1 is the circumference of some circle about the central object. Then if we move inward by some distance $d\ell$, along the spacetime surface, we would find that if we measure the circumference there s_2 we would find that

$$r_2 = \frac{s_2}{2\pi} \neq r_1 - d\ell$$

since as can be seen from the diagram, the length along the spacetime is longer than the dr between the radii.

This also means what might appear to be a "shortcut" path can actually be longer, so that a normal orbit-like path is the shortest path i.e. the path of an un-forced particle.



Can see that this might change elliptical orbits, for example causing them to not close, since the measurement of distances is no longer the same in the radial and non-radial directions.

In reality things are more complex than this, since we have not drawn the time direction, but the "shortest path" is really the through the 4-dimensional spacetime, not just through space.

23.1 Principle of equivalence

While it is fairly clear that frames moving at constant speed are inertial, i.e. physics works as expected in them, things are less clear for frames subject to the force of gravity. One oddity is that the gravitational mass (i.e. the gravitation "charge") and inertial mass appear to be identical to extremely high accuracy. That is compare the coulomb force (student)

$$m_i a = -\frac{qQ}{r^2}$$

and the gravitational force

$$m_i a = -\frac{Gm_g M}{r^2}$$

it appears that the ratio of m_g, the gravitational "charge", and m_i the inertial mass is constant. In this case

$$a = -\frac{GM}{r^2}$$

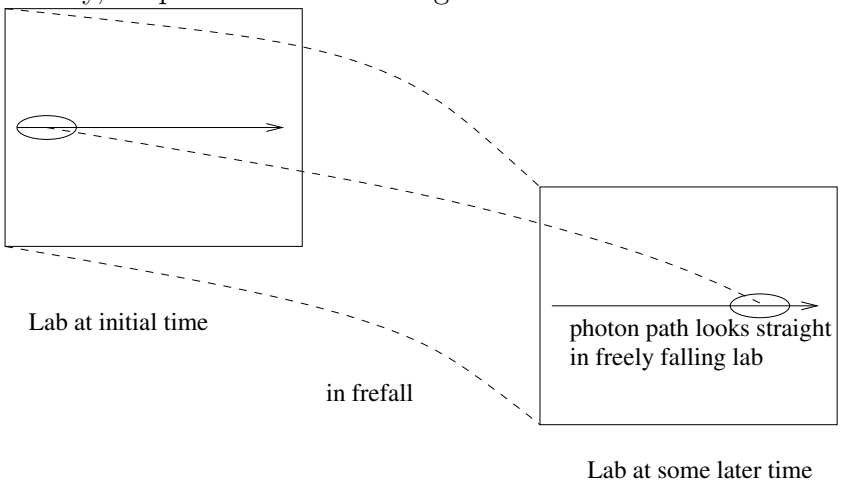
So how can there every be an inertial frame if all frames feel acceleration due to gravity?

The solution to this conundrum is that all objects do feel the same acceleration absolutely – i.e. assume that m_i **and** m_g **are not different things.** This means that all objects in a small region will feel the same a and therefore will move together. i.e. a frame free in a gravitational field acts inertial despite being accelerated.

This principle forms a similar role in general relativity that the constancy of the speed of light does in special relativity. It is the postulate from which we derive the expected behavior of experiments.

23.2 Effects on light - bending

The first effect to consider is the bending of light. If a photon crosses a freely falling laboratory, its path should be straight in the local inertial frame of the laboratory.



As such, an observer outside the laboratory will see the path of the light curved. This is also consistent with **both observers seeing the light follow the shortest path** in their integrated history. The free-floating observer sees the the photon move properly in the local flat space by moving with it, while the external observer sees the photon find the shortest path due to the curved space that is also causing the free-floating lab to accelerate.

We will consider gravitational redshift next time.

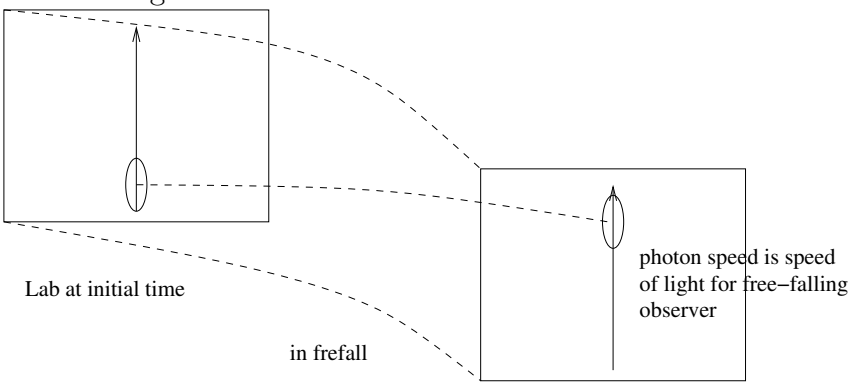
24 Astro notes 2018/10/19 - Fri - General relativity

Last time we saw that an observer outside a freely falling laboratory will see the path of the light curved. This is also consistent with **the light follows the shortest path in both coordinate systems (frames)** in their integrated history. The free-floating observer sees the the photon move properly in the local flat space by moving with it, while the external

observer sees the photon find the shortest path due to the curved space that is also causing the free-floating lab to accelerate.

24.1 Effects on light - gravitational redshift

The postulate that a free-falling frame is inertial can also be used to derive the gravitational redshift. Consider a laboratory in free-fall that emits a photon from the floor and detects it at the ceiling.



For the laboratory to be an inertial frame, the observer must not see a shift in the frequency. However, from the point-of-view of the frame in which the observatory is in free-fall, on detection the observer has gained a downward speed of $v = gt = gh/c$ where h is the height the photon is allowed to travel in the lab. In this case the Doppler blue shift should be
(student)

$$\left(\frac{\Delta\nu}{\nu_0}\right)_{observer\ \ motion} = \frac{v}{c} = \frac{gh}{c^2}$$

But, since the free-falling frame is known to be inertial, there must be a compensating effect of gravitational redshift

$$\left(\frac{\Delta\nu}{\nu_0}\right)_{grav} = -\frac{gh}{c^2}$$

so that, for the lab observer, $\Delta\nu/\nu_0 = 0$. Thus photons crossing small distances in a gravitational field are redshifted if they are moving upward. The actual expression for light escaping from a gravitational field is

$$\frac{\nu_\infty}{\nu_0} = \left(1 - \frac{2GM}{r_0 c^2}\right)^{1/2}$$

where r_0 is the radius at which emission occurs, where ν is ν_0.

But where does this arise from? Clocks tick at a different rate different depths within a gravitational field.

$$\frac{\Delta t_0}{\Delta t_\infty} = \left(1 - \frac{2GM}{r_0 c^2}\right)^{1/2}$$

This is called gravitational time dilation and has been measure extensively. $\Delta t_0 < \Delta t_\infty$.

24.2 Structuring the structure of spacetime

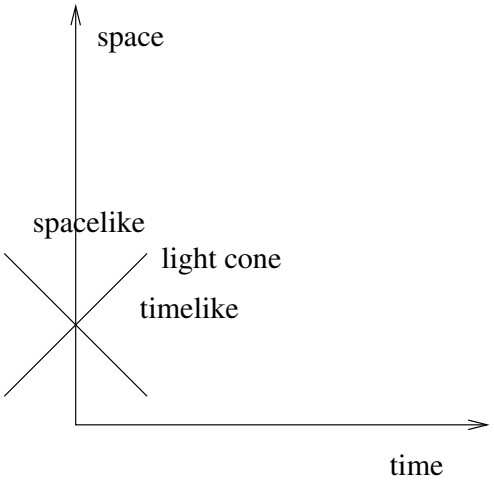
Coordinates vs. spacetime: Before general relativity, it has so far been essentially possible to use coordinates themselves directly to compute things like intervals. For example time differences are simply $t_2 - t_1$.

However, already in special relativity this has become less true. We discussed this as how things were "measured by different observers" but a better way to think about this is that the result of any given "observation" must be computed, and sometimes the coordinates (time for example) in which that result are computed are not the same as the measured analog of that coordinate (e.g. time). For example, using proper intervals (see below), the time interval experienced on a moving spaceship can be computed in the solar system's frame, but it is not the same as the time coordinate in the frame used to compute it.

This drives us to separate coordinate systems from physical intervals, which is closer to reality. Then physical things are generally computed by integrating over a path, rather than differencing coordinates. This is a manifestation of the fact that spacetime is inherently *local*.

Flat space is that special circumstance in which intervals can be measured by differencing for things "at rest" in that frame.

So, if we have given up on space and time coordinate systems what do we have left? Space and time directions! At each point in spacetime, which we usually call an "event", spacetime directions are well defined. This is called the light-cone:



Local spacetime directions are then measured with respect to the light cone, with space-like directions lying outside the light cone and time-like ones within it. Causality, critical for physics, flows strictly within the lightcone.

25 Astro notes 2018/10/22 - Mon - GR: intervals, Schwarzchild metric

25.1 Proper intervals

As we have seen in special relativity, the "distance", if defined as the same-time separation, between separated points can depend on the speed of the observer. i.e. it is not a universal thing. The analog of distance that is not observer-dependent is the spacetime interval.

$$I = (\Delta s)^2 = [c(t_2 - t_1)]^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

(this assumes a flat spacetime) note there is a sign convention ambiguity here, so that some prefer

$$I = -[c(t_2 - t_1)]^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

Whether the interval I is positive or negative determines if it is timelike or spacelike, based on adopted sign convention. Note that in flat space without motion (or without time elapsed), the interval reduces to either the time (or space) separation. "proper time" and "proper distance" are defined this way, in a frame at rest with respect to the object/clock under consideration.

Note that:

- Particle paths are exclusively timelike.
- Events with spacelike separations cannot be causally connected.
- Light moves on zero interval paths.

But we have said that space is local, not global like these intervals imply. This leads us to instead define the local "metric" for how to measure small pieces of intervals:

$$(ds)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2$$

Now intervals are found by integrating along a worldline

$$\Delta s = \int_{path} \sqrt{(ds)^2}$$

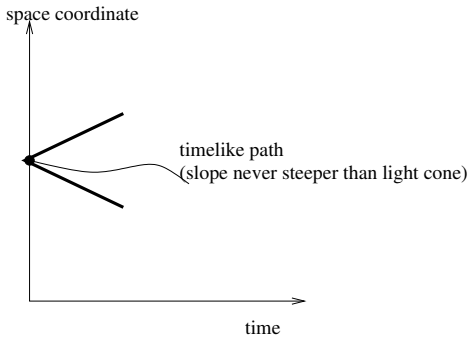
We can now actually finally compute the time interval experienced by a clock moving on an arbitrary path:

$$\Delta t_{proper} = \frac{\Delta s}{c}$$

This works for any path in any coordinate system and does not depend on anything special about the "observer". Computing the time interval elapsed for an observer in flat space is a similar computation, but with a path that has no spatial components (at rest), and therefore matches the "t" coordinate.

If we want to find the path of a free object, that is a bit more complex. A free-falling object follows a geodesic – a path that extremizes the spacetime interval (max or min depending on sign convention). For flat space that is a straight line path.

25.2 Clarity on sign convention



With the + - - - convention, as in the text, positive spacetime intervals are *timelike*. So the proper time (the time measured by a clock moving along a given path) is given by

$$\Delta\tau = \frac{\Delta s}{c} = \frac{1}{c}\int_{path}\sqrt{(ds)^2}$$

where ds^2 is given from the coordinates by the metric. Note that if ds^2 is negative (spacelike) *at any point in the path* the proper time is not defined. For purely spacelike paths the spacelike distance is

$$L = \int_{path}\sqrt{-(ds)^2}$$

In this case also free particle paths are geodesics with maxiumum spacetime interval.

With the - + + + convention,

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

and negative spacetime intervals are timelike. So the proper time is

$$\Delta\tau = \frac{\Delta s}{c} = \frac{1}{c}\int_{path}\sqrt{-(ds)^2}$$

Now if ds^2 is positive (spacelike) the proper time is again not defined. In this case free particle paths are geodesics with minimum spacetime interval (i.e. most negative spacetime interval).

25.3 Curved space - the Schwarzschild metric

I will give one example of curved space - the metric for the space around a spherically symmetric gravitating object like a planet or star or black hole. This is

$$ds^2 = \left(1 - \frac{2GM}{rc^2}\right)c^2dt^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\phi^2$$

Note, for example, that this has gravitational time dilation built into it. The time interval for a clock at some radius r is different than for one further out.

For comparison, the metric of flat space-time (i.e. no gravitational sources) in similar coordinates is

$$ds^2 = c^2dt^2 - dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\phi^2$$

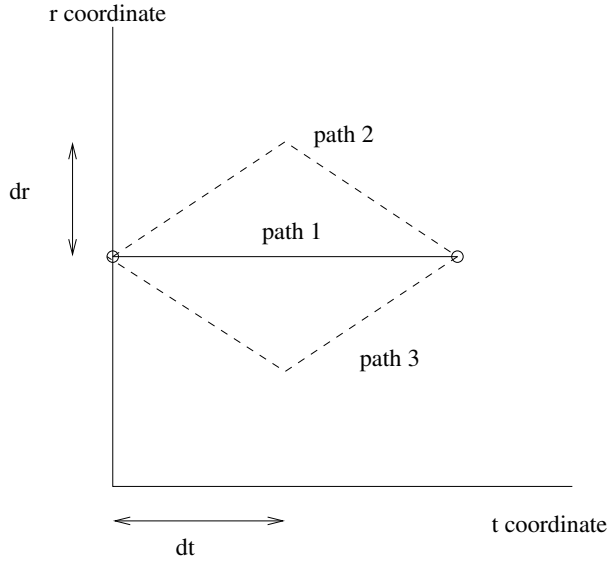
We see that, for the Schwarzschild metric, as represented by the rubber sheet analogy, the distance between r and $r + dr$ in space-time is

$$\Delta s = \frac{dr}{\sqrt{1 - 2GM/rc^2}}$$

which is larger than dr for all radii.

But if we integrate a circle at a fixed r in either metric (fixing $\theta = \pi/2$ and running ϕ from 0 to 2π so $ds = rd\phi$) we obtain the circumference $2\pi r$.

We said that particles move along paths of extremal distance. We will investigate this by first showing that in the flat spacetime metric (also called Minkowski spacetime) a straight path is the longest spacetime interval. Consider the following paths through spacetime (we will only use the radial space coordinate). So this is a free-falling path with the same r at two times.



Now we will compare the spacetime interval Δs integrated along each paths. For flat space, path 1 is

$$\Delta s_1 = 2cdt$$

For path 2 we have

$$\Delta s_2 = 2\sqrt{c^2dt^2 - dr^2}$$

And since the sign of dr doesn't matter due to the square, this is also the length of path 3. So we see that for any $dr > 0$, that path 1 has a larger spacetime interval than path 2 or 3. i.e. it is the extremal path and therefore the one along which a particle would travel between these the initial and final points. Also we find from here that the extremum we want for a particle path is actually the maximum.

next time we will consider the Schwarzchild metric...

26 Astro notes 2018/10/24 - Wed - Schwarzchild metric and Black Holes

26.1 Curved space - the Schwarzschild metric

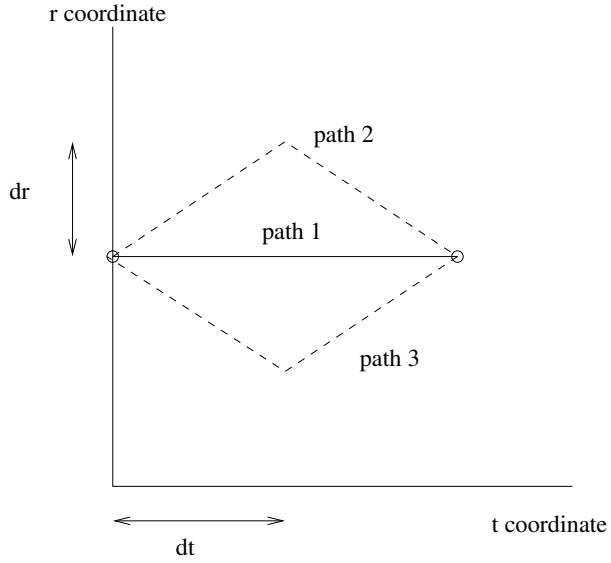
Last time we considered paths in flat space, this time we will use the Schwarzchild metric, which is:

$$ds^2 = \left(1 - \frac{R_S}{r}\right) c^2 dt^2 - \left(1 - \frac{R_S}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2$$

where $R_S = 2GM/c^2$ is called the Schwarzchild radius.

For this sign convention, particle paths, which are timelike geodesics, are paths with maximum interval. (It turns out that this is the analog of the "shortest distance". One way to think about this is that the shortest distance path will have the largest increment of proper time.)

Consider the same paths as before



Now we will repeat the analysis in the presence of gravity, that is, for the Schwarzschild metric. Assume that path 1 is at constant radius R and so has $dr = 0$, $d\theta = 0$, and $d\phi = 0$, then

(student 1)

$$\Delta s_1 = 2\sqrt{1 - \frac{2GM}{Rc^2}} c dt = 2\left(1 - \frac{R_S}{R}\right)^{1/2} c\, dt \ .$$

We now wish to compare this with the distance along paths 2 and 3. Assuming dr is small, and still that $d\theta = 0$ and $d\phi = 0$, the distance along path 2 can be approximated using the coefficient evaluated at the mipoint $r = R + dr/2$:

(grad student 2)

$$\Delta s_2 \simeq 2\left[\left(1-\frac{R_S}{(R+dr/2)}\right)c^2dt^2-\frac{dr^2}{1-\frac{R_S}{(r+dr/2)}}\right]^{1/2}$$

where we have defined the Schwarzschild radius $R_S = 2GM/c^2$.

For small dr we can assume that the first term, the dt term, dominates. Also putting in

(student)

$$(R+dr/2)^{-1} = R^{-1}(1+dr/2R)^{-1} \simeq R^{-1}(1-dr/2R)$$

we get

$$\Delta s_2 \simeq 2\left(1-\frac{R_S}{R}+\frac{R_S}{R}\frac{dr}{2R}\right)^{1/2} cdt > \Delta s_1 \ .$$

Path 3 can be evaluated similarly, but dr is the opposite sign:

$$\Delta s_2 \simeq 2\left(1-\frac{R_S}{R}-\frac{R_S}{R}\frac{|dr|}{2R}\right)^{1/2} cdt < \Delta s_1 \ .$$

Thus we find

$$\Delta s_2 > \Delta s_1 > \Delta s_3$$

So the extremal path will be more like path 2. This is the expected result, since path 2 is like a parabolic trajectory as expected under gravitational acceleration. For large enough dr the second term in Δs_2 will start to dominate, causing it to decrease, and therefore give the actual extremal path - which is the path of a particle.

26.2 Black Holes

The the solution to Einstein's equation for a point mass is the Schwarzschild metric, as we saw last time.

$$ds^2 = \left(1-\frac{2GM}{rc^2}\right)c^2dt^2-\frac{dr^2}{\left(1-\frac{2GM}{rc^2}\right)}-r^2d\theta^2-r^2\sin^2\theta d\phi^2$$

Note that this is only the metric *outside* of a finite-size object like a planet or star. Nominally this means the singularities $(1/r)$ are not so much of a problem. However there is an important radius, called the schwarzschild radius

$$R_S = \frac{2GM}{c^2} \approx 3km\left(\frac{M}{M_\odot}\right)$$

Note that the dr component involves $1/(1-R_S/r)$, which is actually singular at $r = R_S$. It turns out that this is just a "coordinate" singularity, not a physical one.

For any solar-mass object bigger than 3km in radius, this is really not important, because the vacuum solution is no longer valid there. This includes white dwarfs (about the size of the earth, few thousand km in radius) and neutron stars (about 10km in radius).

However for objects with $R < R_S$, we say these are black holes, as light cannot escape. Now to talk about that...

26.3 event horizon

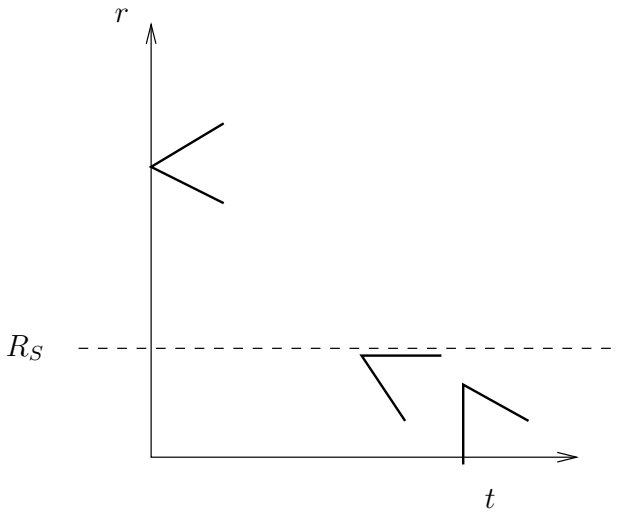
Consider a light geodesic:

$$0 = ds = \left(cdt\sqrt{1 - R_S/r}\right)^2 - \left(\frac{dr}{\sqrt{1 - R_S/r}}\right)^2$$

Then

$$\frac{dr}{dt} = c(1 - R_S/r)$$

so that at the event horizon $dr/dt = 0$ for (outgoing) light. i.e. the light cone of an event at the event horizon only extends along and inward from the event horizotion. Information cannot propagate outward from this radius. This "hides" any internal structure of the black hole from outside.



however things can still pass into the black hole.

Note two things about a path into the black hole. from outside, it takes until infinite time for the infalling stuff to reach R_S. However, time passes fine for the observer falling in.

The problem is actually tidal force - the difference in the force between near and farther, which will "stretch" a finite-size infalling object.

27 Astro notes 2018/10/29 - Mon - Close Binaries

27.1 Gravity in close binaries

Close binaries can be treated mostly by Newton's law of gravity. But it is often worthwhile to treat the potential in the rotating instead of non-ratating frame. The potential has two types of components: The potential from each of the stars:

$$U_s = -G\frac{Mm}{r}$$

one for each star. In order to be in the rotating from, there is an additional force, the centrifugal force $F = m\omega^2 r$. This leads to an overall potential:

$$U = -G\left(\frac{M_1 m}{s_1} + \frac{M_2 m}{s_2}\right) - \frac{1}{2}m\omega^2 r^2$$

where s_i is the distance from star i.

equipotential surfaces are show in figure 18.3

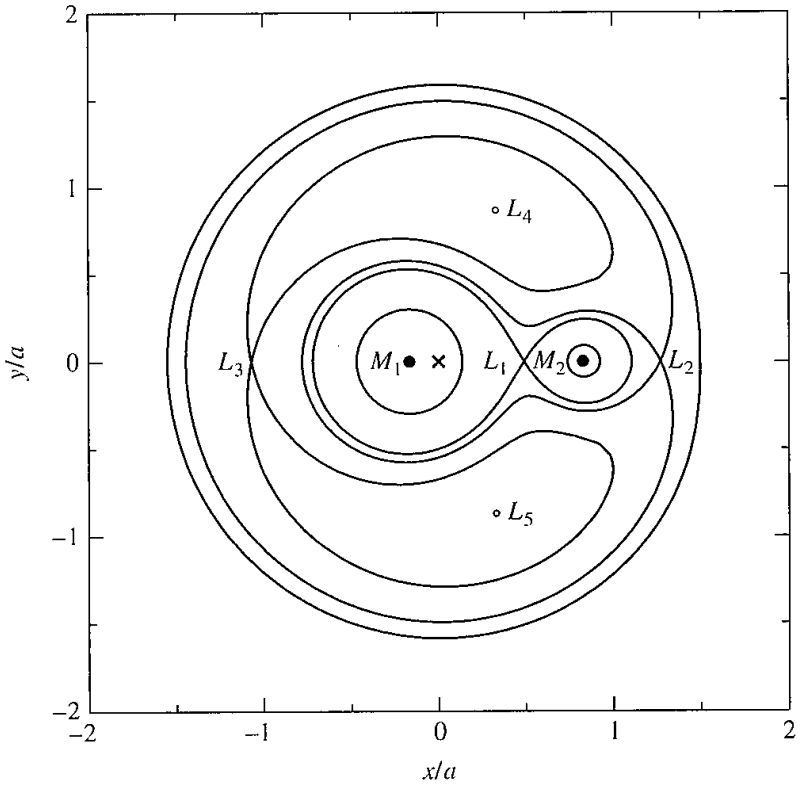
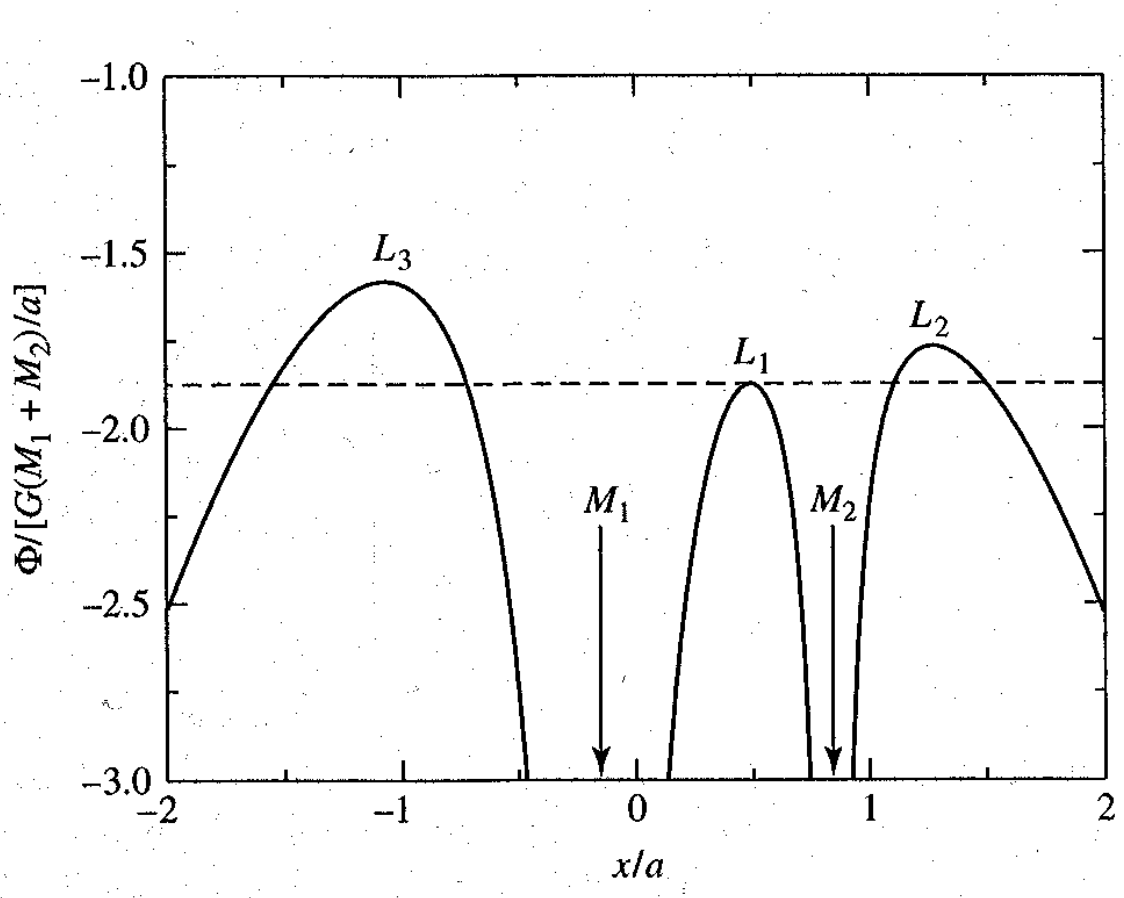


FIGURE 18.3 Equipotentials for $M_1 = 0.85\,M_\odot$, $M_2 = 0.17\,M_\odot$, and $a = 5 \times 10^8\,\text{m} = 0.718\,R_\odot$. The axes are in units of a, with the system's center of mass (the "×") at the origin. Starting at the top of the figure and moving down toward the center of mass, the values of Φ in units of $G(M_1 + M_2)/a = 2.71 \times 10^{11}\,\text{J kg}^{-1}$ for the equipotential curves are $\Phi = -1.875, -1.768, -1.583, -1.583, -1.768$ (the "dumbbell"), -1.875 (the Roche lobe), and -3 (the spheres). L_4 and L_5 are local maxima, with $\Phi = -1.431$.

Along the centerline through both masses this looks like this (fig 18.2):



27.2 Lagrange points

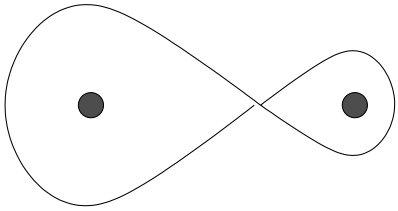
The Lagrange points are points of zero gradient in potential. They are all maxima or saddle points, and therefore unstable. The most important one is L_1 through which material can be transferred from being bound to one star to being bound to the other.

27.3 types of binaries

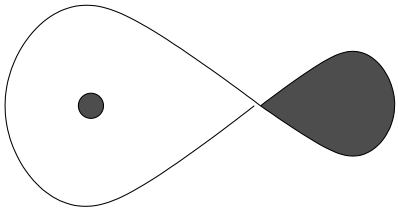
There are several general types of binaries

- Detached - just two stars
- semi-detached - one star filling its Roche lobe
- contact - both stars filling their Roche lobes

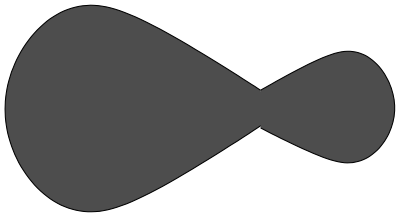
In semi-detached binaries, commonly mass is being transferred from the Roche-lobe filling star to the other star.



Detached



Semi−detached



Contact

(primary and secondary are backwards from typical convention in the textbook. The more massive star has the larger Roche lobe and is also typically called the primary. Also, being more massive, it is typically the first to evolve into a giant star.)

If evolving, the first star to fill its Roche lobe is the more massive one, because it evolves first, but mass transfer from a more massive to a less massive star is unstable. Why? because moving material away from the center of mass causes the binary to shrink - thus causing more mass transfer. This instabality leads to a common envelope phase.

After this the star that evolved becomes a WD, then, if the second star is less massive than the WD when it fills its Roche lobe, the mass transfer process is stable.

27.4 Drivers of mass transfer

Mass transfer can be *driven* by several factors

- Orbital stability as described above - very short-term (10^4 years)
- wind - typically in a detached binary. If the primary is an evolved star with high mass loss, some of that wind can be captured by the other star. With a WD accreting these are one type of "symbiotic" star.
- growth of Roche-lobe-filling star due to nuclear evolution - will evolve on timescale of late evolution, typically tens to hundreds of millions of years (10^7-10^8 yr)

- angular momentum loss as driver of binary evolution - causes binary to shrink
  - magnetically attached wind - as the binary loses mass like the solar wind (i.e. not fast mass loss) that wind can carry away angular momentum. Typically evolves on $\sim 10^8$ year timescales
  - gravitational radiation - very slow, system evolves on 10^9 year or longer timescales.

For gravitational radiation

$$\dot{P} = -\frac{96}{5}\frac{G^3 M^2 \mu}{c^5}\frac{P}{a^4} = -\frac{96}{5}\frac{G^3 M^2 \mu}{c^5}\left(\frac{4\pi^2}{GM}\right)^{4/3}\frac{1}{P^{5/3}}$$

28 Astro notes 2018/10/31 - Wed - Close Binaries: accretion

28.1 Nomenclature of semi-detatched binaries with a compact object

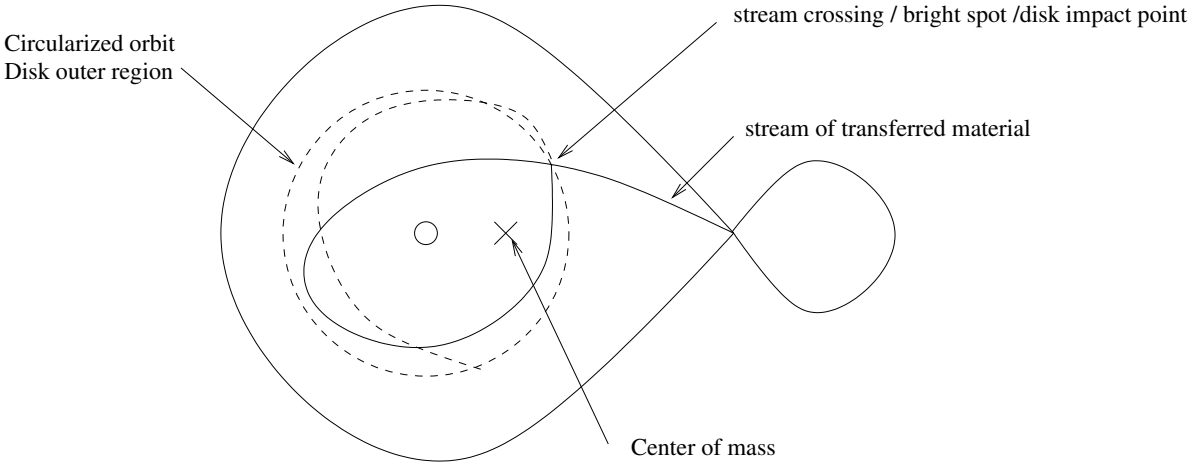
For semi-detached binaries that are transferring matter onto their primary star, the name depends on the type of primary compact object.

- White dwarf star - called Cataclysmic Variables
  Types of Outbursts:
  - "Classical" Novae - Thermonuclear outbursts - hydrogen-rich material accreted on surface leads to re-ignition of shell burning and mass ejection
  - "Dwarf" Novae - Thermal instabity in accretion disk - accretion disk is only actively putting matter on the primary star periodically in bursts
- Neutron star - called low-mass X-ray binaries
  These also have the same two types of "outburst", but the thermonuclear outburst is called an "X-ray burst", and the disk outburst doesn't have a particular name, typically called "high state".
- Black hole - still called an X-ray binary
  no thermonuclear bursts since there is no surface causally connected to the outside spacetime

There are also high-mass X-ray binaries where the primary star is the high-mass star and the lower mass star is the neutron star or black hole. These are typically not roche-lobe filling, so that the mass transfer is not unstable, and therefore are wind-accretion.

28.2 Impact point and circularization

When material leaves the L_1 point, it is not in the disk. It is on a non-circular and non-closing orbit that will cross itself due to the irregular potential in the binary. This leads circularization of the orbits of the material and eventually to a disk impact point – sometimes called the "bright spot" on the disk. This is near the outer edge of the disk and is where material is added to the accretion disk from the companion star



28.3 accretion disk and its temperature profile

Accretion disk moves material toward the star by moving its angular momentum outward. (so material can move inward) Angular momentum is thought to be coupled back into the orbit through unclear means (disk tidal effects). The overall disk luminosity from gravitational potential energy

$$L_{disk} = \frac{GM\dot{M}}{2R}$$

Where R is the inner edge of the disk.

The temperature profile of the disk is given by equating the gravitational energy release to the local energy emitted by the disk for mass flowing through at some rate $\dot{M}$.

$$dL_{ring} = 4\pi r\sigma T^4 dr = G\frac{M\dot{M}}{2r^2}dr$$

This gives a profile that looks like

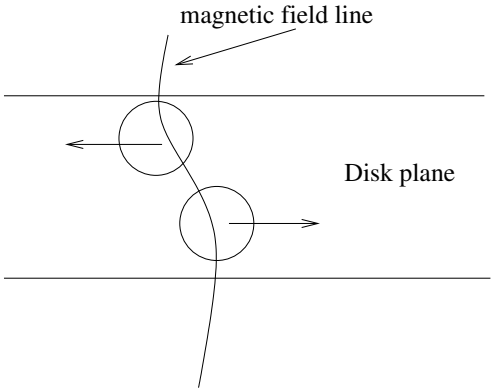
$$T = \left(\frac{GM\dot{M}}{8\pi\sigma R^3}\right)^{1/4}\left(\frac{R}{r}\right)^{3/4}$$

where R is the radius of the central object - the hottest part of the disk.

28.4 Disk viscosity and instability

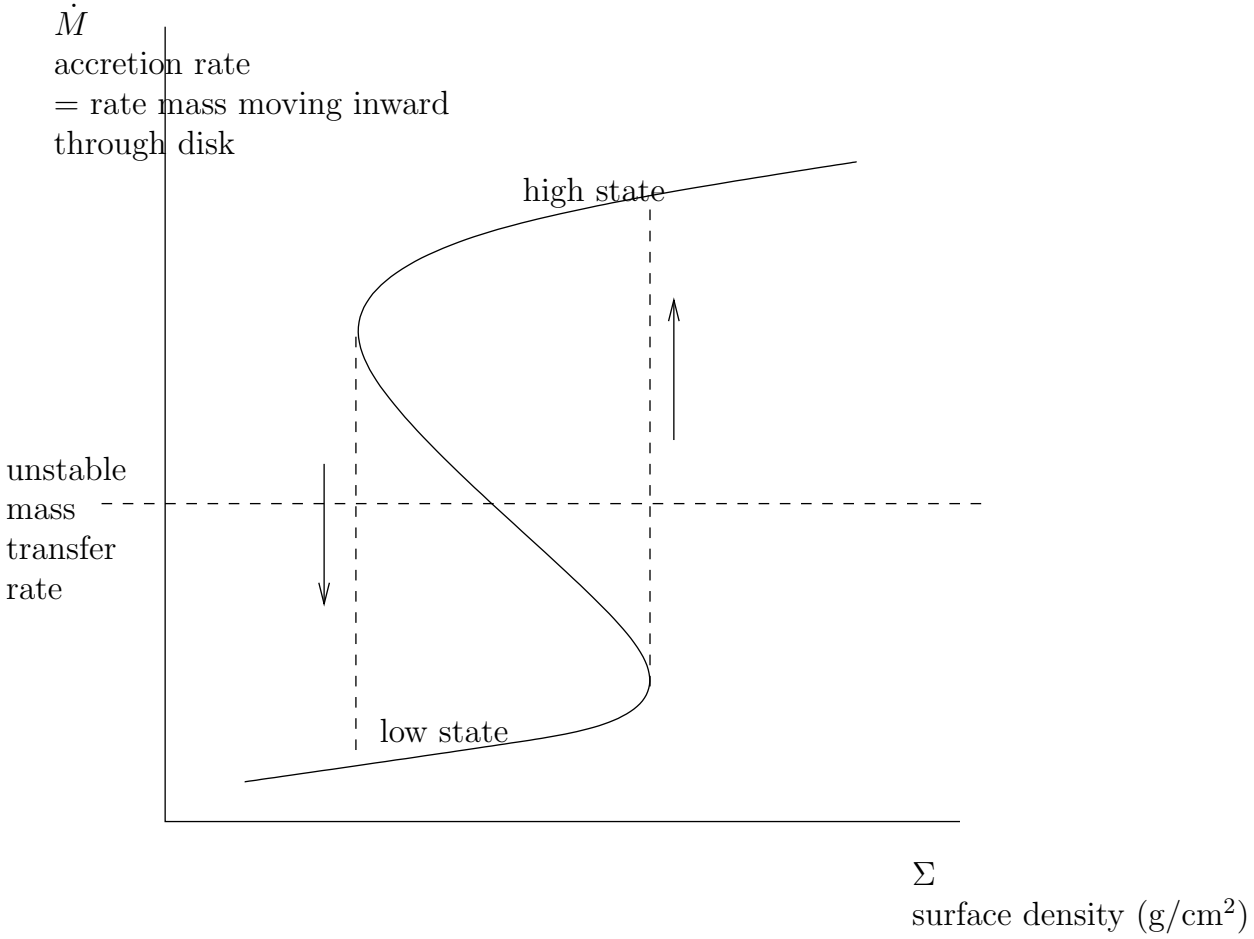
Disk viscosity is what transfers angular momentum in the disk. This is believed to be driven by something called the magneto rotational instability or MRI.

If a fluid parcel in the disk moves to larger radius, it will want to rotate more slowly. If this parcel is magnetically connected to a more inner parcel, this will have two effects: first, it will impart angular momentum to the outer parcel - causing it to move outward even more; second it will stretch and strengthen the magnetic field lines.



However, the MRI requires a hot disk in order to sustain the plasma that hosts the magnetic field. This can lead to instability, which leads to dwarf novae.

The disk needs enough surface density Σ to stay optically thick and "hold" heat to maintain an ionized plasma. Here Σ is the mass per area when considering a column passing vertically through the accretion disk. When Σ is large, the disk is optically thick and can stay hot and ionized, therefore hosting a magnetic field and having a high effective viscocity. This is often called the "high" state.



If the mass transfer rate from the companion star into the outer disk is in the intermediate (unstable) range, this leads to instability and limit-cycle behavior. The disk will build up matter, increasing surface density Σ until it is able to hold heat (become optically thick) and then become ionized. At that point the disk becomes magnetically active and the fields begin to allow material to lose angular momentum and move inward. This continues until Σ is too low to sustain the temperature for ionized plasma, at which point the disk switches back to the low state and begins accumulating surface density again.

For a WD primary star, this type of outburst, caused by a thermal instability in the disk, and the accompanying magnetic activity, is called a dwarf nova. For neutron star systems it doesn't have a particular name but is observed as a transience of the mass transfer onto the neutron star.

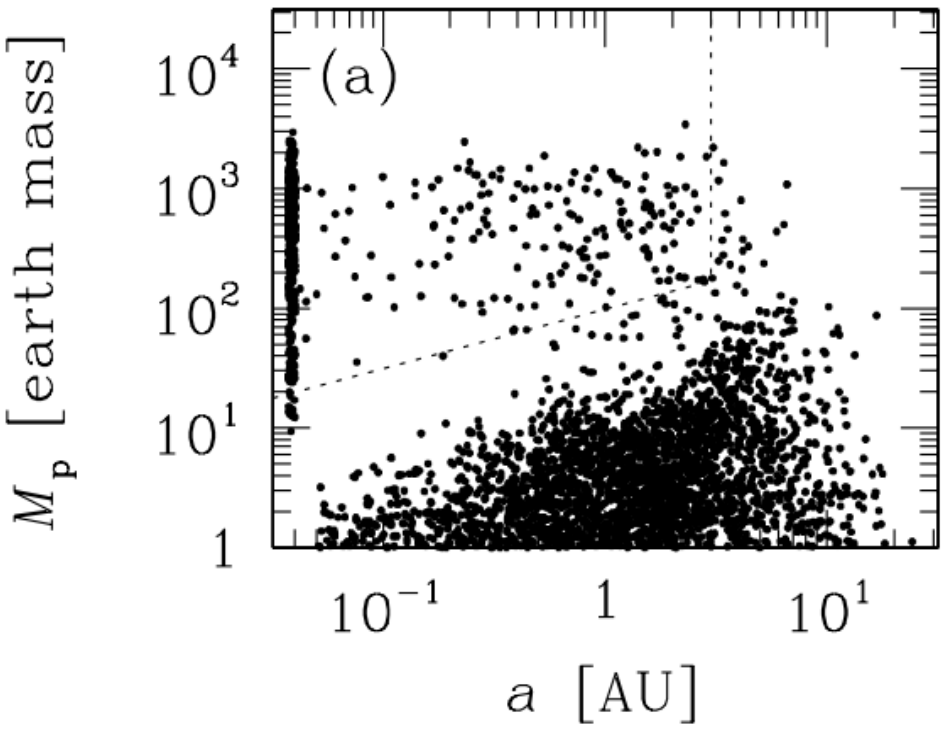
29 Astro notes 2018/11/2 - Fri - planet formation

29.1 Exoplanet systems and Planet formation

Planets are thought to be formed by accretion/merger of small objects. This kind of "bottom-up" planet formation scenario is quite different from the collapse and fragmenta-

tion model of star formation.

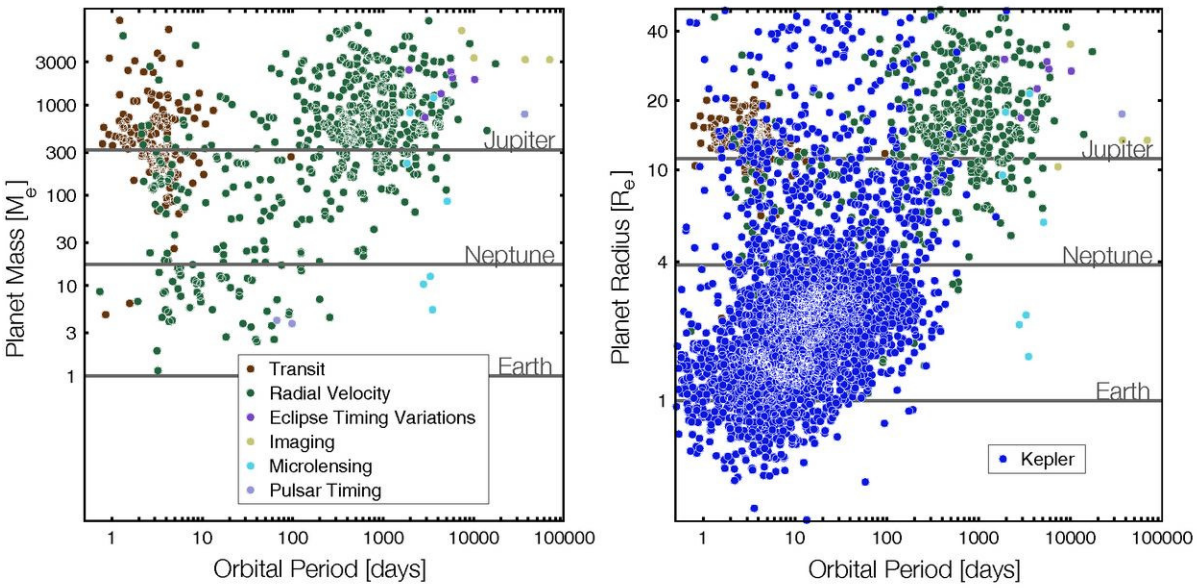
The bottom-up type model, however, makes specific predictions. It predicts a large number of small planetary bodies that are the "leftovers" of the bottom-up process. The truncation of the process has to do with the lifetime of the protoplanetary disk, and so planet formation can be stopped at different times, giving different leftovers in different star systems. Predicted planet distrubition from Ida & Lin (2004, http://adsabs.harvard.edu/abs/2004ApJ...616..567I):



The dashed lines indicate the area of parameter space (high mass, short period) accessible to investigation with measurements of stellar radial velocity (i.e. the gravitation effect on the parent star of the planet). This is hard to detect for small, low-mass planets, which is the bulk of the distribution.

The Kepler satellite mission was designed to prove or disprove the bottom-up model for planet formation by pushing detection limits down to low enough masses to measure the predicted plethora of low-size, low-mass planets. This was done by using a different method - transits of the planets across the star.

The distribution of planets found by Kepler shows the characteristic increase of planetary bodies to small sizes. (Note the vertical scale here is now is size rather than mass, but the two are somewhat related.)



Above figure from Batalha (2014PNAS..11112647B)

The actual main Kepler discover paper Borucki et al. (2011, http://adsabs.harvard.edu/abs/2011ApJ...728..117B) (see figure 2) compares the number distribution of planet candidate radii, and shows that it follows an approximate $1/R^2$ trend (where R is the radius of the planetary body) of the type expected from the bottom-up formation theory.

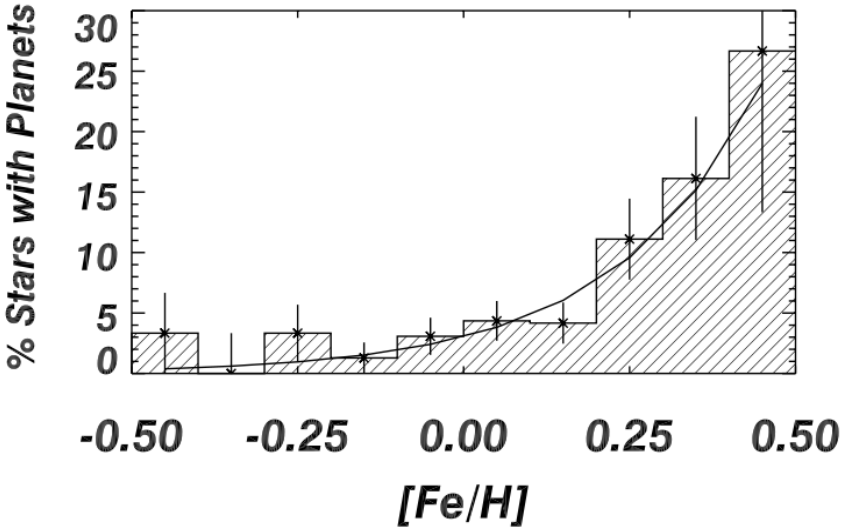
The Kepler result was not unexpected, but was critical to confirm in order to understand how much like our own other planetary systems might be. In advance of the Kepler mission, there were good reasons already to believe the bottom-up scenario. The metallicity distribution of giant planets indicates that planets form from bottom-up.

Giant planets form when a growing planet becomes large enough to capture gas gravitationally. Higher metallicity systems have more material to build planet cores out of, and therefore a better chance at a core large enough to undergo core accretion stage.

Recall metallicity here means all material other than H and He, so it includes carbon and oxygen, which can make water and hydrocarbon ices. The conventional notation for metallicity is so-called "bracket notation"

$$[Fe/H] = \log_{10}\left(\frac{n_{Fe}/n_H}{n_{Fe,\odot}/n_{H,\odot}}\right)$$

where n is the number density (typically in the photosphere). so zero is the same fraction of Fe, and therefore other elements (other than H or He) as that in the Sun.



This is from Fischer & Valenti (2005, http://adsabs.harvard.edu/abs/2005ApJ...622.1102F) and is also discussed in our textbook because it was printed before 2011, when Kepler made its measurements. This was the best indication of the bottom-up theory at the time.

29.2 Basics of theory of planet formation

Want to consider how a "big body" can grow into a planet-sized body by accreting smaller bodies from the protoplanetary disk. The basic assumption is that there is naturally a distrubution of many small bodies and a few slightly larger ones that can start the formation process.

Hill radius is radius at which a big body's gravity is stronger than Sun's tidal forces. occurs when keplerian period for orbit of sun and orbit of body are similar

$$\sqrt{\frac{a^3}{M_\odot}} = \sqrt{\frac{R_H^3}{M}}$$

where M is the mass of the "big body" that will grow and a is the distance from the sun. This gives

$$R_H = \left(\frac{M}{M_\odot}\right)^{1/3} a$$

or conventionally written

$$R_H = \frac{R}{\alpha}$$

where

$$\alpha = \left(\frac{\rho_\odot}{\rho}\right)^{1/3} \frac{R_\odot}{a}$$

Where R is the radius of the big body. Note that $\rho_\odot/\rho$ is not too different from 1 because both typical matter and the sun have an average density of about 1 g/cc. The latter factor

$R_\odot/a$ is the angular size of the sun from an orbit of size a. From earth $\alpha \sim 10^{-2}$ and from the Kuiper Belt $\alpha \sim 10^{-4}$. R_H is thus bigger further from the sun.

Show Hill radius and comparison of hill to disk thickness: Goldreich, Lithwick, Sari (see notes webpage)

Actual collision rates depend on density of particles in the disk (σ/m) and how their velocity dispersion compares to R and R_H.

There is a Trade-off: lower velocity dispersion, u in this paper, means more efficient capture but from a smaller interval of orbits (this leads to smaller planets). Modeling seeks to find right u and possible mechanisms to establish it. This can lead to different "scenarios" for any given outcome. For example Uranus and Neptune may have formed further inward and migrated out, implying they formed more quickly than if they were farther out where they are now, thus requiring different assuptions about velocity dispersion.

Also revisit predicted planet distrubition from
Ida & Lin (2004, http://adsabs.harvard.edu/abs/2004ApJ...616..567I)

Note that this gives $M(t)$ and its dependence on various assumed parameters.

30 Astro notes 2018/11/5 - Mon - Galaxies

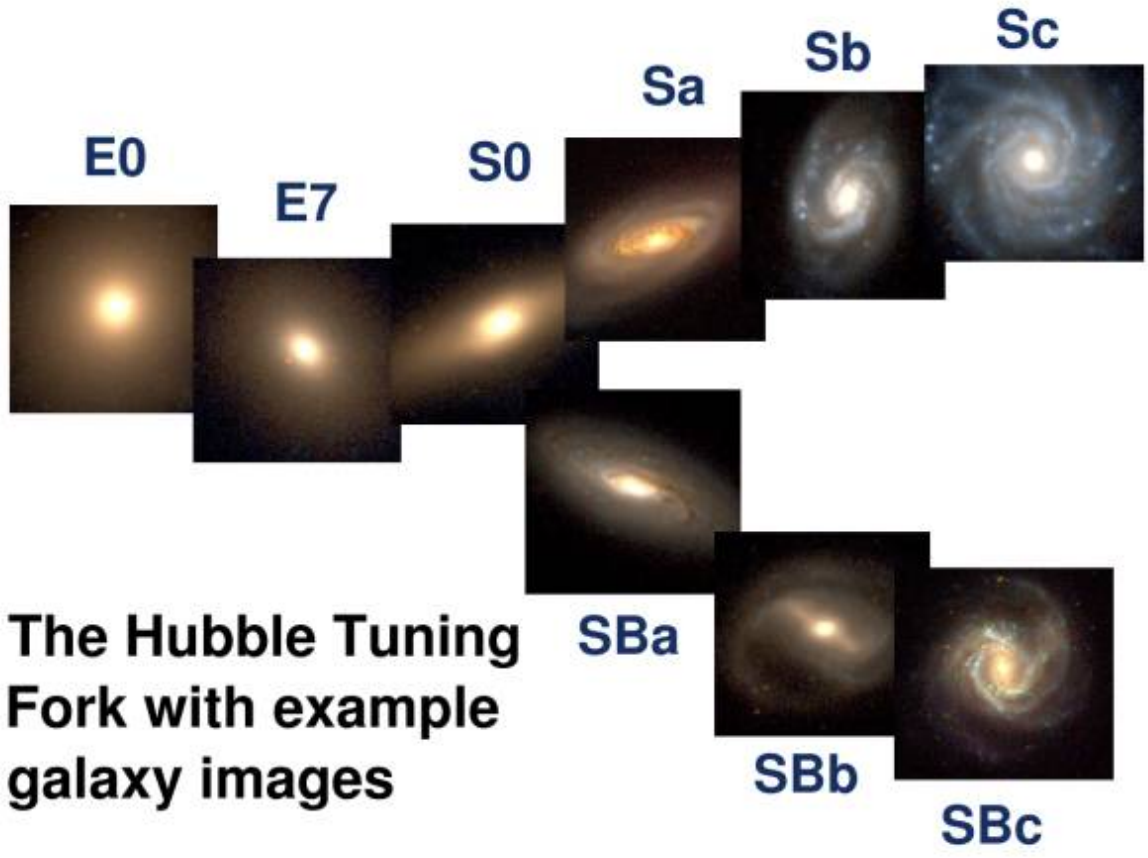
30.1 Galaxy fundamental parameters

The two major physical parameters determining the properties of galaxies are their **total mass** and their **specific star formation rate** (or more generally star formation history.

Note that SSR is related / corellated with many other properties: luminosity-weighted stellar age, color, presence of disk and spiral structure (called morphology, often quantified as bulge/disk ratio), presence of cold gas. Exactly which of these is the cause - and which others are the result - is not entirely clear. Also it is apparent that galaxies can change state and morphology in response to their evolution and environment. e.g. mergers, gas accretion, stripping, depletion of cold gas. Note that with mergers the identity of "a galaxy" is no longer well defined, and mergers are ubiquitous.

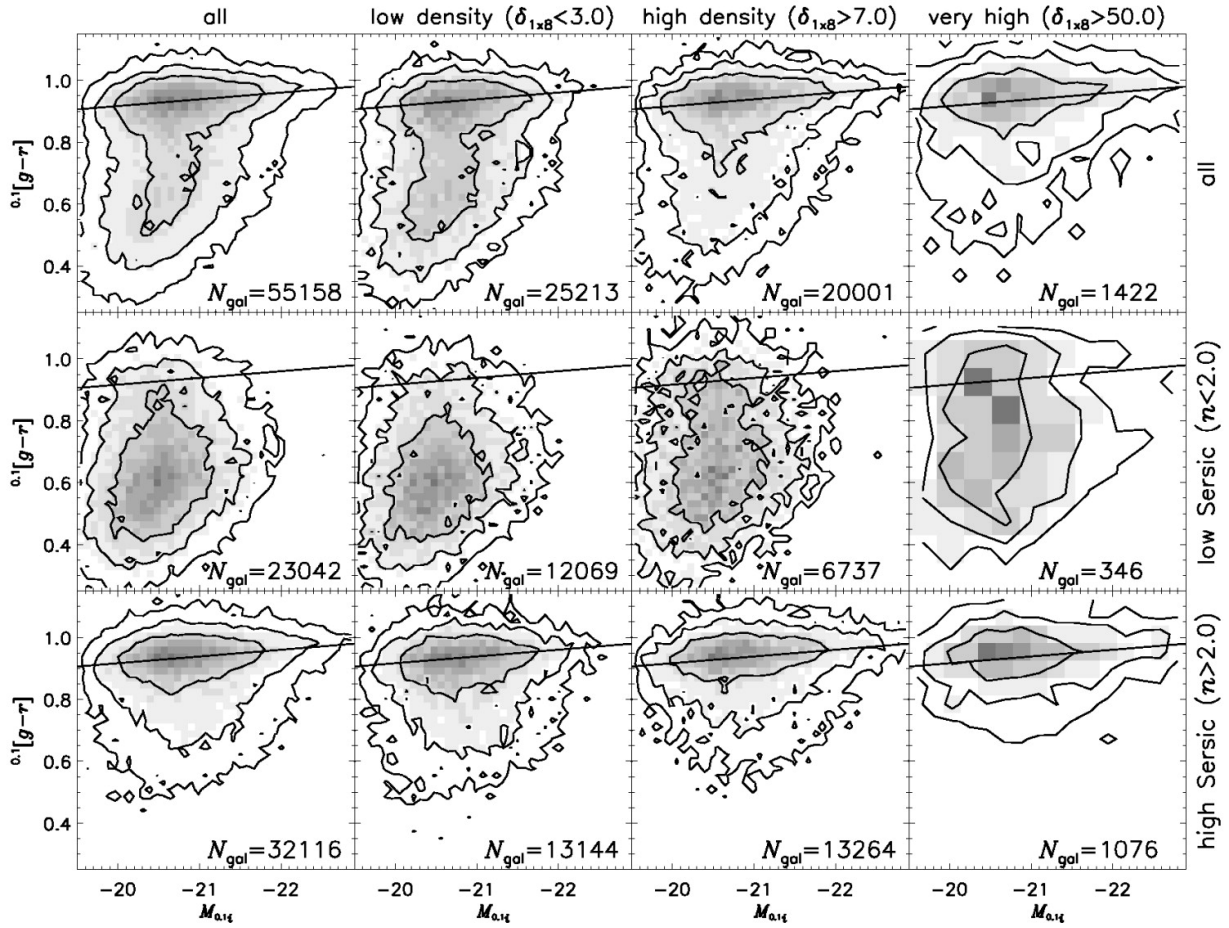
Two views of galaxies:

1. galaxy shape / morphology typically discussed via the hubble type



From left to right this sequence correlates with SSR and all the other properties mentioned above.

2. Color-magnitude, from Hogg et al. (2004, (http://adsabs.harvard.edu/abs/2004ApJ...601L..29H)

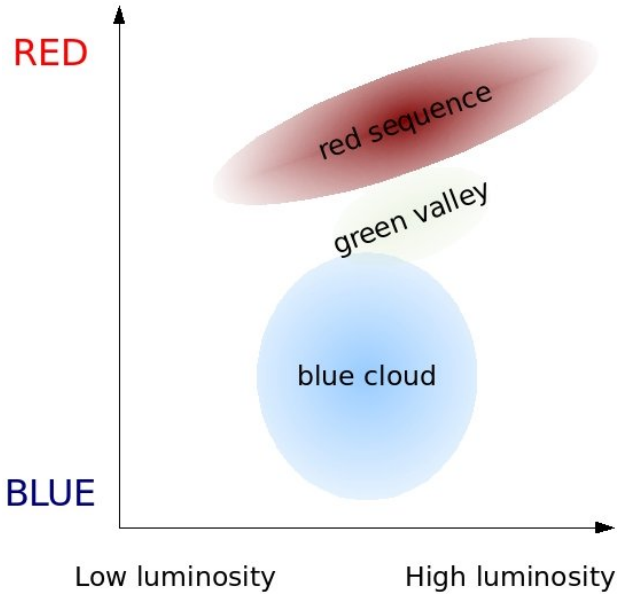


A measure of morphology can be obtained by fitting a Sersic profile of surface brightness:

$$\mu(r) = \mu_e + 8.3268 \left[\left(\frac{r}{r_e} \right)^{1/n} - 1 \right]$$

where n, r_e and μ_e are fitting parameters. r_e is called the "effective radius" and is the half-light radius. n, when fit, is **about 4 for an elliptical galaxy** (which is an exponentially falling spheriodal distribution), and **about 1 for an exponential disk**. The plots below are split by Sersic index to identify disk or spheroid-dominated galaxies.

Blue colors (toward the bottom on the vertical scale) indicate younger stellar populations (more high-mass, high T_{surf} stars). This shows that the population is split into two groups by their profiles, quantified here by Sersic index.



(from `http://en.wikipedia.org/wiki/Galaxy_color%E2%80%93magnitude_diagram`)

From this we find that galaxies with a significant disk fraction are blue but have a wide variety of colors. On the other hand spheroid-dominated galaxies for a fairly tight sequence in color vs. luminosity.

The third fundamental parameter for galaxies is their dynamical state. For example their specific angular momentum, which can determine more about details of their morphological type. e.g. the presence of a bar.

30.2 Spiral galaxies - Mass-Luminosity

One of the major features of spiral galaxies is their rotation curve. (draw example) This increases to some V_{max} toward the edge of the star/gas component, as a result of the surrounding dark matter halo in which all galaxies are imbedded. This is one of the major evidences for dark matter, as the rotation curve should fall off as $1/r$ away from the visible component without dark matter.

A relation between the galaxy's absolute brightness (luminosity) and its V_{max} is observed. Called the Tully-Fisher relation.

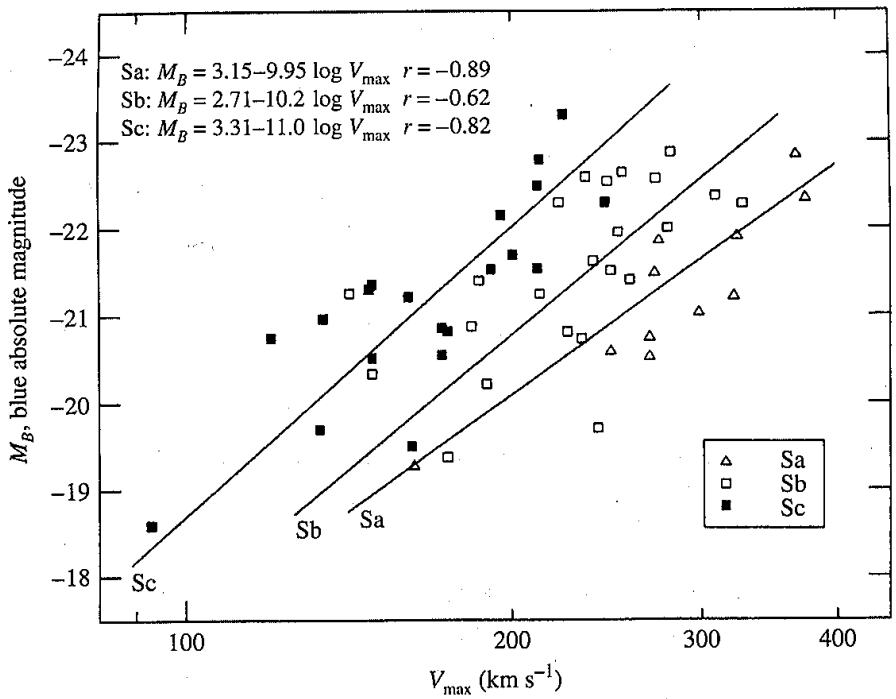


FIGURE 25.10 The Tully–Fisher relation for early spiral galaxies. (Figure adapted from Rubin et al., *Ap. J.*, *289*, 81, 1985.)

from Carroll & Ostlie

31 Astro notes 2018/11/9 - Fri - Galaxies, Distance

31.1 Disk galaxies - Mass-Luminosity - Tully-Fisher

One of the major features of spiral galaxies is their rotation curve. (draw example) This increases to some V_{max} toward the edge of the star/gas component, as a result of the surrounding dark matter halo in which all galaxies are imbedded. This is one of the major evidences for dark matter, as the rotation curve should fall off as $1/r$ away from the visible component without dark matter.

A relation between the galaxy's absolute brightness (luminosity) and its V_{max} is observed. Called the Tully-Fisher relation.

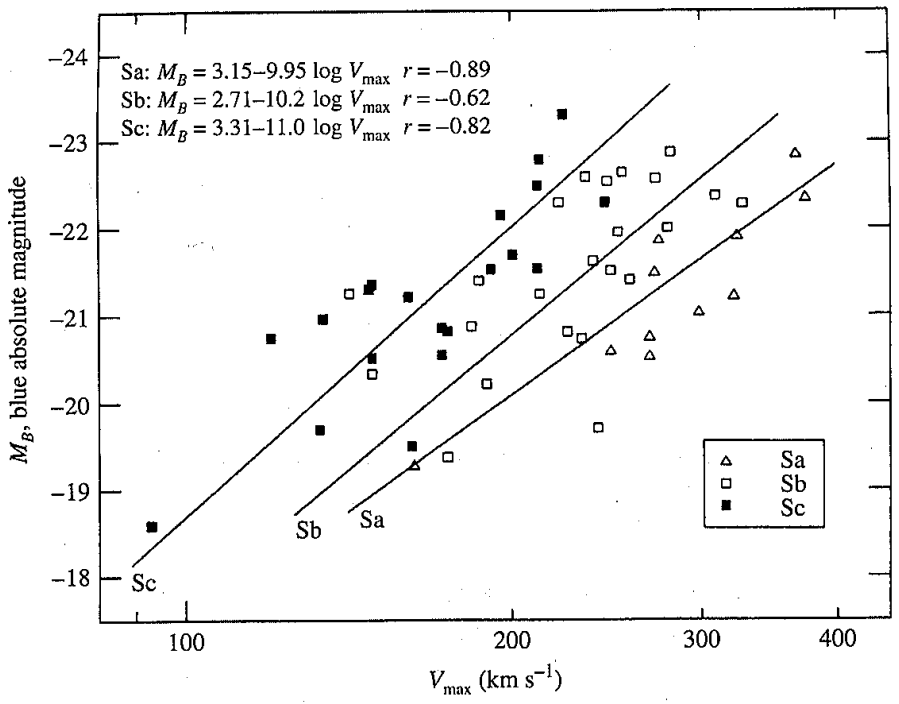


FIGURE 25.10 The Tully–Fisher relation for early spiral galaxies. (Figure adapted from Rubin et al., *Ap. J.*, *289*, 81, 1985.)

from Carroll & Ostlie

A good argument for this relation between the orbital property V_{max} and the overall Luminosity is having them both arise from the overall mass of the galaxy. The virial relation tells us that

$$\frac{GM}{R} \sim V_{max}^2$$

(or just this relation for a keplerian orbit). Where M is the mass of the visible matter and R its radial size. We need to make two assumptions to turn this into a relation to L and eliminate R. First of all, we assume that the mass-to-light ratio is not strongly dependent on galaxy mass

$$\frac{L}{M} \sim C_{ML}, \text{a constant}$$

Also, it is observed (at least for a given class of spiral), that the central surface brightness is relatively universal. From this we can assume

$$\frac{L}{R^2} \sim C_{SB}, \text{a constant}$$

Putting these together we find that
(student: eliminate R and M to get an L-V relation)

$$L \propto M \propto RV_{max}^2 \propto L^{1/2}V_{max}^2$$

which gives

$$L \propto V_{max}^4$$

The observed Tully-Fisher relation in infrared is

$$M_H^i = -9.50(\log_{10} W_R^i - 2.50) - 21.67 \pm 0.08$$

where W_R^i is a measure of the rotation speed from the radial velocity difference across the galaxy in emission lines. Converting L above to magnitudes

$$M = M_{Sun} - 2.5 \log_{10}(L/L_\odot) = -2.5 \log_{10} V_{max}^4 + ... = -10 \log_{10} V_{max} + ...$$

which matches the observed TF relation.

31.2 Spheroidal galaxies mass-luminosity relation - Faber-Jackson, fundamental plane

Elliptical galaxies are generally "simpler" but also follow a relation between their Luminosity and mass. This is much like the relation derived for the spirals, but instead of rotation velocity in the disk, the velocity dispersion σ_0 appears in the virial theorem, so that

$$L \propto \sigma_0^4$$

The observed relation is not quite this well behaved and is more like

$$L \propto \sigma_0^\alpha$$

with α between 3 and 5. This is called the Faber-Jackson relation. A better relation is the fundamental plane, which includes the radial size of the elliptical,

$$L \propto \sigma_0^{2.65} r_e^{0.65}$$

This is fairly well followed, and galaxies live in a "plane" defined by this relation in the space of L, σ_0 and r_e. Some of this additional variation parameterizes the surface brightness, which is not universal as it is for spirals. Also the mass-to-light ratio depends more significantly on overall mass for elliptical, making the exponents slightly different than would otherwise be expected.

31.3 Spiral structure

The basic tenet of spiral structure is that the spiral structure is related to spiral density waves in the disk. Stars are formed in the dense regions, and the brightest (high-mass) stars only live for a small fraction of an orbit around the galaxy center. Thus the bright, short-lived stars highlight the regions near the current location of the spiral density wave.

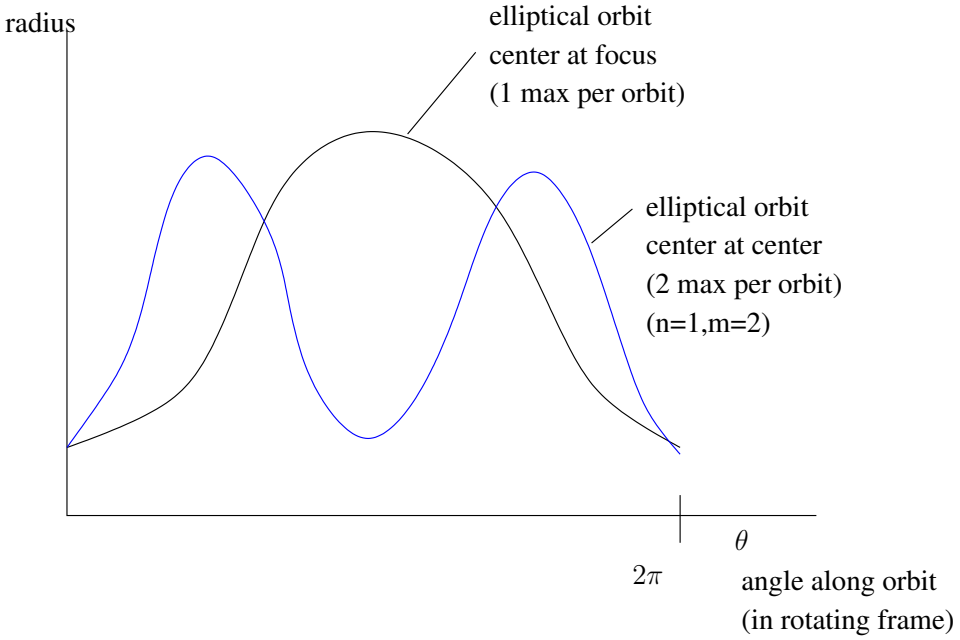
Maintaining a long-lived spiral **pattern** requires that orbits behave in a few particular ways made possible by the non-keplerian potential provided by the presence of the disk

itself. This creates a local effective potential in which stars can oscillate in radius (and height above the plane) as they orbit the galaxy.

In general, in a non-rotating frame the orbits of stars will not be closed because the period of their orbit around the galaxy and the period of their oscillation in radius will not be the same. However, orbits turn out to be able to close in a rotating frame over a wide range of distances from the center of the galaxy. This allow a pattern to develop when

$$m(\Omega(R) - \Omega_{\ell p}) = n\kappa(R)$$

where Ω is the local orbital frequency around the galaxy, and κ is the local frequency of oscillations in radius in the effective (non-newtonian) potential of the galactic disk. It turns out that for $n = 1$ and $m = 2$, a single value of $\Omega_{\ell p}$ satisfies this equation over a wide range of radii. This frequency defines a pattern speed.



For this n and m (1 and 2) in the frame of the pattern, each star will oscillate twice in radius for each time it goes around the galaxy. This effectively creates something like an orbit that is elliptical, but with the center of the ellipse at the center rather than its focus. If these ellipses are rotated for different orbits, a spiral pattern will develop.

From our text:

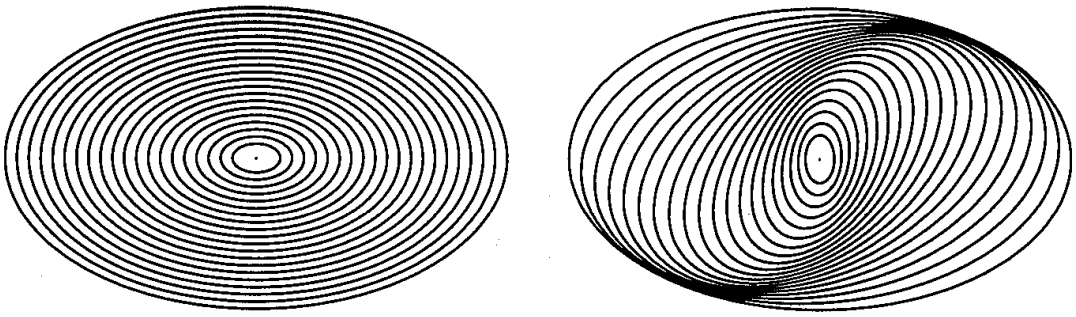


FIGURE 25.28 (a) Nested oval orbits with aligned major axes, as seen in a reference frame rotating with the global angular pattern speed ($n = 1, m = 2$), or $\Omega_{gp} = \Omega - \kappa/2$. The result is a bar-like structure. (b) Each oval is rotated relative to the orbit immediately interior to it. The result is a two-armed grand-design spiral density wave.

Note this is in the frame of the pattern, which is itself rotating at some frequency $\Omega_{gp} = \Omega_{\ell p}$.

32 Astro notes 2018/11/12 - Mon - Distance, Universe Expansion

32.1 Distance in the "nearby" universe

There is a wide variety of distance measures of varying accuracy. You have done main sequence fitting in your homework. Also some spectral types of stars on the main sequence have known luminosity based on similar nearby stars. By using this to infer luminosity based on just spectrum, this can be used to measure distance. Also the brightest red giant stars appear to have a fixed luminosity distribution, which can be used to infer distance.

See http://en.wikipedia.org/wiki/Milky_Way for spatial structure in our galaxy and just beyond. (see the Structure and Environment sections.)

In the early 20th century, it was not known that there were other galaxies until the distance to Andromeda was measured by measuring the Cepheid variables in it.

Cepheid variables are giant stars that vary periodically on 10-90 day timescales in a way that is related to their brightness. (will show this shortly)

32.2 The expansion of the Universe

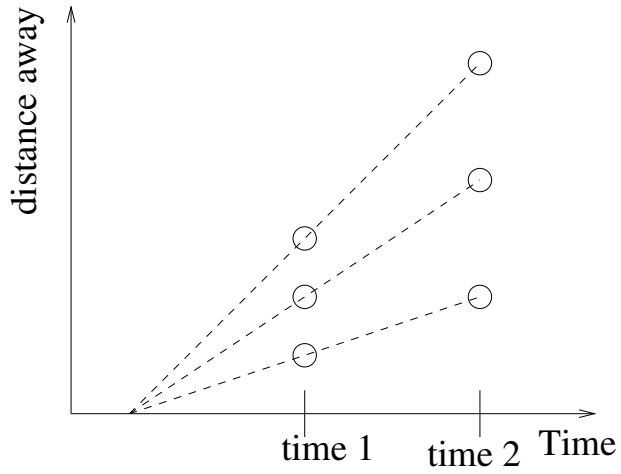
Once distances to galaxies could be measured it was found that the distance and recession speed of galaxies were related:

$$v = H_0 d$$

where H_0 is called the "Hubble constant" or "Hubble parameter", which is measured to be about 70 km/s / Mpc. The latter because it is not actually thought to be constant. The

Hubble constant is often rewritten as $H_0 = 100h$ km/s / Mpc, so that $h = 0.7$ is the value quoted. Cosmological values that depend on H_0 are then phrased in terms of h, since it was historically difficult to measure, and continues to be not quite clear.

Note that this observed relation leads to the conclusion that the universe is *expanding*. That is, all galaxies appear to be moving away from each other.



This also demonstrates the cosmological principle: we are not in a special place in the large-scale universe. The recession of galaxies that we see would also be observed from any other galaxy.

One convenient thing about this relation is that it makes it straightforward to estimate distance from the recession velocity

$$d = \frac{v}{H_0}$$

Note that this leads to a simple estimate for the age of the universe. If an object is at the distance above, moving away at speed v, the time it took to reach that point is

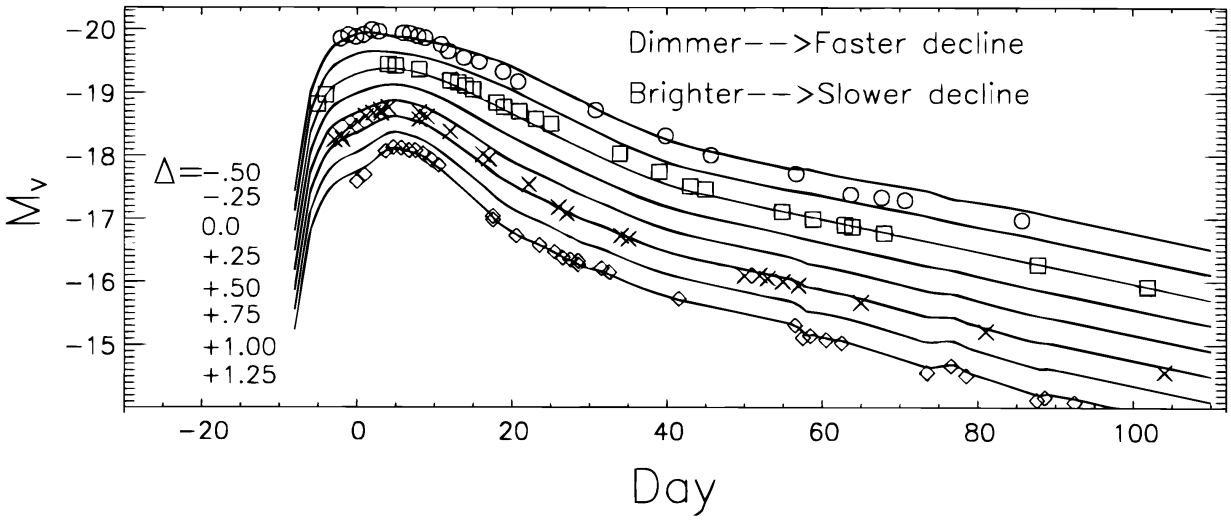
$$\Delta t = \frac{d}{v} = \frac{v/H_0}{v} = \frac{1}{H_0}$$

Which is called the Hubble time, and is about 14 billion years.

Another thing to keep in mind is that typically we don't talk about galaxies "moving away from us" but instead discuss space expanding. This also creates a subtle distinction between Doppler shift and cosmological redshift. Cosmological redshift arises because spacetime has expanded while the photon was in flight from its source, which is not quite the same as having the source be moving. Much of this ambiguity is due to the equivalence relations in relativity - by assuming that gravitational acceleration is the same as normal acceleration some dynamics get rephrased in terms of spacetime structure. Thus the recession of galaxies is rephrased as the expansion of space, though the two are equivalent.

32.3 SNIa

Brightness-decline rate relation:



SN Ia are spectroscopically well-defined by their Si-dominated spectra at maximum light. They then follow a relation between their peak luminosity and their decline time that is related to how much radioactive ^{56}Ni is produced in the explosion. This determines the overall energy release as well as the opacity that determines how slowly that energy can escape from the expanding ejecta.

32.4 Extragalactic Distance Scale

Phrased as a "ladder". There is no single distance indicator that spans all the way from nearby to distant galaxies. Nearby distances can be measured with parallax. However this is only good to maybe a kpc until Gaia (except in special cases, which I will discuss one of.)

Thought of as a ladder because each method of measurement must be used to calibrate the next.

A modern ladder is used in Riess et al. (2016, http://adsabs.harvard.edu/abs/2016ApJ...826...56R A summary of the ladder indicating which measurements overlap is shown in figure 10. This is used to measure the current expansion rate of the universe, H_0.

- – Parallax - used to measure distance to nearby stars - including some Cepheid variable stars
  - – Cepheids - period-luminosity relationship allows measurement in galaxies nearby enough for individual stars to be resolved. Includes some galaxies that have had a Type Ia Supernova. (Show paper figure 3 for example of isolating Cepheid stars)
- – Water Maser - special type of galaxy with an emission region of known size.
- Type Ia supernovae - Can be observed out to redshift 1 and somewhat beyond. But are fairly rare and so haven't been seen nearby.

Figure 3 shows Cepheids identified in other galaxies
Figure 4 shows light curves during each period (phase-folded).
Will continue discussion next time

33 Astro notes 2018/11/14 - Wed - Distance, Large-scale Structure

33.1 Extragalactic Distance Scale

Discussing modern ladder as constructed in Riess et al. (2016, http://adsabs.harvard.edu/abs/2016ApJ...82
A summary of the ladder indicating which measurements overlap is shown in figure 10.

Continuing from last time...

Cepheids - slowly pulsating giant stars. Bright - show period-luminosity relation from Figure 6. Each galaxy has multiple cepheids of various periods. Ones towards the end are those with other measurements of their distance. (Cepheids for the MW, eclipsing binaries for LMC and M31, and the maser)

Note distant objects have mostly long-period Cepheids, though we mostly have parallax distances for short-period Cepheids.

Hooking the ladder together - Figure 10 and other figures. Determines the match-up between distance scales.

Piecing together a distance scale like this makes systematic errors very important. Everything depends on how things were put together. (show figure 1)

Figure 13 shows current tension between measurement of H_0 from local distances and from CMB.

CMB method measures the size of structures in the early universe, but requires various other parameters about the cosmological expansion to get H_0.

33.2 Large-scale structure

On the largest scales, the universe is thought to be homogeneous and isotropic. As far as we know this is true. But obviously on smaller scales this is not true. Here we discuss structure on scales above galaxy cluster scales but below the scale at which the universe is homogeneous and isotropic. These are the scales of filaments, voids and superclusters.

Movie showing the 3D filamentary structure of the universe out to tens of Mpc. Note that these simulations are *dark matter only* which is most of the matter anyway:

http://astronomy.ua.edu/townsley/ay101/lectures/vid/cr.avi

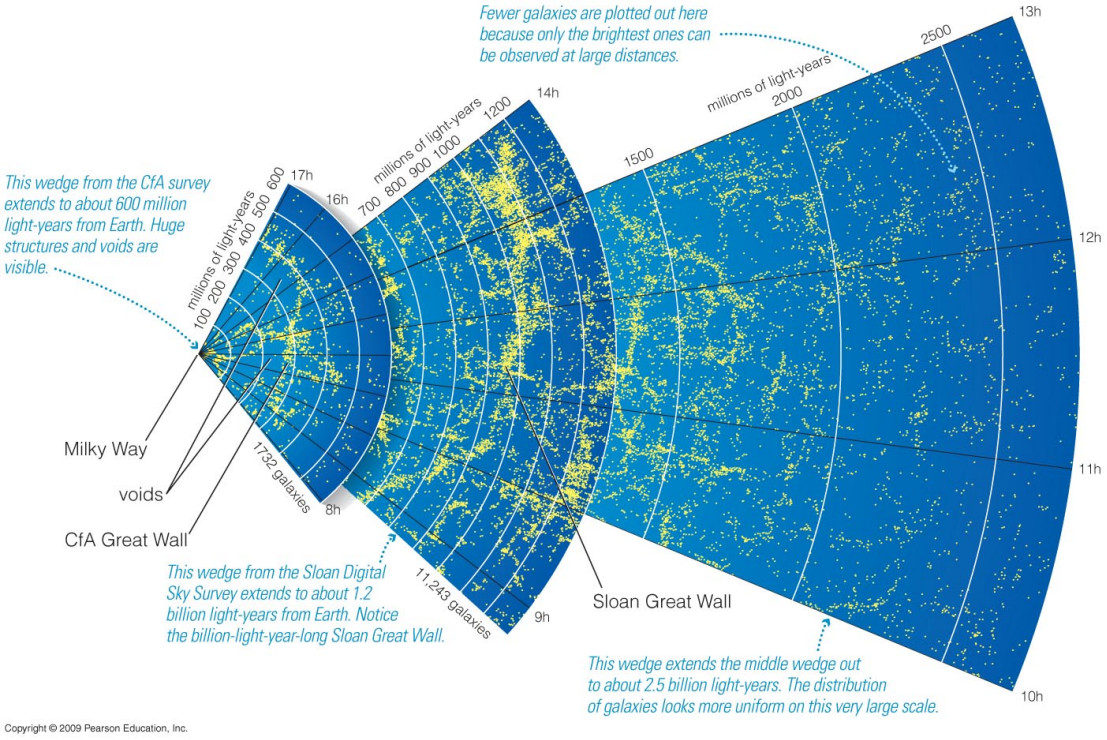
Clusters grow by forming galaxies and having them group together. Creates voids and filaments. The following is a simulation of structure formation in the universe on very large scales. We will discuss the initial conditions later when we discuss the cosmic microwave background (CMB) and again when we discuss inflation. The CMB is the earliest emitted light that we can see from the first time that the universe became transparent to photons, and so contains information about the early structure of matter in the universe. These

simulations begin after the formation/emission of the CMB. So we know what the initial conditions are.

http://astronomy.ua.edu/townsley/ay101/lectures/vid/lcdm_color1_highres_divx.avi

33.3 Large-scale structure

How is large-scale structure measured? Large-scale galaxy surveys show exactly the structure seen in these simulations. Galaxy positions from the CfA survey and the Sloan Digital Sky Survey (SDSS): (figure from Bennett et al. The Essential Cosmic Perspective)



The connection between the cosmic microwave background and this later epoch means that we understand the matter dynamics during the history of the universe very well. This constrains how much dark matter there is.

(there are several similar figures from various surveys including CfA redshift survey and the 2DF survey in Carroll & Ostlie)

From these we can constrain the distribution of galaxies. The pattern seen above means galaxy positions are correlated in a specific way. This can be measured with the 2-point correlation function. In a uniform distribution, the probability of finding a galaxy in some volume dV would be

$$dP = ndV$$

where n is just the uniform number density. In a correlated structure this becomes

$$dP = n[1 + \xi(r)]dV$$

where $\xi(r)$ is the two point correlation function, and r is the distance from any given galaxy. if $\xi > 0$ it means that galaxies are more likely than average to be found at that separation, and if $\xi < 0$ they are less likely than average to be found at that separation. Once galaxy positions are known, this is fairly straightforward to calculate by averaging over the sample in radial bins. We find that

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}$$

where $r_0 \approx 6h^{-1}$ Mpc and $\gamma = 1.8$. This is the structure of the universe on large scales, and the power 1.8 is characteristic of the filament and void structure. Note that generally galaxies are correlated (i.e. small r is favored).

It is also found that when averaged over larger and larger regions r_0 does not change. This means that there is a characteristic scale for structure in the universe, but despite being correlated, the universe is still isotropic on large scales $dP \to ndV$ for averaging over $r \gg r_0$. (note this means the universe is not fractal)

From Carroll & Ostlie:

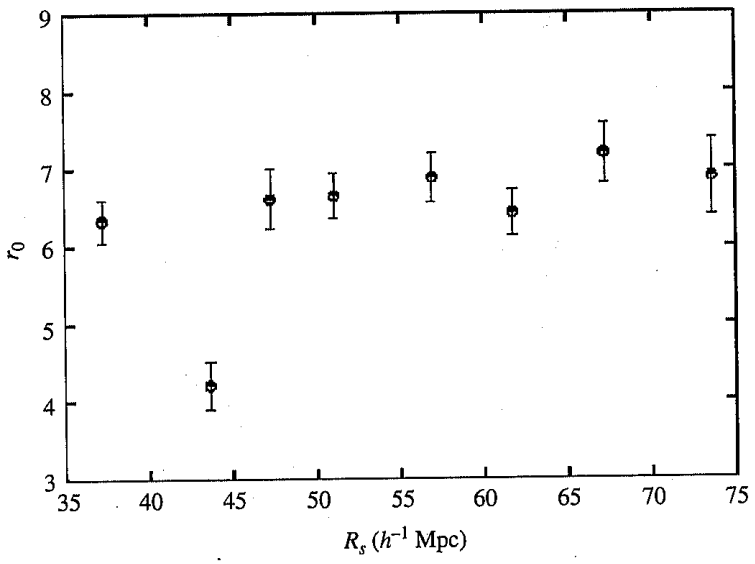


FIGURE 27.29 The correlation length r_0 as a function of the sample depth R_s for the CfA-II catalog (a galactic redshift catalog compiled by the Harvard Center for Astrophysics). The flat plateau on the right indicates a transition to homogeneity for this study at about $R_s \approx 60$–$70h^{-1}$ Mpc. (Figure adapted from Martinez et al., *Ap. J.*, 554, L5, 2001.)

34 Astro notes 2018/11/16 - Fri - AGN: Accreting Supermassive Black Holes

34.1 AGN Essential Phenomenology and Nomenclature

Quasars:
One of the original phenomena was "Quasi-stellar Radio Sources" – QSRs or "Quasars". These were sources of radio emission with counterparts at other wavelengths did not show extension - therefore being star-like instead of galaxy-like. Can emit hundreds of times as much energy as a galaxy. First high-redshift objects.

Seyferts:
There are also more nearby galaxies called Seyferts that showed a compact bright emission line source embedded in a galaxy. Again very bright and "star-like" (unresolved)

All types of sources can show narrow or broad+narrow emission lines.

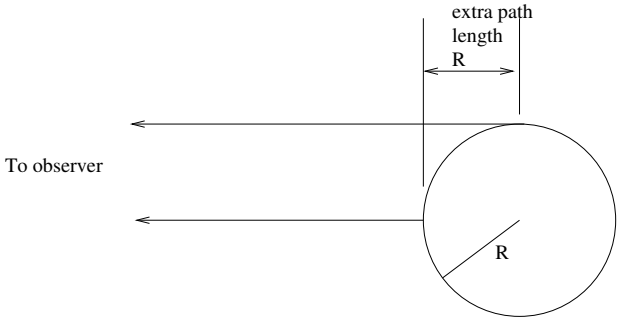
Note that, as mentioned before, redshift arises from the expansion of the universe:

$$\frac{R_{obs}}{R_{emit}} = 1 + z$$

where R is the "scale factor" that will appear in the spacetime metric when we get to cosmology. This allows the expansion rate of the universe to change, since the spectral shift doesn't depend on the motion of the galaxies, and so can change after the light is emitted. Quasars are found out to redshift 5 or 6, when the average distance between galaxies was 6 times smaller than it is today.

34.2 Size of central engine

Rough size of the emission region can be inferred from variability timescale. Consider a sphere of radius r whose brightness varies all together. There will be a light delay for the brightening. i.e. light from the closer edge of the sphere will reach the observer before light from other parts. The difference in distance is approximately R.



So the time delay for brightening/dimming is related to the size by

$$\Delta t = \frac{R}{c}$$

Since this light source might be moving away at relativistic speed, for high redshift, it is important to include a possible time dilation factor, so that

$$R = \frac{c\Delta t}{\gamma}$$

Using $\Delta t \approx 1$ hour, a typical value of AGN variability, and $\gamma \sim 1$, The radius is (AU=1.5×10^{11} m)

$$R \simeq 3 \times 10^8 \text{ m/s}(3600 \text{ s}) = 10^{12} \text{ m} = 7 \text{ AU} .$$

So there must be some object that would fit well within the solar system that can steadily emit a hundred times the luminosity of a galaxy, of order 10^{39} W. i.e. it must be relatively small on galaxy scales. Not much bigger than a giant star.

34.3 Luminosity Limit

There is a limit to radiated energy above which any object will start to lose matter Above this limit, the radiation pressure on a particle (electron-proton pair) is larger than the gravitational pull. Eddington limit:

$$L_{\text{Edd}} \simeq 10^{31} \text{ W} \left(\frac{M}{M_\odot}\right)$$

For the high luminosity of a quasar, $L = 5 \times 10^{39}$ W, this gives as *lower* limit for the mass:

$$M > \frac{L}{10^{31} \text{ W}} M_\odot \simeq 3 \times 10^8 M_\odot$$

Such a large mass in such a small space strongly implies the presence of a supermassive black hole. If we assume that the object is a black hole of this mass, we get the Schwarzchild radius:

$$R = \frac{2GM}{c^2} \simeq 10^{12} \text{ m} = 7 \text{ AU}$$

So the mass inferred from the luminosity limit in combination with the size imply a supermassive black hole, of typical mass $10^8 M_\odot$.

34.4 Efficiency

What is the most efficient way to convert matter into energy? Nuclear fusion of H to Fe releases a little under 10MeV per proton. That's about 10 MeV per 1000 MeV of rest mass energy or around 1%. By comparison a maximally spinning black hole can have material release energy down to $r = 0.5R_S$. This would give an energy release, for a mass m that makes it to the inner edge of the accretion disk

$$\frac{GM}{2r} m \simeq \frac{GMm}{GM/c^2} = mc^2$$

in reality its not quite the whole amount. What one gets is about 40% of rest mass enery. This leads to an expression for the accretion luminosity in terms of the mass-energy of the infalling material and an efficiency parameter, η,

$$L_{\text{disk}} = \eta \dot{M} c^2$$

where $\eta \simeq 0.4$. This is far mor efficient (in terms of energy release per mass) than nuclear fusion.

35 Astro notes 2018/11/19 - Mon - Active Galactic Nuclei - as observed

35.1 Disk temperature

By using the efficiency parameter we were able to write the luminosity of the disk as

$$L_{\text{disk}} = \eta \dot{M} c^2$$

But what about the temperature of the disk? Paradoxically, AGN actually have lower temperature disks than neutron stars or stellar mass black holes. The reason for this can be seen by first assuming that the characteristic disk size is R_S. Then the disk energy balance:

$$2\pi R^2 \sigma T_{disk}^4 = \frac{GM\dot{M}}{R} \quad \Longrightarrow \quad T^4 \propto \frac{M\dot{M}}{R^3}$$

becomes

$$T^4 \propto \frac{\dot{M}}{M^2} \quad \text{for } R = R_s \propto M$$

Then if we assume that the accretion rate is some fraction of the edddington accretion rate $L_{disk} = f_{ed} L_{Ed}$ where

$$L_{Edd} = \frac{4\pi G c}{\kappa} M$$

then we can form the eddington accretion rate:
(Student: equate and solve form $\dot{M}$)

$$\dot{M}_{Edd} = \frac{f_{\text{Edd}}}{\eta} \frac{4\pi G}{\kappa c} M$$

where κ is the average opacity, related to the amount of photon pressure on a g/cm^2 of material. This then gives

$$T_{disk}^4 \propto \frac{1}{M}$$

so that the disk temperature actually decreases with increasing black hole mass.

Supermassive black hole accretion disks tend to have temperatures around 10^5 so that emission peaks in the UV.

35.2 The Unified Model

While this gives us a central engine, we still need to explain how it can lead to the variety of phenomenology that is seen. Most notably the narrow/broad line dichotomy, and the correlation of broad lines with bright objects. This is explained by the so-called "Unified Model", which basically says that most AGN are surrounded by a dusty torus at large radii that can obscure the AGN, so that the accretion disk is only visible from certain angles.

AGN that show broad lines are ones for which we can actually see the disk, with its large range of orbital speds.

AGN that only show narrow lines are ones for which the accretion disk itself is obscured. However, gas at high distances that is ionized by the AGN is still visible, and gives the narrow emission lines.

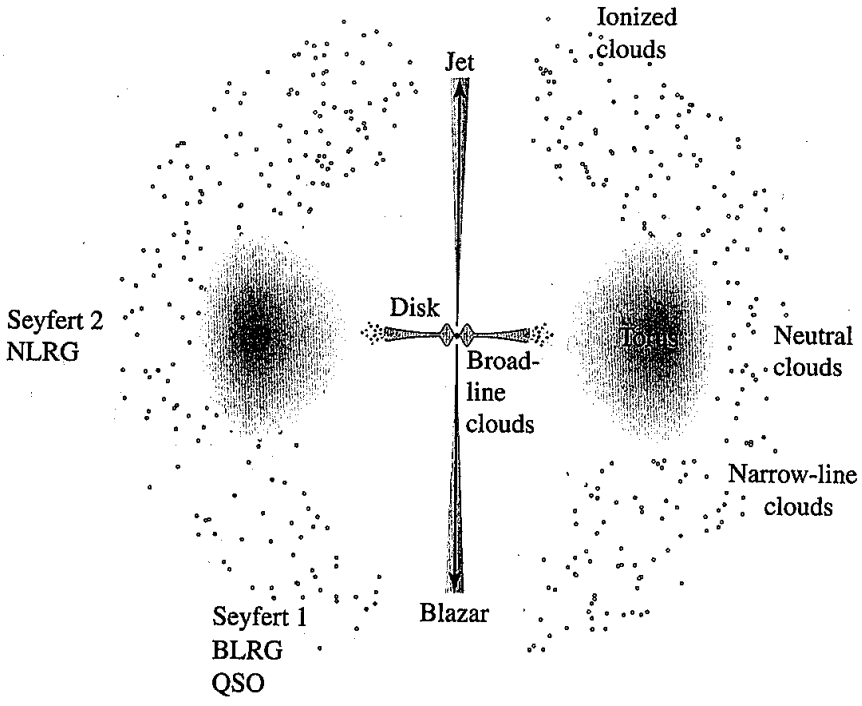


FIGURE 28.25 A sketch of a unified model of an active galactic nucleus. The jets would be present in a radio-loud AGN. A typical observer's point of view is indicated for AGNs of various types.

35.3 The broad view of AGN spectra

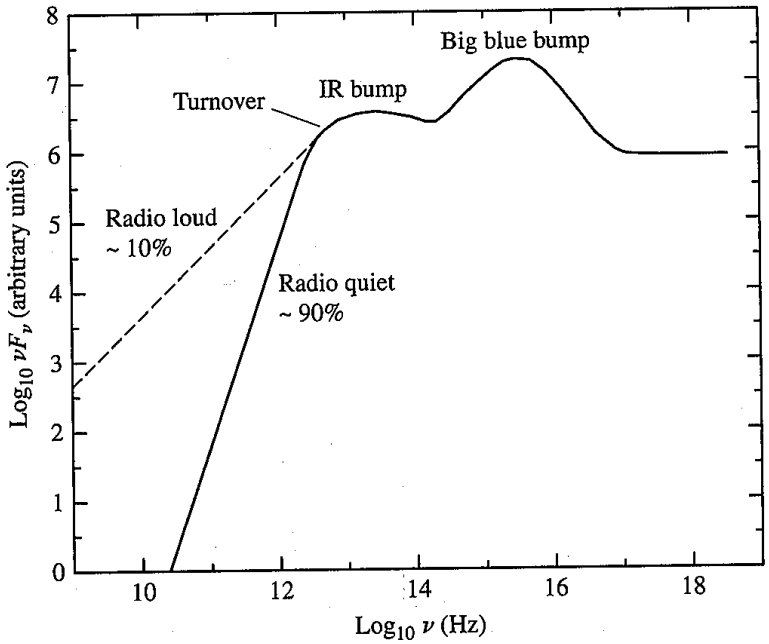


FIGURE 28.4 A sketch of the continuum observed for many types of AGNs.

AGN spectra are typically plotted in νF_ν vs $\log \nu$. This is done because it allows integration by eye:

$$L_{interval} \propto \int_{\nu_1}^{\nu_2} F_\nu d\nu = \int_{\nu_1}^{\nu_2} \nu F_\nu \frac{d\nu}{\nu} = \ln 10 \int_{\nu_1}^{\nu_2} \nu F_\nu d \log_{10} \nu$$

The major features in the spectrum include:

- The big blue bump, from the accretion disk (most likely)
- The IR bump, from the dust torus
- The underlying continuum from synchrotron emission

Syncrotron emission, powering the underlying continuum, comes from high energy (high speed) electrons spiralling in a magnetic field. The power-law slope of the spectrum is derived from the power-law distribution of electron energies.

35.4 Jets and Radio Lobes

AGN accretion disks or black holes appear capable of accelerating particles to relativestic speeds in the form of jets. These then interact with the intergalactic medium to create huge radio lobes. Sizes can range up to thousands of kpc or nearly a Mpc.

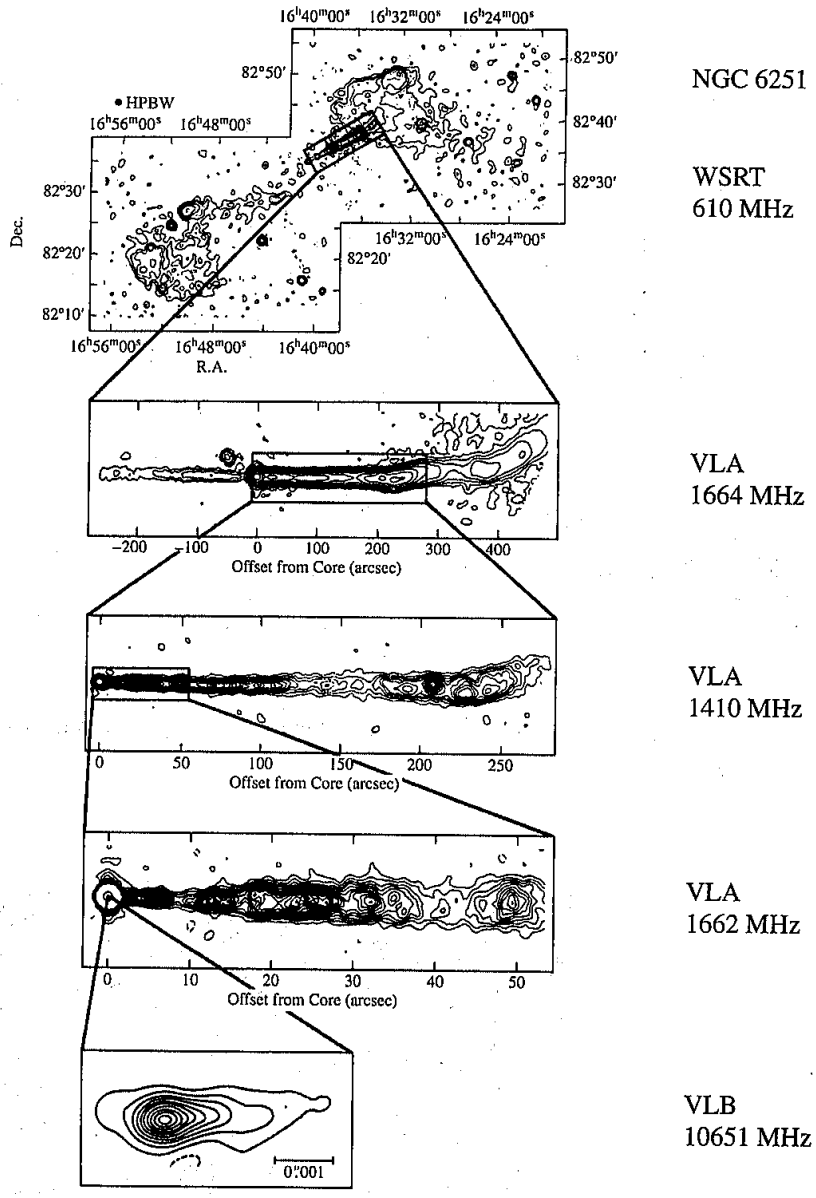


FIGURE 28.8 The jet and counterjet (second panel) of the radio galaxy NGC 6251. (Figure adapted from Bridle and Perley, *Annu. Rev. Astron. Astrophys.*, *22*, 319, 1984. Reproduced by permission from the Annual Review of Astronomy and Astrophysics, Volume 22, ©1984 by Annual Reviews Inc.)

The mechanism of launching these jets and accelerating the particles in them is not well understood.

The high speeds of the outflows are confirmed by superluminal motion of knots in the jets of some sources. This is caused by material moving at relativistic speeds toward the observer such that the projected motion on the sky is higher than c.

35.5 Quasars and the mass structure of the universe

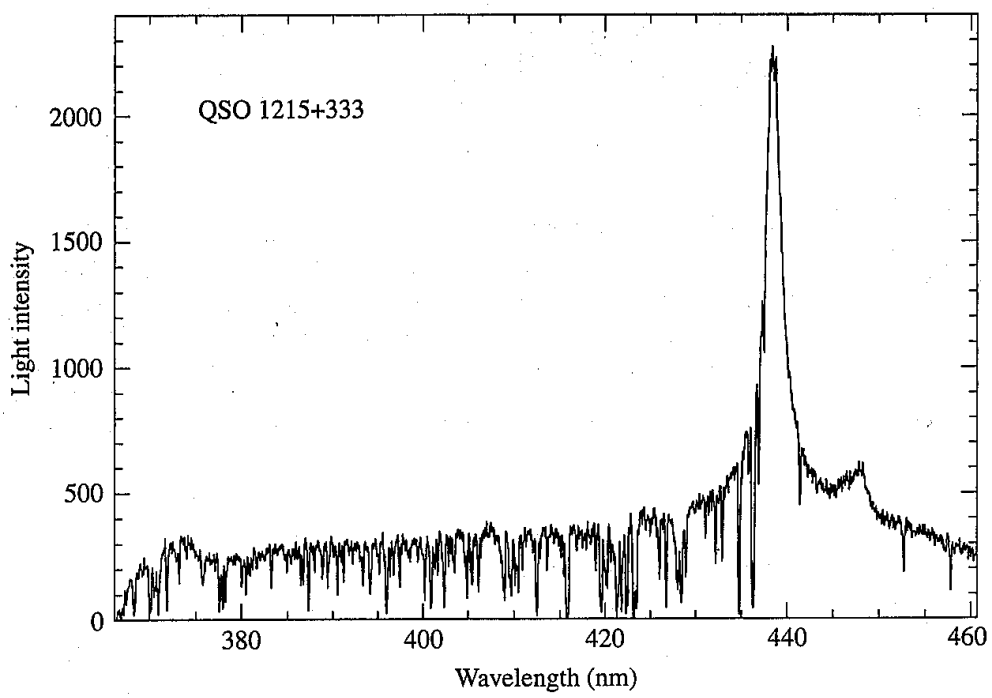
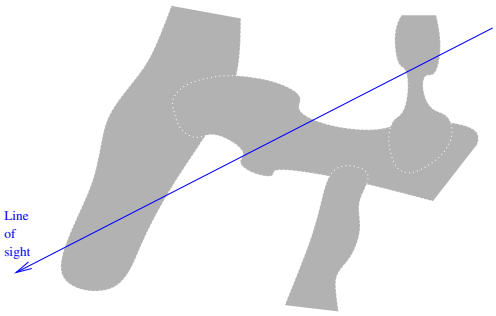


FIGURE 28.41 The strong Lyα emission line in the spectrum of QSO 1215+333, with the Lyα forest of absorption lines at shorter wavelengths. (Adapted from a figure courtesy of J. Bechtold, Steward Observatory, University of Arizona.)

The spatial structure of gas in the universe can be inferred partially from the absorbtion features in the spectra of distant quasars. The large emission feature here is a broad Lyman α emission line. As this light passes through the universe between the quasar and us, if the gas density is high enough it will absorb some light at the wavelength corresponding to the redshift of the absorbing cloud. Since Ly α is very easily obsorbed, this provides a sensitive probe of the structure (spatial distribution) of gas along the line of sight to the quasar.

This can also be extended to metal species (carbon, magnesium) to understand which parts of the intergalactic medium have been "polluted" by outflows from places where stars have formed. This provides an ongoing puzzle since the metal-enriched regions extend much farther than was expected, but modern simulations are improving the match.



Light records th expansion factor (R) at the time of passage through the absorbing gas in the redshift of the line.

36 Astro notes 2018/11/26 - Mon - Cosmology - structure, metric, expansion factor

We're going to mostly follow a relativistic formalism, but I will relate it to the newtonian formalism at one point. (The text follows a mostly Newtonian formalism.) The basic problem is that in order to discuss the cosmological constant or dark energy, we need the relativistic formalism. People used to think the question was whether the universe would recollapse or not, and what its curvature is, but it turns out even the question (or the possible answers) is different because of the cosmological constant.

36.1 Writing down the structure of spacetime

The fundamental assumption, which appears to be true, is that the distribution of mass-energy in the universe is homogeneous and isotropic. This means that the spacetime is also homogeneous and isotropic, since the spacetime structure is created by the presence of material. However even under this condition it is possible for the universe to expand, and do so at a rate that is not constant, or to have nonzero curvature, i.e. to not be flat.

We also know that the universe is currently expanding.

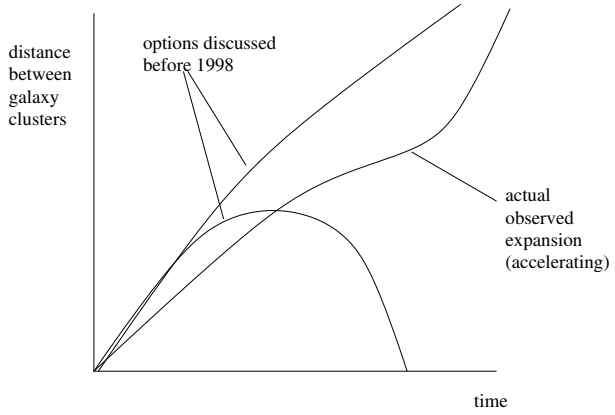
It turns out that the universe does appear to have zero curvature, but in order to understand what that means we must understand what it would mean to have nonzero curvature.

We will write down the structure of the universe as its metric, but it helps to briefly consider the newtonian limit to have an intuitive crutch for its dynamics and to make touch with what is in the textbook.

36.2 Newtonian and Relativistic cosmology

Newtonian cosmology uses the homogeneous, isotropic assumption to think of the mass in the universe much like dust particles evenly spread out in space. These dust particles are collisionless. This is approximately how galaxy distributions look on very large scales.

These particles are all attracted to each other, but since the distribution is uniform and fills all space, this attraction just pulls everything together. Since the universe is currently expanding, we can then think of how the dynamics of gravity may or may not cause the expansion of the matter in the universe to slow.



However, I will not frame our discussion in Newtonian terms. One reason is because it is extremely clumsy to introduce dark energy, which is now known to be the largest current component of the universe.

In relativistic cosmology, we will define something called the scale factor and that will quantify how the universe expands. Let's get on to that.

36.3 Relativistic cosmology

Like any other general relativistic problem, most of the game is about choosing coordinates. So lets try to choose some sensible time and space coordinates for the universe.

The assumption of homogeneity and isotropy creates a huge simplification of cosmology because it makes it well defined to define and use a global time coordinate. i.e. the history of the universe can be sliced in time in a way that is well-defined at every point – which gives a well-defined reference clock.

Isotropy also means that, for example, x, y, and z directions can't have different scale factors associated with them. There must be just one scale factor for all space directions.

Consider the flat spacetime metric (Minkowski space)

$$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)$$

How does this metric differ from what we know about the universe? Geodesic paths in this metric don't naturally match those we see for particles in the universe (not naturally moving apart. We would have to put galaxies at higher and higher speeds far away. This would not be consistent with Einstein's equation since the spacetime does not follow the mass-energy. It would also lead to infinite energies at large distance.

We can instead define coordinates in which galaxies stay at the same coordinates. These are usually called "co-moving coordinates". With these a galaxy at a coordinate distance ξ should have a steadily increasing actual space distance. This is accomplished by introducing

a "scale factor" $R(t)$, so that the spacetime metric is

$$ds^2 = c^2dt^2 - R^2(t)(d\xi^2 + d\eta^2 + d\zeta^2)$$

then $R(t)$ can be chosen so that on large scales galaxy clusters are fixed in the ξ, η, ζ coordinates. It is common, in the text for example, to write this using spherical coordinates

$$ds^2 = c^2dt^2 - R^2(t)\left[dr^2 + (rd\theta)^2 + (r\sin\theta d\phi)^2\right]$$

Again remember that the coordinate distance between galaxies Δr does not change. However the distance between galaxies, $R\Delta r$, does because the scale factor R increases with time.

This reduces the question of the history of the universe to understanding or specifying $R(t)$, i.e. the expansion history. We will come back to this soon.

36.4 the Hubble parameter in relativistic cosmology

Note that the current expansion rate of the universe is measured by the hubble constant H_0. This appeared in the relation, for any galaxy,

$$v_0 = H_0 d_0$$

where the 0 subscript indicates evaluated at the current time. More generally the expansion rate of the universe is measured by what is called the "Hubble parameter" defined similarly

$$v(t) = H(t)d(t)$$

where the distance for a fixed co-moving coordinate is

$$d(t) = R(t)r$$

(note that light travel distance is actually a bit more complicated)

Will use this next time to get $H(t)$.

37 Astro notes 2018/11/28 - Wed - Cosmology - curvature, parameters

37.1 the Hubble parameter in relativistic cosmology

Note that the current expansion rate of the universe is measured by the hubble constant H_0. This appeared in the relation, for any galaxy,

$$v_0 = H_0 d_0$$

where the 0 subscript indicates evaluated at the current time. More generally the expansion rate of the universe is measured by what is called the "Hubble parameter" defined similarly

$$v(t) = H(t)d(t)$$

where the distance for a fixed co-moving coordinate is

$$d(t) = R(t)r$$

Also the recession velocity for the same co-moving coordinate is
(student:)

$$v = \frac{dd}{dt} = \frac{dR}{dt}\, r$$

And thus

$$v(t) = \frac{dR}{dt}\, r = \frac{dR}{dt}\, \frac{d(t)}{R(t)} = H(t)d(t)$$

so that we can identify

$$H(t) = \frac{1}{R(t)}\, \frac{dR(t)}{dt} = \frac{\dot{R}}{R}$$

Also if we choose $R(t_0) = 1$ we see that H_0 is just the current rate of expansion $H_0 = \dot{R}(t_0)$.

37.2 Curvature

The metric presented above is spatially flat. While this appears to be how our universe is to good accuracy based on observations of the cosmic microwave background, we should consider what it would mean for the universe to be non-flat, i.e. curved.

First it is good to keep in mind that curvature may not mean what you think it means. The best example of this is that while a sphere has positive curvature, a cylinder is flat. (you can turn a flat piece of paper into a cylinder).

Curvature is related to whether distances along certain related paths compare. For example if you draw a "circle" by following a path equidistant from some point in space, how does the length of the path compare to the integrated distance from the "central" point. We will use this to demonstrate different kinds of curvature:

Positive curvature: like in a sphere. The circumference of the path is shorter than it would be for a circle of the same radial distance. This can be seen by consideringa the circumference of a line of lattitude θ on a sphere which is given by 2π times the distance to the axis through the poles. Whereas the distance from the pole to this latitude is the longer distance from the pole down along the surface of the sphere

(draw diagram 29.16)

Negative curvature: is like a saddle point. because the circle is not flat but is flexed upward and downward it is actually longer in circumference than a flat circle. This is like a gathered piece of fabric that has had an extra panel sewn into it - the edge is longer than 2π times its radius, and therefore connot be laid flat.

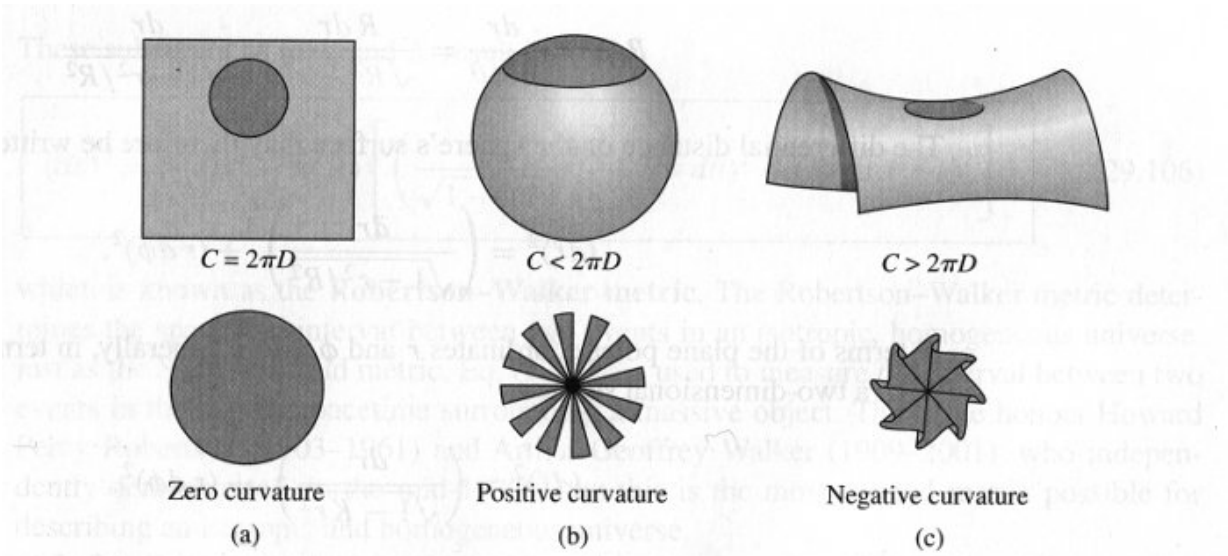


FIGURE 29.17 Calculating the curvature of a surface in three geometries: (a) a flat plane, (b) the surface of a sphere, and (c) the surface of a hyperboloid.

It is possible to add curvature to our metric by adding a term to the radial distance:

$$ds^2 = c^2dt^2 - R^2(t)\left[\left(\frac{dr}{\sqrt{1-kr^2}}\right)^2 + (rd\theta)^2 + (r\sin\theta d\phi)^2\right]$$

This is called the **Robertson-Walker Metric** and is the most general metric that describes a homogeneous and isotropic universe. This therefore essentially sets the way in which we can write down the structure of the universe, and now we must determine R and k for our universe.

37.3 Dynamics of the expansion factor

Einstein's equation (which I will not discuss directly) specifies the dynamical relationship between the spacetime metric and the mass-energy contained in the spacetime. This is the analog of Newton's gravity because it uses the mass-energy configuration to determine the geodesics that free-floating particles will follow. For the Robertson-walker metric and a uniform mass-energy density, the result of Einstein's equations is

$$\left[\left(\frac{1}{R}\frac{dR}{dt}\right)^2 - \frac{8}{3}\pi G\rho - \frac{1}{3}\Lambda c^2\right]R^2 = -kc^2$$

where ρ is the mass-energy density and Λ is the so-called cosmological constant, also called dark energy. This is the Friedmann equation plus the cosmological constant term.

It is customary, and a good idea at early times, to write the density as $\rho_m + \rho_{rel}$ where ρ_{rel} is the mass-energy density of relativistic particles like photons or neutrinos.

Note that

$$\rho_m \propto R^{-3}$$

$$\rho_{rel} \propto R^{-4}$$

$$\rho_{\Lambda} = constant$$

38 Astro notes 2018/11/30 - Fri - Cosmology - dynamics, fate

38.1 Dynamics of the expansion factor (continued)

Last time I stated the differential equation for the expansion history $R(t)$ is the Friedmann equation

$$\left[\left(\frac{1}{R}\frac{dR}{dt}\right)^2 - \frac{8}{3}\pi G\rho - \frac{1}{3}\Lambda c^2\right] R^2 = -kc^2$$

Where the mass-energy has several components (note that Λ can be considered either a density component or a separate constant.)

$$\rho_m \propto R^{-3}$$

$$\rho_{rel} \propto R^{-4}$$

$$\rho_{\Lambda} = constant$$

Since different components dominated the dynamics at different times, there are three "epochs" of expansion: an early radiation-dominated-era, a matter dominated era, and we are now moving into an era dominated by what appears to be vacuum energy or cosmological constant.

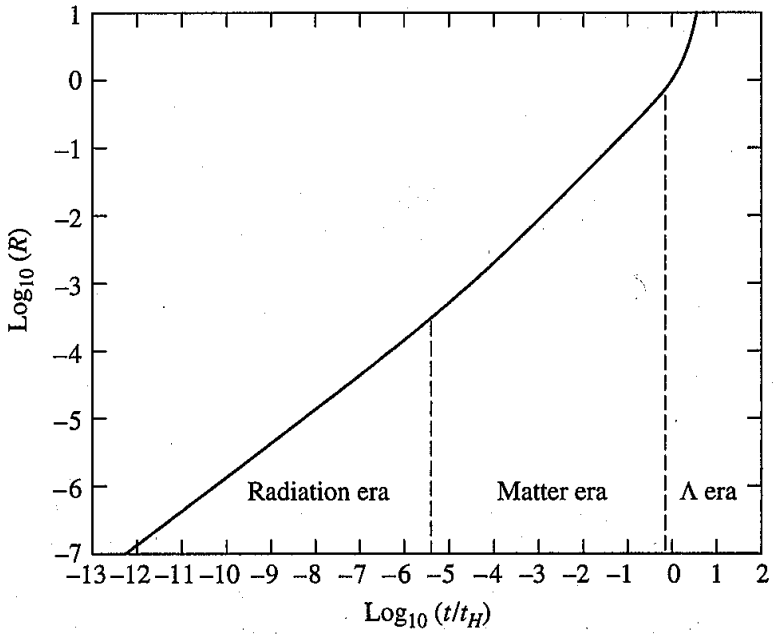


FIGURE 29.19 A logarithmic graph of the scale factor R as a function of time. During the radiation era, $R \propto t^{1/2}$; during the matter era, $R \propto t^{2/3}$; and during the Λ era, R grows exponentially.

or on a linear scale

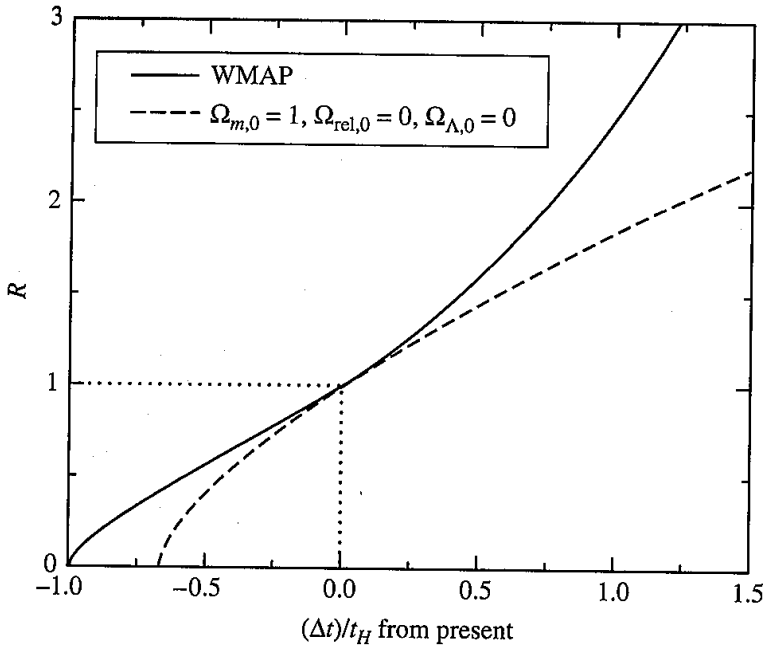


FIGURE 29.20 The scale factor R as a function of time, measured from the present, for a WMAP universe with $t_0 \simeq t_H$, and a flat, one-component universe of pressureless dust with $t_0 = 2t_H/3$ (Eq. 29.44). The dotted lines locate the position of today's universe on the two curves.

38.2 Cosmological Parameters

With the structure of the homogeneus, isotropic universe specified by the Robertson-Walker metric, we are now discussing how the scale factor $R(t)$ varies in time and how it is related to the curvature k. We found the evolution of $R = 1/(1+z)$ is given by:

$$\left[\left(\frac{1}{R}\frac{dR}{dt}\right)^2 - \frac{8}{3}\pi G(\rho_m + \rho_{rel}) - \frac{1}{3}\Lambda c^2\right] R^2 = -kc^2$$

Recall that the Hubble parameter is defined such that $H = \dot{R}/R$.

We can divide all mass-energy density components by

$$\rho_c = \frac{3H^2}{8\pi G}$$

which is the so-called "critical" density. One can see that if $\Lambda = 0$ and $\rho = \rho_c$, from the definition of H, $k = 0$. This was called "critically expanding" universe before the cosmological constant was discovered, since it was the border between matter-dominated solutions that recollapse and those that expand forever. With these parameters it is possible to simplify the Friedmann equation to

$$H^2[1 - (\Omega_m + \Omega_{rel} + \Omega_\Lambda)]R^2 = -kc^2$$

Where

$$\Omega_i = \frac{\rho_i}{\rho_c}$$

and for dark energy

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

Then the densities can be written like so:

$$\rho_m = \Omega_m \rho_c$$

so that, for example,
(Student: write $\rho_m(R)$)

$$\rho_{m,0} = \Omega_{m,0}\rho_{c,0} \quad \text{and} \quad \rho_m = \rho_{m,0}R^{-3} = \Omega_{m,0}\rho_{c,0}R^{-3} = \Omega_{m,0}\frac{3H_0^2}{8\pi G}R^{-3}$$

and similarly for ρ_{rel} with R^{-4}.

The above dynamics can be solved for $H = \dot{R}/R$ in terms of the Ω's where $\Omega = \Omega_m + \Omega_{rel} + \Omega_\Lambda$, to give an expression for the expansion rate history of the universe in terms of the current values of the component fractions:

$$H = H_0(1+z)\left[\Omega_{m,0}(1+z) + \Omega_{rel,0}(1+z)^2 + \frac{\Omega_{\Lambda,0}}{(1+z)^2} + 1 - \Omega_0\right]^{1/2}$$

We currently know that approximatey $\Omega_{m,0} = 0.3$, $\Omega_{rad,0} = 8 \times 10^{-5}$, and $\Omega_{\Lambda,0} = 0.7$. But since ρ_m, ρ_{rel} depend differently on R from each other and Λ, the balance of these evolves with time as R changes. Initially photons were dominant, then matter, and finally "dark" vacuum energy. (see plots above)

Since radiation is well constrained and only important at very early times. The state and fate of the universe is generally shown in a sort of "phase diagram" for Ω_m and Ω_Λ. Note that these are the *current* values of these factors.

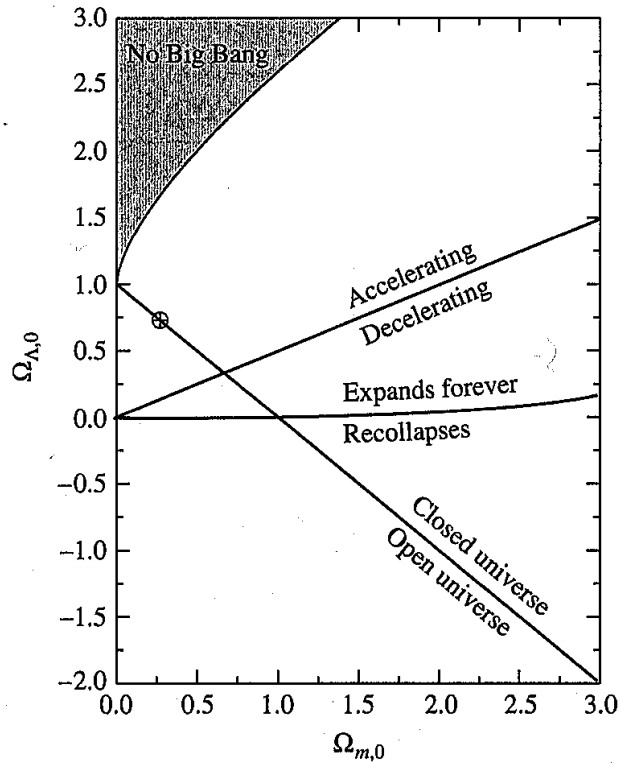


FIGURE 29.21 Model universes on the $\Omega_{m,0}$–$\Omega_{\Lambda,0}$ plane. Every point on this plane represents a possible universe. The point ($\Omega_{m,0} = 0.27$, $\Omega_{\Lambda,0} = 0.73$) is indicated by the circle.

Can see that with no cosmological constant, the expanion is always declerating and the open/closed boundary and the expands forever/recollapses boundary are the same point, $\Omega_{m,0} = 1$. (note open/closed here actually means negatively or positively curved, the naming is a bit archaic. A positively curved universe *might* loop around and "close" like a sphere, but it also may not. It is possible to define an infinite, positively curved 2D surface, it is just not possible to embed such a surface in 3D.)

39 Astro notes 2018/12/3 - Mon - Cosmology - Cosmic Microwave Background

39.1 Cosmic Microwave Background

The cosmic microwave background originates from the first epoch the universe cooled enough to become **transparent to photons**. This occurred when the hydrogen filling the universe became neutral – an event often called "recombination". The physics is related to what sets the surface temperature of cool stars, since both are set by the ionization transition of H.

The temperature at which the CMB was emitted is thus similar to that of the surface of a star, about 3000 K. These photons have been stretched by the expansion of the universe

so that we now see the blackbody spectrum as having $T = 2.725 \pm 0.002$ K. The spectrum is very close to a blackbody.

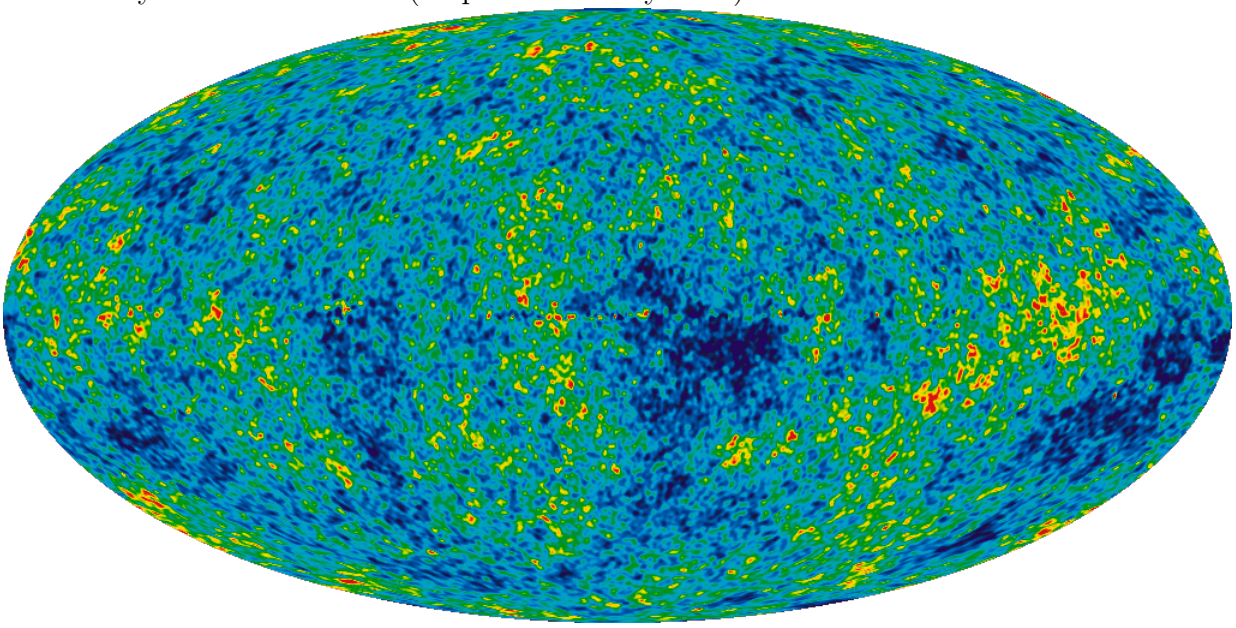
The initial discovery of this effectively proved the "big bang theory". That is the universe has expanded from an earlier hot state into what we see today. The CMB are remnant photons from the first epoch in which the universe became optically thin.

39.2 CMB anisotropy

But the anisotropy in the CMB can be used for much more. The large scale structure of the universe (over- and under-densities) are imprinted as small variations in the temperature of the CMB. This can be used to infer properties of universe, specifically its geometry and also the existence of *non-baryonic* dark matter.

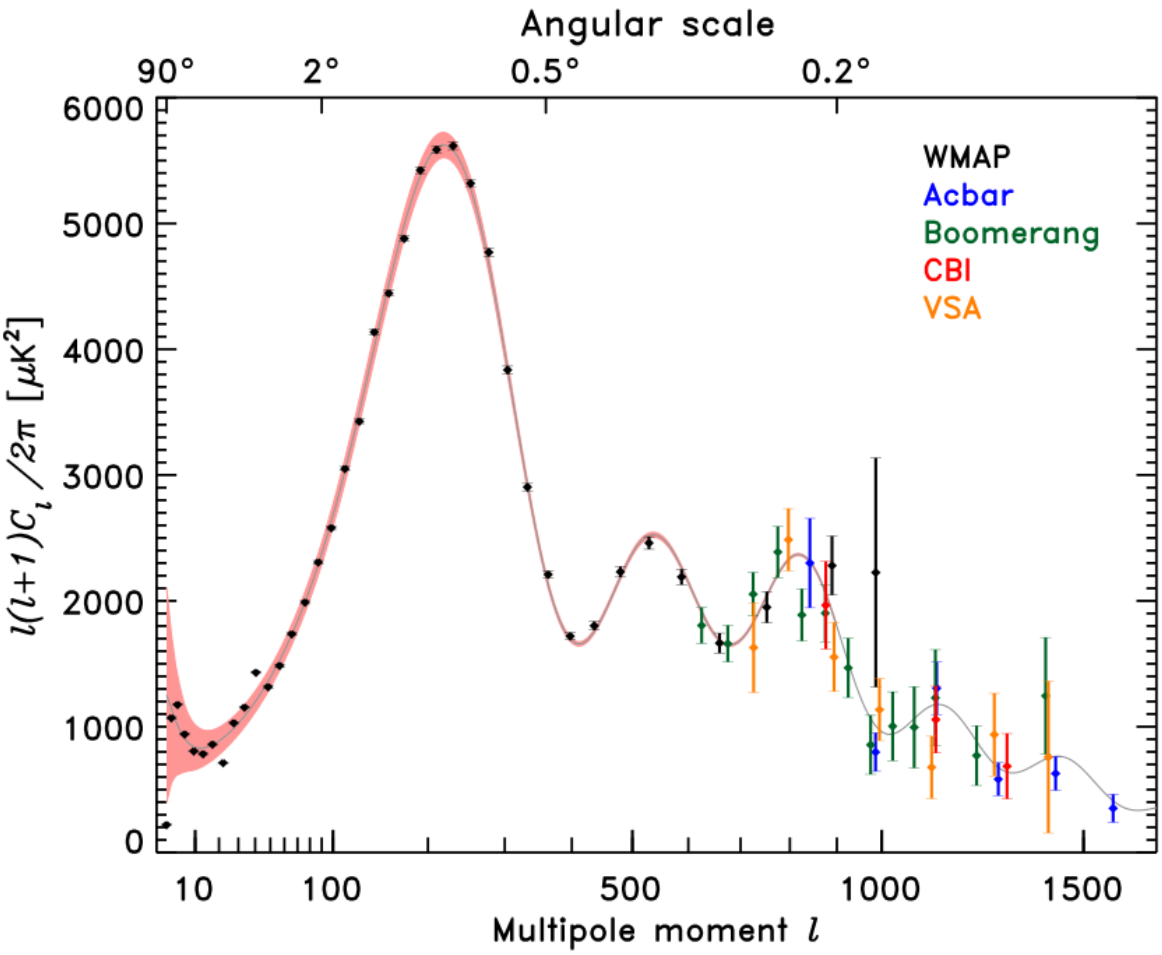
Components want to collapse (fall) but this is resisted by pressure. Each component starts falling at a different time.

On the sky these look like so (amplitude of only 10^{-5}):



(from http://en.wikipedia.org/wiki/Cosmic_microwave_background)

Acoustic oscillations, driven by random noise, in the early universe had preferred scales due to the different components involved - radiation, neutrinos, dark matter, baryonic matter. In order to study this we take the **angular power spectrum**. This has a peak at about 1 degree angular scale, among other peaks.



Here the red bands are **Cosmic variance** – the uncertainty because we only have one universe to average over, and this is based on randomness in the early universe. The first peak is the acoustic signature of the baryon-photon coupling - i.e. the preferred scale for inhomogeneities in the universe at the time CMB is emitted. But dark matter has a different (smaller) preferred scale because it had already started to collapse well before the photons started to decouple. So it appears as smaller structure evidenced in the next peak in the power spectrum.

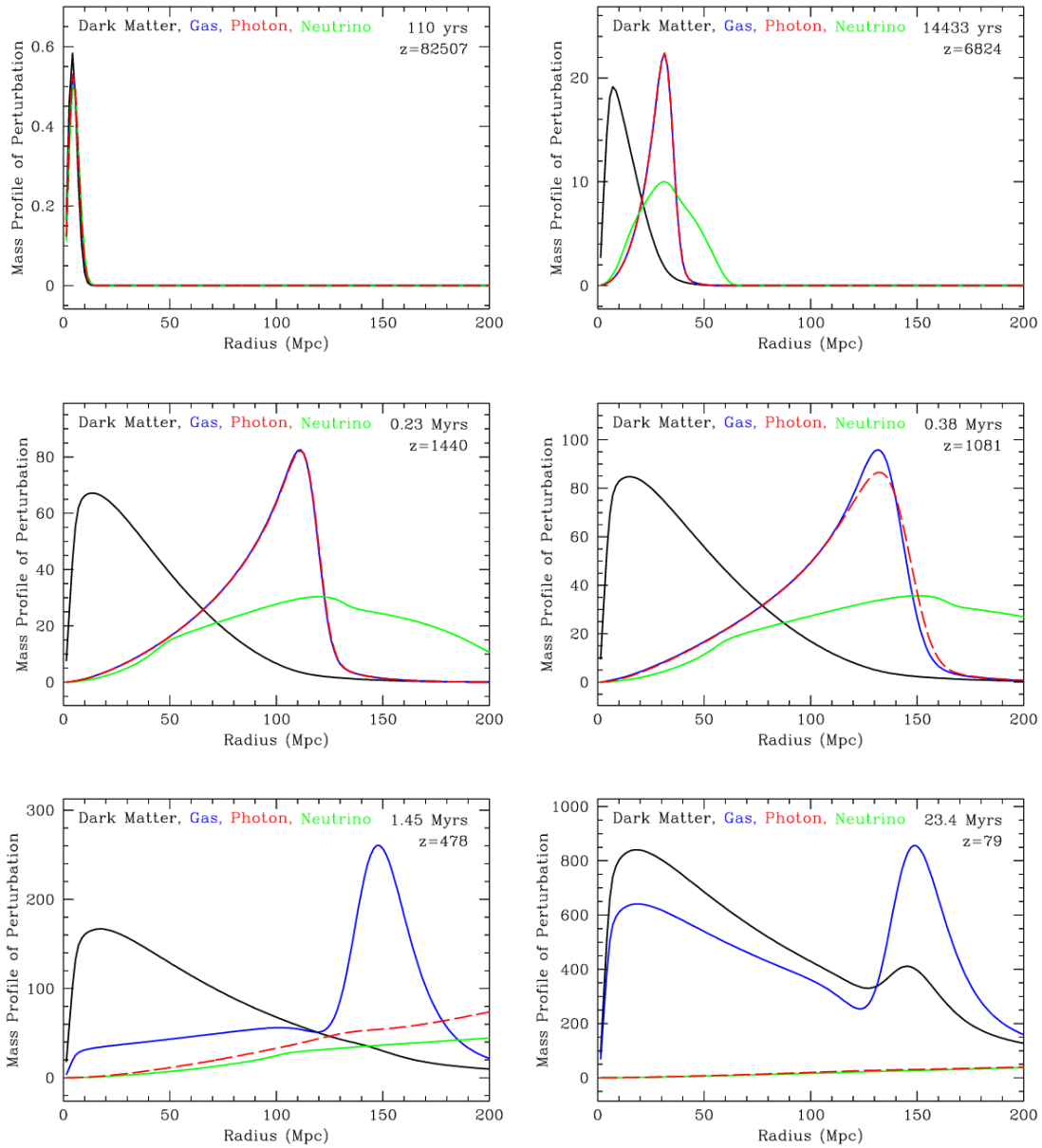
From the structure in the CMB it is possible to conclude that **Dark matter must be electromagnetically neutral** even at high temperatures. Usually this is called non-baryonic, i.e. it can't be anything that contains protons (a baryon, which is therefore charged).

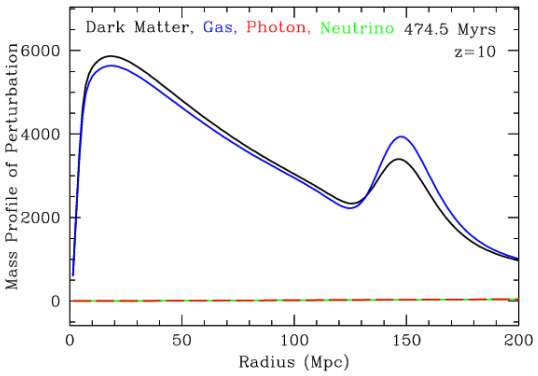
Again, how do we know this? because dark matter decoupled from the hot photons very early in the universe. Photons, due to their high speed, tend to smooth out inhomogeneities. But once a component decouples from photons, its inhomogeneities are no longer smoothed out.

The following sequence of figures demonstrates the decoupling of the differen compo-

nents for a impulse propagating away from the origin.

(these are from http://www.ias.u-psud.fr/Dark_energy/presentations/castanderBAO_081124.pdf but are very commonly shown when discussing Baryon Acoustic Oscillations. Originally created by Daniel Eisenstein)





In the second pane, can already be seen that the dark matter has decoupled from the "Gas" (Baryons) and photons, which are still coupled. Also neutrinos can also decouple, and begin to stream away.

The CMB IS the analog of the photon component here. It becomes decoupled from the gas just after the third pane ($z \simeq 1000$) and begins to free-stream, leaving the gas behind. The gas begins to then fall back to where most of the mass is, where the dark matter is located. The CMB provides a snapshot of the universe at this time. Thus in the overdensities evident in the CMB we can already see that most of the matter decoupled much earlier, and has already fallen toward overdense regions, and thus must not interact electromagnetically. This has long been a very strong arument for non-baryonic dark matter.

The dark matter also has to be "cold", which means not high velocity due to a large particle mass. Neutrinos are an example of "hot" dark matter.

Toward the end, the dark matter distribution is actually modified by the normal matter, so that there are two preferred scales for structure. This can be found in the distribution of galaxies in space. Since we know the scale from this physics (and the parameters obtained from the CMB) this can be used to construct a standard ruler to measure distances. These are called Baryon Acoustic Oscillations (BAO), and can fulfill a similar role to supernovae - measuring the expansion history of the universe.

40 Astro notes 2018/12/5 - Wed - Cosmology - acceleration

40.1 Homework discussion

Big blue bump problem

40.2 Expansion history

So far we have seen the metric for a homogenous, isotropic universe, as well as the dynamical equation for how the expansion factor depends on the makeup of the universe as

parameterized by $\Omega_{M,0}$, $\Omega_{rel,0}$ and $\Omega_{\Lambda,0}$, the current ratios of the density of each component to the critical ration ρ_c. Frequently, including below, the "0" will be dropped. You will need to infer whether Ω_M is a parameter or a scale-factor-dependent quantity based on whether it is multiplied by factors of $(1+z)$ or not. If the factors of $(1+z)$ are put in explicity, it is generally the current value of the parameter. Also, I, and others, will sometimes use $\Omega_{\rm rad}$, for the "radiation" component, rather than Ω_{rel}, but they are generally the same for cosmological purposes.

In order to *measure* the values of these parameters, we need some observables that depend on them. It turns out that the relation of luminosity distance to redshift measures the expansion history of the universe. Luminosity distance is just the distance given by a naive interpretation of observed flux, i.e.

$$F = \frac{L}{4\pi D_L^2} \qquad D_L = \sqrt{\frac{L}{4\pi F}}$$

where L is the known isotropic luminosity of an object, F is its observed flux (observed brightness) and D_L is then the luminosity distance. By integrating through a spacetime like we have been discussing, we find that the luminosity distance is given by:

$$D_L = \frac{c(1+z)}{H_0} \int_0^z \left[\Omega_M (1+z')^3 + \Omega_\Lambda (1+z')^{3(1+w)}\right]^{-1/2} dz'$$

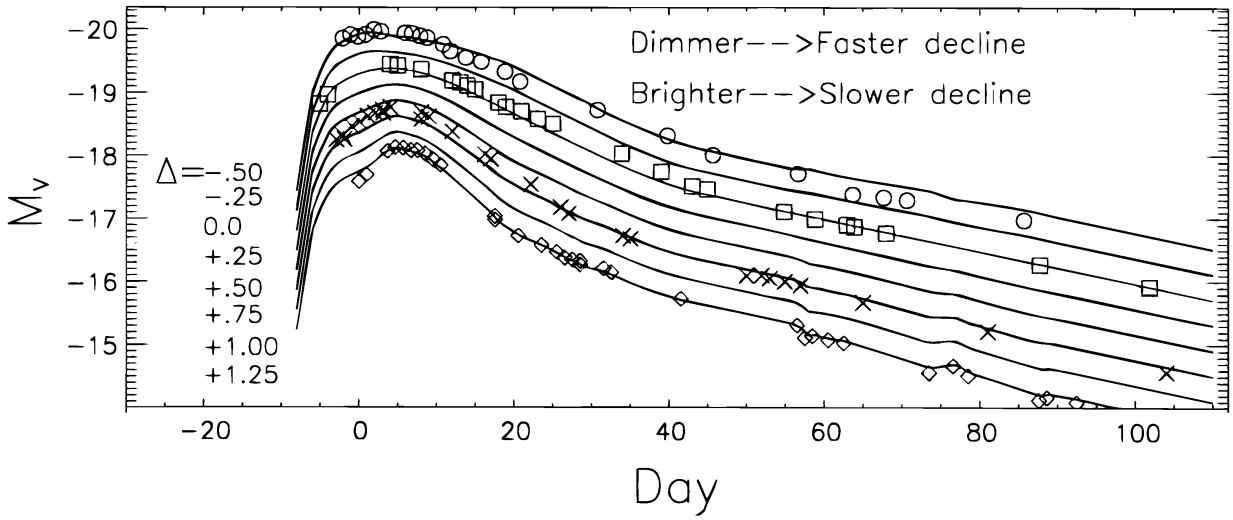
(For more information see: Carroll, Press, Turner 1992, ARA&A, 30, 499; Riess et al. 1998, AJ, 116, 1009; Garnavich et al. 1998, ApJ, 509, 74. The last gives this relation for luminosity distance.)

So if we have D_L and z for objects over a range of z we can fit the parameters of this relation.

40.3 measuring with supernovae

The luminosity of a Type Ia supernova can be determined with decent accuracy from its fading rate. This relation is determined (calibrated) using supernovae that are nearby enough to have known distance.

(From Riess et al. 1996, ApJ, 473, 88:



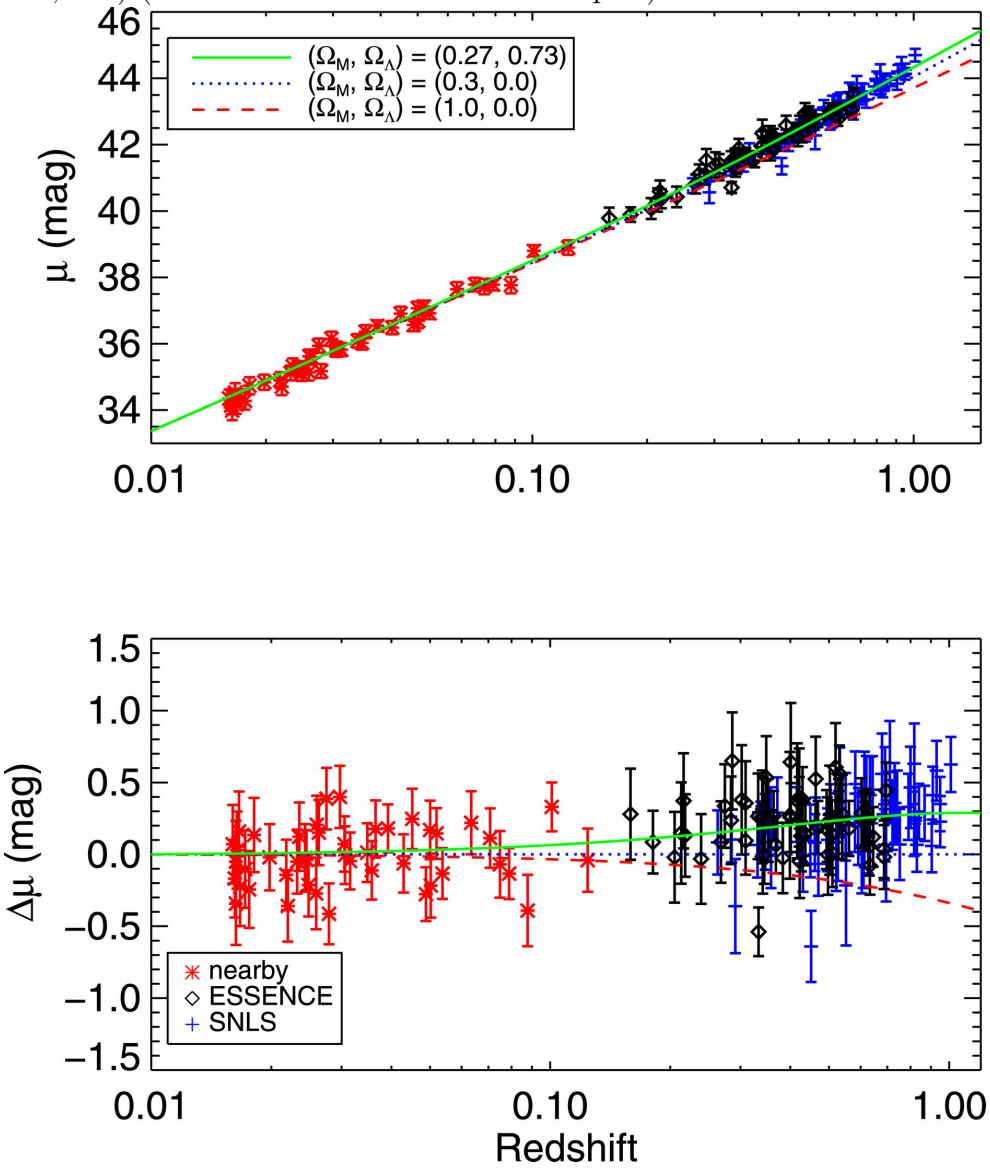
What is actually measured is luminosity distance, often shown as "distance modulus" μ in magnitudes as a function of redshift. where m is the apparent magnitude (from observed flux) and M is the absolute magnitude (from Luminosity)

$$\mu = m - M = 2.5 \log_{10} \left[\left(\frac{d}{10 \text{ pc}} \right)^2 \right]$$

Recall that redshift is basically just the ratio of the expansion factor now vs. that when the light was emitted. So this measures the expansion history.

The acceleration of the expansion was discovered in 1998, but I will show more recent work that has better statistics.

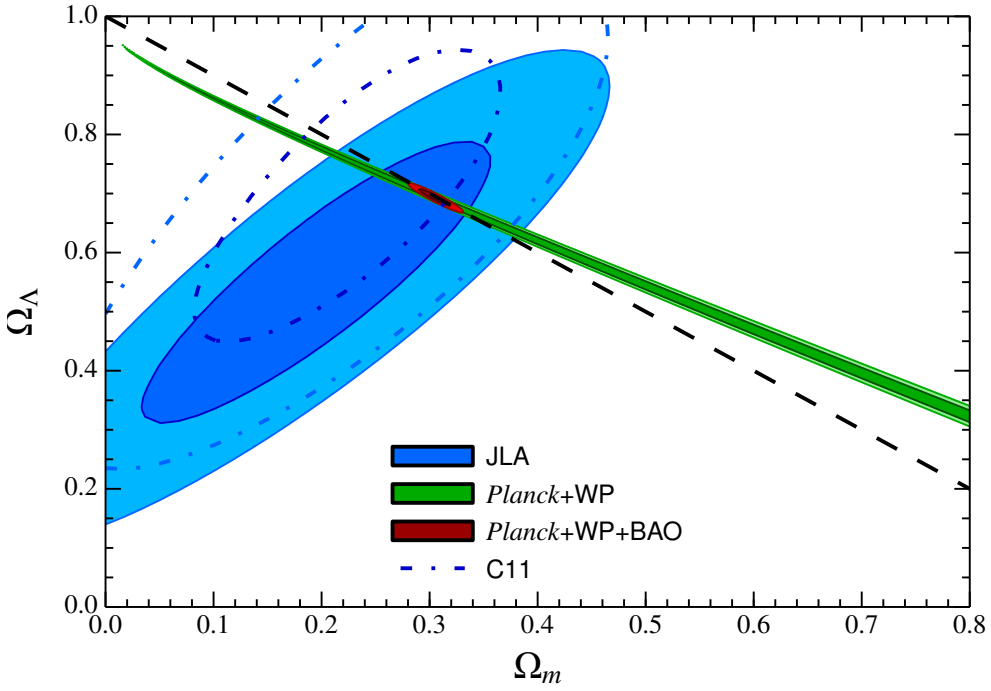
One example from a study from about a decade ago (Wood-Vasey et al. 2007, ApJ, 666, 694) (shown here because it is a clear plot):



Here the difference is shown from a flat cosmology with $\Omega_M = 1$ and an open cosmology with $\Omega_M = 0.3$, the value favored by measurements of galaxy clusters.

40.4 More state of the art

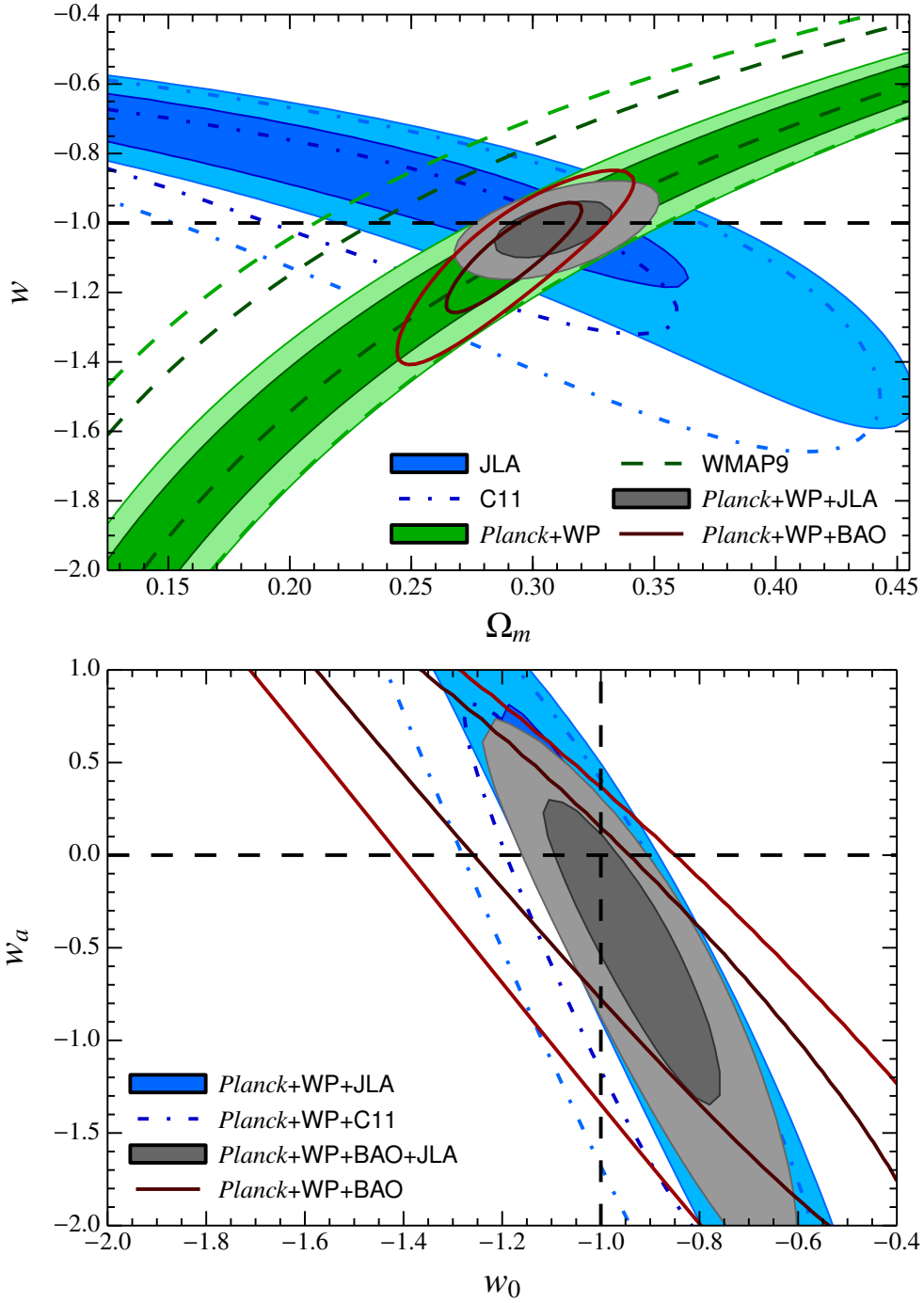
Currently the Ω parameters are very well measured. From a more recent reference that selects the best understood data - SDSS, SNLS and HST, Betoule et al. 2014, A&A, 568,



A22:

These are contours of 1 and 2 sigma likelyhood (basically probability that a particular parameter combination is the correct one, given the uncertainty in the data and measurement and the finite number of measurements made. There is a 68% chance the correct value is in the 1σ region and a 95% chance that it is in the, larger, 2σ region) SNIa ("JLA" = "Joint Light-curve Analysis") is consistent with Planck combined with BAO (Baryon acoustic oscillations, a standard ruler technique.)

But state-of-the-art goes on to measure both the parameter w, which allows the Dark Energy to have some other dependence on the scale factor other than being a cosmological constant ($w = -1$), and also the first derivative of w, w_a:



Things mostly consistent, but will be interesting as Planck and more BAO results are released.

40.5 Course reviews

.

41 Astro notes 2018/12/7 - Fri - Cosmology - nucleosynthesis and inflation

41.1 Big Bang Nucleosynthesis

As the hot, early universe cooled, after protons and neutrons were formed, after most neutrons were integrated into He4, there was time for a small amount of nucleosynthesis. The reactions involved are shown here in a figure from Carroll & Ostlie:

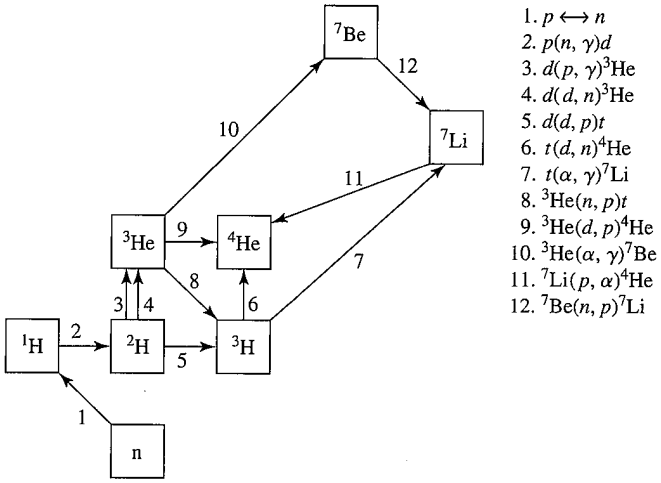
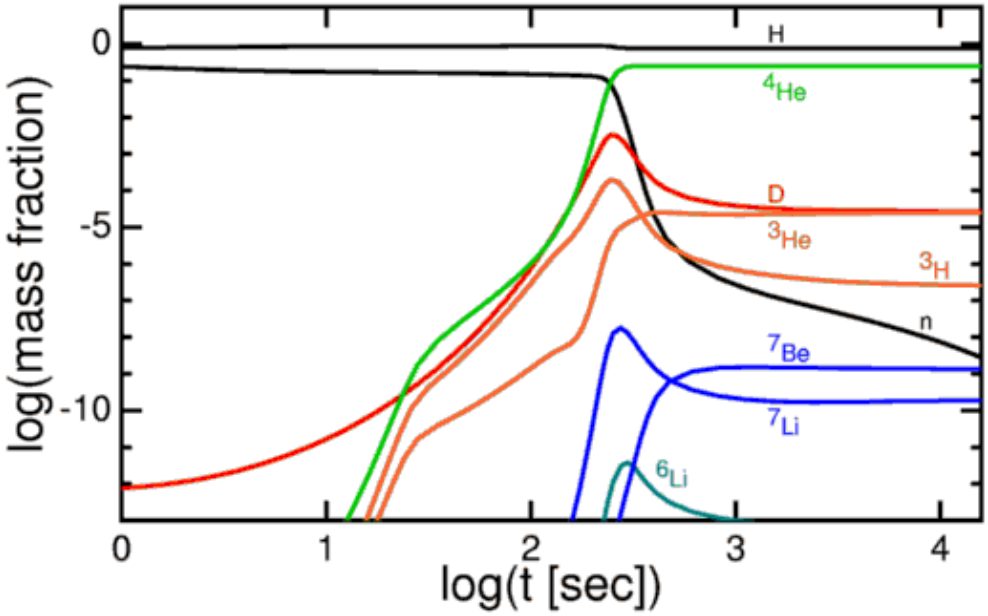
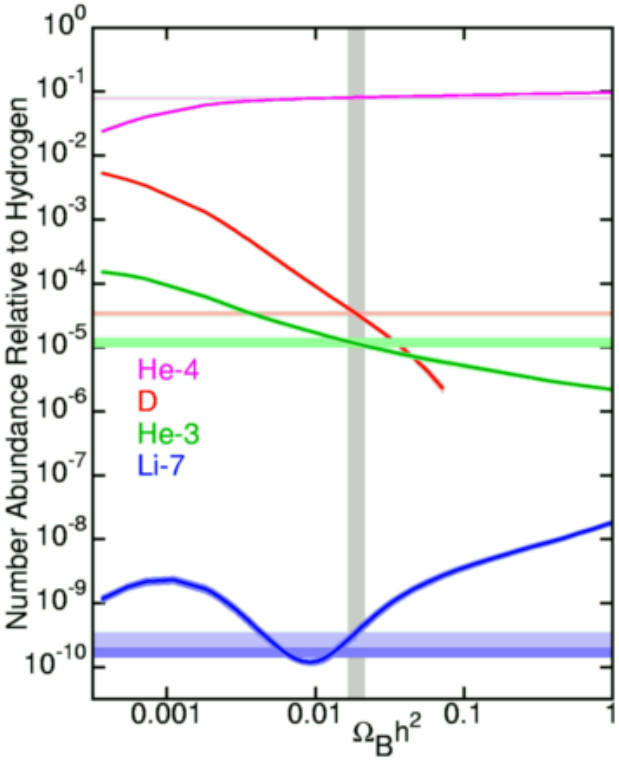


FIGURE 29.13 The reaction network that is responsible for Big Bang nucleosynthesis. The letter "*d*" stands for deuterium, and "*t*" stands for tritium. (Figure adapted from Nollett and Burles, *Phys. Rev. D, 61*, 123505, 2000.)

Basically a small amount of Deuterium, lithium, and Beryllium are made.
(from `http://www.astro.ucla.edu/~wright/BBNS.html`)
In time:



The measured values can be used to infer the amount of baryons in the universe, parameterized as $\Omega_b h^2$. This is the current mass-energy density of the universe in the form of Baryons, and is approximately 0.021. Getting the pristine gas to measure this in is tricky, and is done typically with quasar spectra passing through the IGM.



Here horizontal bands correspond to measured abundance ranges (with uncertainty) and vertical is the consistent value for $\Omega_b h^2$. Note that the lithium measurement is not quite consistent with the others, but is close.

41.2 Physics in the early universe

Start with physics now:

The standard model of physics governs physics today. Typically phrased as "scattering" of particles:
leptons: electron, muon, tau; and a corresponding neutrino for each.
quarks: up/down, charmed/strange, top/bottom
Force carrying bosons: photon, Z,W, gluons, Higgs

basically three generations of particles, 4 forces, and the Higgs field

The proton is a combination of 2 ups and a down quark, along with the gluon field that holds them together (where most of the mass is). Neutron is udd.

Physics is typically phrased in terms of scattering particles, where force carriers can be emitted and absorbed to change trajectories. These are represented schematically with Feynmann diagrams, which represent a scattering matrix calculation.

Example for electron-electron scattering: (feynman diagram with e scattering from e)

but also particles can change type with sufficient energy: example electron-positron annihilation, and pair creation.

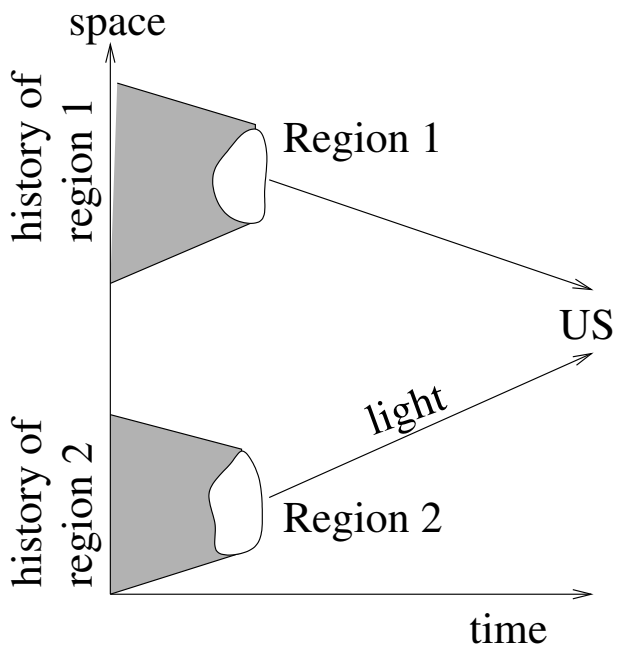
This can happen for ANY particle as long as there is enough energy available for $E = mc^2$.

Thus for high energies, energies above the rest masses of other particles, the forces appear less distinct due to the creation/destruction of intermediary particles.

41.3 The need for inflation – oddities of the Universe

There are several interesting aspects of the large-scale universe that are hard to explain. I will discuss the following two.

1. How can the CMB be so uniform? The horizon at the time the CMB was created, i.e. the the total distance that information could have propagated, is about 2 degrees on our sky. So the CMB temperature is almost uniform across portions of the universe that seem to have no way to be corellated.



2. Why is spacetime flat? Since Omega depends on time, it is somewhat surprising that the total Omega is so close to 1 even at this late time in the universe, when it has had plenty of time to evolve away from 1. Thus it must in fact be very close to 1. What mechanism would set this?

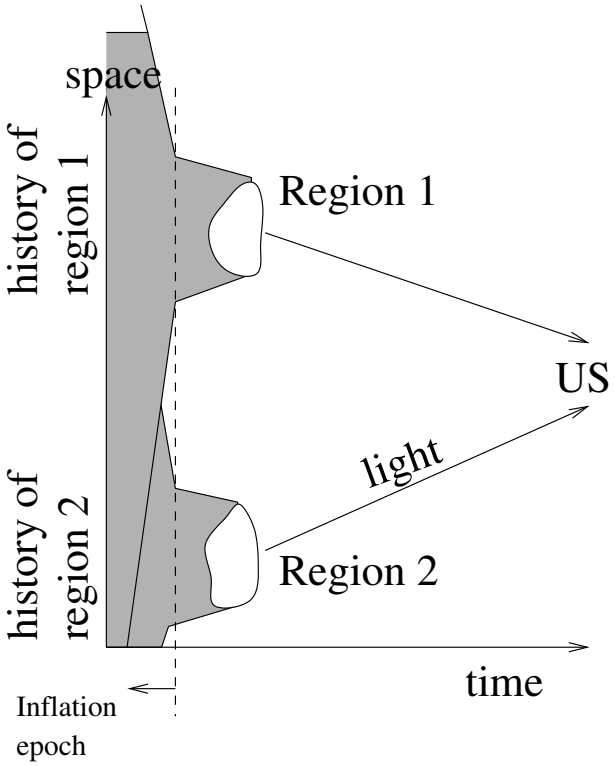
There are two others relating to monopoles and the power spectrum of structure that I won't discuss.

Note that these are observed properties of the universe. But we would like to know why or how they came about.

41.4 Inflation as an explanation

If, in the early universe there were a "false vacuum" that then decayed, there would have been a period of fast expansion (inflation). This false vacuum is analogous to the vacuum energy we now know as the cosmological constant - energy of empty space that is constant density as space expands. Just like the cosmological constant is now moving us toward exponential expansion, this early inflationary phase would have had $R \propto \exp^{(..)t}$. This explains the causal connection between places in the CMB.

This solves the causal structure problem because due to this early inflation, the causal structure of the universe is not what we extrapolated from the current expansion. There was an earlier time when the expansion vastly expanded the light cones, with the expansion of space carrying them beyond the light horizon.



Eventually this false vacuum must decay into our own vacuum (otherwise the exponential expansion would continue forever). Having a higher energy vacuum than we have today is actually quite natural from particle physics, as the natural vacuum you would estimate based on what we know of particles and the uncertainty principle is about 10^{120} times more energy per volume than the density equivalent of the cosmological constant. So what is more puzzling really is why our current vacuum is such low energy, as the natural vacuum energy seems like it should be much higher.

Since we don't know how physics works at the highest energies, it is possible to "make up" a variety of candidate physics for how this epoch of the universe has worked. There are many candidate models already developed.

Also this early epoch of expansion would have stretched the universe into flatness. So that it solves both problems. This is a very nice result of the inflation proposal.